FloodNet: low-cost ultrasonic sensors for real-time measurement of hyperlocal, street-level floods in New York City

C. Mydlarz[1], P. Challagonda[2], B. Steers[1], J. Rucker[1], T. Brain[3], B. Branco[4,5], H. Burnett[6], A. Kaur[1], R. Fischman[7], K. Graziano[8], K. Krueger[2], E. Hénaff[3,1], V. Ignace[4], E. Jozwiak[7], J. Palchuri[1], P. Pierone[4], P. Rothman[9], R. Toledo-Crow[2], and A. I. Silverman[10]

[1]Center for Urban Science and Progress, NYU Tandon School of Engineering, 370 Jay Street, Brooklyn, NY 11201, USA.
[2]Advanced Science Research Center, City University of New York, 85 St Nicholas Terrace, New York, NY 10031, USA.
[3]Integrated Design and Media, NYU Tandon School of Engineering, 370 Jay Street, Brooklyn, NY 11201, USA.
[4]Science and Resilience Institute at Jamaica Bay, Brooklyn College, 2900 Bedford Ave, Brooklyn, NY 11210, USA.
[5]Earth and Environmental Sciences, CUNY Graduate Center, 365 5th Ave, New York, NY 10016, USA.
[6]Department of Technology, Culture, and Society, NYU Tandon School of Engineering, 2 Metrotech Center, Brooklyn NY 11201, USA.
[7]New York City Mayor's Office of Climate and Environmental Justice, New York, NY, USA.
[8]New York Sea Grant, Cornell University, 112 Rice Hall, Ithaca, NY, USA.
[9]New York City Office of Technology and Innovation, New York, NY, USA.
[10]Department of Civil and Urban Engineering, NYU Tandon School of Engineering, 6 Metrotech Center, Brooklyn NY 11201, USA.

Key Points:

- Low-cost, ultrasonic sensors were designed and built to monitor the profiles of hyperlocal, street-level floods
- Sensor hardware, network architecture, and data ingestion, processing, and visualization tools were designed to maximize data usability
- The FloodNet project is installing flood sensors across New York City to collect data for community, city agency, and research stakeholders

Corresponding author: Andrea I. Silverman, as10872@nyu.edu

Abstract

Flooding is one of the most dangerous and costly natural hazards, and has a large impact on infrastructure, mobility, public health, and safety. Despite the disruptive impacts of flooding and predictions of increased flooding due to climate change, municipalities have little quantitative data available on the occurrence, frequency, or extent of urban floods. To address this, we have been designing, building, and deploying low-cost, ultrasonic sensors to systematically collect data on the presence, depth, and duration of street-level floods in New York City (NYC), through a project called FloodNet. FloodNet is a partnership between academic researchers and NYC municipal agencies, working in consultation with residents and community organizations. FloodNet sensors are designed to be compact, rugged, low-cost, and deployed in a manner that is independent of existing power and network infrastructure. These requirements were implemented to allow deployment of a hyperlocal, city-wide sensor network, given that urban floods often occur in a distributed manner due to local variations in land development, population density, sewer design, and topology. Thus far, 87 FloodNet sensors have been installed across the five boroughs of NYC. These sensors have recorded flood events caused by high tides, stormwater runoff, storm surge, and extreme precipitation events, illustrating the feasibility of collecting data that can be used by multiple stakeholders for flood resiliency planning and emergency response.

1 Introduction

Of the many hazards that are expected to increase with climate change, flooding is one of the most dangerous and costly (National Academies of Sciences, Engineering, and Medicine, 2019), and can have a large influence on public health, infrastructure, and mobility in urban areas (Galloway et al., 2018). For example, climate change is projected to increase the frequency and magnitude of precipitation events (Lenderink & Van Meijgaard, 2010; Sillmann et al., 2013), which can overwhelm urban rivers, streams, and drainage systems, resulting in pluvial or fluvial flooding (National Academies of Sciences, Engineering, and Medicine, 2019). Additionally, sea level rise associated with climate change has been observed to increase coastal flooding during high tides. In New York City (NYC), for example, there is an increasing trend in the number of days with cumulative rainfall over 1.75 inches since 1950 (Depietri & McPhearson, 2018), and in community-reported evidence of coastal flooding due to high tides (Science and Resiliency Institute, Jamaica Bay, NYC, 2023). While flooding from extreme events - such as hurricanes, typhoons, and cyclones - can cause catastrophic damage, it has been estimated that floods from smaller, yet more frequent, storms and high tide events have a large cumulative impact as well (Moftakhari et al., 2017; National Academies of Sciences, Engineering, and Medicine, 2019).

Urban flooding is affected by a number of local factors including upstream land use and development; the size of vegetated and impervious areas; population density and human contributions to the sewer system; drainage and sewer design; and physical characteristics, including elevation and watershed topography. The combination of these factors make flood events highly specific to a city block, street corner, or other subsection of a neighborhood (*New York City Stormwater Resiliency Plan*, 2021). The highly-localized and ephemeral nature of urban flooding make it challenging to measure, as monitoring needs to happen at a hyperlocal resolution and in real-time. As such, very little data exist on the frequency and extent of urban surface flooding (Galloway et al., 2018), and there is an unmet need for hyperlocal information on the presence, depth, and duration of street-level floods.

There is a range of stakeholders that require real-time, accurate, and reliable data on flood events, including flood frequency, depth, and profile in flood-prone locations (Silverman et al., 2022). In discussions with city agencies, community members, and researchers,

we have identified needs that include: data for the development and validation of hydrological models that predict flooding (city agencies and researchers; e.g., pluvial flood models produced as part of the *New York City Stormwater Resiliency Plan* (2021)); flood data for resilience and transportation planning, emergency response, issuing flash flood warnings, and tracking flood frequency over time (city agencies); data that can be used for day-to-day decision making and longer-term advocacy when faced with living with flood water (community members); and data that can signal when to collect samples when evaluating flood water quality (researchers)(Silverman et al., 2022). While some data sources exist, such as crowd-sourced flood data and US Geological Survey (USGS) stream gauges, each has limitations that prevent them from meeting all these needs (Helmrich et al., 2021).

Crowd-sourced data, including social media posts and reports from community members or citizen scientists - such as those made to the Community Flood Watch Project in NYC (Science and Resiliency Institute, Jamaica Bay, NYC, 2023) and NYC's 311 service request line (NYC 311, 2023) - are limited in that they require that someone observes the event and that the observer creates a report (Helmrich et al., 2021; Smith & Rodriguez, 2017). Additional challenges with crowd-sourced data are that the data are not always correctly or precisely geotagged, reports rarely include flood depth readings over the duration of an event, and observers are discouraged from collecting data during hazardous conditions that might be present during high tides or extreme rain events. Existing water level sensors, such as USGS stream gauges, overcome some of the challenges above by collecting continuous, real-time data on water depths in the locations in which they are deployed. However, many of these sensors are deployed to monitor water bodies as opposed to flooding at the street-level, and are too expensive and bulky for large scale implementation in urban areas, which precludes their ability to be deployed in a network to capture hyperlocal flooding across a city scale.

Therefore, our aim was to develop real-time flood sensors that overcome the limitations described above. In particular, the goals of this work were to design a sensor that is low-cost (allowing deployment of a large sensor network and increasing accessibility for community science), accurate (enabling detection of flood depths as low as 25 mm), flexible for multiple use cases and installation scenarios (not requiring power or connectivity infrastructure), robust to withstand long-term deployment in the urban environment, and easy to construct with an open-source design. Community engagement and collaboration with NYC municipal agencies are also core components of the FloodNet project that have been instrumental in assisting project implementation, including selection of sensor installation locations, development of tools for meaningful data access and visualization, and collation of growing knowledge about the experience and impacts of flooding within communities most at risk.

In this paper, we describe technical aspects of our project, FloodNet, including the design of FloodNet's ultrasonic-based sensor hardware, network architecture, data ingestion and analysis pipeline, and data visualization tools.

2 Prior Research on Flood Sensing

Previous research has been conducted to design water level sensors using various sensing modalities, each of which has opportunities and limitations for measuring flood depths. Moreover, most previous flood sensing efforts have focused on monitoring water bodies, and have not been deployed in an urban street-level setting where the street is dry most of the time. Here we discuss the strengths and weaknesses of prior flood monitoring strategies, which have provided motivation and justification for the FloodNet sensor design and hardware choices.

Generally, water level sensors can be divided into contact and non-contact designs (Kang et al., 2021). Contact sensors include bubble or float gauges (Kang et al., 2021),

pressure sensors (Garcia et al., 2015), and capacitive or conductivity-based sensors (Chetpattananondh et al., 2014). Pressure sensors placed within drainage infrastructure, for example, have been used to detect street-level flooding through monitoring the surcharging of sewers (Gold et al., 2023). Contact sensors, however, have a number of limitations that make them ill-suited for wide-spread and long-term sensing of street-level floods. For one, contact sensor readings are potentially influenced by the composition, temperature, and turbidity of the water being monitored (Chetpattananondh et al., 2014). Moreover, given that these sensors must contact floodwaters that often have poor water quality, they are subject to fouling and require frequent maintenance, which are challenges for long-term installation. Additionally, the logistics of deploying contact sensors on the sidewalk or street in the urban environment in a manner that allows water contact are challenging and can make them susceptible to damage or vandalism.

Much recent development of water level sensor technologies has focused on non-contact sensing modalities such as camera (Lo et al., 2015; Filonenko et al., 2015; Hiroi & Kawaguchi, 2016), ultrasonic (Mousa et al., 2016; Loftis et al., 2018), and LiDAR-based sensors (Loftis et al., 2018; Paul et al., 2020), given their relative low cost, ease of installation, compatibility with wireless connectivity, and lower maintenance requirements when compared with sensors that contact flood waters.

Visual sensing techniques that utilize camera-based technologies (such as preexisting video surveillance cameras (CCTVs) or traffic cameras) paired with image analysis algorithms and computer vision have been used to generate real-time flood data and predictive flood alerts (Lo et al., 2015; Filonenko et al., 2015; Hiroi & Kawaguchi, 2016; Sabbatini et al., 2021; Moy de Vitry et al., 2019; Jan et al., 2022; Jafari et al., 2021; Arshad et al., 2019). Camera-based sensors can be mounted to existing infrastructure and provide quantitative complexity unmatched by one-dimensional data capture, but they come with a host of challenges. For one, image analysis and computer vision approaches can struggle with low-quality images, and cameras affected by heavy rain, fog, or glare may not be able to provide useful data. These sensing systems also have limited success in processing imagery in low light, limiting their nighttime effectiveness (Lo et al., 2015). Data from camera-based sensors may not include flood depth measurements, and existing cameras are not always positioned in locations that are most at risk for flooding (Helmrich et al., 2021). Moreover, these technologies use a large amount of power and are highly reliant on availability of existing power and communication infrastructure, and therefore less independent and flexible than low-powered counterparts described below (Lo et al., 2015). This also renders them vulnerable to extreme weather events, when they may be needed the most. Finally, the use of cameras raises ethical concerns relating to privacy, surveillance and consent. Collected images would need to be anonymized, risks associated with unintended uses of imagery would need careful consideration, and stringent data security protocols would be required.

Another camera-based approach to flood monitoring is the use of satellite imagery to track flood water inundation. The unparalleled field of view and global coverage offered make satellite-based sensing particularly suitable for large-geographical-scale flood event tracking (Li et al., 2018). However, its shortcomings - including the high cost of long term image acquisition, low temporal and spatial resolution, difficulty in gathering water depth data, and, importantly for pluvial flood events, its inability to obtain imagery through cloud cover (Olthof & Svacina, 2020) - make it impractical for urban flood monitoring.

The increasing affordability of ultrasonic and LiDAR time-of-flight sensors has yielded other novel flood monitoring approaches. Each of these sensor modalities detects water levels by emitting a pulsed sound or light wave and measuring the round trip travel time of the pulse after it reflects from a surface and returns to the sensor; this travel time can be converted to distance based on the speed of the ultrasonic or LiDAR pulse.

LiDAR has become a popular tool to measure the physical world and is used for oceanography, archaeology, topographical modeling, and urban planning. Paul et al. recently used a LiDAR-based flood sensor to measure water levels with relative error around 0.1% and at angles as small as 40 degrees to the planar surface (Paul et al., 2020). However, while LiDAR is becoming more affordable, it is still too expensive to deploy a large number of sensors for spatial coverage at the city-scale, and it can be affected by both the relative calmness and opacity of the water being measured (Paul et al., 2020). Additionally, LiDAR is best used to measure complex surfaces and is difficult to scale in a wireless environment, given the amount of raw data it produces.

Alternatively, several flood monitoring projects have had success in using ultrasonic distance sensors given their low cost, low power consumption, and opportunities for flexible deployment due to their ability to use solar power and transmit data over wireless networks (Mousa et al., 2016; Loftis et al., 2018; Kang et al., 2021; Bartos et al., 2018). Examples of flood monitoring projects that utilize ultrasonic sensors include StormSense, an IoT water-level monitoring system in Hampton Roads, Virginia that employs wireless ultrasonic sensors that are deployed over water bodies and connected to a Long Range Wide Area Network for data transmission (Loftis et al., 2018), Open Storm with installations in Ann Arbor, Michigan and Dallas-Fort Worth, Texas (Bartos et al., 2018), and an ultrasonic sensor project carried out by a research team from King Abdullah University of Science and Technology in Saudi Arabia that utilized ultrasonic sensors installed over street traffic, to monitor and model flood events and traffic flow (Mousa et al., 2016).

Due to the opportunities presented by ultrasonic-based sensors, we employed ultrasonic range finders as the basis for the FloodNet sensor design described below. While prior studies have used ultrasonic sensor networks for riverine (Bartos et al., 2018) and urban flood monitoring (Loftis et al., 2018; Kang et al., 2021; Lo et al., 2015; Hiroi & Kawaguchi, 2016; Mousa et al., 2016), there are limited examples of such systems being deployed at street-level on a city-wide scale, let alone in a metropolitan area as dense and granularly complex as New York City. FloodNet, as a result, offers a novel approach to developing a scalable, reliable, and informative flood-sensing network to aid in urban resilience efforts.

3 FloodNet Sensor Network

The FloodNet project is a collaboration between academic researchers at New York University and the City University of New York with NYC municipal agencies: NYC Mayor's Office of Climate & Environmental Justice, NYC Office of Technology and Innovation, and NYC Department of Environmental Protection. FloodNet has been focused on (1) the design, construction, and deployment of sensors to record street-level floods in NYC; (2) the development of data analysis tools to accurately process flood data; and (3) community engagement, data sharing, and the communication of flood data to various stakeholders in meaningful ways. The FloodNet sensors (Figure 1) are novel in that they were designed to be compact, rugged, low-cost, and independent of existing urban power and network infrastructure, allowing the project to go beyond the limitations of some previously developed ultrasonic flood sensing systems. These specifications make the deployment of a hyperlocal, city-wide sensor network possible.

More specifically, the FloodNet sensor network was designed to meet the following criteria:

1. Sense water depth within an accuracy of $< \pm 25$ mm
2. Collect and transmit data to a central server every $\approx$1 min
3. Operate autonomously in the environment for long periods of time
4. Operate independent of existing power and networking infrastructure
5. Comprise low-cost components for sensor network scalability

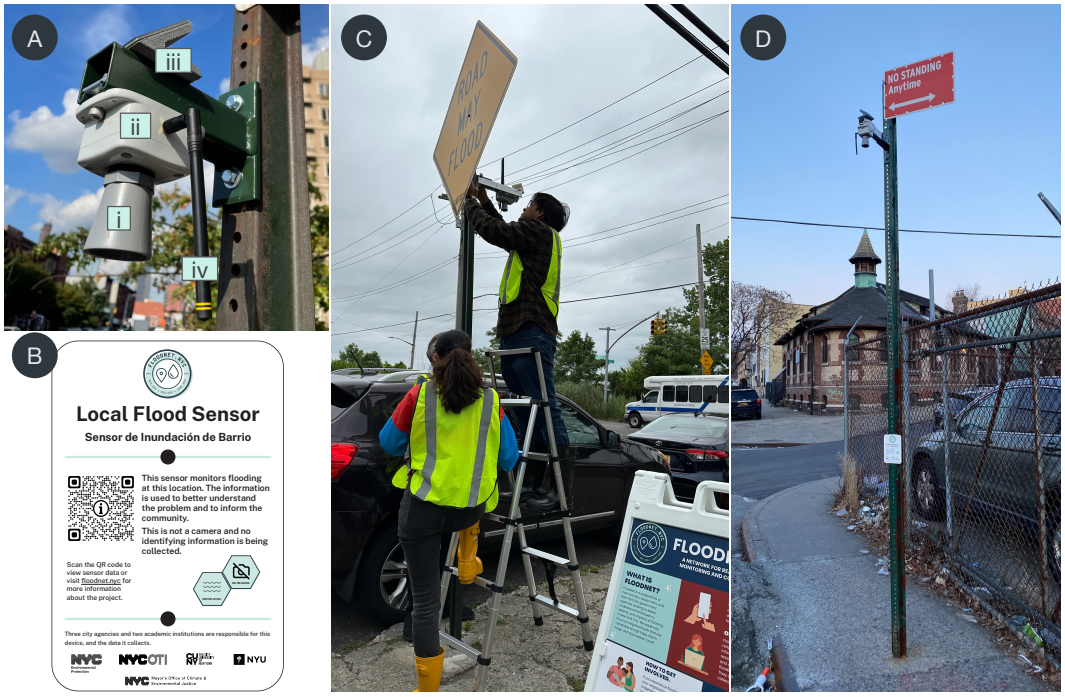

Figure 1. (A) Closeup of the FloodNet sensor, showing the (i) ultrasonic sensor cone, (ii) sensor housing, (iii) solar panel for battery charging, and (iv) antenna for data transmission. (B) Signage that is installed with each sensor. (C) FloodNet engineers installing a sensor in the field. (D) FloodNet sensor and sign installed on a U-channel pole in the Bronx, NY

In addition to sensor hardware, the FloodNet system includes a data ingestion and analysis pipeline, user-facing dashboards that display sensor readings in real-time, and a tool that can send automated flood alerts to subscribing government agencies, community members, or other stakeholders when a flood is detected. All information on the sensor design, including build instructions, quality control practices, and data analysis pipelines, is open-source and provided in a GitHub repository (FloodNet, 2022b). While FloodNet is currently focused on NYC, FloodNet sensor designs could be utilized in other locations.

The following sections explain how FloodNet sensors meet the design criteria listed above, with technical details on the sensor design, operation, deployment scenarios, and networking infrastructure.

3.1 Sensor Core Components

FloodNet sensors (Figure 1) use the Maxbotix MB7389 ultrasonic range-finder to measure the distance to the horizontal surface below the sensor, with a detection range between 30-500 cm, and 1 mm resolution. The sensor transmits 42 kHz ultrasound pulses, which it detects on return to the sensor after reflection from any large surface directly below. Sensor accuracy is a key design criterion for FloodNet, and the MB7389 sensor's datasheet specifies the accuracy as being within 1% of the measured distance, which equates to a range of 30 mm (i.e., ± 15 mm) with a typical mounting height of 3 m, satisfying design criteria 1; data illustrating this accuracy are described in Section 5.2.

Data are collected by the sensor and transmitted every 60 s, satisfying design criterion 2. Data transmission occurs through the use of Long Range Wide Area Network

(LoRaWAN), which is a low power, wide area (LPWA) networking protocol designed to wirelessly connect battery operated devices to the internet. The microcontroller unit (MCU) platform at the core of the sensor is the Heltec HTCC-AB02 development board that uses the Heltec HTCC-AM02 as its MCU/LoRaWAN radio module, which integrates a Semtech SX1262 LoRaWAN radio and PSoC 4000 series ARM Cortex M0 MCU. The MCU is programmed to record range measurements and transmit them to a nearby internet connected gateway using LoRaWAN networking technology, discussed below.

The power consumption of the sensor averages 2 mA at 3.3 V, peaking at $\approx$ 50 mA when collecting measurements, making it suitable for long-term battery-powered field operation. Between measurements, the MCU and radio are set to a low power sleep mode with the power to the ultrasonic sensor switched off using the in-built transistor switch of the Heltec development board; current consumption is $\approx$ 20 μA when in sleep mode. This allows for long-term operation with a modest battery capacity. The sensor uses a lithium ion polymer (LiPo) 400 mAh single cell battery. To satisfy design criterion 3, a UV resistant 0.6 W solar panel is integrated to charge the battery, allowing for indefinite operation. A custom solar panel mount was designed and 3D printed in resilient ABS plastic (Figure 1). The sensor's power consumption is discussed in more detail in Section 3.5.

For ease of assembly and to securely accommodate the development board and battery, a custom breakout printed circuit board (PCB) was designed and fabricated with headers for the development board to be mounted to, and screw terminals for the ultrasonic sensor and solar panel wiring.

A common constraint in the deployment of sensor networks is a reliance on existing infrastructure such as power and networking. By designing a sensor network that makes use of ultra-low power sensing, solar energy harvesting and wireless networking technologies, we have alleviated these requirements, thereby greatly expanding the potential deployment locations for these sensors and meeting design criterion 4. As street flooding occurs in myriad and distributed urban locations, this flexibility is critical for urban flood detection.

The bill of materials for the sensor, excluding mounting hardware costs, is listed in Table 1. At the time of writing (September 2023), the total cost of the sensor parts was <$200 USD, allowing the sensors to meet design criterion 5. The build process for the sensor is included on the project's public Github page (FloodNet, 2022a).

3.2 Sensor Deployment

The FloodNet sensor network was designed to monitor and collect data from various urban flood scenarios, and is indifferent to the cause of the flooding or water level change. For example, the sensors can measure the profiles of floods caused by stormwater runoff, high tides, storm surge, water infrastructure failures, or compound events, and are able to measure water depths over ground-level or water bodies, as long as the sensor is placed above and at a 90° angle to the surface to be monitored. Nonetheless, sensor deployment locations have a few key requirements: (1) mounting infrastructure (e.g., sign post, pole, wall, overhang, etc.) located directly above the surface to be monitored; (2) a LoRaWAN gateway within range for ingestion of the transmitted data; and (3) sunlight exposure to recharge the sensor's battery. This section describes how these requirements inform the sensor installation process. Note that we plan for new versions of the FloodNet sensor to transmit data over a cellular network (described in Section 5.3), which would eliminate the need for a nearby LoRaWAN gateway.

U-channel posts, used to mount street signs, are some of the most abundant pieces of street hardware in NYC (Figure 1). The NYC Department of Transportation (DOT) granted FloodNet permission to mount sensors on these sign posts. The posts are typ-

Table 1. Bill of materials (BOM) for prototype sensor unit including cost in USD (at time of writing)

Item	**Cost (USD)**
Ultrasonic sensor	$99.95
Mounting hardware	$22.96
Heltec AB02 dev. board	$14.40
Antenna	$12.85
Breakout components	$11.44
Solar panel 0.6W	$9.00
LiPo battery - 400 mAh	$6.95
Solar mount	$6.00
Custom breakout board	$0.50
Cable glands	$0.45
Total	**$184.50**

ically >3 m tall, providing a mounting height out of reach of passersby, and include 5 cm spaced mounting holes running vertically up the U-channel. FloodNet sensors must be mounted such that the ultrasonic beam is transmitted perpendicular to the ground surface. This configuration is important because ultrasonic pulses transmitted at an angle greater than 5° from vertical may not be returned to the sensor following reflection.

While plentiful, U-channel sign posts present some sensor installation limitations. First, sign posts are typically located on the sidewalk (i.e., not the roadway), and therefore may not be located directly above the lowest elevation in an area of interest (i.e., the location most likely to flood first). As such, in many locations, sensors are unable to measure flooding before the water depth exceeds the street curb height (usually ≈15 cm in NYC); this offset is noted in the metadata for each sensor, to aid the usability and interpretation of measured flood data. U-channel posts with vegetation beneath pose another challenge, as the shifting vegetation can produce variance in sensor measurements during non-flooded conditions. U-channel posts can also be susceptible to damage (i.e., being pushed off-vertical) or removal; these changes can be detected in real-time sensor readings and must be monitored as part of ongoing operation and maintenance. Despite the limitations of U-channel sign posts, their plenitude in NYC make them great candidates for sensor mounting. In situations where U-channel posts are unusable or unavailable, sensors can also be mounted on other types of street infrastructure, such as street lights or utility poles, using modified mounting hardware.

All installed sensors are accompanied by signage that provides information about the FloodNet project, the type of data collected by the sensor and where to access it, and an explanation that no identifying information is collected by the sensor (Figure 1). Sensor maintenance occurs on a regular and ad hoc basis. Regular visits to visually inspect functioning sensors currently occur yearly; the frequency may change as we expand the network and learn more about specific maintenance needs for the hardware. The health of each sensor is monitored remotely using its battery charge, data packet upload success ratio, frequency of non-returned distance measurements, and ratio of noise to valid distance measurements as metrics. Poorly functioning sensors are visited on an ad hoc basis and replaced with new functioning sensors.

3.3 Data Collection

For each data point collected by the sensor, seven distance measurements are recorded at 150 ms intervals, based on the time recorded between sending and receiving the ultrasound pulse. A median of these seven measurements is taken as a means of internal filtering, to exclude any erroneously small or large range measurements that could be reflections from smaller surfaces, such as the base of a light pole, and ensure that ranging is to the largest surface below the sensor, such as the street, sidewalk, or floodwater.

The sensor also includes an internal temperature sensor that it uses to correct distance measurements for the temperature-dependent speed of the ultrasound pulse through air (i.e., speed of sound increases with an increase in air temperature). Of note is that this temperature compensation can over-correct if, for example, the sensor is exposed to direct sunlight, causing it to absorb heat and measure warmer temperatures than ambient air. In this case, the sensor assumes that the ultrasonic pulse travels faster and for a longer distance than it actually does and will record an exaggerated distance to the ground (this phenomenon is further described in Section 5.1).

After determining the median distance value, this measurement is transmitted to a nearby internet-connected gateway using LoRaWAN networking technology, described in the next section.

3.4 Networking and Data Transmission

Every sensor using LoRaWAN for networking must be within range of an internet-enabled gateway for data transmission and ingestion. FloodNet team members install these LoRaWAN gateways in flood prone areas to receive sensor data payloads; gateways can receive data from any sensors installed within a ≈2 km radius. The internet-connected gateways forward these small (≈28 byte) data payloads to our LoRaWAN network provider, The Things Network, which packages the data and delivers them via the MQ Telemetry Transport (MQTT) lightweight publish/subscribe messaging protocol to the project servers. The gateways themselves can be mounted inside or outside, but their 1 m antennas are mounted outside as high as possible for optimum coverage. Typically, gateways are mounted on the roofs of buildings on mounting points such as railings or existing vertical poles to ensure that the gateway and antenna are securely fastened.

Gateways require a continuous power source, such as an AC power outlet. The gateway typically used (i.e., MikroTik LtAP) measures 17 x 17 x 4 cm, weighs ≈500 g, and consumes around 5 W of power (similar to a small phone charger). The gateway also requires an internet connection to upload sensor payloads. A wired ethernet connection is preferred, but a gateway that incorporates a cellular modem and SIM card with low data rate plan can be used where ethernet connectivity is difficult to obtain. The data throughput of the device is minimal, at around 5 MB/day, which would not exceed a data plan bandwidth of 64 kbps. FloodNet has mounted gateways on the rooftops of local businesses, community organizations, academic institutions, public schools, and apartment buildings.

Each sensor transmits its distance measurement payload (i.e., the median of seven independent distance measurements, as described in Section 3.1) via LoRaWAN to the nearest gateway, which forwards it to The Things Network that then routes the data to our project servers. To allow for an extensible platform for sensor data ingestion, storage, analysis, visualization, and sharing, FloodNet uses a server setup that runs a set of open source tools. More specifically, the FloodNet data pipeline is composed of a combination of Docker containers including: (1) a load-balanced HTTP reverse proxy for the efficient routing of secure web traffic from a set of subdomains to each service running on the server; e.g., https://dataviz.floodnet.nyc (Traefik); (2) a layer for data ingestion,

routing, and processing (NodeRed); (3) a time-series database for data storage (InfluxDB); (4) a dashboard platform for data visualization (Grafana); and (5) a public facing dashboard built on the FieldKit platform at https://dataviz.floodnet.nyc. The data pipeline is illustrated in Figure 2.

Of note, and described further in Section 5.3, is that the logistics of operating a system requiring internet-connected, continuously powered gateways for data transmission are challenging, and has led our team to decide to transition to the use of a cellular network for data transmission in subsequent versions of the FloodNet sensor.

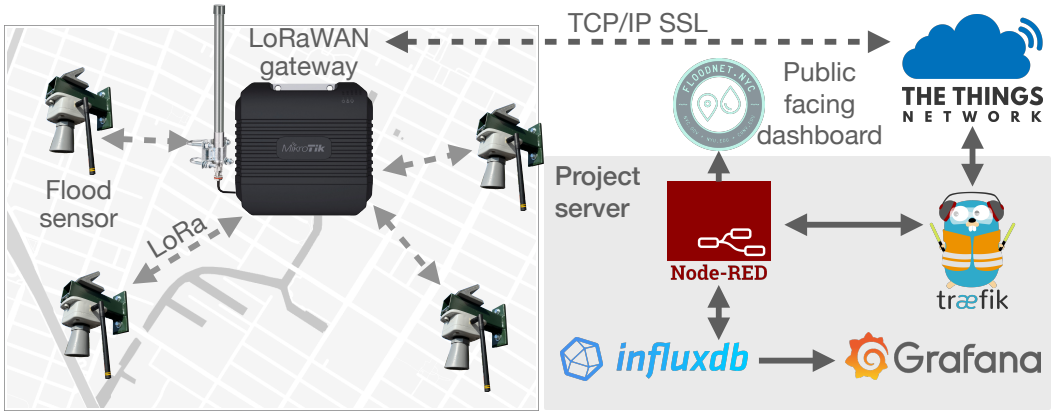

Figure 2. FloodNet data pipeline including deployed sensors, gateway, backend services, and data dashboards.

3.5 Sensor Power Consumption

Figure 3 shows the sensor's current consumption during all possible states of the sensor operating in LoRaWAN Class A, the default LoRaWAN class in which the end device (i.e., the sensor) always initiates the communication. The sensor's data collection and transmission periodicity is ≈1 min; the total duration of each cycle varies by a few milliseconds due to the varying receive windows after a transmission, allowing the opportunity for bi-directional communication. In Figure 3, (a), (b), (d), and (f) are wake periods, and (c), (e), and (g) are sleep periods. The current consumption spike at the beginning of period (a) is when the sensor awakens from a deep sleep state. During every sensing period, seven measurements are taken from the ultrasonic range sensor, with seven corresponding current consumption peaks observed during period (a) in Figure 3. After the measurements are collected, a median is applied to these seven measurements and the result is transmitted in period (b). A short receive window (Rx1; period (d)), follows the transmission period (Tx) and a sleep period (c). Next, a second receive window Rx2, period (f), is typically opened one second after Rx1. The sensor enters a low power state between the receive windows to save power, period (e). The sleep period (g) lasts until the next sensing state (a) occurs.

Table 2 details the average current consumption of different sensor states during a single cycle illustrated in Figure 3. For one such sleep-wake cycle, the wake-time (i.e., the sum of periods (a), (b), (d), and (f)) was ≈2.7 s and sleep-time (i.e., the sum of periods (c), (e), and (g)) was ≈60.5 s (i.e., the duty cycle was approximately 4.5%). The average consumption of one sleep-wake cycle is ≈ 708 μA. This average varies from cycle to cycle but is within the range of 700-750 μA. Based on these observations, with a fully charged 400 mAh battery with no additional power or means of charging, the estimated lifetime of the sensor can be up to 22 days and 17 hours. This headroom in power

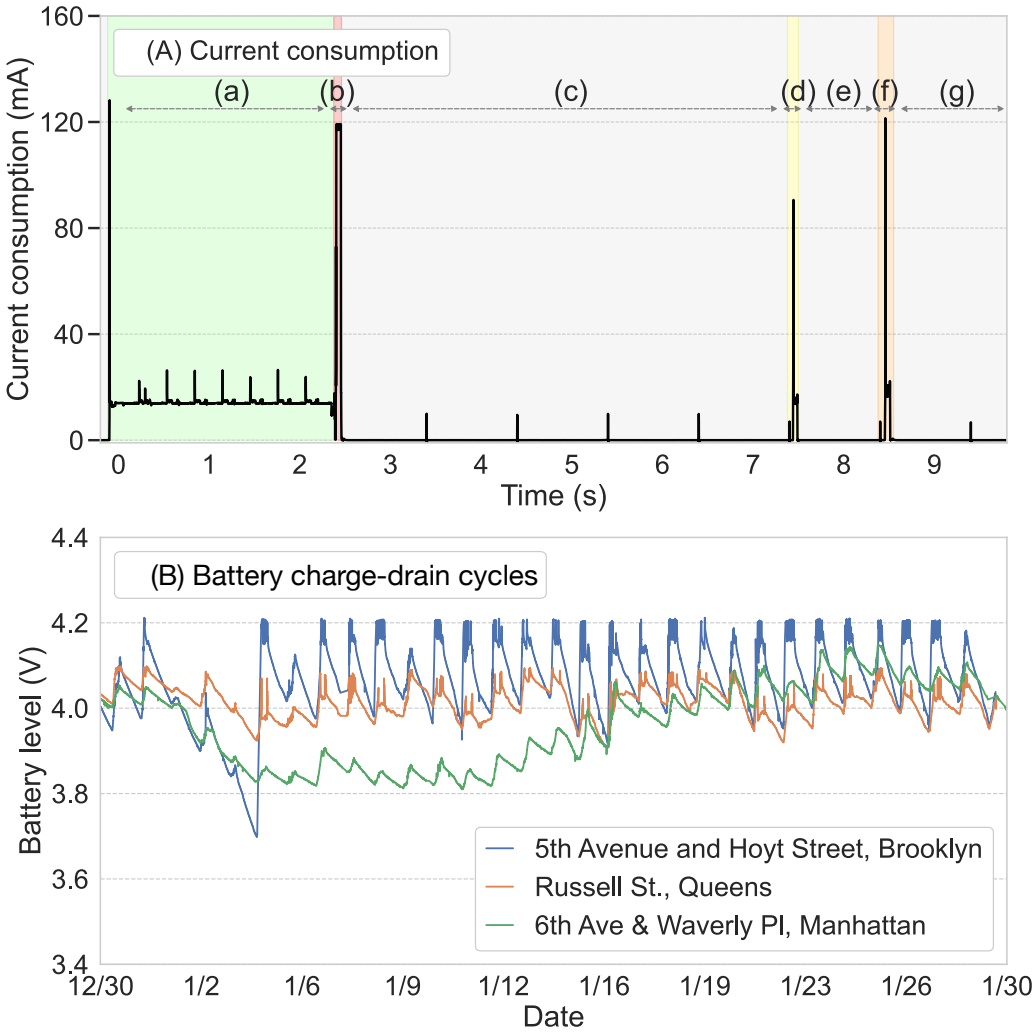

Figure 3. (A) Sensor current consumption in milliamperes (mA) during different operational states. The total cycle is approximately 60 s; only the first 10 s is presented, the remainder of the time the sensor is in sleep mode. (B) Battery charge-drain cycles for select sensors, over the month of January 2021.

consumption can accommodate periods of reduced sunlight, prolonged cover on the solar panel (e.g., snow), or a panel failure.

Battery operation and solar charging of the deployed sensors have been successful thus far. The sensors utilize a 0.6 W solar panel, mounted at an angle of 45°. Figure 3 illustrates the battery levels of three sensors deployed in locations with different cloud coverage, shade, and mounting conditions, which contribute to the differences in their battery trends. The three sensors also displayed differences in average voltage levels due to expected variations in lithium polymer battery condition. The sensor deployed at 5th Avenue and Hoyt Street had the best power harvesting of the three sensors illustrated in Figure 3 due to less shade from buildings and trees. All three sensors had minimal battery charging from January 1 to 3, 2021, due to consistent heavy cloud cover. Nonetheless, the sensor batteries were able to recharge in the following days as the cloud cover reduced. In conditions of full sun exposure, a single day of sun is sufficient to charge

Table 2. Typical duration and average current consumption of different sensor operational states detailed in Figure 3. These values vary from cycle to cycle but on average are below 1 mA average current consumption

Period	Sensor State	Duration (ms)	Avg. current consumption
(a)	Sensing	2500	14.1 mA
(b)	Up-link transmission (TX)	67	115 mA
(c)	Sleep after TX until RX1	5000	22.4 μA
(d)	Receive window 1 (RX1)	52	18.3 mA
(e)	Sleep between RX1 and RX2	1000	18.5 μA
(f)	Receive window 2 (RX2)	59	19.3 mA
(g)	Sleep until next sensing cycle	54500	21.3 μA

the 400 mAh battery completely. However, the sensor deployed at 6th Avenue and Waverly Place had less favorable power harvesting conditions, as observed by the slower upward charging trend after the heavy cloud cover period. Nonetheless, given the headroom provided through the battery size and low power consumption design, the sensors were able to operate through adverse weather conditions and maintain a healthy battery voltage level.

4 Data Processing and Visualization

4.1 Flood Sensor Data Ingestion and Calibration

There are a number of data processing stages implemented to convert distance measurements collected by the sensors to clean profiles of flood depth; this data processing pipeline is illustrated in Figure 4. First, any data point where the ultrasonic pulse did not return to the sensor, which the sensor records as the maximum range value = 5000 mm, is labeled as 'undefined', which is the designation for a missing data point. The remaining distance measurements collected at time t (z_t) are then converted to flood depth (D_t); this calculation occurs in NodeRed. To calculate D_t, the measured distance between the sensor and the ground below must be known under stable and non-flooded conditions. This value is determined through a dynamic calibration procedure that occurs at 5 am daily, in which z_t collected over the previous three nights (between 10 pm and 5 am; $n \approx 1260$) are analyzed to determine the median value ($z_{night-median}$). Daytime measurements are excluded from the calibration because of their potential temperature related variance, as discussed in Section 3.1. If the standard deviation of z exceeds 5 mm (signifying either a flood or large variance related to noisy or unstable conditions), the previous day's $z_{night-median}$ calculation is used. To calculate D_t, the sensor's distance measurement at that time (z_t) is subtracted from $z_{night-median}$ (Equation 1).

$$D_t = z_{night-median} - z_t \tag{1}$$

This dynamic calibration approach allows the project to adapt to seasonal variation in baseline $z_{night-median}$ readings or changes in sensor height, caused, for example, by a shift in U-channel post position.

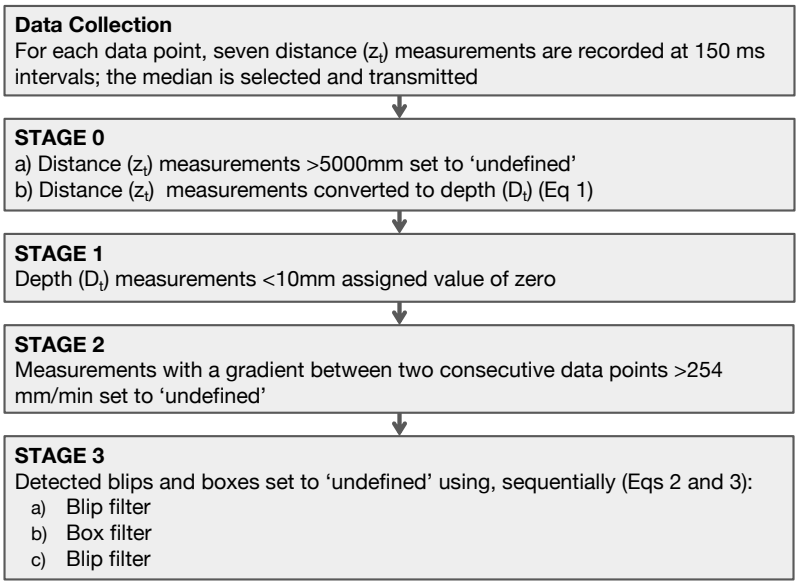

Figure 4. Sequential stages used to collect and filter flood sensor data. Data processing during data collection occurs on the sensor, whereas filter stages 0 to 3 occur on the FloodNet servers in NodeRed.

4.2 Flood Depth Data Processing

After converting distance measurements to depths, the sensor data are processed through a series of filter stages to filter anomalous data related to: (1) low-level measurement noise within our acceptable range of error, (2) the impact of direct sunlight on temperature compensation, and (3) objects located below the sensor, which is a type of noise more likely to be experienced by water level sensors monitoring flooding on streets and sidewalks than those installed over water bodies, making data filtering critically important for the FloodNet project. In the following, a measured 'event' is defined as a series of one or more depth readings greater than 10 mm.

The first filter stage is designed to target the first two anomalies listed above. To correct for low-level measurement noise, all reported depths less than 10 mm are assigned a value of zero. This filter stage also works to filter measurements impacted by solar irradiance incident on the sensor housing, causing the ultrasonic sensor's internal temperature sensor to read greater than ambient temperatures. When this occurs, the ultrasonic sensor's calculation of distance is slightly increased as the speed of sound in air used in the calculation is exaggerated. As such, measured $z_t > z_{night-median}$, resulting in negative calculated depth values (i.e., measurements that dip below ground level). This phenomenon has also been observed by others utilizing ultrasonic sensors for water level monitoring (Mousa et al., 2016). Assigning all depths less than 10 mm a value of zero corrects for this anomaly. Regardless, the resulting measurement error caused by this effect is typically within the range of our target accuracy (i.e., ±25 mm; see Section 5.1).

The second filter stage applies a gradient-based filter to target non-flood depth measurements caused by objects located under the sensor. When people, animals, or large objects – such as trash, bicycles, vehicles, etc. – pass beneath or are placed under a sensor, there is an immediate increase in depth measured from ground surface to the height of the object. These non-flood measurements can be distinguished from floods by their profile, given that floods have a gradual onset (Figure 5); this difference in profile allows the use of a filter that assesses the change in depth values between two time-adjacent

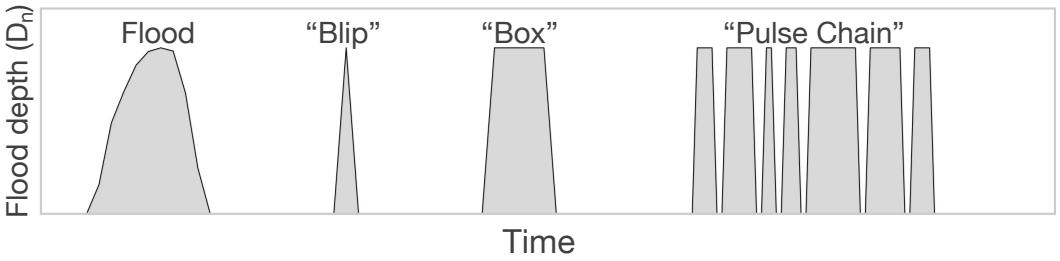

Figure 5. Examples of flood and non-flood event data recorded by FloodNet sensors. An 'event' is defined as a series of one or more depth readings greater than 10 mm.

depth measurements (Δdepth/Δtime) on a rolling basis. The upper threshold for allowable change in depth over time was set at 254 mm per minute, which is a rate six times greater than the fastest rate of flood onset we have measured thus far in NYC (i.e., during Hurricane Ida). Any data point that exceeds the allowable Δdepth/Δtime threshold is labeled as 'undefined' and is not included in the data visualization or alerting platforms. Of note, however, is that this filter stage can miss non-flood events when a sensor is experiencing weak network connectivity and there are longer periods of time between adjacent depth measurements, reducing the chance that the gradient threshold is exceeded.

To address non-flood data points that are not captured by the gradient filter, we characterized three primary types of noise events that appear in the sensor data - (1) blips, (2) boxes, and (3) pulse chains (described below) - and designed a series of real-time filters to target them. This third filter stage applies three filters in the following order: a blip filter, followed by a box filter, followed by another blip filter to clean non-flood data points that the box filter can leave behind. Each filter is described below.

Blips occur when sensor readings increase for a single measurement then return to the original (or similar) value at the next time point. These single, anomalous measurements are likely caused by a person, vehicle, or object temporarily located beneath the sensor in the moment a measurement is collected. For sensors that are consistently uploading data (i.e., collecting and transmitting data every $\approx$ 1 min), these readings can be confidently filtered from the dataset because they are clearly transient. Blip events are characterized by three consecutive readings with values of D_1=D, D_2=D+ΔD, and D_3 =$D \pm 0.1\Delta D$, where D is any depth measurement ≥ 0, and ΔD is a change in depth greater than 2 mm (i.e., D_2-$D_1 > 2$ mm).

To recognize this condition, a blip metric for depth measurement D_2 is calculated following Equation 2. To apply the filter in real-time, any depth measurement that is at least 2 mm greater than the previous measurement is held in memory and not released to the database or visualization platform until the next measurement (D_3) arrives, to assess whether the depth measurement is a blip. If the blip metric is less than 0.1, point D_2 is characterized as a blip and is set to 'undefined', otherwise, the measurement remains unmodified.

Under certain conditions, a low frequency of available data points caused by poor network connectivity could cause a flood event to only be measured in a single data point, mimicking the characteristics of a blip. To reduce the risk of filtering out flood events with few data points, if blips are detected in instances where D_1 and D_3 span more than 6 minutes, the filter will query our weather database for reports of precipitation from any rain gauge located in NYC, and will only filter if there has been no reported pre-

cipitation in the past hour. Incorporation of tide data to check the possibility of coastal floods is still under development.

$$BlipMetric(D_2) = |\frac{D_3 - D_1}{D_2 - D_1}| \tag{2}$$

Boxes occur when sensor readings increase immediately and then remain at a constant (or near constant) value for a sustained period of time. This can happen when a static object (e.g., parked vehicle, garbage, etc.) is placed beneath the sensor. The box filter was designed to detect a sharp increase in depth measurement, followed by one or more measurements that remain within 10% of the initial increase in depth. Concretely, given a sequence of depth values of length N, box events are characterized as measurements where D_1=0, D_2=ΔD, and D_n=$D_2 \pm 0.1\Delta D$ for each sequential measurement n in [3, N], until some measurement D_N deviates from the original increase in height (ΔD) by more than 10%. To recognize this condition, a box metric for measurement D_n is calculated following Equation 3, with values less than 0.1 indicating that D_n is part of a box. In real-time application, if the conditions of D_1 and D_2 are met, then D_2 is held in memory until the following point D_3 is received. If D_3 and subsequent depth measurements (until D_N) meet the stated condition of a box, then D_2 through D_N will be set to 'undefined'. To prevent false characterization of floods as boxes, such as mistakenly filtering out the top of a flood given that standing water may have the appearance of a plateau in the flood profile, we restrict the box filter to only filtering non-flood events that start at a depth of zero (i.e., $D_1 = 0$). In cases when poor network connectivity causes sparse data points, the box filter will follow the logic of the blip filter and will only continue to filter points if there was no precipitation within the past hour of a measured data point.

$$BoxMetric(D_n) = |\frac{D_n - D_2}{D_2}| \tag{3}$$

Pulse chains are characterized as a noisy series of blips and boxes (and blips on top of boxes) that cannot be easily categorized into the other two categories. These anomalous measurements are more challenging to filter and are not addressed by the filters explicitly, but can be captured partially by a combination of both the blip and box filters. Thus, some pulse chains may remain after the application of the three filter stages. Ongoing research is being conducted to optimize data analysis strategies to recognize and filter out this measurement noise.

4.3 Data Visualization and Alerting

Raw data and filtered flood depth measurements are stored in a database that is optimized for large-scale, time-series data ingestion and retrieval. Currently, the FloodNet project utilizes two real-time data visualization platforms that retrieve data from the database: a Grafana-based platform that the FloodNet team uses internally, and a public-facing dashboard openly available on the web (https://dataviz.floodnet.nyc). The design of the public-facing dashboard incorporated feedback collected from city agency personnel and community residents to ensure data visualization is useful and meaningful to both stakeholder groups.

Following feedback from stakeholders, the data visualization platforms enable users to view sensor data alongside other sources of information that contextualize how flood data were collected and provide important touchpoints for understanding the impacts of flooding. For example, users can learn about the infrastructure on which a sensor is mounted, when it was installed, and whether it collects measurements over a sidewalk, road, or waterway. Additionally, tide gauge and precipitation data are provided to contextualize how environmental factors impact street-level flooding. We are conducting on-

going research to learn how to best incorporate qualitative documentation of flood events (such as photos, videos, and written accounts) into our visualization platforms. A core tenet of the FloodNet project is open access to collected data. To facilitate this, a public facing API is under development, with an accompanying data download portal.

Flood sensor data present several opportunities for real-world application. One key application for the data is a real-time flood alerting system, triggered when measured flood depths reach a specified depth threshold. This alerting system has been identified as a critical feature by some of our stakeholders, including both community members and municipal employees involved in emergency management and response. During pilot testing of the alerting system during summer 2021, flood alerts were sent via email or messaging app to registered users when measured flood depths exceeded 7.6 cm. During Hurricane Ida (1 Sept 2021), the FloodNet system alerted NYC Emergency Management of flooding recorded by the sensors installed in the Gowanus neighborhood in Brooklyn 50 minutes before they received other notifications of the event.

5 Results and Discussion

FloodNet sensor installation and data collection are ongoing. To date, a total of 87 FloodNet sensors have been installed in flood-prone areas across the five boroughs of NYC (Figure 6). Sensor installations have been staggered, with some sensors installed for years and others installed more recently. We plan to expand the network across NYC in the coming years. Nonetheless, analysis of sensor data from flood events and static baseline measurements collected thus far provides insight about the functionality of ultrasonic sensors and the particularities of flood monitoring in general.

5.1 Sources of Sensor Measurement Error

The main sources of potential error in the depth measurements made by the sensor are the noise floor of the ultrasonic range finder, exposure to direct sunlight, shifts in ambient temperature, and the mount angle of the sensor. The noise floor of the ultrasonic range finder is a function of the manufacturer stated accuracy of 1% or better (i.e., $\pm$15 mm at a typical mounting height of 3 m) and measurement deviations caused by variations in surface profile beneath the sensor. To assess the extent of the noise floor error, data collected by 46 sensors from February 14 to 21, 2024 between the times of 10 pm and 5 am were assessed for variability in distance readings; other sensors were not included due to non-flood measurements (e.g., blips and boxes described in Section 4.2) that would have confounded analysis. Night time measurements were used to avoid the influence of direct sunlight on measurements. Of 98,013 data points where the water depth measurements should have been 0 mm, the minimum measurement was -14 mm and the maximum measurement was 19 mm, resulting in a error range of 33 mm (Table 3), which is within our accuracy tolerance of $\pm$25 mm.

The largest contributor of measurement error was observed when direct sunlight was incident on the ultrasonic range finder. To isolate the effects of direct sunlight and calculate its impact, data collected by the same sensors and time period as above, but during the hours of 8 am and 5 pm, were assessed. Although a maximum error of 73 mm was found, this occurred only during times of peak sunlight intensity. Further, the mean (-12.2$\pm$10.7 mm) and median (-9 mm) measurements were within our allowable accuracy tolerance, and fewer than 12% of daytime measurements exceeded the allowable error.

As previously mentioned, there is a known effect of ambient weather conditions on ultrasonic distance measurements (Mousa et al., 2016). To evaluate this effect on FloodNet sensor readings, daily mean distances to the ground were calculated for data collected by a FloodNet sensor (located at the intersection of Hoyt Street and 5th Street in Brook-

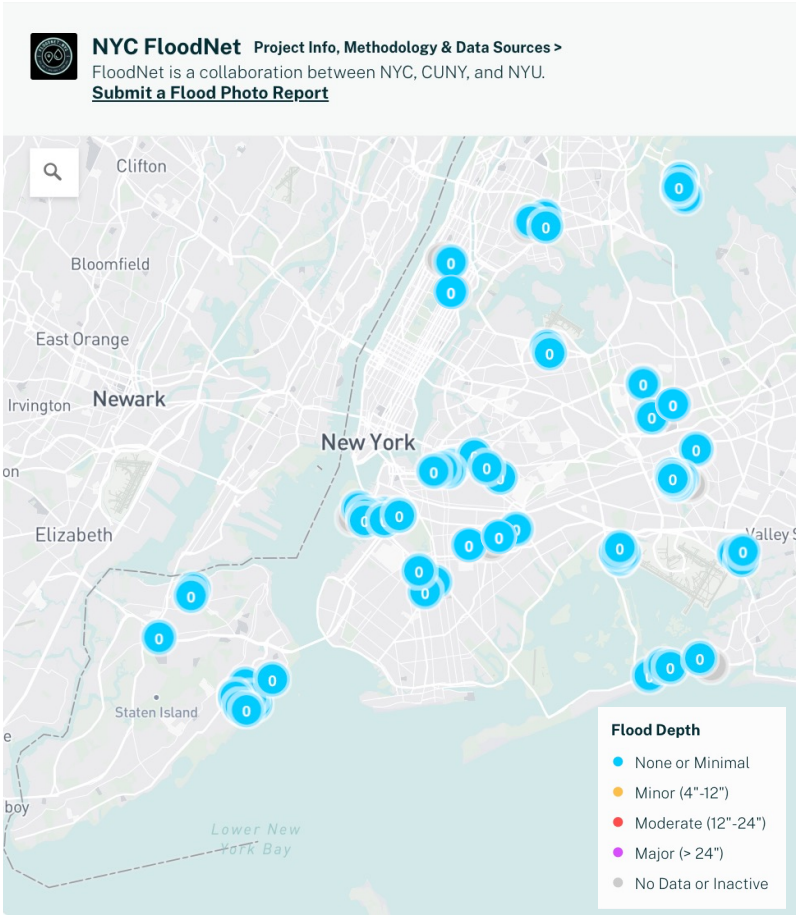

Figure 6. FloodNet's public facing data dashboard. The map view of the dashboard, shown here, includes the locations of the 87 sensors installed across the five boroughs of NYC as of February 2024. Each circle indicates a sensor location and is color coded corresponding to the real-time measured flood depth; the depth value in inches is indicated by the number located in the center of each symbol.

lyn) between 10 pm and 5 am each day, over the course of a year. Night time values were evaluated alone to assess seasonal-dependence of sensor readings due to ambient temperature, without the confounding impact of direct sunlight on the sensor during the day. To understand baseline (i.e., non-flood) conditions, known flood and snow events were excluded from the dataset, as were absolute values greater than two standard deviations from the mean.

Despite no change in the physical height of the sensor as installed, there was a statistically significant shift in the baseline distance measurements made by the sensor over the course of a year (Figure 7), with greatest distances measured in January and smallest measured in August. Colder ambient temperatures in the winter decrease the velocity of ultrasonic pulses, resulting in longer times of return and exaggerated distance measurement if the temperature was not fully accounted for in the distance calculation. Conversely, warmer temperatures in summer would lead to calculation of shorter distances due to faster velocity and shorter time of return of the ultrasonic pulse. As such, it is possible that the internal temperature sensor housed within the casing of the ultrasonic sensor erroneously detected lower than ambient temperatures during the winter and higher

Table 3. Sources of potential error in sensor measurements and summary statistics of their contribution to potential measurement inaccuracy. For all data points evaluated, the water depth measurement should have been 0 mm (i.e., no flood was occurring). Data used for evaluation of error related to ultrasonic noise floor and ambient temperature were limited to measurements collected between 10 pm and 5 am; data used to assess the impact of direct sunlight exposure were limited to those collected between 8 am and 5 pm.

Source of error	Ultrasonic noise floor	Direct sunlight	Ambient temperature	Mount angle
Number of sensors evaluated	46	46	1	10
Sample duration (days)	7	7	365	-
Number sensor measurements assessed (n)	98013	120739	525543	6
Mean depth measurement (mm)	-0.4	-12.2	0.7	5.7
Standard deviation (mm)	2.6	10.7	4.2	4.3
Median depth measurement (mm)	0	-9.0	1.7	5.7
Minimum depth (mm)	-14.0	-73.0	-5.1	0
Maximum depth (mm)	19.0	0	6.1	11.5
Error range = maximum-minimum (mm)	33.0	73.0	11.2	11.5
% of measurements within error tolerance	100%	88.7%	100%	100%

than ambient during the summer. Nonetheless, the magnitude of difference between greatest and smallest distance measurement was only 11.2 mm, which is within the allowable error of the sensor. Additionally, the dynamic calibration procedure described in Section 4.1 accounts and corrects for this temperature-dependent phenomena.

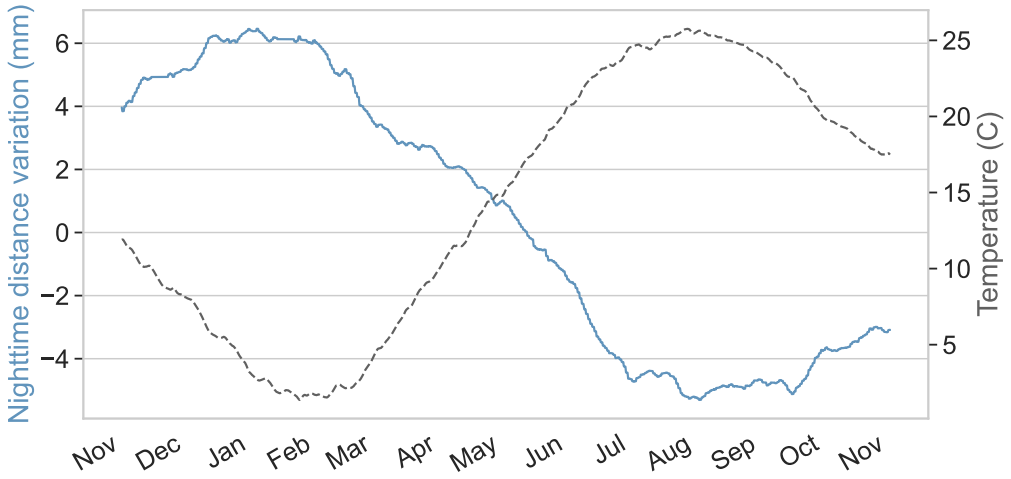

Figure 7. Difference in daily nighttime mean distance measurements (under non-flooded conditions from 10 pm to 5 am) from the annual mean for a flood sensor located at Hoyt Street and 5th Avenue (Brooklyn) over 12 months from 2020 to 2021. Temperature data represent ambient conditions and were sourced from a nearby New York State Mesonet managed weather station at the Brooklyn Navy Yard (ID: bknyrd), which is located ≈ 3 km away from the sensor.

The ideal mounting angle between the ground and the ultrasonic range finder is 90°. If the mounting angle deviates from 90° by an angle of θ then measured flood depths will be exaggerated, given that the pathlength of the ultrasonic pulse would be along the hypotenuse, with a distance equal to the the height of the sensor multiplied by $cos^{-1}(\theta)$. Ten different ultrasonic range finders were evaluated in the laboratory to assess the maximum θ at which ultrasonic pulses reflected off the ground were still able to return to

the sensor, which was found to be 5°. At angles greater than this, the majority of reflected ultrasonic pulses are not received, resulting in an invalid measurement. When mounting sensors, a two-axis spirit level is used to ensure the ultrasonic range finder is at 90° to the surface beneath, reducing the likelihood of this error. If, however, the mounting angle deviates from 90° by the 5° maximum θ, then the maximum error in flood depth measurement would be 11.5 mm at a typical sensor mount height of 3 m (Table 3), which is within our accuracy tolerance of ±25 mm.

When assessing all data points collected by the sensors between February 14 and 21, 2024, significant majority (95.8%) of the sensor measurements (n= 329,124) were within our allowable range of error.

5.2 Validation of Sensor Measurements

Sensor measurements were validated in three ways: (1) in-laboratory testing before deployment; (2) comparison with data from tide gauges operated by the National Oceanographic and Atmospheric Administration (NOAA); and (3) comparison with manually collected flood depth measurements during a street-level, tidal flood event.

The flood sensor assembly process follows a detailed quality assurance (QA) procedure (FloodNet, 2022a). After the assembly, and before deployment, sensors undergo a data validation test to ensure accurate measurements. During this in-laboratory quality control (QC) testing, each sensor is tested for accuracy at known mounting heights that are similar to actual deployment scenarios. These heights are measured with a standard scale, and the mounts are aligned at a perpendicular angle to the ground surface using a spirit level. The duration of this test is one hour, and a median of seven measurements are collected every minute. The observed sensor measurements at all mounting heights must be within $h_i \pm n_{allowed}$ to pass the QC test, where h_i is the known mounting height and $n_{allowed}$ is the acceptable noise floor, which is the ultrasonic range finder's manufacturer stated 1% of the measured distance. All of the FloodNet sensors constructed thus far have met this QC criteria.

To test the ability of FloodNet sensors to detect water level changes under real-world conditions, sensors were mounted over tidally-influenced water bodies (Gowanus Canal in Brooklyn and the Jamaica Bay estuary in Queens). Changes in water level due to the daily tidal cycle measured by the FloodNet sensors were compared with measurements collected by nearby tide gauges operated by NOAA, and found to be similar with similar periodicity (Figure 8).

Finally, the FloodNet sensors demonstrated accuracy in measuring water depths during an actual street-level flood event. A validation experiment was conducted during a high tide flood event in Hamilton Beach, Queens, on July 23, 2021. A standard measuring scale was installed on the same pole used to mount a FloodNet sensor, and flood depth measurements were manually recorded at regular time intervals during the flood, concurrent with flood sensor measurements. The FloodNet sensor readings were in agreement with the manual flood depth measurements, with an accuracy within a few millimeters (Figure 8).

5.3 Network and Data Acquisition Performance

FloodNet sensors collect distance data every minute and transmit these data payloads to our servers via LoRaWAN. However, one factor affecting data transmission (also referred to as sensor uptime) is the signal strength between the sensor transmitting the data and the gateway receiving it. Low signal strength caused, for example, by a gateway being out of range, non-ideal topography between the sensor and gateway, or objects obstructing signal transmission (e.g., buildings, trees), can result in the loss of data payloads. The sensor design described herein does not store data on the device - a de-

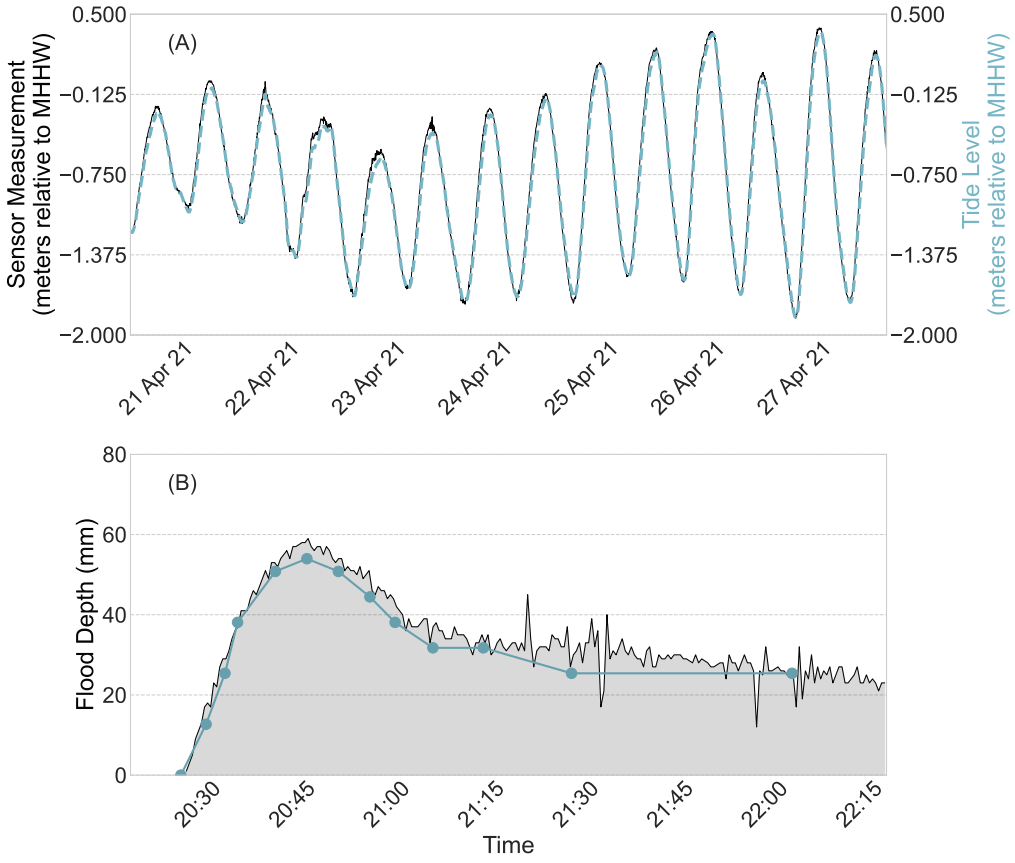

Figure 8. (A) Comparison of water level data collected by a FloodNet sensor installed above the Gowanus Canal in Brooklyn (black line), which is tidally influenced, with the tide level measured by a NOAA tide gauge located at the Battery, NY (NOAA station ID: 8518750; dashed blue line). The FloodNet sensor measurements are very similar to those from the NOAA tide gauge with a slight temporal delay due to being installed in different locations in New York Harbor. (B) FloodNet sensor measurements (black line) collected during a high tide flood in Hamilton Beach, Queens on July 23, 2021, compared with measurements collected intermittently by manually reading the depth value off of a standard ruler (blue data points and line).

cision made to keep the sensor small and low-cost - and data can be lost if not received by the gateway.

Much effort has been made in installing sensors and gateways in locations that ensure good signal strength. Nonetheless, FloodNet sensors deployed thus far have had transmission efficiencies of less than 100% (Figure 9), which is typical for LoRaWAN networks. For example, in analyzing the transmission efficiencies of 21 sensors that were installed and operating during the full six month period between October 2022 and April 2023, eight had transmission efficiencies that were greater than 75%, seven were between 50 and 75%, and the remaining six were between 35 and 46%. Given that the sensors collect and transmit data every minute, a transmission efficiency of greater than 50% is relatively good, and means that, on average, data packets sent by a sensor were received at a frequency of at least every other minute.

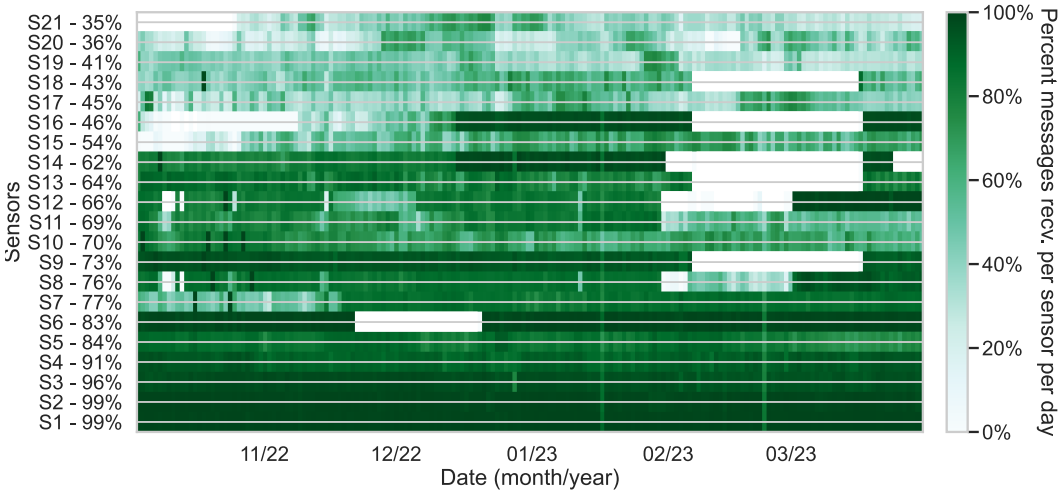

Figure 9. Data transmission efficiencies of a subset of FloodNet sensors, showing percent of payload uploads per day, where 100% equals 60 successful uploads/hour over 24 hours. The percentage provided on the left hand side of the figure is the average daily transmission efficiency across all days in the period between October 2022 to April 2023. The 21 sensors included here are a subset of the total number of sensors installed thus far, and were selected because they were installed and operated during the full duration of the six month period from October 2022 to April 2023.

Poor sensor uptime can be caused by a variety of factors. First, if a sensor is observed to have consistently low upload rates, it is often caused by poor or inconsistent gateway coverage. This can be a function of pure distance between the sensor and the gateway - as seen for sensor 19 (Figure 9), which was located 1.8 km from the nearest gateway - or can be due to seasonal or environmental effects that could impede the signal, such as increased tree foliage in summer or high voltage electrical or heavy metal infrastructure located nearby. Sensor 18, for example, was mounted next to the underpass of a steel railway bridge, which impeded signal transmission.

Second, downtime can manifest in an otherwise well-performing sensor as a sudden and complete loss of connectivity for a period of time, which can be caused by the temporary malfunction of a gateway or sensor. A gateway malfunction can cause an outage across multiple sensors that rely on that gateway for connectivity. For example, in February 2023, two gateway outages occurred, the first affecting sensors 8, 11, 12, and 14, and the second affecting sensors 9, 13, 16, and 18. During the first gateway outage, sensor 8 initially had a total outage, but was able to regain connectivity, albeit weaker in signal strength, with another gateway shortly after, and sensor 11 saw a drop in connectivity, but never completely lost connectivity due to proximity to another, further gateway. Improved gateway connectivity was gained through deployment of two new gateways in March 2023, first reconnecting sensors 8 and 12, and later reconnecting sensors 9, 11, 13, 14, 16, and 18. Additionally, vertical bands can be observed across all sensors in January and February 2023 related to short term system-wide outages caused by server infrastructure maintenance, when downtime was less than a day.

To reduce the loss of data during transmission over LoRaWAN, one could use a stronger antenna on the sensor to improve signal strength and increase the packet success rate. For the next iteration of the FloodNet sensor, however, we have decided to pivot to the use of a cellular network for data transmission instead of LoRaWAN. In addition to im-

proving the packet success rate, switching to a cellular network will alleviate other challenges associated with using LoRaWAN, including the need to identify and gain permission for gateway installation locations that have access to power, and gateway downtime due to power outages or equipment malfunctions. We also plan to implement onboard flash memory to store data on the sensor. In the event of a network disruption, the sensor will re-transmit stored data once the disruption is resolved, minimizing the loss of data packets.

5.4 Data Filter Performance

Between October 2020 and May 2023, a total of 6,641 events were recorded by the sensors (Table 4); each event is defined as a series of one or more depth readings greater than 10 mm. Of these, 360 were manually identified as flood events, 7 as snow events (i.e., snow accumulation during winter storms), and the remaining 6,274 as non-flood events. The recording of non-flood events (described in Section 4.2) resulted from the nature of the sensors being located over active streets and sidewalks in NYC, and included blips, boxes, pulse chains, and complex noise, the latter being a catch-all category for events that did not fit into the other three categories. Most of the non-flood events (74%) were identified as blips.

As described in Section 4.2, we developed and implemented a series of automated data processing stages to filter non-flood events from the depth data, thereby removing false positive events from the data visualization and alerting platforms. As shown in Table 4, the use of all three filter stages described above resulted in an overall reduction in reported non-flood events of 76%, whereas only 19% of the non-flood events were captured after the gradient filter stage (i.e., using only the first two filter stages). Blip events were the most common non-flood events and the easiest to filter, with 92% of blips correctly classified. Boxes and pulse-chains occurred less frequently, but were more challenging to capture, as shown by the correct identification of 65% of boxes and 32% of pulse chains. All flood and snow events remained unfiltered, indicating that the filters did not result in false negative results.

There are some challenges to automatically distinguishing flood events from non-flood events, including having sparse data points at times for some sensors. More specifically, when data are not received from the sensor regularly (e.g., every minute), it is difficult to distinguish a possible flood from a non-flood event. For example, with frequent data points, large gradients of flood depth per time can be readily identified as unlikely coming from a flood event. However if the data points are separated by long or irregular time intervals, a large increase in depth that is physically unlikely at a 1 minute resolution could be possible at a 5 or 10 minute resolution because the large instantaneous gradient is now spread over a longer period of time. In an extreme case, if only one measurement is received over the duration of a flood, it could appear as just a blip, which is why blip and box filters are both limited by the length of the time window they're allowed to operate on; this challenge will be alleviated with engineering improvements to increase data transmission.

Additionally, there are events that do not match the aforementioned non-flood or flood event patterns. For example, some non-flood events have two points that meet neither the characteristics of a box nor a blip, some boxes don't start at zero and are not identified because the box filter requires the first point to have a depth of zero, and sometimes the tops of boxes have a large variance or have boxes or noise on top of boxes, both of which would bypass the box filter.

Our goal is to release data to stakeholders as close to real-time as possible, ideally within 5-10 minutes of the start of a flood. Therefore, frequent data collection ($\approx$ 1 minute interval) is needed for the filters to be able to rapidly recognize flood events. Additionally, we are conducting ongoing research to improve the data filters' ability to distinguish

Table 4. Counts of flood, snow, and non-flood events that were recorded by FloodNet sensors between October 2020 and May 2023, in which an event is defined as a series of one or more depth readings greater than 10 mm. Total event numbers presented for the unfiltered data were manually counted. The numbers of non-flood events identified after the gradient filter stage and after application of all three filter stages are presented; the percentages provided are calculated as the number of events identified by the filters divided by the number of events identified in the unfiltered data

	Unfiltered Data	**After Gradient Filter (Stage 2)**		**After All Filter Stages (Stage 3)**	
Event Type	**Number Events Identified Manually**	**Number Non-flood Events Identified**	**Percent of Total from Unfiltered Data**	**Number Non-flood Events Identified**	**Percent of Total from Unfiltered Data**
Flood	360	0	0%	0	0%
Snow	7	0	0%	0	0%
Blip	4641	990	21%	4298	93%
Box	812	235	29%	528	65%
Pulse Chain	597	58	10%	193	32%
Complex Noise	224	4	2%	19	8%
Total Non-Flood Events	**6634**	**1287**	**19%**	**5038**	**76%**

floods from non-flood events. One avenue being explored is training machine learning models to identify the patterns of flood event data, with a goal of reducing the number of false positive floods reported for non-flood events, which is important for a reliable public flood alert system. Additionally this will help in post-hoc flood labeling and boundary regression allowing computation of flood event statistics.

5.5 Community Engagement with Flood Sensor Data

Community engagement is an important component of the FloodNet project, given our goal to meaningfully collect and share data that can contribute to flood risk mitigation and community flood resilience. FloodNet's ongoing outreach and collaboration with local community members has assisted with project design and implementation. In addition, though community engagement activities we have learned that flood sensor data have the potential to support community residents' decision making when faced with chronic floods; for example, a sensor placed on an important roadway that floods semi-monthly has helped residents know if it is passable in real time. In addition, the sensor data offer opportunities to assist in increasing residents' understanding of how flooding impacts their neighborhoods. For example, community members who participated in a focus group on FloodNet data and public engagement noted that sharing summary flood data with their neighbors could be a good way to communicate whether they live in a flood zone, in what time of year floods are most likely to occur, and what resources are available to prepare for flood events. Through presentations, community meetings, educational workshops, community walk-throughs, and community-collected feedback on flood sensor placement, we have also heard a desire among community members to use flood sensor data to validate community experiences of flooding in the eyes of elected officials and other people in power. This can serve to support action plans and advocacy that connect flooding to other relevant community issues in flood-prone areas. Given the various potential uses of flood data, a goal for future research is to assess how community residents actually use the data when it is available.

5.6 Flood Sensor Data Examples: Hurricane Ida and Tidal Flooding

Even with ongoing upgrades to the sensor design, network performance, and data analysis pipeline, the FloodNet sensors deployed thus far have been able to collect a rich dataset, capturing the profiles of pluvial and tidal floods in NYC neighborhoods. The

arrival of the remnants of Hurricane Ida in NYC on September 1, 2021 was a landmark event, with record rainfall (up to 79 mm/h recorded in Central Park) resulting in unprecedented flooding across NYC. All three FloodNet sensors deployed at the time in the Gowanus neighborhood in Brooklyn measured flood profiles during the storm, with the most extreme flooding occurring at the intersection of Carroll Street and 4th Avenue (Figure 10). A peak flood depth of 890 mm above the sidewalk was recorded (water depths in the roadway were deeper), as well as a rapid rate of onset: up to 91 mm/min for the first 5 minutes of the flood. See Silverman et al. (Silverman et al., 2022) for a description of flood data collected by the same sensor during Tropical Storm Henri on August 21, 2021.

Flood events have been measured more frequently by sensors installed in coastal areas susceptible to regular tidal flooding. The northernmost block of Beach 84th Street on the Rockaway peninsula in Queens is one such example (Figure 10). In the 17 months between December 2021, when the sensor was installed, and April 2023, the sensor located on this block measured 121 distinct flood events, each between 29 and 560 mm in depth, during relatively high, high-tide events, which occur on a semi-monthly basis, typically coinciding with the full moon or new moon. For example, in September 2022, the FloodNet sensor recorded 13 distinct flood events on 7 consecutive days, following the approximately 12-hour period of the tidal cycle. Other FloodNet sensors located in additional coastal areas have measured a similar degree of flooding.

6 Conclusions and Future Work

The design of the FloodNet sensor allows accurate measurement of street-level floods as shallow as 10 mm, at a frequency of one measurement per minute under optimal conditions. The associated sensor network transmits measured data to our servers and the FloodNet data analysis pipeline, which enables flood data visualization and provision of alerts to users in near real-time.

Given the success of FloodNet sensors in monitoring floods thus far, and the potential utility of flood sensor data (Silverman et al., 2022), we plan to expand the sensor network in NYC. To accomplish the expansion, ongoing research is being conducted to iterate the sensor design to make it more manufacturable for scale-up, to improve data transmission rates through the use of a cellular network, and to update the data analysis pipeline to improve the detection and logging of flood events. All sensor design files are located on the FloodNet Github page (FloodNet, 2022b), which is updated as new designs are implemented.

A variety of stakeholders, including NYC residents, are being consulted as we iterate the design of the FloodNet web-based data dashboard and other data sharing tools, such as printable data sharing reports, to improve the meaningful sharing and communication of collected flood data. Additionally, given the thousands of locations that are at risk of flooding in NYC (*New York City Stormwater Resiliency Plan*, 2021), it is unlikely that a flood sensor network will be distributed enough to monitor every one. As such - and given the relevance of the data to municipalities for uses such as emergency management, urban planning, decision making for capital improvement projects and other resource allocation - an additional line of ongoing research is to create a risk and equity based framework to help prioritize flood sensor deployment locations in a systematic and equitable way. In addition, we have ongoing collaborations with hydrologists to assess the ways in which flood sensor data can be incorporated to hydrologic and hydraulic models to improve flood prediction and risk assessment.

In conclusion, FloodNet sensors were specifically designed to overcome challenges related to measuring floods in a distributed manner across a complex urban environment, and have a demonstrated ability to collect street-level flood data, regardless of the flood

typology (e.g., pluvial, fluvial, tidal, storm surge, infrastructure-related). While the FloodNet project is based in NYC, the sensor design is flexible for other locations and contexts, with an open-source design available for others who would like to build and deploy flood sensors in their own communities.

Open Research

All sensor hardware designs and associated software are available on our publicly accessible Github organization (FloodNet, 2023b). Data used in the analyses presented herein can be found on Zenodo (FloodNet, 2023a). Time series visualizations of all FloodNet sensor data can be found on our public facing data dashboard (FloodNet, 2023c).

Acknowledgments
Funding for this work was provided by the New York City Department of Environmental Protection, the Connected Cities for Smart Mobility towards Accessible and Resilient Transportation (C2SMART) University Center (U.S. Department of Transportation award number 69A3351747124), the Alfred P. Sloan Foundation, the New York State Empire State Development's New York Smart Cities Innovation Partnership, and seed grants from the NYU Marron Institute of Urban Management and the CUNY Office of Research. We thank colleagues at the City of New York for feedback and assistance in the development and installation of the sensor network, and our community partners for valuable insight regarding their experiences with flooding. We also thank student researchers at the high school, undergraduate, and graduate levels for assisting in aspects of this work. There are no conflicts of interest to declare.

References

Arshad, B., Ogie, R., Barthelemy, J., Pradhan, B., Verstaevel, N., & Perez, P. (2019). Computer vision and iot-based sensors in flood monitoring and mapping: A systematic review. *Sensors*, *19*(22), 5012. Retrieved from `https://www.mdpi.com/1424-8220/19/22/5012` doi: 10.3390/s19225012

Bartos, M., Wong, B., & Kerkez, B. (2018). Open storm: a complete framework for sensing and control of urban watersheds. *Environ. Sci.: Water Res. Technol.*, *4*(3), 346-358.

Chetpattananondh, K., Tapoanoi, T., Phukpattaranont, P., & Jindapetch, N. (2014). A self-calibration water level measurement using an interdigital capacitive sensor. *Sensors and Actuators A: Physical*, *209*, 175–182.

Depietri, Y., & McPhearson, T. (2018). Changing urban risk: 140 years of climatic hazards in new york city. *Climatic change*, *148*(1), 95–108.

Filonenko, A., Hernández, D. C., Seo, D., Jo, K.-H., et al. (2015). Real-time flood detection for video surveillance. In *Iecon 2015-41st annual conference of the ieee industrial electronics society* (pp. 004082–004085).

FloodNet. (2022a). *FloodNet - Sensor Assembly with Quality Assurance Procedure.* `floodnet-nyc.github.io/quality-management/sensor-assembly-qap`. ([Online; accessed 10-Mar-2022])

FloodNet. (2022b). *FloodNet - Sensor repository.* `github.com/floodnet-nyc/flood-sensor`. ([Online; accessed 02-Jun-2022])

FloodNet. (2023a). *FloodNet flood sensor data - October 2020 to October 2023* [dataset]. `zenodo.org/records/10211443`. ([Online; accessed 27-Nov-2023]) doi: 10.5281/zenodo.10211443

FloodNet. (2023b). *FloodNet - Github organization* [collection]. `github.com/floodnet-nyc`. ([Online; accessed 02-Jun-2023])

FloodNet. (2023c). *FloodNet - Public facing dashboard* [website]. `dataviz.floodnet.nyc`. ([Online; accessed 27-Nov-2023])

Galloway, G. E., Reilly, A., Ryoo, S., Riley, A., Haslam, M., Brody, S., ... Parker, S. (2018). *The Growing Threat of Urban Flooding: A National Challenge.* Web. Retrieved from `https://cdr.umd.edu/urban-flooding-report`

Garcia, F. C. C., Retamar, A. E., & Javier, J. C. (2015). A real time urban flood monitoring system for metro manila. In *Tencon 2015-2015 ieee region 10 conference* (pp. 1–5).

Gold, A., Anarde, K., Grimley, L., Neve, R., Srebnik, E. R., Thelen, T., ... Hino, M. (2023). Data from the drain: A sensor framework that captures multiple drivers of chronic coastal floods. *Water Resources Research*, *59*(4), e2022WR032392.

Helmrich, A. M., Ruddell, B. L., Bessem, K., Chester, M. V., Chohan, N., Doerry, E., ... Zahura, F. T. (2021). Opportunities for crowdsourcing in urban flood monitoring. *Environmental Modelling & Software*, *143*, 105124. doi: https://doi.org/10.1016/j.envsoft.2021.105124

Hiroi, K., & Kawaguchi, N. (2016). Floodeye: Real-time flash flood prediction system for urban complex water flow. In *2016 ieee sensors* (pp. 1–3).

Jafari, N. H., Li, X., Chen, Q., Le, C.-Y., Betzer, L. P., & Liang, Y. (2021). Real-time water level monitoring using live cameras and computer vision techniques. *Computers & Geosciences*, *147*, 104642. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0098300420306191` doi: https://doi.org/10.1016/j.cageo.2020.104642

Jan, O. R., Jo, H. S., Jo, R. S., & Kua, J. (2022). Real-time flood monitoring with computer vision through edge computing-based internet of things. *Future Internet*, *14*(11), 308. Retrieved from `https://www.mdpi.com/1999-5903/14/11/308` doi: 10.3390/fi14110308

Kang, S., David, D. S. K., Yang, M., Yu, Y. C., & Ham, S. (2021). Energy-efficient ultrasonic water level detection system with dual-target monitoring. *Sensors*, *21*(6), 2241.

Lenderink, G., & Van Meijgaard, E. (2010). Linking increases in hourly precipitation extremes to atmospheric temperature and moisture changes. *Environmental Research Letters*, *5*(2), 025208.

Li, X., Lu, J., Song, X., Sun, Y., Li, L., Lei, T., & Qu, W. (2018). Application of the gf satellite data in flood disaster monitoring. In *Igarss 2018 - 2018 ieee international geoscience and remote sensing symposium* (p. 7293-7296). doi: 10.1109/IGARSS.2018.8519548

Lo, S.-W., Wu, J.-H., Lin, F.-P., & Hsu, C.-H. (2015). Visual sensing for urban flood monitoring. *Sensors*, *15*(8), 20006–20029. doi: 10.3390/s150820006

Loftis, D., Forrest, D., Katragadda, S., Spencer, K., Organski, T., Nguyen, C., & Rhee, S. (2018). Stormsense: A new integrated network of iot water level sensors in the smart cities of hampton roads, va. *Marine Technology Society Journal*, *52*, 56–67.

Moftakhari, H. R., AghaKouchak, A., Sanders, B. F., & Matthew, R. A. (2017). Cumulative hazard: The case of nuisance flooding. *Earth's Future*, *5*(2), 214-223. doi: https://doi.org/10.1002/2016EF000494

Mousa, M., Zhang, X., & Claudel, C. (2016). Flash flood detection in urban cities using ultrasonic and infrared sensors. *IEEE Sensors Journal*, *16*(19), 7204–7216.

Moy de Vitry, M., Kramer, S., Wegner, J. D., & Leitão, J. P. (2019). Scalable flood level trend monitoring with surveillance cameras using a deep convolutional neural network. *Hydrology and Earth System Sciences*, *23*(11), 4621–4634. Retrieved from `https://hess.copernicus.org/articles/23/4621/2019/` doi: 10.5194/hess-23-4621-2019

National Academies of Sciences, Engineering, and Medicine. (2019). *Framing the challenge of urban flooding in the united states.* Washington, DC: The National Academies Press. doi: 10.17226/25381

New york city stormwater resiliency plan. (2021, May). New York City, NY. Retrieved from https://www1.nyc.gov/assets/orr/pdf/publications/stormwater-resiliency-plan.pdf

NYC 311. (2023). *NYC 311 - non-emergency government services in NYC.* portal.311.nyc.gov. ([Online; accessed 23-Jul-2023])

Olthof, I., & Svacina, N. (2020). Testing urban flood mapping approaches from satellite and in-situ data collected during 2017 and 2019 events in eastern canada. *Remote Sensing*, *12*(19), 3141. Retrieved from https://www.mdpi.com/2072-4292/12/19/3141 doi: 10.3390/rs12193141

Paul, J. D., Buytaert, W., & Sah, N. (2020). A technical evaluation of lidar-based measurement of river water levels. *Water Resources Research*, *56*(4), e2019WR026810.

Sabbatini, L., Palma, L., Belli, A., Sini, F., & Pierleoni, P. (2021). A computer vision system for staff gauge in river flood monitoring. *Inventions*, *6*(4), 79. Retrieved from https://www.mdpi.com/2411-5134/6/4/79 doi: 10.3390/inventions6040079

Science and Resiliency Institute, Jamaica Bay, NYC. (2023). *Community Flood Watch Project.* srijb.org/jbfloodwatch. ([Online; accessed 23-Jul-2023])

Sillmann, J., Kharin, V. V., Zwiers, F. W., Zhang, X., & Bronaugh, D. (2013). Climate extremes indices in the cmip5 multimodel ensemble: Part 2. future climate projections. *Journal of Geophysical Research: Atmospheres*, *118*(6), 2473-2493. doi: https://doi.org/10.1002/jgrd.50188

Silverman, A. I., Brain, T., Branco, B., sai venkat Challagonda, P., Choi, P., Fischman, R., ... Toledo-Crow, R. (2022). Making waves: Uses of real-time, hyperlocal flood sensor data for emergency management, resiliency planning, and flood impact mitigation. *Water Research*, *220*, 118648. Retrieved from https://www.sciencedirect.com/science/article/pii/S0043135422006017 doi: https://doi.org/10.1016/j.watres.2022.118648

Smith, B., & Rodriguez, S. (2017). Spatial analysis of high-resolution radar rainfall and citizen-reported flash flood data in ultra-urban new york city. *Water*, *9*(10), 736. doi: 10.3390/w9100736

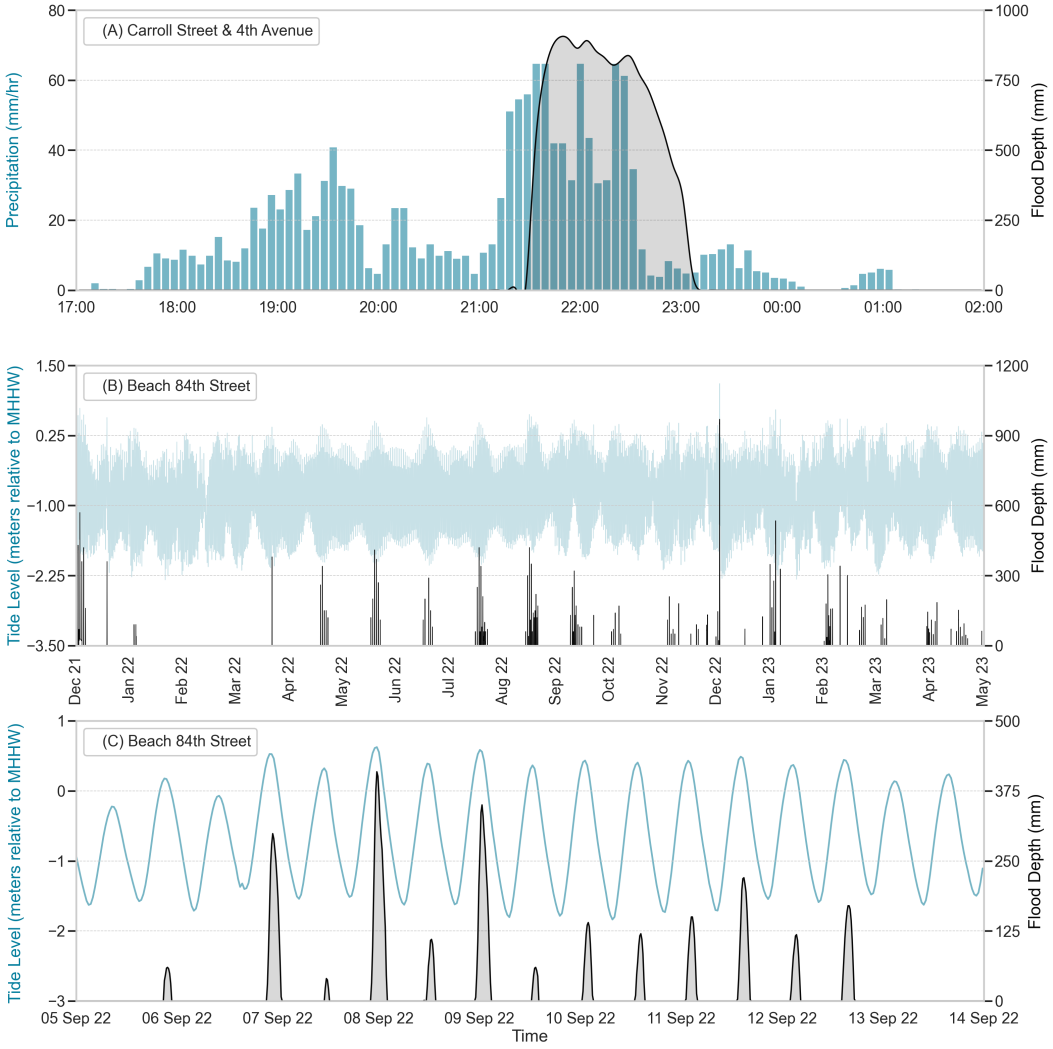

Figure 10. Examples of flood data collected by FloodNet sensors: (A) Data collected by the sensor located at the intersection of Carroll Street and 4th Avenue in Brooklyn during Hurricane Ida (September 1, 2021). The black line represents the flood depth data (secondary y-axis); blue bars represent rainfall intensity (primary y-axis; precipitation data was sourced from a rain gauge at a New York State Mesonet managed weather station at the Brooklyn Navy Yard (ID: bknyrd) located ≈ 2.7 km from the sensor). The flood sensor is located over a sidewalk, therefore flood depths were greater in the adjacent roadway. (B) Data collected by the sensor located on Beach 84th Street in Rockaway, Queens over the course of 18 months (December 2021- May 2023). The Black line represents the flood depth data (secondary y-axis); the blue line (primary y-axis) represents the tide height above the mean higher high water (MHHW) level. Tide data are from NOAA for the North Channel, NY tide station (NOAA station ID: 8517201). (C) Same data as in Panel B, but zoomed in to a week period from September 5 to September 14, 2022, when high tide flooding caused by the full moon on September 10 resulted in 13 consecutive flood events, with floods occurring every 12 to 24 hours.