Evaluation of Road Network Resilience to Natural Hazards using Network Analysis

NCDOT Project 2024-56
FHWA/NC/2024-56
April 2025

Katherine Anarde, Ph.D. et al.
North Carolina State University
Department of Civil, Construction, and Environmental Engineering
Campus Box 7908, Raleigh, NC 27695-7908

STATE OF NORTH CAROLINA DEPARTMENT OF TRANSPORTATION

RESEARCH & DEVELOPMENT

Evaluation of Road Network Resilience to Natural Hazards using Network Analysis

FINAL REPORT

Prepared by:

Katherine Anarde, Ph.D.
Evan Goldstein, Ph.D.
Elizabeth Farquhar, MS
Ryan McCune

A report on research sponsored by:

THE NORTH CAROLINA DEPARTMENT OF TRANSPORTATION

April 2025

1. Report No. FHWA/NC/2024-56	2. Government Accession No.	**3.** Recipient's Catalog No.
4. Title and Subtitle Evaluation of Road Network Resilience to Natural Hazards using Network Analysis		5. Report Date April 2025 6. Performing Organization Code
7. Author(s) Katherine Anarde, Evan Goldstein, Elizabeth Farquhar, Ryan McCune, Thomas Thelen		8. Performing Organization Report No.
9. Performing Organization Name and Address NORTH CAROLINA STATE UNIVERSITY Department of Civil, Construction, and Environmental Engineering Campus Box 7908, Raleigh, NC 27695-7908		10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency Name and Address North Carolina Department of Transportation Research and Development Unit 1549 MSC Raleigh, North Carolina 27699-1549		13. Type of Report and Period Covered Final Report August 1, 2022 - December 31, 2024 14. Sponsoring Agency Code RP2024-56
Supplementary Notes:		

Abstract:
In the research herein we use network science and network analysis to understand the connectivity of coastal road networks during extreme events. The network analysis focuses on flood hazards in coastal settings – including barrier islands and more inland communities along the Neuse and Pamlico estuaries (Oriental, NC and the larger Pamlico County; Davis, NC and the larger Down East region) – but the methodology is broadly applicable to other regions of North Carolina and additional natural hazards (e.g., landslides). We provide maps of the critical nodes – the intersections that break network connectivity when not functioning – for each location. These maps can be used to inform locations for longer-term monitoring by NCDOT during disruptions to ensure public safety. The network analysis is supplemented with data from new land-based flood sensors deployed at critical intersections in Down East Carteret County. The sensors are providing data on road closures for comparison to the network model, and for improved knowledge of flood frequency and durations. We will continue to maintain the sensors for use by NCDOT through leveraged projects. Additional research products from development of the flood sensors include new automated methods for extracting information about flooded roads and road closures from imagery using machine learning and computer vision techniques. The new land-based flood measures have also revealed that tide gauge data are poor indicators of flooding: floods occur 25-125 days annually along coastal roadways in North Carolina, an order of magnitude greater than what regional tide gauges suggest in some places. Improving the accuracy of coastal flood measures is critical for identifying the impacts of sea-level rise and developing effective adaptation strategies for transportation networks.

17. Key Words Network analysis, coastal flooding, machine learning, computer vision	18. Distribution Statement		
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 53	22. Price

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized**

DISCLAIMER

The contents of this report reflect the views of the author(s) and not necessarily the views of the University. The author(s) are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of either the North Carolina Department of Transportation or the Federal Highway Administration at the time of publication. This report does not constitute a standard, specification, or regulation.

EXECUTIVE SUMMARY

Roadway vulnerability assessments are often used to predict which routes are currently, or may in the future, be subject to natural hazards. However, these assessments are often conducted for individual roadways and therefore do not assess to what degree road closures affect the connectivity of road networks – i.e., the ability for a user to access other roads in the network. A consequence of this is that future roadway retrofits, such as raising the elevation of roadways, could alter network connectivity in a way that has cascading impacts on community accessibility during extreme events.

In this project, we use network science and network analysis to understand the connectivity of coastal road networks during extreme events. By treating road intersections as 'nodes' and road segments as 'edges', we successively remove nodes based on increasing elevations (akin to flooding or another extreme event) to identify the threshold where the entire network starts to break apart. The network analysis in this project focuses on flood hazards in coastal settings – including barrier islands and more inland communities along the Neuse and Pamlico estuaries (Oriental, NC and the larger Pamlico County; Davis, NC and the larger Down East region) – but the methodology is broadly applicable to other regions of North Carolina and additional natural hazards (e.g., landslides). The associated research products from the network analysis include maps of the critical nodes – the intersections that break network connectivity when not functioning – for each location. These maps can be used to inform locations for longer-term monitoring by NCDOT during disruptions to ensure public safety. A comparison of network analysis and critical nodes for small communities (e.g., Davis) versus larger regions (e.g., Down East) highlighted that the scale, size, and geography of the region of interest impact the location of the critical node.

The network analysis was supplemented with data from new land-based flood sensors deployed at critical intersections. In Davis, NC, the location of the critical node informed the placement of two sensor systems developed as part of this project: the Sunny Day Flood Sensors (SuDS) and the Tiny Machine Learning Cameras (TinyCamMLs). Elsewhere in Down East Carteret County, the placement of sensors was determined by community-identified hotspots (Sea Level and Cedar Island) and the needs of NCDOT (North River Bridge). The sensors are providing data on road closures for comparison to the network model, and we will continue to maintain the sensors for use by NCDOT moving forward through leveraged projects. Additional research products from the flood sensors include new automated methods for extracting information about flooded roads and road closures from imagery using machine learning and computer vision techniques. In this report we also illustrate use cases of the flood data for improved understanding of flood frequency and durations along coastal roadways, and how this differs from marine-based estimates of flooding. For example, using SuDS data from the NCDOT-sponsored sensor in Sea Level, NC, we found that the roadway in Sea Level flooded 122 days over a year-long study period. Comparatively, flood thresholds derived from the tide gauge in Beaufort indicate 9-31 days of flooding over the same period (depending on which flood threshold is used). More accurate information on coastal flooding can inform where and how we invest resources in building more resilient transportation networks.

Table of Contents

1. Introduction and research objectives

Extreme events such as hurricanes, elevated river levels, and landslides cause road closures across North Carolina by causing flooding, damage to roadways, debris deposits, and traffic accidents. Roadway vulnerability assessments are often used to predict which routes are currently, or may in the future, be subject to natural hazards (e.g., Velasquez-Montoya, Sciaudone, Smyre, and Overton, 2021). However, these assessments are often conducted for individual roadways and therefore do not assess to what degree road closures affect the connectivity of road networks – i.e., the ability for a user to access other roads in the network. In coastal settings such as narrow barrier islands, network connectivity may hinge on a single roadway (e.g., NC Highway 12) or a series of critical roadway intersections (e.g., Carolina Beach). As low-lying environments, barrier islands are particularly exposed to coastal hazards such as storm-driven flooding and sea-level rise, which climate change will only exacerbate (Melillo et al., 2014; Wong et al., 2014; Williams, 2013; Moser et al. 2012; Zhang and Leatherman, 2011; McGranahan et al. 2007). ***It is unclear how roadway connectivity on barrier islands will change with climate impacts, or moreover, with subsequent roadway retrofits.***

Over the coming decades, population projections estimate that the number of people living in coastal areas, including barrier islands, will continue to increase (Neumann et al., 2015). Thus, continuous population growth and coastal development in hazard-prone zones will increase the risk of environmental disasters in these areas. Increased exposure to extreme coastal events extends to transportation networks, on which developed barrier islands depend to allow the mobility of residents and tourists and ensure the supply of goods and services. Due to their limited number, coastal roadways also constitute the main terrestrial route of evacuation, emergency response and recovery operations during and after a disaster in these islands (Velasquez-Montoya et al., 2021; Anarde et al., 2018, Frazier et al., 2013). The disruption, congestion or destruction of sections of the road network can therefore lead to major socio-economic impacts, isolating entire neighborhoods, compromising evacuations, and preventing people from accessing critical services (Jenelius and Mattson, 2015; Mattson and Jenelius, 2015; Spanger-Siegfried et al., 2014). Moreover, the restoration of other critical networks (electrical system, water supply, communications, etc.) also depends on the ability to transport people and equipment to the damaged site (Chen et al., 2002). In the barrier islands of the USA, unforeseen road network disruptions – mechanisms that cause reductions in mobility or increases in the costs necessary to maintain the desired levels of mobility (Markolf et al., 2019) – are not uncommon during hurricanes, tropical storms and nor'easters (Velasquez-Montoya et al., 2021; Spanger-Siegfried et al., 2014; Nordstrom, 2004; Dolan and Lins, 2000; Nordstrom and Jackson, 1994). For instance, in NC, coastal routes have a documented history of closures during storms and hurricanes (Velasquez-Montoya et al., 2021; Spanger-Siegfried et al., 2014; Hardin et al., 2012; Krynock et al., 2005).

To mitigate impacts that affect the serviceability of transportation infrastructures, especially during disasters, a growing body of research has focused on assessing the functionality of these networks under different extreme events such as earthquakes (Lam & Shimizu, 2021;

Khademi et al., 2015) or floods (Abdulla & Birginsson, 2020; Kasmalkar et al., 2020; Abdulla et al., 2019; Anarde et al., 2018). Official organizations have also published technical guidance and methods to assess the exposure, sensitivity, and adaptive capacity of the transportation system to extreme weather and climate effects (Filosa et al., 2017; Douglass and Krolak 2008). ***Graph theory is one of the most commonly used mathematical frameworks to assess road network susceptibility to failure.***

Reducing the original network to a two-dimensional matrix of 'nodes' (road intersections) connected by 'edges' (road segments) simplifies the connectivity analysis within the network. However, this approach ignores the crucial importance of the third dimension (elevation) for identifying road sections susceptible to flooding and assessing network functionality under real-world events (Versini et al., 2010). Roadway elevations, for instance, can be linked to extreme water levels, which allow the study of network disruptions under a wider range of feasible flooding events (Kaskalmar et al., 2020; Anarde et al., 2018). The altitude of the network elements can also be used to estimate the likeihood of flooding and predict the propagation of the flood through the system (Abdulla and Birgisson, 2021; Yuan et al., 2021). Nevertheless, disruptions in 3D infrastructure networks are still an area of emerging research (Wang et al., 2019), and very few studies have attempted to assess road susceptibility to coastal flooding in dynamic regions such as the barrier islands of the USA (Velasquez-Montoya, 2021). Floods are a special type of network disturbance considered locally more destructive than random damage (Wang et al., 2019), and are a current frontier of research.

1.1. Research objectives

Current roadway vulnerability frameworks used by NCDOT – and throughout the United States – do not assess the susceptibility of the road network as a whole to coastal flood hazards. A consequence of this is that future roadway retrofits, such as raising the elevation of roadways, could alter network connectivity in a way that has cascading impacts on community accessibility during future flood events. Resilient decisions for roadway retrofits should account for alterations to network connectivity. Hence, in this project, ***our principal goal is to improve predictions of roadway vulnerability to natural hazards by using network science and network analysis to evaluate the vulnerability of whole roadway networks to failure***, both now and in the future. This goal supports the objectives outlined in the NCDOT Resilience Strategy Report (2021) to support robust vulnerability assessments of coastal roadways to assist with future planning efforts. The secondary goal of this project ***is to develop new methods that utilize computer vision and machine learning to automate the identification of flooded roadways from imagery.***

This final report outlines the steps that were taken to meet the stated goals, including:

1. **Evaluation of network connectivity on NC barrier islands for a range of node removal proxies.** We adapted an existing network model, developed by Co-PI Goldstein and colleagues for node removal by elevation.

2. **Extension of the network model and analysis to more inland locations**. This allowed us to test the viability of the method in different (non-barrier island) settings that may be subject to other flood drivers (e.g., wind, rainfall, river runoff) or geographic influences.
3. **Extend data of road closures using land-based measures of flooding.** Low-cost cameras and *in situ* sensors were developed and deployed at flood hotspots to enrich the data available for testing the network model. New automated methods for extracting information about flooded roads and road closures from imagery were developed using machine learning and computer vision techniques.

The final report is structured to first provide an overview of the research methodology (Section 2), followed by results related to the road network analysis (Section 3), two use cases of land-based measures of flooding (Section 4), and an analysis of the utility of machine learning tools for flood identification (Section 5).

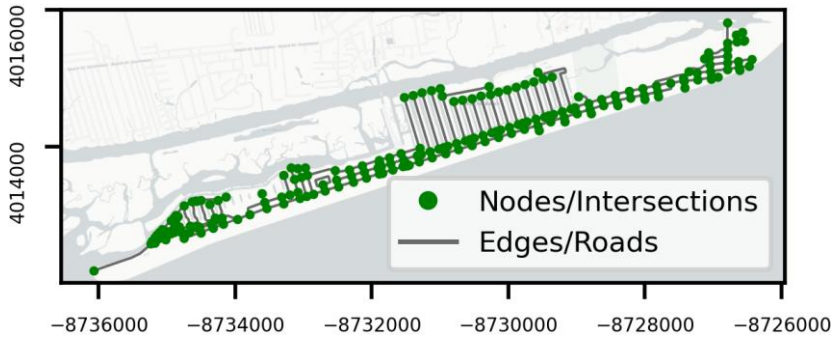


Figure 1. An example network model or 'graph' of Ocean Isle, NC. In graph theory, 'nodes' correspond to roadway intersections and 'edges' correspond to roadways.

2. Research methodology

2.1. Road networks

Road networks can be schematized as a 'graph', which is the mathematical construct of a network: in this framework, roadway intersections are 'nodes' and streets are 'edges' that link nodes together (Figure 1). Thinking of road networks as general nodes and edges allows us to use the same ideas from network science (and the study of social networks) to study roadway connectivity. For example, a key theme in network analysis is examining the behavior of the 'giant component', which is the largest group of connected nodes. In a fully functioning road network, the giant component is the entire network — all roads are connected. The susceptibility of a network to the failure of its components is typically explored by deactivating or removing nodes and calculating metrics that address network functioning, such as measuring a percolation process (Newman, 2010; Schneider et al., 2011; Iyer et al., 2013; Li et al., 2014; Wang et al., 2019). By sequentially deactivating nodes, the functionality of the entire network decreases. For example, travel between two nodes (intersections) becomes impossible or requires long travel distances (and time) on the network. Successively removing road network nodes (intersections) based on some criteria (such as increasing elevations, akin to flooding or another extreme event) can eventually lead to a threshold where the entire network starts to break apart. The specific threshold crossing — the breaking — is quantified by looking at the number of nodes in the giant component and the second largest component as nodes are removed.

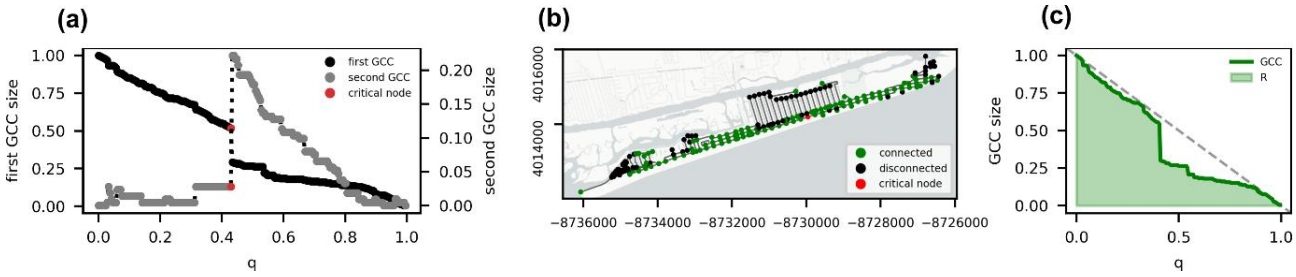


Figure 2. Examples illustrating the methodology used to **(a)** explore the size decay of the first and second giant-connected components (GCC), **(b)** identify the critical node that leads to the fragmentation of the network, and **(c)** quantify overall network robustness to elevation-based node removal. The barrier example shown here is the drivable network at Ocean Isle, North Carolina, USA. In **(a)**, the vertical axes show the first (left) and second (right) GCC size as a fraction of nodes in the original network, as a function of the fraction of nodes removed (q). The red dot in panels **(a)** and **(b)** marks the critical node in the GCC and real physical space, respectively. In panel **(c)**, robustness R is taken as the area (light green) under the decay curve for the first GCC (bold green). The dashed gray line shows the inverse 1:1 reference line, indicating the theoretical maximum for R = 0.5. Maps like the example shown in **(b)** for all 72 barrier road networks with >100 nodes can be found in the data repository.

As an example, we can visualize this process using a network model developed by PI Goldstein and colleagues for Ocean Isle, NC. Nodes were sequentially removed based on a ranked list of elevations, from lowest to highest, mimicking a simplified "bathtub" flooding scenario (i.e., other topological metrics such as road grades or average street/'edge' elevations were not explored). The first panel in Figure 2 tracks both the size of the first giant component (black) and the second giant component (grey) as a certain fraction of nodes (q) are removed. Notably, once the lowest 40% of nodes are removed (q = 0.4), the first giant component suddenly drops in size (the shift downward in the black line) and the second giant component suddenly increases in size. This is the threshold crossing when the network breaks into two large pieces, and strands part of the road network. The network crosses a critical threshold at q_c, when the first giant component completely disintegrates and the size of the second giant component becomes maximal (Li et al., 2014; Wang et al., 2019). Figure 2b visualizes this moment of breakup into geographic space. The active (dry) intersections are shown in green, the flooded intersections are shown in black, and red is the intersection that causes catastrophic breakage when it becomes flooded, stranding the west side of the island from the east side.

Aldabet, Goldstein, and Lazarus (2022) also explored the sensitivity of roadway connectivity to random node removal. They found that road networks on barrier islands are much more prone to breaking when nodes are randomly removed than when removed sequentially by elevation. Our workflow for investigating North Carolina barrier island road networks is shown in Figure 3. We discuss each step in the sequence below.

To isolate barrier island road networks, we used digitized perimeters of all NC barrier islands from the open dataset of Mulhern et al. (2017) & (2021) and extracted the drivable road networks from OpenStreetMap (OSM) with OSMnx (Boeing, 2017). Cast as networks, road intersections are encoded as nodes and road segments are edges. We excluded other possible transportation pathways such as bikeways and walkways.

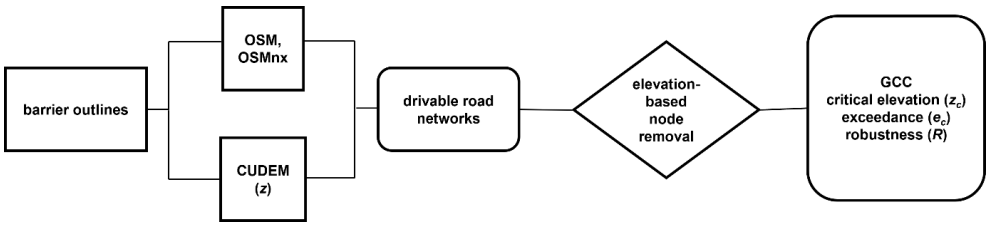


Figure 3. Methodological workflow for assessing robustness to flood-induced failures in road networks on US Atlantic and Gulf barrier islands. Abbreviations are as follows: OSM is Open Street Map; OSMnx is an analytical toolbox (Boeing, 2017). CUDEM is the NOAA Continuously Updated Digital Elevation Model (Amante et al., 2021; CIRES, 2014). GCC is the giant-connected component of a networkor the large cluster of nodes connected in the original network.

Of all NC barrier islands considered, all overlap with tiles currently available in the NOAA Continuously Updated Digital Elevation Model (CUDEM), a set of 1/9 Arc-Second resolution

bathymetric and topographic tiles for the coastal USA (Amante et al., 2021; CIRES, 2014). For statistically meaningful metrics of network structure, we restricted our analysis to barriers with drivable road networks of at least 100 nodes. We determined the elevation of each node (road intersection) in each network by spatially querying the CUDEM dataset.

The susceptibility of a network to the failure of its components is typically explored by nullifying or removing nodes and calculating metrics that reflect network functioning (Abdulla & Birgisson, 2021; Iyer et al., 2013; Li et al., 2015; Newman, 2010; Schneider et al., 2011; Wang et al., 2019). For example, when enough of the network is removed, travel between any two nodes (intersections) becomes impossible or requires long travel distances (and time) on the network. We removed nodes from a network based on a ranked list by elevation – from lowest to highest – in contrast to removing nodes randomly (a common approach, e.g., Albert & Barabási, 2002). Node removal in this way mimics a simplified "bathtub" flooding scenario (e.g., Abdulla & Birgisson, 2020; Wang et al., 2019), which assumes that nodes become nullified because they are actively flooded, damaged by flooding, and/or unusable because of debris and/or sand deposited on the road. We assumed that the removal of a node causes the immediate disconnection of all its connected edges. This work thus considered node removal exclusively; edge removal could also be explored, with the inclusion of other contextual physical metrics such as road grade, lowest street elevation, or average street elevation. Network metrics were calculated using NetworkX (Hagberg et al., 2008).

For road networks, the original network is connected in a single large cluster – the giant-connected component (or giant component). As nodes in the original network are serially removed, the network breaks into smaller networks. Here, we tracked the size of these subnetworks relative to the size of the giant component. Specifically, as the fraction of nodes removed (q) increases and the first giant component degrades, we tracked the size of the second-largest cluster – the second giant-connected component (Figure 2a). The network crosses a critical threshold at q_c, when the first giant component fragments and the size of the second giant component becomes maximal (Li et al., 2015; Wang et al., 2019). Generally, the higher q_c – that is, the more nodes that can be removed before the giant component fragments – the less prone the network is to failure (Newman, 2010). The critical threshold (q_c) can be linked to a specific node that causes the failure of the network (Figure 2b) and to the elevation of that node, which we refer to as the critical elevation (z_c).

2.1.1. Extreme water levels

Comparison of coastal barrier islands solely based on topographic elevation (i.e., one barrier is higher or lower than another) is not meaningful unto itself because of local differences in tidal forcing and extreme water level statistics. For example, road networks on higher-standing barriers subject to frequent extreme storms might be more prone to flooding than road networks on lower-lying barriers subject to fewer storms. To provide meaningful comparisons among the broad geospatial distribution of barriers in our sample, we recast all node (intersection) elevations to local annual exceedance probabilities of extreme water events.

Extreme water levels have been used to examine the direct and indirect impacts of coastal floods on transportation systems and assess the susceptibility of the network to flood-induced failure (Fant et al. 2021; Habel et al., 2020; Jacobs et al., 2018; Pezza & White, 2021). Annual exceedance probabilities and average recurrence intervals are commonly applied for infrastructure design and assessment of flood risk (Apel et al., 2004, 2006; Hackl et al., 2018; Haigh et al., 2014; Sweet & Park, 2014; Vitousek et al., 2017; Wahl et al., 2017). Average recurrence intervals, also known as return periods, provide an estimation of the time elapsed between events of the same magnitude; annual exceedance probability refers to the likelihood that high-water levels exceed a certain elevation in any given year (Haigh et al., 2014). For example, a flood with an annual exceedance probability of 0.01 corresponds to an event that has a 1% chance of annual occurrence, or an average recurrence interval of 100 years. (Return period can be understood as the inverse of exceedance probability.)

Extreme value analysis (EVA) – the branch of statistics that deals with the estimation and prediction of rare values within a series (Coles, 2001) – has been applied broadly to analyses of observed and simulated extreme high-water levels to quantify the probability of occurrence (and/or return period) of extreme events (Vitousek et al., 2017; Wahl et al., 2017; Zervas, 2013). One of the most common EVA methods is block maxima, which considers the maximum of all recorded values within a block of time (i.e., days, months, or years) and approximates extreme values using a Generalized Extreme Value distribution (GEV) (Coles, 2001; Zervas, 2013). The GEV distribution is described by three parameters – location (μ), scale (σ), and shape (ξ) – that refer, respectively, to the center of the distribution, the deviation around the mean, and the tail behavior of the distribution. The shape parameter determines the extreme distribution used: Gumbel ($\xi = 0$), Frèchet ($\xi > 0$) or Weilbull ($\xi < 0$). Using long-term monthly tide gauge records from the 112 US stations operated by the Center for Operational Oceanographic Products and Services (CO-OPS), Zervas (2013) followed a GEV approach to characterize the distributions of extreme high and low water levels and produce exceedance probability curves for each station.

For each barrier island in this analysis, we generated extreme high-water level annual exceedance probability curves by sampling the Gumbel distribution described by the three reported GEV parameters (Zervas, 2013) for the tidal station closest to that barrier by straight-line distance. We then estimated annual exceedance probabilities for the critical node of each barrier network, which we refer to as the critical exceedance, e_c. We thus linked each critical node to a specific annual exceedance probability. All calculations were done using the Python ecosystem, e.g., Scipy (Virtanen et al., 2020) and Numpy (Harris et al., 2020). Note that the choice of extreme value analysis applied to a data set has the greatest effect on events with the lowest likelihood of occurrence (Wahl et al, 2017). Because high-likelihood events are of particular interest to us in this analysis, the Gumbel distributions that we use to reproduce the estimates reported by Zervas (2013) are sufficient: a different method of extreme value analysis would result in different probabilities for the low-likelihood events from these tide gauges but estimates for high-likelihood events will be effectively the same.

2.1.2. Network robustness

Having focused on identifying a single critical node for each island and defining a critical threshold for each barrier road network in terms of elevation (and exceedance probability), we next examined the overall network robustness of each barrier. The purpose of this step is to provide a summary metric for network functioning that includes but is not limited to the occurrence of the critical threshold: for example, determining how much of the original road network is still connected when any given percentage of the nodes is removed. Calculating whole-network robustness permitted us to compare barrier road networks in terms of their entire architecture, rather than solely by comparing aspects of a single critical node (e.g., its elevation and the related exceedance value).

We used the robustness metric *R* proposed by Schneider et al. (2011), which measures the summed size of the giant-connected component as nodes are removed (Figure 2c):

$$R = \frac{1}{N}\sum_{Q=1}^{N} s(Q) \quad (1)$$

where *N* refers to the total number of nodes in the network, *Q* to the number of nodes removed and *s(Q)* is the fraction of nodes in the giant component after removing *Q* nodes. The normalization factor 1/*N* allows comparison between networks of different sizes. The resulting *R* values range between 1/*N* (for a star graph) to 0.5 (a fully connected network; Schneider et al., 2011). Note that we evaluated network robustness in two ways: by removing nodes in rank order of elevation (lowest to highest) and by random node removal (e.g., Wang et al., 2019). Other studies have investigated how *R* changes with non-random but abstracted network disruptions (Iyer et al., 2013), and how *R* varies in transportation networks, specifically, with different types of disruptions (Dong et al., 2020b; Wang et al. 2019).

All scripts associated with the network model can be found on the Github repository (https://github.com/NCRoadNetworks/NC_barrier_roadnetworks).

2.1.3. Compare mathematical models of network failure to real-world examples

To test the efficacy of the network models developed as part of Objective 1, we ground check the network analyses and location of critical nodes using both existing NCDOT data on roadway floods on NC barrier islands. Existing data on the actual range of critical nodes for each NC barrier island – that is, which roadways close or isolate sections of barrier islands during extreme events due to flooding – came from NCDOT data of past roadway incidents via TIMS and informal interviews with regional members of NCDOT.

2.1.4. Extend the network model and analysis to an inland location

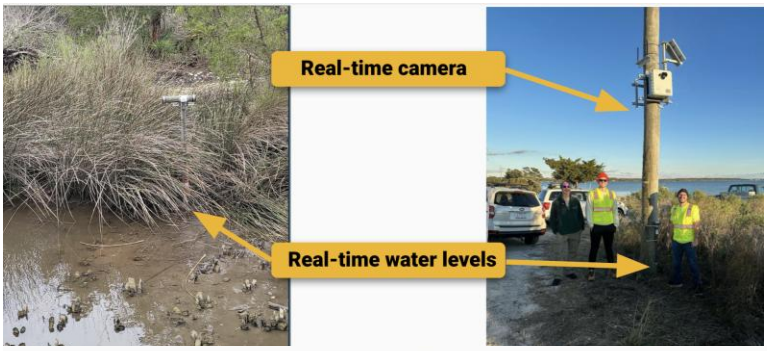
We further extended the network analysis to 4 inland locations: Kinston (as recommended by NCDOT personnel); Davis and the larger Down East region; and Oriental and the larger Pamlico County. Extending the analysis to inland locations allows us to evaluate the broader applicability of the method to other settings.

2.2. Flood sensors

Objective 3 of this project involved the development and deployment of new, land-based flood sensors along vulnerable roadways to extend data of road closures for comparison to the network model and also to better understand the frequency and causes of flooding. Here we describe two low-cost sensors that were deployed at flood hotspots in Down East Carteret County as part of this project, and elsewhere throughout the state (as part of leveraged projects): the Sunny Day Flood Sensors (SuDS) and Tiny Machine Learning Cameras (TinyCamMLs).

2.2.1. Sunny Day Flood Sensors (SuDS)

PI Anarde maintains a network of flood sensors, coined SuDS, in five coastal communities in North Carolina (Figure 4). The flood sensors consist of pressure sensors for measuring *in situ* water levels and subaerial cameras for monitoring flood extent. Pressure sensors are located in storm drains or adjacent to roadways in community-identified hotspots and relay data in real time to the camera, which also contains a cellular gateway. Flood data is displayed on a webapp for communication of flood hazards (Gold et al., 2023). In some locations, flood data has been collected every 6 minutes since 2022.



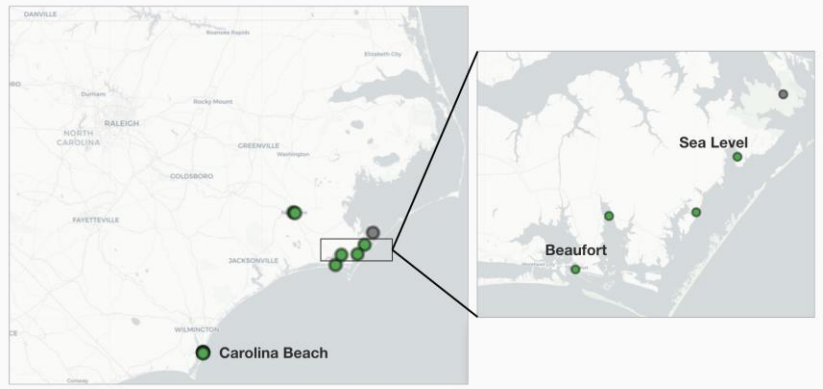


Figure 4. Sunny Day Flood Sensors (SuDS) in Down East Carteret County. (top) Example images from the sensors at Sea Level (left) and at the North River Bridge (right). (bottom) The location of all SuDS sensors in NC.

As part of this project, we deployed 4 SuDS in Down East Carteret County: one each in Sea Level, Davis, and Cedar Island and one at the North River Bridge. The Sea Level sensor is our longest-running sensor (installed April 2023) and Cedar Island is our most recent sensor (installed November 2024). The sensor location at the North River Bridge was selected by NCDOT personnel as being particularly important for assessing network connectivity. The sensor in Davis, NC was selected based on the network analysis and is located at the location of the critical node. For privacy, only the most recent image is displayed on the web app and images are blurred outside

of public spaces. Images are archived for scientific purposes, including the development of automated flood detection algorithms (described below). As described in Section 4 below, we also generate annual flood counts for each sensor to better understand flood frequency, durations, and the seasonality of flood events.

2.2.2. Tiny Machine Learning Cameras (TinyCamMLs)

The causes of flooding can be highly specific to the region in which it's occurring; therefore, it can be difficult to predict the full extent of land-based floods using only tide gauges. We developed a low-cost (<$400), easily maintainable camera system that we deploy in coastal regions to monitor flooding. Our device - a Tiny Camera with Machine Learning, or TinyCamML - is a small, solar-powered, microcontroller-based camera that uses on-device machine learning to classify images taken every 6 minutes as containing a “Flood” or “No Flood.” TinyCamMLs transmit only the classifications to a website in real time, providing updates during flood events. Images of roadways are never transmitted (and can be set up to not even be saved), protecting the privacy of the local communities in which they are deployed.

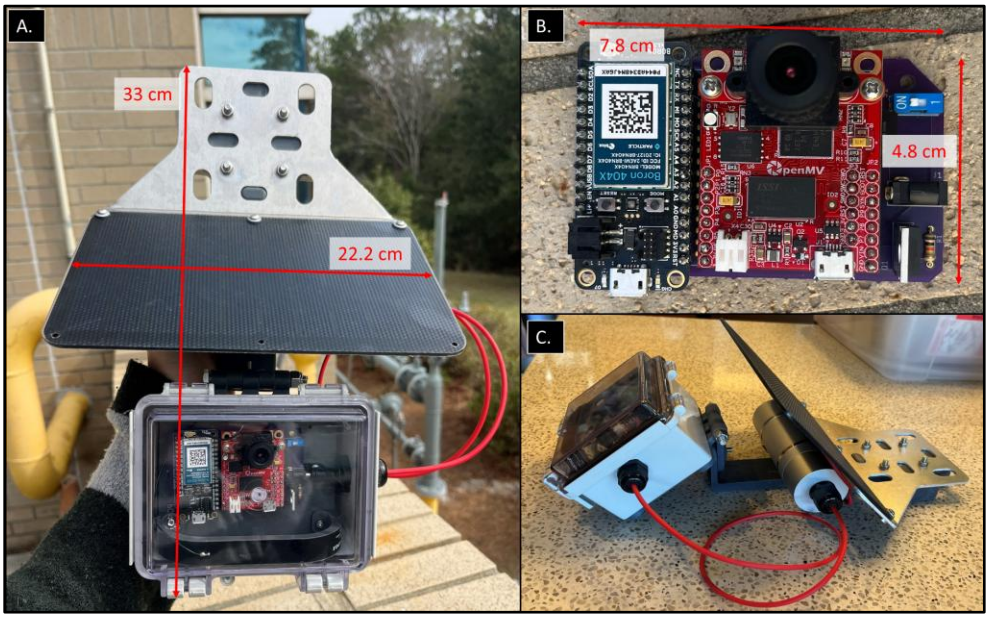


Figure 5. **(a)** Fully assembled TinyCamML. **(b)** The Boron and OpenMV Cam microcontrollers mounted onto the custom PCB board. **(c)** Side view of the fully assembled TinyCamML.

2.2.2.1. Electronics and housing

The TinyCamML hardware is composed of four parts: a 5W 6V solar panel and 5V 36Wh external battery (Voltatic Systems, Brooklyn, NY, USA), a Boron microcontroller (Particle, San Francisco, CA, USA), an OpenMV Cam H7 Plus microcontroller (OpenMV, Atlanta, GA, USA), and a MOSFET switch. The solar panel powers the Boron, which in turn powers the OpenMV Cam via the MOSFET switch. All electrical components are soldered onto a custom PCB board, which is

fixed onto a 3D-printed electronics tray and secured inside a Polycase enclosure (Polycase, Avon, OH, USA). The polycase is fastened to the solar panel by a variety of 3D-printed parts. We use a Formlabs 3 SLA printer and Tough 2000 resin for all parts (Formlabs, Boston, MA, USA). Specific part numbers for all components, 3D designs, and a circuit diagram can be found on the Github repository (https://github.com/TinyCamML).

2.2.2.2. Software

TinyCamML software is designed so the system takes an image, classifies it as "Flooded" or "Not Flooded" according to the onboard machine learning model, and then transmits the classification over cellular to a website for logging. This is accomplished with both microcontrollers. Once woken, the OpenMV takes a photo and runs the ML model. The classification is then sent to the Boron over a universal Asynchronous Receiver/Transmitter (UART), and the Boron publishes it to a Google sheet in real time using its cellular capabilities. The Boron sends its timestamp data to the OpenMV Cam over UART, so the OpenMV Cam saves a timestamp with each classification (i.e., flooded or not flooded) in its data log. If enabled, the OpenMV Cam can also save images to its onboard microSD card, which can be used to further train the model. Images are never transmitted, only classification scores. This saves on bandwidth and maintains privacy. Currently, the OpenMV Cam does not classify images when the scene brightness is below a threshold, which prevents unnecessary operation during low- and no-light conditions (i.e., nighttime) when accurate classifications are not currently possible. The source code for the Boron microcontroller was developed using the Particle extension in Visual Studio Code version 1.95.3 and device OS firmware 6.1.0. OpenMV Cam source code was developed on the OpenMV IDE version 4.1.5 and device OS firmware 4.5.5. All source code is available on the Github repository.

2.2.2.3. ML model: binary classification

The first step to developing a generalizable ML model is to assemble a larger training dataset. We use over 14,000 images from deployed TinyCamML devices as well as imagery from the SuDS cameras (Gold et al., 2023), and publicly available datasets of flooding. Images are taken from different locations, vantages, and times of day. We use a single labeler for consistency but note that we perform an interrater experiment with non-training data, and these results can be seen in section 4.2 below. The initial distribution of labeled images is strongly skewed toward nonflooded imagery, so we use ClimateGAN (Schmidt et al., 2021) on unflooded images to produce synthetic flood imagery to balance the dataset. We are then left with 17,448 total images, with a 50/50 split between flood and no flood images.

We train a deep-learning-based classifier model using TensorFlow (Abadi et al., 2016) leveraging the MobilenetV2 architecture (Sandler et al., 2018) and relying on ImageNet weights. We replace the last classification layers of the model with a global average pool, a 1024-neuron dense layer, and a 2-neuron dense layer (corresponding to our two classes, flood and no flood). We use 224px images, and all 3 channels (RGB).

We train the model with a 60/25/15 training/validation/test split with early stopping, using categorical cross-entropy as the loss function, Adam as the optimizer (learning rate of 3x10 -3), a dropout rate of 50% between the dense layers, a batch size of 16, an early stopping callback (patience of 10 epochs), image augmentations on the training data (rotation, width and height shift, shear, zoom, horizontal and vertical flips), and calculating the binary accuracy and a confusion matrix as a performance metric. Note that during training only the new classification layers are trainable. The trained model is then integer quantized using TensorFlow Lite micro (David et al., 2021) for inference on the OpenMV. The quantized model has a binary accuracy of 84% with the test set.

2.3. Machine learning methods for automated image classification

New automated methods for extracting information about flooded roads and road closures from SuDS imagery were developed using machine learning and computer vision techniques. Here we detail two methods: image segmentation and depth mapping. All processing scripts for the image rectification and depth mapping workflow can be found on the Github repository (https://github.com/rtmccune/depth_mapping).

2.3.1. Image segmentation

We applied deep learning techniques to analyze Imagery from the SuDs sensors, specifically image segmentation. In recent years, Deep Learning models based on the UNet (Ronneberger et al., 2015) and the Residual UNet (Liu et al., 2019; Zhang et al., 2018) have become the standard in state-of-the-art Earth science applications involving image segmentation.

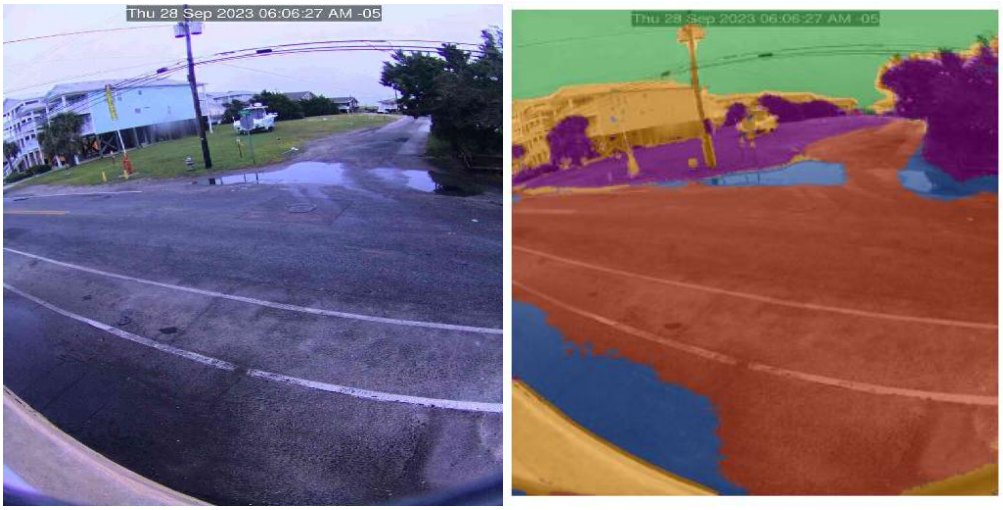


Figure 6. A scene from the Carolina Beach SuDS camera (left) and the segmentation label overlay created by a human labeler and used as training data (right): red is road, blue is water, yellow is building, purple is vegetation, and green is sky.

In the case of image segmentation, training data consists of label imagery where each pixel is categorized into any number of predetermined discrete classes (i.e., water, road, vegetation, building, sky, etc). We use a state-of-the-art Human-in-the-loop system (Buscombe, et al., 2022) to develop the training data and label 500 images from 3 different scenes (SuDS cameras from Carolina Beach, Beaufort, and Down East). Each pixel in each scene corresponds to one of 5 classes (road, water, building, vegetation, and sky). An example of a SuDs image and the corresponding segmentation mask (overlain) is shown above (Figure 6).

With this training data, we develop several machine-learning models using Segmentation Gym (Buscombe and Goldstein, 2022; https://github.com/Doodleverse/segmentation_gym). Our goal is to automatically and rapidly analyze imagery using a computer program, versus needing time-consuming analysis by an operator. We use the Segformer Model, using 512 x 512-pixel input images, a batch size of 8, and train the model for 150 epochs using various augmentations on the training samples.

2.3.2. Depth mapping

A series of images from a given flood event are first passed through a trained image segmentation machine learning model as described above. The original images and their predicted segmentations are then converted to real-world coordinates on the NC State Plane following the methodology and photogrammetry techniques outlined by the developers of the Quantitative Coastal Imaging Toolbox (Bruder & Brodie, 2020). This requires the image, calibration parameters for the camera, and a grid of real-world coordinates to project the image onto.

The first step was to obtain camera calibration parameters, including both intrinsic (i.e. internal characteristics of the camera like distortion) and extrinsic (i.e. external characteristics of the camera like tilt) parameters. Intrinsic parameters were calculated using a series of images of a printed checkerboard pattern captured using a camera and lens of the same make and specification as the one deployed in the field. These images were processed using the Open Computer Vision Library (Bradski, 2000) in Python to obtain the intrinsic parameters including the focal length, principal point, radial and tangential distortion coefficients.

Extrinsic parameters were calculated using a set of known points in 3D coordinates that are present in the camera's field of view, which we refer to as virtual ground control points (GCPs). These GCPs were first selected from the imagery as static objects (e.g. corners of storm drains or permanent signposts) and then their real-world coordinates in NAD83(2011)/North Carolina (ft US) were collected with a real-time kinematic (RTK) GPS. The surveyed GCPs were manually matched to 2D pixel coordinates in a base image and the field of view was then assumed to be static for each subsequent image. Each surveyed point and its pixel coordinates were then processed using OpenCV to obtain the extrinsic parameters including the rotation matrix (i.e. camera's roll, pitch, and yaw) and translation vector (i.e. camera's real-world location).

After collecting the camera calibration parameters, it was necessary to generate a grid of real-world points. We used a National Oceanic and Atmospheric Administration (NOAA) LiDAR

point cloud dataset with a resolution of 20 cm interpolated to a 5 cm grid in Northing and Easting on the North Carolina state plane.

The images, segmentation predictions, camera calibration parameters and real-world grid were then processed using a Python translation of the Quantitative Coastal Imaging Toolbox. This process maps pixel values to their corresponding locations in real-world coordinates. The result is an array for both the original captured image and its prediction labels corresponding to real-world coordinates at 5 cm resolution.

To analyze the inundation of the roadway, the geo-rectified labels were masked to pixels predicted as water by the segmentation model. An important step before calculating the depth of the inundation present in the image is to group ponds of water. Without treating each pond separately, the later steps of edge extraction and depth calculation could falsely assign greater depths to disconnected ponding on the roadway than is accurate. The ponding algorithm groups inundated points and removes groups below a minimum size to exclude minor segmentation errors. It then assigns a pond number to each group for further processing.

To calculate depth, the edges of the pond are extracted using the Scikit Image Library (van der Walt et al., 2014). We then select the 95^{th} percentile elevation of these identified pond edges as the flood elevation. We assume that any points in the interior of the pond group below the flood elevation are inundated and any values above the flood elevation are set to zero. This same process is completed for every pond in the image. These pond depths are then combined to create one cohesive depth map for the entire image.

3. North Carolina road network analysis

In the results that follow, we present maps that identify the critical nodes – that is, the elevation of network failure – for both barrier islands and more inland coastal communities. These maps can inform locations for monitoring by NCDOT during disruptions to ensure public safety.

3.1. Barrier islands

We analyzed all NC barrier islands and show illustrative results for the 3 barrier islands below: Wrightsville Beach, Carolina Beach, and Ocean Isle. Each figure below shows the critical node that leads to the fragmentation of the network (red), connected nodes (green), and disconnected nodes (blue/black).

Figure 7. Illustrative results and road network diagram for 3 barrier islands - Wrightsville Beach, Carolina Beach, and Ocean Isle.

Beyond the geographic space, we can examine these results in elevation space (relative to the NAVD88 datum) and also explore the size decay of the first and second giant-connected components (GCC). In these next figures, the vertical axes show the first (left) and second (right) GCC size as a fraction of nodes in the original network, as a function of the fraction of nodes removed (q). Red dots mark the critical node in the GCC and link this to the elevation (the bottom panel).

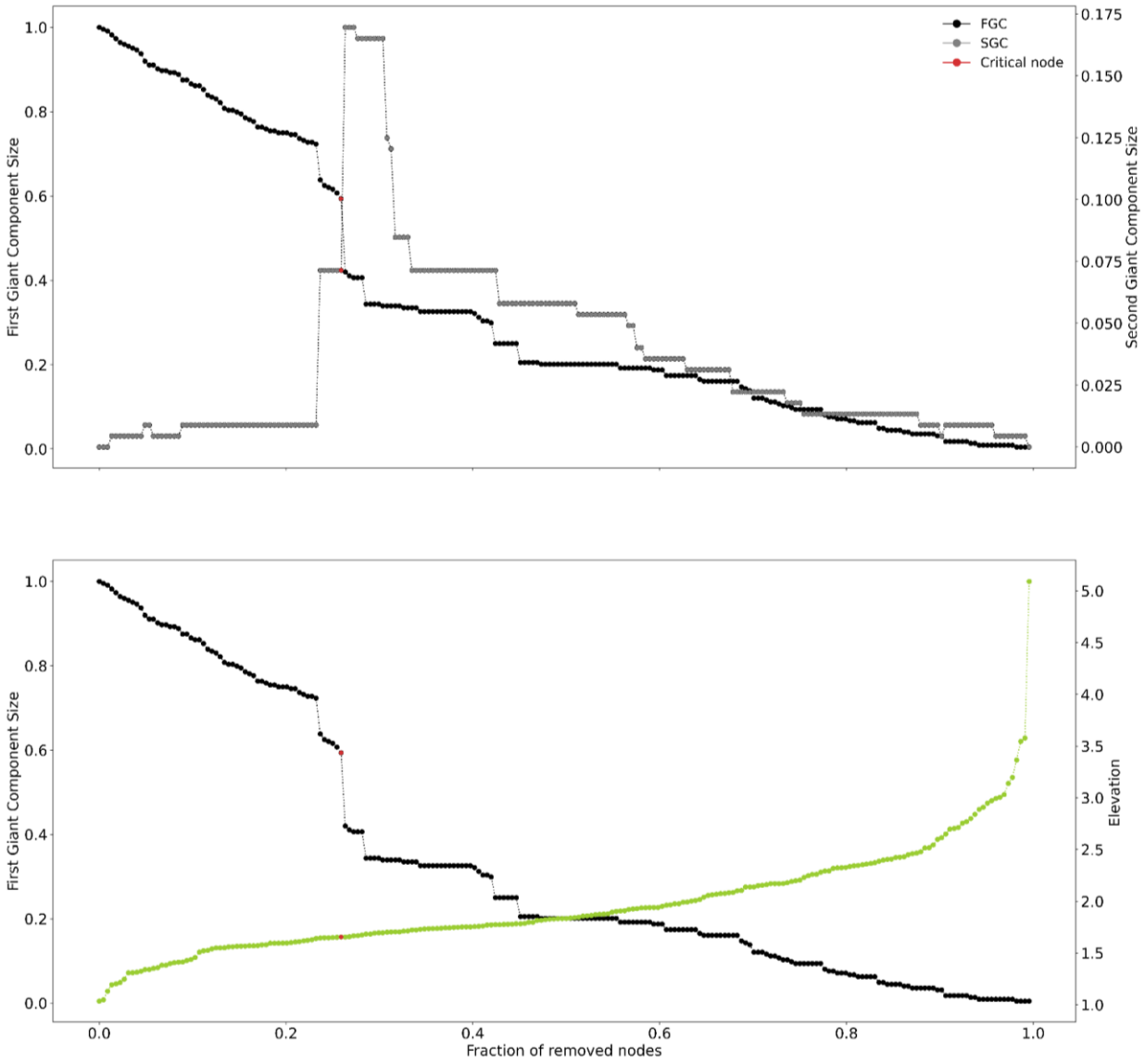


Figure 8. First and second giant-connected component size for Wrightsville Beach. Note that the critical node is located at 1.6 m NAVD88 in elevation. Flooding of this elevation would cause the network to fracture.

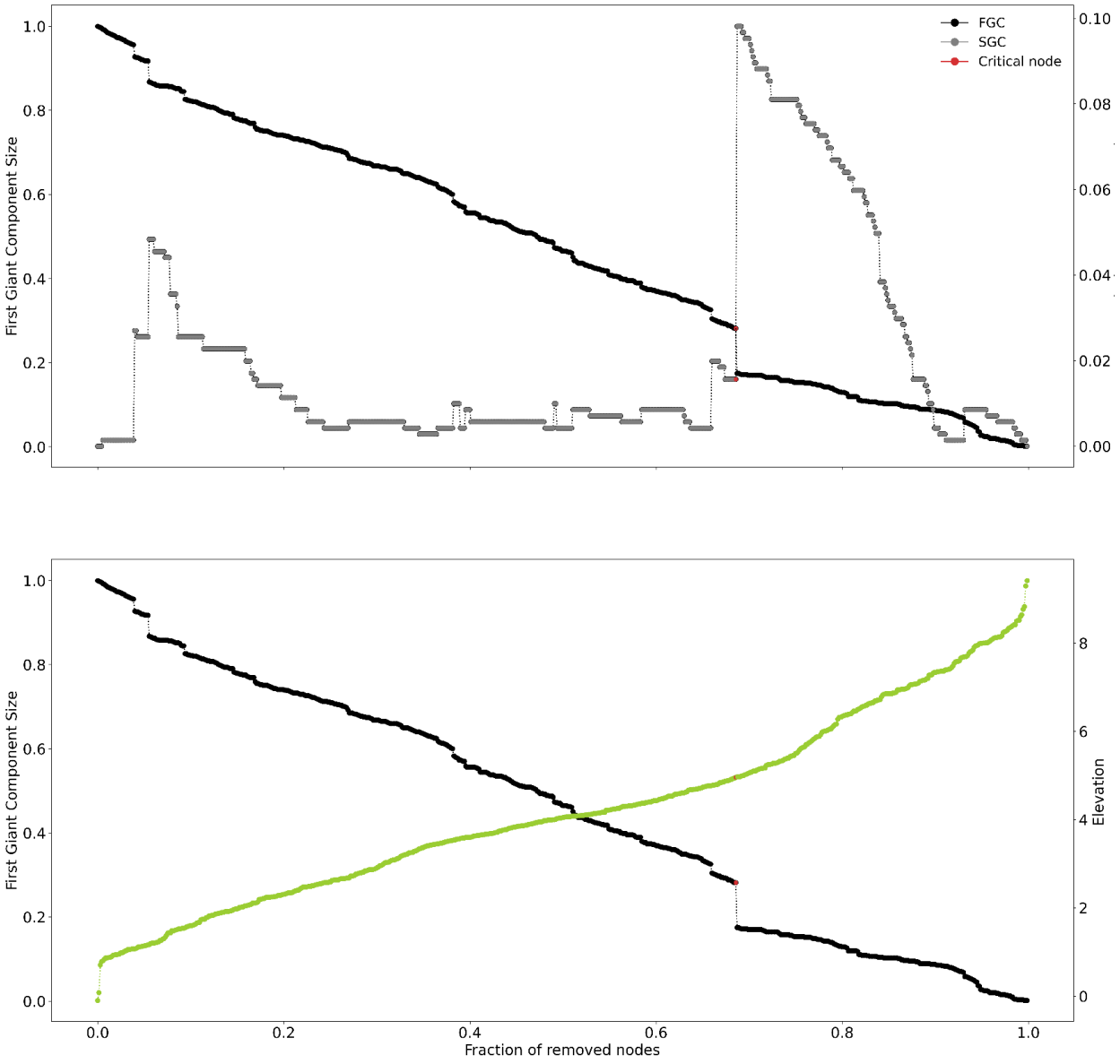


Figure 9. First and second giant-connected component size for Carolina Beach. Note that the critical node is located at 4.5 m NAVD88 in elevation. This road system is robust to failures.

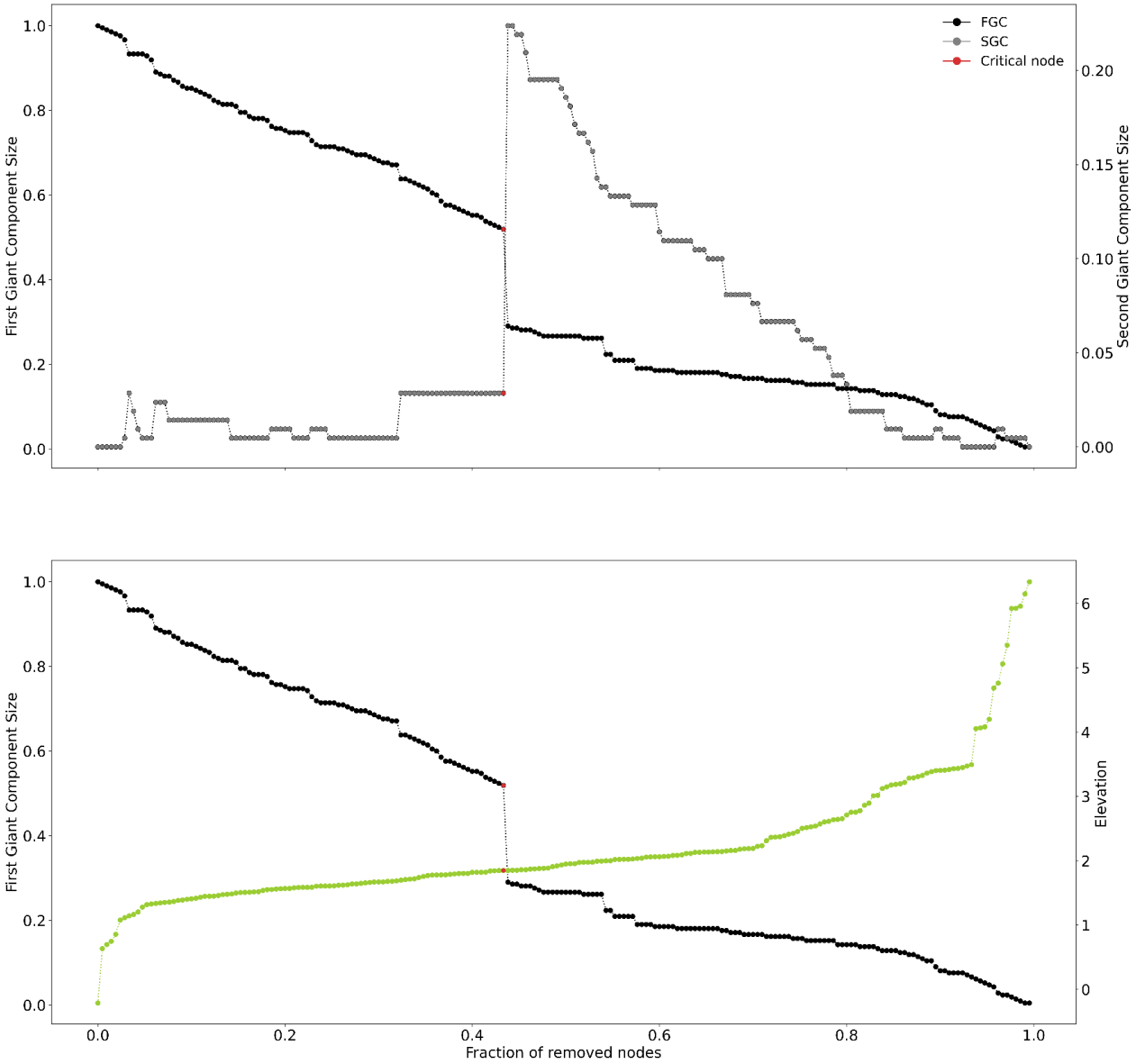


Figure 10. First and second giant-connected component size for Ocean Isle Beach. Note that the critical node is located at 2 m NAVD88 in elevation. Flooding of this elevation would cause the network to fracture.

3.2. Oriental and Pamlico County

We analyzed Pamlico County and the area around Oriental, NC and showed the results below. We pick two different areas of interest, first a smaller area around Oriental and then a larger area that makes up most of Pamlico County. Each figure below shows the critical node that leads to the fragmentation of the network (in red), connected nodes (green), and disconnected nodes (blue/black).

When analyzing a large area it is important to remember the definition of the critical node, which is outlined in Section 2: the critical node is the threshold crossing when the network breaks into two large pieces, and strands part of the road network. Specifically, it is when the size of the second giant component reaches its maximum proportion of the network. In the larger Pamlico and Oriental region analyzed (Figure 11 & 12) are many nodes that – when disconnected – lead to large jumps in the second giant connected component (see the peaks in the grey line of Figure 13 & 14). These peaks occur when important nodes become disconnected. Our analysis is focused on identifying the largest jump, but for such a large area there could be multiple nodes that would be worth monitoring in the future, because when deactivated they cause large breaks in the network.

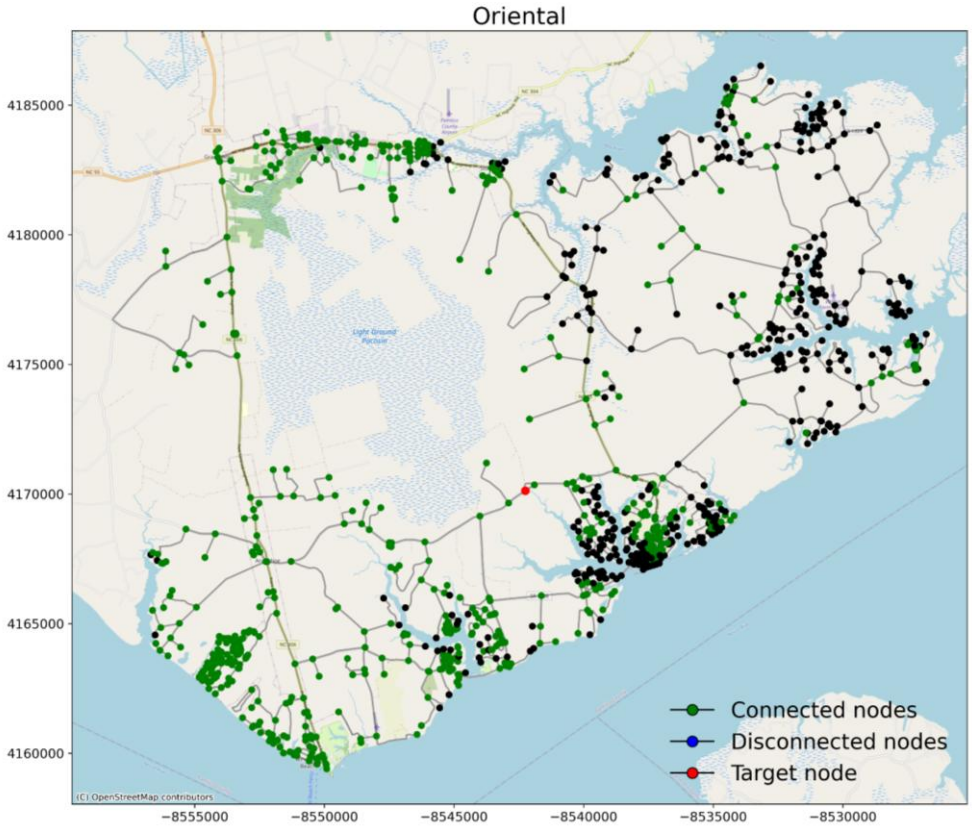


Figure 11. Geographic results and road network diagram for a zoomed in area around Oriental, NC. The critical node is shown in red.

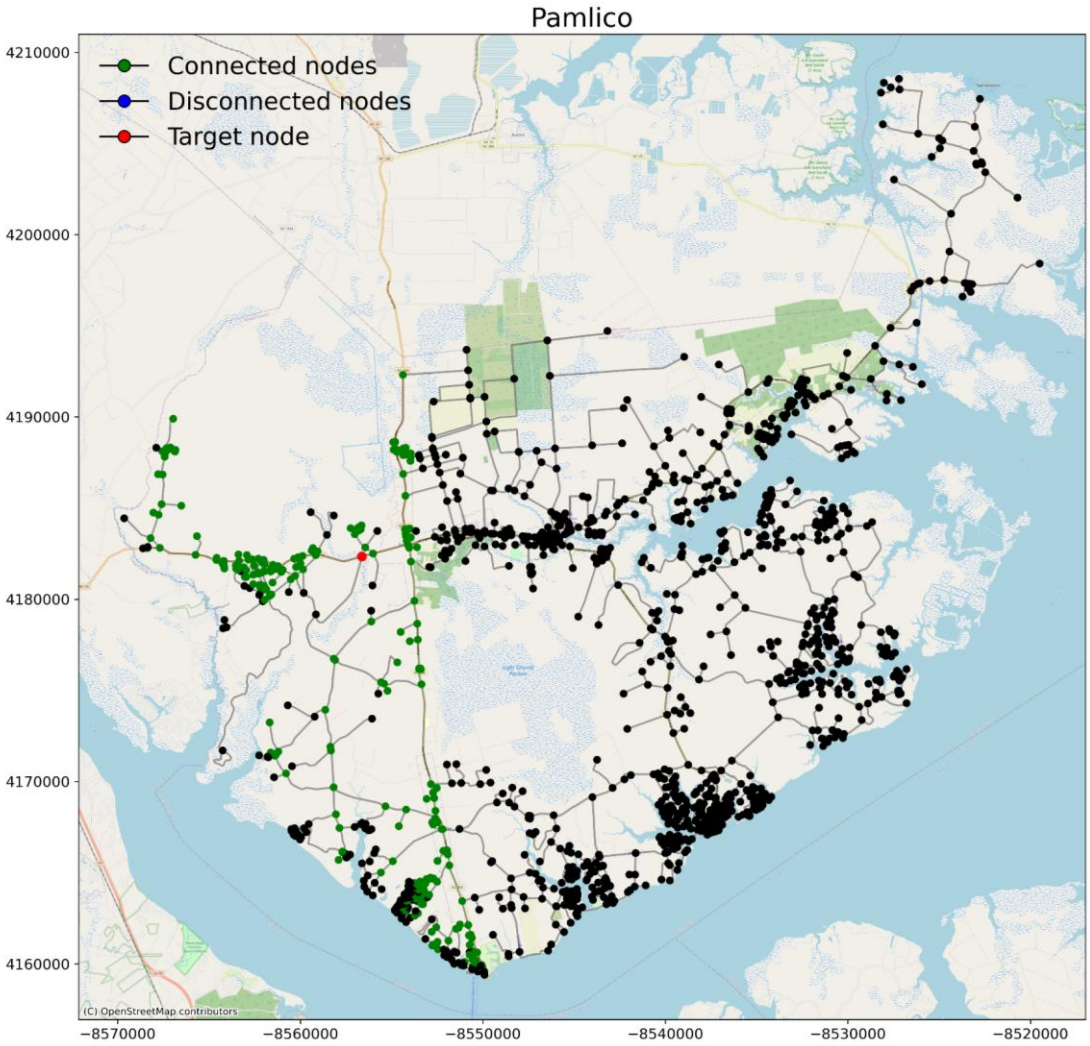


Figure 12. Geographic results and road network diagram for a larger region of Pamlico County. The critical node is shown in red. Note the critical node is different from the zoomed-in region above, highlighting that the scale, size, and geography of the region of interest impact the location of the critical node.

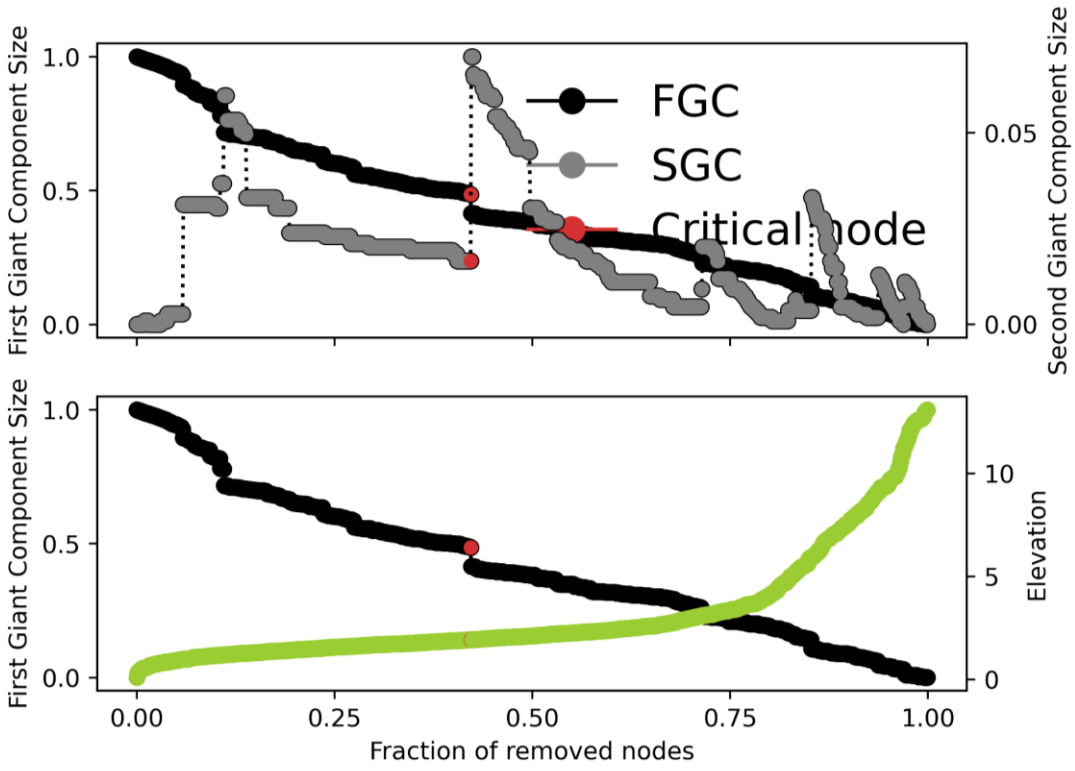


Figure 13. First and second giant-connected component size for the smaller, zoomed-in region around Oriental, NC. Note that the critical node is located at 2 m NAVD88 in elevation. Flooding of this elevation would cause the network to fracture.

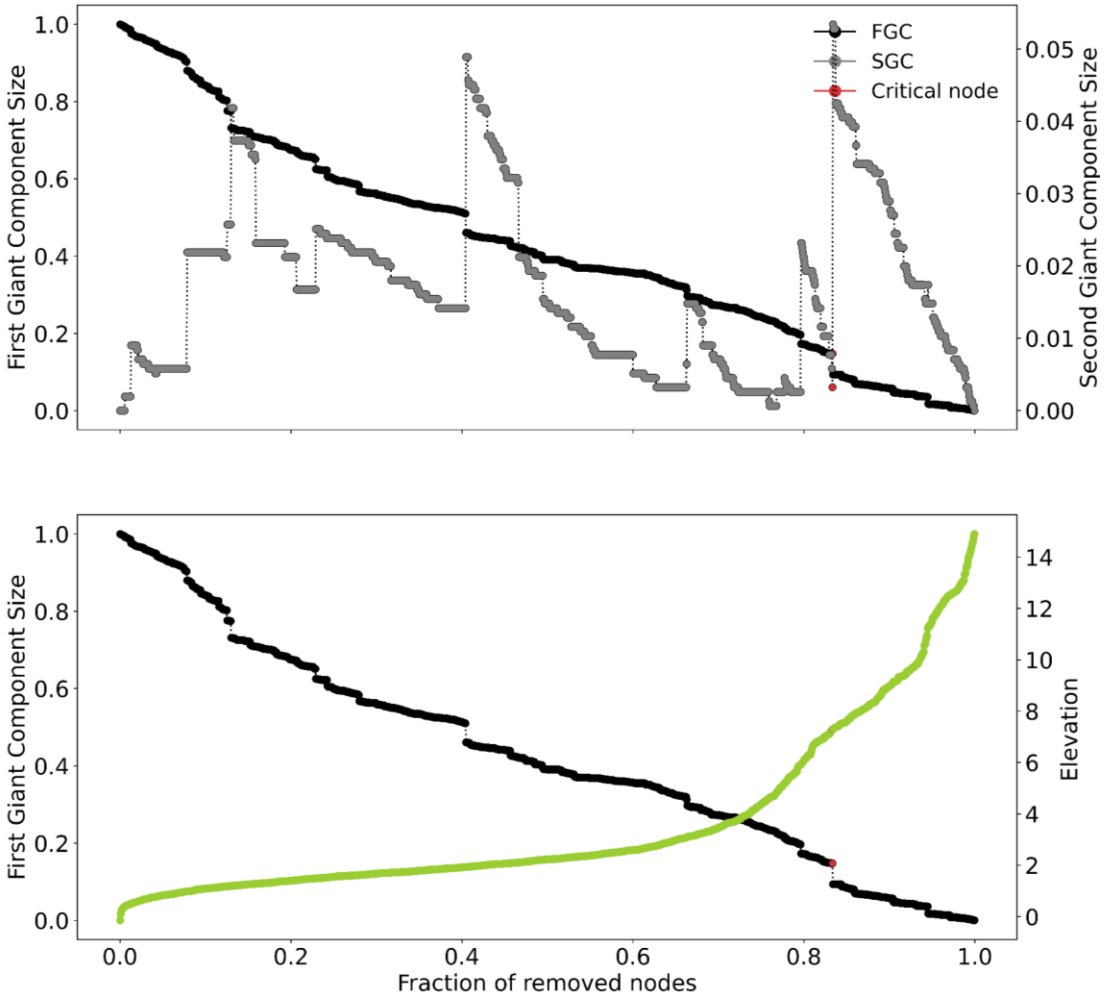


Figure 14. First and second giant-connected component size for the larger Pamlico County region of interest. Note that the critical node is located at 7 m NAVD88 in elevation. Flooding of this elevation would cause the network to fracture. This larger region is much more resilient to flooding, as more roads connect disparate locations in the larger area of interest.

3.3. Davis and Down East Carteret County

We analyzed the Down East region of Carteret County, as well as a much smaller area around Davis, NC and show results below. Each figure below the critical node that leads to the fragmentation of the network (in red), connected nodes (green), and disconnected nodes (blue/black).

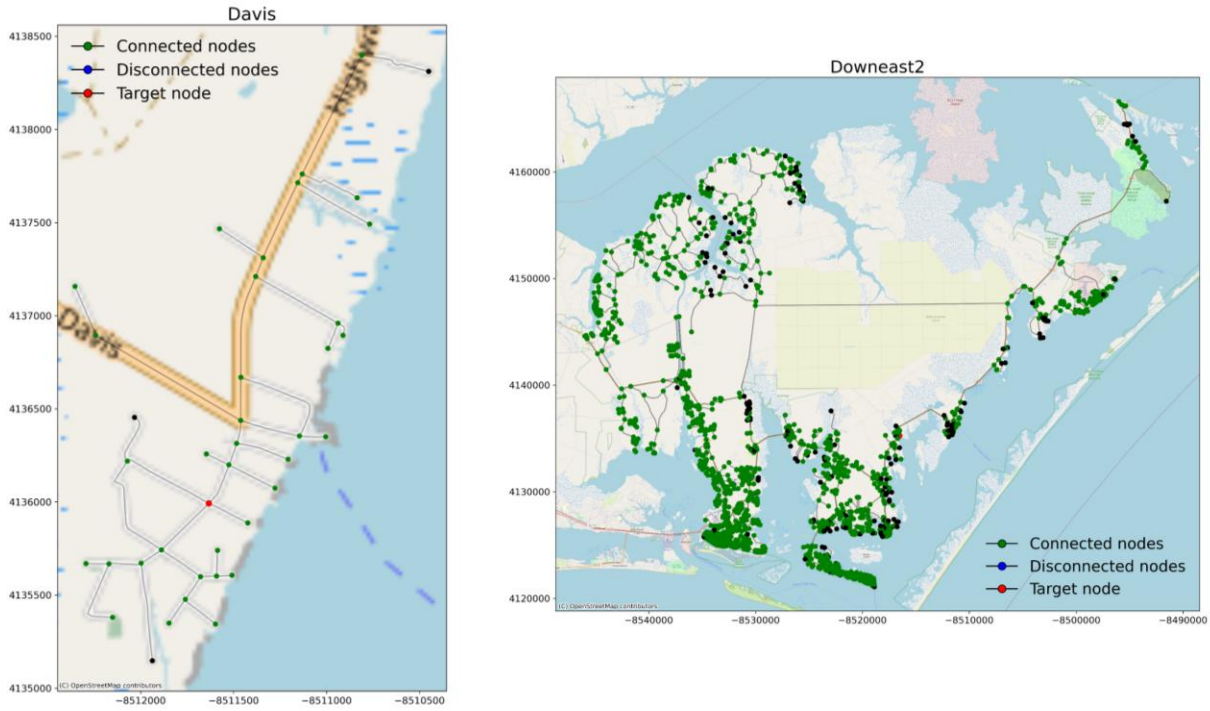


Figure 15. Geographic results and road network diagram for Davis, NC and the larger Down East region. The critical node is shown in red on both maps. Note the critical node is different than the zoomed-in region, highlighting that the scale, size, and geography of the region of interest impact the location of the critical node. The location identified in Davis informed the placement of our SuDS system and TinyCamML device.

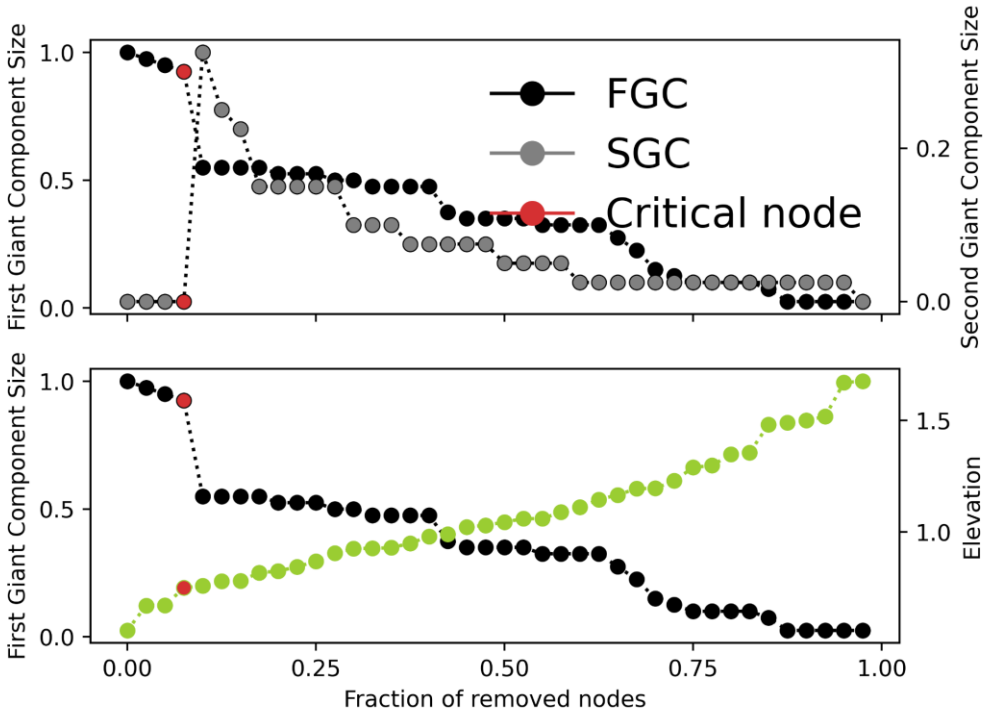


Figure 16. First and second giant-connected component size for the road network around Davis, NC. Note that the critical node is located quite low, ~0.5 m NAVD88 in elevation. Flooding of this elevation would cause the network to fracture.

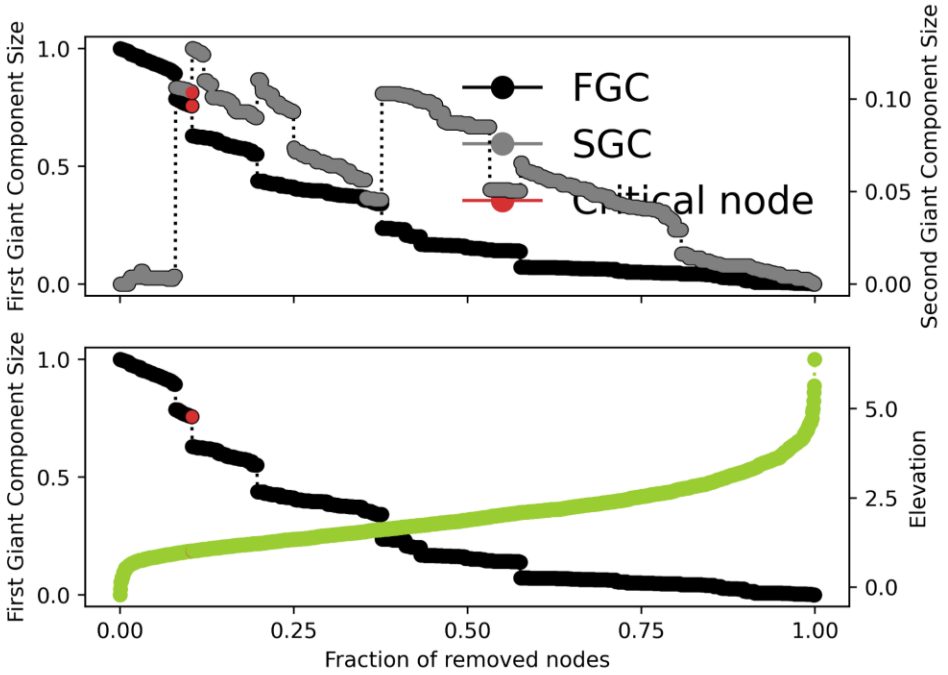


Figure 17. First and second giant-connected component size for the entire Down East region. Note that the critical node is located quite low again, around 1.5 m NAVD88 in elevation. Flooding of this elevation would cause the network to fracture.

3.4. Kinston

While the Kinston region was identified during NCDOT meetings as a potential study site, analysis of the region was not possible with this technique. More work needs to be done to understand how to apply this method to more inland regions that do not have natural boundaries, such as coastal edges or other bodies of water that constrain the region of interest. One possible option identified is to use travel time metrics or drainage area/basin boundaries, but this is beyond the scope of this project.

3.5. Comparison with TIMS

TIMS data was not successfully able to be used to validate the network models. Since this data is human-reported, it is not a surprise that it is too sparse in space and time to validate the sort of very dense network modeling.

3.6. Evaluation of other vulnerability metrics

As discussed during our final Team meeting with NCDOT personnel, due to Dr. Beth Sciaudone's departure from the project, our ability to evaluate other vulnerability metrics, including the distance to the ocean shoreline, could not be met. Additionally, it was deemed that these metrics were not desired by NCDOT personnel as part of the network analysis.

4. Land-based measurements of roadway flooding

Due to the timeline of the project, continuous data sets are just now becoming available in Down East Carteret County for testing the network model. Here, we describe two additional use cases of the flood data collected by sensors supported by this project.

4.1. Sunny Day Flood Sensors: statewide comparison against tide gauges

The following text comes from the paper "Land-based sensors reveal a high frequency of coastal flooding" by Hino, Anarde, et al. (2025), which is *in press* in *Nature Communications Earth and Environment*. All methods related to the statistical analysis of flood and tide gauge data are included in the supplement of this publication and not restated here for simplicity.

In this paper, we provide new estimates of flood frequency in three coastal communities based on a custom, open-source network of sensors on land (Gold et al., 2023) and compare our measures against tide gauge-based estimates. The flood sensors, deployed either in stormwater drains or just adjacent to roadways, measure the presence and depth of water on land from all sources, including tides, wind, waves, rain, local infrastructure, and groundwater. By capturing the full range of floods regardless of source, these measures enable a novel perspective on the frequency of coastal flooding and the relative accuracy of commonly used tide gauge threshold exceedances. Given the growing scientific use of tide-gauge data as a measure of coastal flooding (e.g., Dahl et al., 2017; Hauer et al., 2023; Hino et al., 2019), this analysis provides critical empirical grounding for researchers.

We measured flooding over one year using the SuDS in three coastal communities spanning a range of physical and social settings. Beaufort, North Carolina (NC), USA and Carolina Beach, NC are small urban communities with subterranean drainage infrastructure, while Sea Level, NC is rural and relies on drainage ditches. The three sites also vary in their proximity to tide gauges. The sensors were deployed in flood hotspots identified by community members as causing disruptions. By comparing data across three sensors in a single region, we characterize how flood frequency varies with small changes in geography.

We compare the *in-situ* flood measures (water depth and photographs) to data from the nearest tide gauge to assess how flood frequency varies from tide gauge threshold exceedances. The two thresholds we use are the National Weather Service's minor flood threshold, which is often used to issue coastal flood advisories and warnings, and NOAA's high-tide flood (HTF) threshold. NOAA's HTF thresholds provide a nationally consistent approach based on local tidal ranges but are not calibrated to match flood incidence on land (Sweet et al., 2018). We define flooding as any amount of water on the road, and similarly, any instance of water level above the relevant tide-gauge threshold is classified as an exceedance. The tide-gauge thresholds differ from road elevations because the thresholds are used for purposes other than flood monitoring (e.g., statistical analysis of nationwide trends in Sweet et al., 2018). By analyzing how our sensors and gauge-based indicators diverge, we assess how the impacts of tides, wind, rain, and local infrastructure vary across our study sites.

Over one year from May 1, 2023, to April 30, 2024, we detected 26, 65, and 128 days of flooding on roadways in Beaufort, Carolina Beach, and Sea Level, respectively. Very few of these floods were associated with extreme storm events. Excluding floods coinciding with extreme storms – specifically, Tropical Storm Idalia on Aug 30-31, Ophelia on Sep 22-24, and the Nor'easter on Dec 17-18 – our annual flood frequencies decrease to 20, 60, and 122 days. Herein, we focus on floods and threshold exceedances outside of extreme storms.

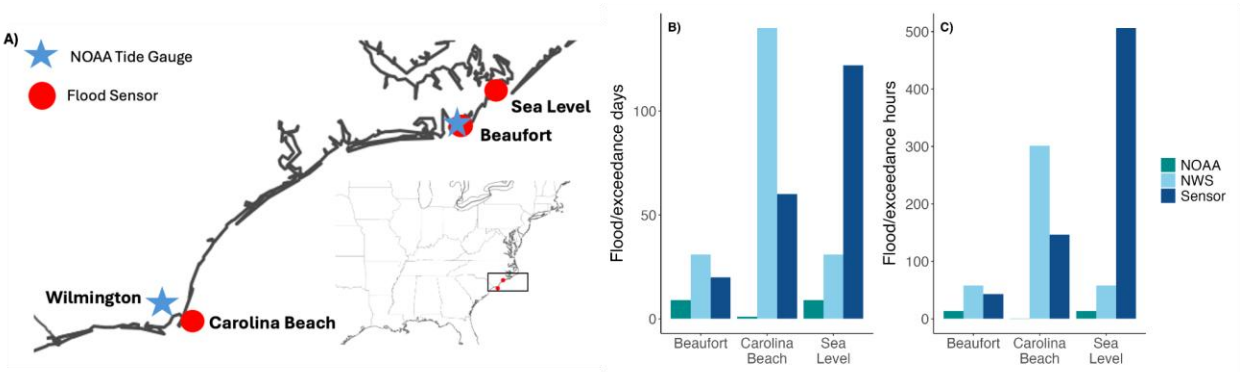


Figure 18. Exceedances of tide gauge thresholds do not accurately reflect flood frequency on land. Exceedances of the NOAA HTF threshold are consistently lower than actual flood events at our three sites, while exceedances of the NWS threshold are more frequent than flooding in Beaufort and Carolina Beach and less frequent than flooding in Sea Level. This pattern applies when measured in days (Panel B) and hours (Panel C). Plots show data for one year from May 1, 2023 - April 30, 2024.

Documented flood frequencies diverge substantially from tide gauge threshold exceedances (Figure 18b-c). At all three locations, flooding occurred more often than NOAA HTF exceedances. Water levels exceeded HTF thresholds on just nine days at the gauge nearest to Beaufort and Sea Level and on one day at the Wilmington gauge near Carolina Beach. While NOAA HTF exceedances consistently underestimate flood frequency, exceedances of the NWS minor flooding threshold show no consistent relationship with observed flooding. In Beaufort and Carolina Beach, the NWS threshold was exceeded more often than flooding occurred, but the opposite was true at Sea Level. Depending on the chosen threshold (NWS or NOAA HTF), tide gauge threshold exceedances could substantially overestimate OR underestimate flood frequency. For example, the Wilmington NWS threshold was exceeded on 140 days, more than twice as often as we documented flooding in Carolina Beach. On the other hand, Sea Level flooded 122 days, compared to 31 days for NWS exceedances and 9 days for NOAA HTF exceedances (using the Beaufort gauge).

Actual flood durations are generally longer than the duration of threshold exceedances for both NWS and NOAA HTF thresholds (Figure 19). Flood durations are also longer in rural settings than in urban settings for the locations examined here. Our average flood duration was 1.71 hours (standard deviation: 1.2 hours) in Beaufort (urban) and 3.35 hours (SD: 2.19) in Sea Level (rural),

compared to average threshold exceedances of 1.25 hours (SD: 0.70) for the NOAA threshold and 1.70 hours (SD: 0.79) for the NWS threshold at the Beaufort tide gauge. Floods averaged 2.28 hours (SD: 0.84) in Carolina Beach (urban), compared to 0.35 hours (SD: 0.35) and 1.91 (SD: 0.99) hours for an average NOAA HTF or NWS exceedance at the Wilmington tide gauge. The differences in measured durations and tide gauge threshold exceedances are due to 1) thresholds not being exceeded during long flood events (discussed in more detail below), and 2) relatively slow drainage of floodwaters on land (over low-sloping landscapes, through stormwater infrastructure, and groundwater infiltration) versus marine water bodies. The slow recession of floodwaters is particularly pronounced in rural settings. For the same flood depths, flood durations at Sea Level are generally 2-3 times longer.

The discrepancies between flooding on land and tide gauge threshold exceedances shown here are not solely due to tide gauge thresholds being too high or too low; rather, they are caused by marine water levels varying over small distances and land-based flood drivers that cannot be captured at tide gauges. The effect of land-based drivers is particularly evident in Beaufort, where the tide gauge and sensor are just 0.6 miles apart. There, the NOAA HTF threshold is 3.23 ft, the NWS minor flood threshold is 2.92 ft, and the elevation of the road is 2.58 ft (all referenced to NAVD88). Using the NOAA HTF threshold, as shown in Figure 19, several flood events are only detected by the sensor (purple dots), suggesting the threshold may be too high. However, there are also five "gauge-only" events (i.e., instances where the tide gauge exceeds the threshold but no flooding is observed at the sensor location). If we lower the threshold to match the road elevation, the number of gauge-only events increases to 76. Some of these gauge-only events are likely due to a backflow prevention device at the outfall of the monitored storm drain, which limits water moving through the pipes onto the road when marine water levels are high. The effects of stormwater infrastructure cannot be captured at gauges, so modifying the tide gauge threshold does not necessarily improve the fit between threshold exceedances and actual flooding. Modifying the tide-gauge threshold to match the road elevation in Sea Level, which is approximately 20 miles from the nearest tide gauge, causes similar mismatches described in more detail below.

In addition to local infrastructure, non-tidal flood drivers contribute substantially to the mismatch between actual flooding and tide gauge threshold exceedances. Our site-specific studies that utilize data from the same flood sensors in Beaufort (Gold et al., 2023) and Carolina Beach (Thelen et al., 2024) show that rain, falling on stormwater networks impaired by high marine water levels, contributes significantly to flooding (for the periods analyzed, 30% of floods in Beaufort and 35% in Carolina Beach had contributions from rainfall). The impacts of rain and local infrastructure on flooding cannot be captured at tide gauges, so modifying the tide gauge threshold would not resolve this discrepancy.

Finally, variation in marine water levels across small distances also drives discrepancies between observed flooding and tide gauge threshold exceedances. Wind commonly increases water levels in the northeast U.S. (Li et al., 2022), and in NC, small changes in geography modify the relative importance of wind to local water levels. The Beaufort sensor and tide gauge are approximately 19 miles from the Sea Level sensor, but flood frequency differs by an order of

magnitude: 20 days in Beaufort and 122 days in Sea Level. Sea Level is on a low-lying peninsula within an elongated estuary where water levels can pile up from wind forcing from several directions (Figure 18a). During two long floods at Sea Level that were *not* accompanied by a NOAA HTF exceedance, wind speeds averaged over 20 mph from the southwest during one event and the northeast during the other.

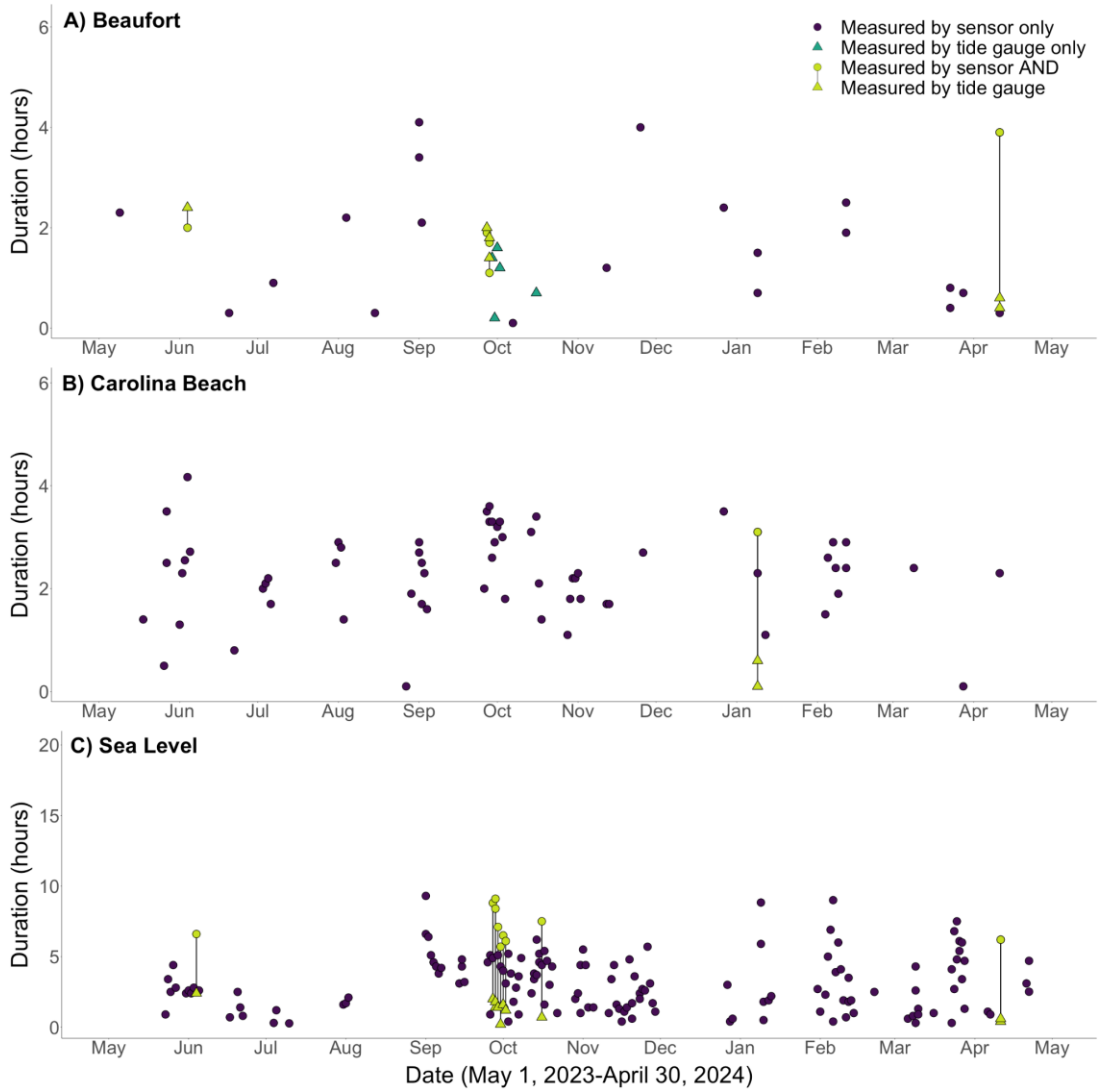


Figure 19. A comparison of flood events and NOAA HTF threshold exceedances shows frequency and duration mismatches between data across all seasons in (a) Beaufort, (b) Carolina Beach, and (c) Sea Level. Plots compare the incidence and duration of events as documented by the flood sensors (circles) and as inferred by tide gauge exceedances using the NOAA HTF thresholds (triangles). Circles are purple for events that were only captured by the flood sensors. Connected light green circles and triangles represent events that were captured by the flood sensors and also inferred by NOAA HTF exceedances. Triangles are dark green for events only inferred by NOAA HTF exceedances (not measured by flood sensors).

Because of local variation in water levels, modifying the tide-gauge threshold to match the Sea Level roadway elevation does not resolve the mismatch between tide-gauge threshold exceedances and flood events. The road at Sea Level is at 1.62 ft NAVD88, significantly lower than the tide-gauge thresholds (2.92 ft for NWS and 3.23 ft for NOAA). The tide gauge exceeded 1.62 ft NAVD88 on 316 days during our one-year study period (excluding extreme storm days), while flooding only occurred on 122 days. Hence, marine-based contributions to floods vary substantially even over small distances, and they are not adequately captured by a sparse gauge network.

Our results demonstrate the many benefits of measuring water levels on land rather than relying on tide gauge-based estimates for understanding community exposure to coastal flooding. Data from custom sensors show that tide gauge threshold exceedances can underestimate flood frequency by an order of magnitude and that both the timing and duration of events are not accurately reflected in tide gauge data. Studies from other regions using land-based observational data have found similarly high frequencies of coastal flooding, suggesting that tide gauge threshold exceedances are likely problematic proxies for floods elsewhere as well (Mydlarz et al., 2024; O'Donnell et al., 2024; Kang et al., 2024). Neither the NOAA nor NWS thresholds are designed to perfectly reflect flood incidence; they are used by the respective agencies for different purposes, and regional tide gauges cannot capture hyper-local variation in flood frequency. Moreover, regardless of the threshold, tide gauges will not reflect the influences of rain and local infrastructure on flood incidence. Our findings highlight that tide gauge-based indicators are problematic as measures of coastal flood frequency, not because of the specific threshold, but because of the other flood drivers that are not captured by tide gauges.

Land-based measures of flooding can capture the effects of local drainage infrastructure, rain, and groundwater, all of which are integral to accurate flood depth and duration measures. Such measures can also provide critical input to build on existing coastal water level modeling efforts, many of which capture several, but not all of these drivers (Serafin and Ruggiero, 2014; Rinaldo et al., 2020; Rose et al., 2024; Thelen et al., 2024). Remotely sensed observations of non-storm flooding can enable characterization of extent at regional scales, building on the point-source measures described here. At the same time, other observational data streams to capture marine water levels, wind, and rain are needed to decompose location-specific flood drivers. For example, it is possible that due to the condition of local infrastructure, flooding is driven entirely by rain ponding and has no relationship to tides or other marine drivers. Richer observational records can also help inform the development of more accurate coastal flooding thresholds (e.g., Mahmoudi et al., 2024).

Coastal flooding will become much more frequent and widespread due to sea-level rise (Sweet et al., 2022), and evidence of the incidence and impacts of such floods is critical for effective adaptation. Developing and using land-based flood measures is needed to more accurately reflect the true exposure and experiences of coastal residents and to identify areas most affected by chronic flooding.

4.2. TinyCamML proof-of-concept: November 2024

The following text comes from the manuscript "Detection of coastal flooding with TinyCamML: a low-cost, privacy-preserving camera with onboard ML" by Farquhar, Goldstein, et al. (in prep) which will be submitted to *Earth and Space Science* in May 2025. All methods related to this publication are included in Section 2.

We demonstrate the utility of the TinyCamML in sensing roadway flooding during a series of floods in Carolina Beach, NC, that took place from November 14–18, 2024. Four TinyCamMLs were deployed along a single roadway: Canal Drive. The floods were driven in part by perigean-spring tides (when the moon is either new or full and closest to Earth in its orbit), as well as rainfall. From November 11–18, 1.8 cm of precipitation fell across the region (as measured by the Coastal Ocean Research Monitoring Program Masonboro Island weather station, located about 5 km from Canal Drive), with 1.63 cm inches of rain falling on November 14 alone, which contributed to several compound flood events.

We compare the ML classification (e.g., "flood" or "no flood") of each TinyCamML against user (visual) confirmation of flooding from each photo. As was the case in the training data, an image was labeled as "flooded" when the water extended across the crest of the road such that a car could not use the roadway without driving through water. The TinyCamMLs were deployed at three locations along a 1 km stretch of Canal Drive where there are existing water level sensors located in storm drains (i.e., part of the SuDS network; Gold et al., 2023). The water-level data from each gauge is plotted against the TinyCamML data in Figure 21 to provide context for flood depth and duration.

Example imagery from the TinyCamMLs during various conditions throughout the deployment is shown to demonstrate the effect that environmental conditions (e.g., rain, sunlight) may have on the image classification (Figure 20). The photo in I shows a correct classification of "flood" by TinyCamML 2 during a clear sunny day (tidal flood event). In II, the image was taken at the transition between "no flood" and "flood," where the floodwaters had just barely extended over the crest of the roadway on the right side of the photo (the yellow solid line). TinyCamML 1 classified this image as "no flood," when labeler identification would consider this a "flood." In III, there are puddles on the roadway that we do not consider to be a "flood" because they do not extend across the crest of the road, but TinyCamML 3 classified it as one. Image IV demonstrates how water droplets look when impeding the TinyCamML's view (compound rainfall event); however, despite this, TinyCamML 4 correctly classified this image as "no flood."

We compare time series of water levels from SuDS sensors (mounted within storm drains) against the classifications of the TinyCamMLs and visually-confirmed user-labeled images (Figure 21). There are large gaps in the reporting records of TinyCamMLs 1 and 2 because they were earlier designs and therefore had issues with reporting consistency. Since then, all TinyCamMLs have been updated to the most recent design and are performing as expected. We include example imagery from TinyCamMLs 1 and 2 here because their field of view is distinctly different from the other cameras. Collectively, for the data shown in Figure 21, the TinyCamMLs

classified images as “flooded” when there was a visually-confirmed flood 90% of the time, excluding “too dark” imagery. When there was not a visually-confirmed flood, the TinyCamMLs reported “flood” 26% of the time (i.e., a “false positive”).

Figure 20. Example imagery from the TinyCamMLs in different flood conditions during November 14-18, 2024 along three different locations on Canal Drive in Carolina Beach, NC. TinyCamML 1 and 2 were at the same location facing different directions

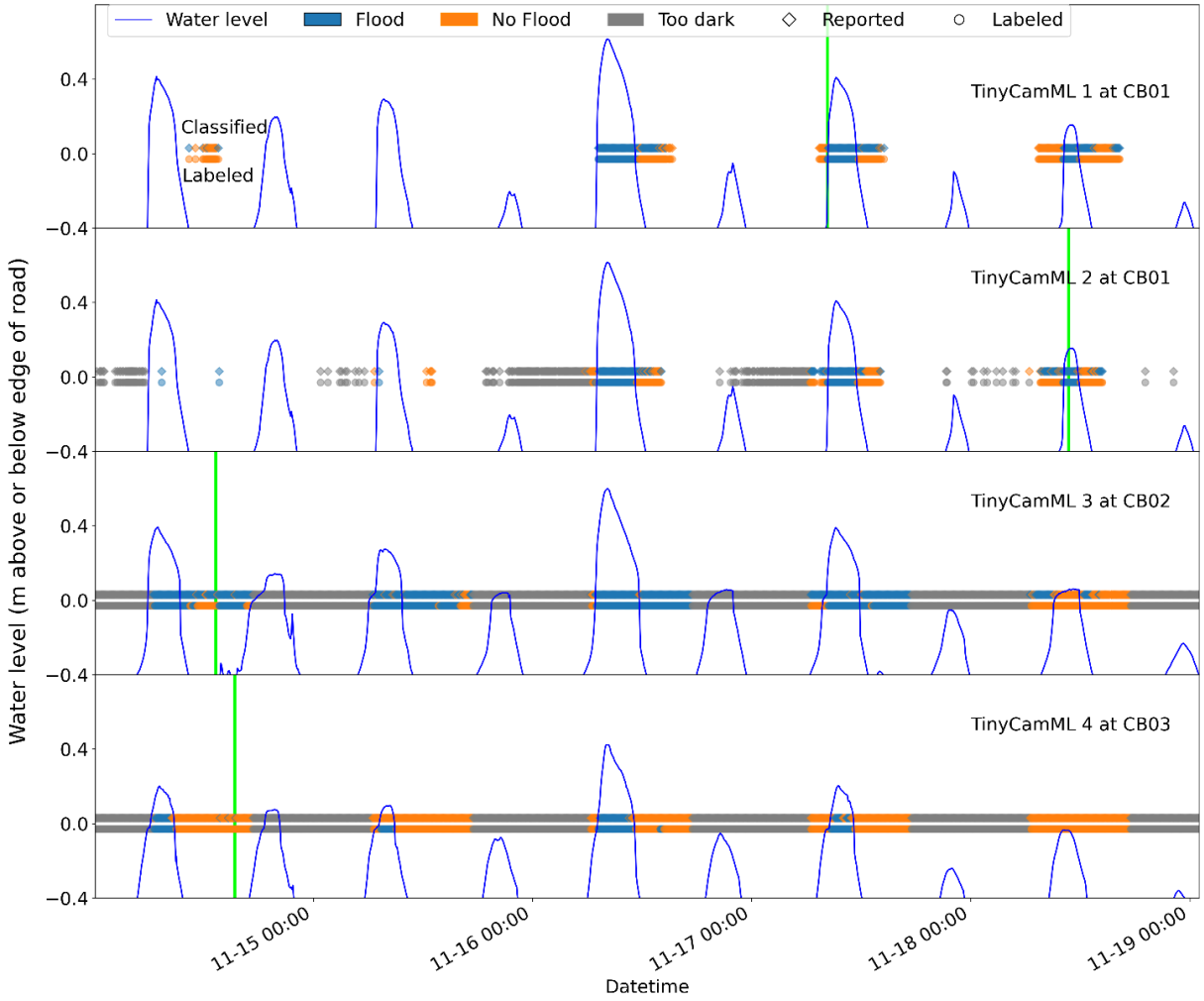


Figure 21. TinyCamML classifications compared with the user-labeled images, plotted against water levels above and below the edge of the pavement, recorded within storm drains (Gold et al., 2023). The green vertical lines indicate when the images in Figure 20 were taken.

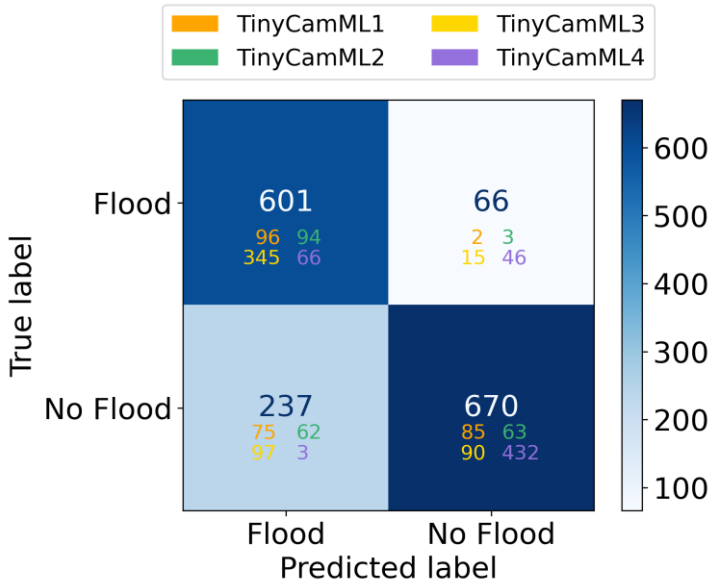


Figure 22. Confusion matrix of the classifications by the TinyCamMLs during the November 14-18, 2024 deployment. The larger number in each box indicates the sum of the number of classifications by each individual TinyCamML, which is color coded.

We use a confusion matrix to further assess the performance of the ML model for the 1574 images taken during the November 14-18, 2024, deployment and show aggregate data as well as individual camera data (Figure 22). Across all four cameras, this deployment had an 81% accuracy, a 72% precision for detecting floods, and a 90% recall for detecting floods. When looking at all four cameras, the model was generally strong at predicting floods with only 66 false negatives (i.e., instances where there was flooding but the model predicted "no flood"). However, the model tended to overpredict flooding with 237 false positives (i.e., instances of no flooding where the model predicted "flood"). This general trend was seen in the data from TinyCamMLs 1-3. TinyCamML 4 had many false negatives and very few false positives.

Lastly, we performed an inter-rater agreement experiment with two labelers for this dataset to better understand consistency in labels. This dataset and its labeling was not used for training the classification model. We used Krippendorff's Alpha to calculate agreement, where a score of -1 indicates no agreement, a score of 1 indicates total agreement, and a score of 0 being chance agreement (Krippendorff, 1970). TinyCamMLs 1-4 had a Krippendorff value of 0.98, 0.92, 0.59, and 0.94, respectively.

Here we demonstrated the ability of the TinyCamML, a new open-source ML camera that both preserves privacy and offers low data transmission cost, in classifying roadway flooding under variable environmental conditions. Four TinyCamMLs deployed along a single road in Carolina Beach, NC, successfully identified flooding over a four day period 90% of the time (as compared to visually-confirmed user-labeled images; Figure 21). However, the TinyCamMLs also

reported false positives; the classification model reported flooding when there was not visually-confirmed flooding 26% of the time. The on-device ML model is continually being trained and improved upon, so we expect the number of false positives (and false negatives) shown in Figure 22 to decrease as more data is captured and used for training. Based on the confusion matrix, the models are generally performing in a manner where they do not miss detecting a flood (low false negatives), but tend to have false alarms (higher false positives). Depending on the specific deployment scenario, this may be more or less preferable.

The accuracy of the TinyCamML in predicting roadway floods was only calculated for a single definition of flooding – that is, when the water extended across the crest of the roadway such that a car could not pass without going through water. Using our definition of flooding, there were instances where the TinyCamMLs did detect water on the road and reported "flood," but because the water was not fully covering the roadway, the labeler determined those instances as "no flood," such as in image III of Figure 20. The TinyCamMLs reported false negatives 10% of the time, but again, many of these instances occurred during the transition between the roadway being dry to fully inundated at the centerline (Figure 20, image II). While the model performance is indeed sensitive to our chosen definition of flooding, the definition used here provides insight into when a roadway potentially becomes a depth hazard for vehicles (e.g., in 30-40 cm of water, a vehicle can float; Martínez-Gomariz et al., 2016), or for the case of tidal floods, when damage to vehicles is likely due to splashing from saltwater. Images that showed persistent puddles and/or this transition seem to be edge cases for the TinyCamML that will likely improve with further ML model training. Importantly, model accuracy, as well as the number of images needed to train the model to reach that accuracy, may differ for other definitions of flooding (e.g., any amount of ponding on the road, as used by Hino et al., 2025).

Overall model accuracy is also highly dependent on the deployment location and field of view. TinyCamML 3 had the highest individual percentage of false positives, but its field of view contained many large persistent puddles. As shown in Figure 21, TinyCamML 3 reported "flood" even after the water levels had receded because these puddles stayed on the roadway. For our interrater experiment with non-training data, TinyCamML 3 also had the lowest Krippendorff value (0.59) between labelers, which is likely also due to persistent puddles and the difficulty of determining exactly when the puddles connect enough to become classified as a flood. Despite this, a Krippendorff value of 0.59 is still comparable to other interrater agreement experiments in the coastal sciences (Goldstein et al., 2021). For the images from TinyCamMLs 1, 2, and 4, the labelers showed near complete agreement (0.98, 0.92, and 0.94), and these fields of view were typically devoid of persistent puddles.

More work is needed to determine how well our ML model generalizes to other environments that experience roadway flooding (e.g., in rural communities or more urban settings), and how the model accuracy might change under different environmental forcing conditions. Here, we showed that the TinyCamMLs can successfully classify roadway floods even during rainfall events when droplets obstruct the images (e.g., for TinyCamML 3 and 4 on

November 14, 2024; Figure 21). It is unknown how changes in solar glare, shadows from large buildings, or vegetation changes may influence our model classifications and accuracy. Future deployments will be tailored toward development of a large, generalized training dataset with many different fields of view and spanning more environmental variables.

We envision several use cases for edge ML devices like the TinyCamML beyond spatial identification of roadway floods. First, since our device reports its classifications in real time to a website, and because the devices are low-cost, a network of TinyCamMLs can be used for real-time monitoring of transportation hazards. In the context of flooding, this type of sensor network could be used for real-time routing of emergency vehicles around flooded areas, and provide data to validate models of the impacts of flooding on road networks (e.g., Aldabet et al., 2022). The TinyCamML hardware and classification model can also be adapted to other monitoring and measuring tasks focused on difficult-to-observe, or ephemeral, environmental phenomena. In coastal settings, this could include binary classification of dune erosion, extreme run-up, impacts to structures, and other storm driven processes. In non-coastal settings, TinyCamMLs may be modified to observe extreme water levels in streams (similar to Latham et al., 2025 and Loftis et al., 2018), identify landslides, and early detection of wildfires (e.g., Shi et al., 2020), avalanches (e.g., Fox et al., 2024), or a range of other processes.

As coastal communities will experience more chronic flooding with rising sea levels, data on flood incidence, extent, and duration will become increasingly important for informing risk assessments and developing flood mitigation strategies (Albano et al., 2017; Van Alphen et al., 2009). The TinyCamML is an important advancement in low-cost, privacy preserving monitoring technology that enables identification of flooding in places where investment in in-situ sensors may not have been historically prioritized: on land or in residential areas, where people interact with floodwaters most.

5. Machine learning tools for flood identification from imagery

The following text comes from the manuscript "Quantification of chronic coastal flooding: a machine-learning driven approach to water level extraction" by McCune, Anarde et al. (in prep) which will be submitted to *Water Resources Research* in August 2025. All methods related to this publication are included in Section 2.

5.1. Segmentation model

Results from training the segmentation model demonstrated that the model showed good performance - for the validation subset, the Mean Intersection over Union (IoU) score was 0.692 (varies from 0-1, where 1 is perfect). The mean Dice scores were 0.739 (varies from 0-1, where 1 is perfect), and the Mean KLD scores were 0.991 (varies from 0-1, where 1 is perfect).

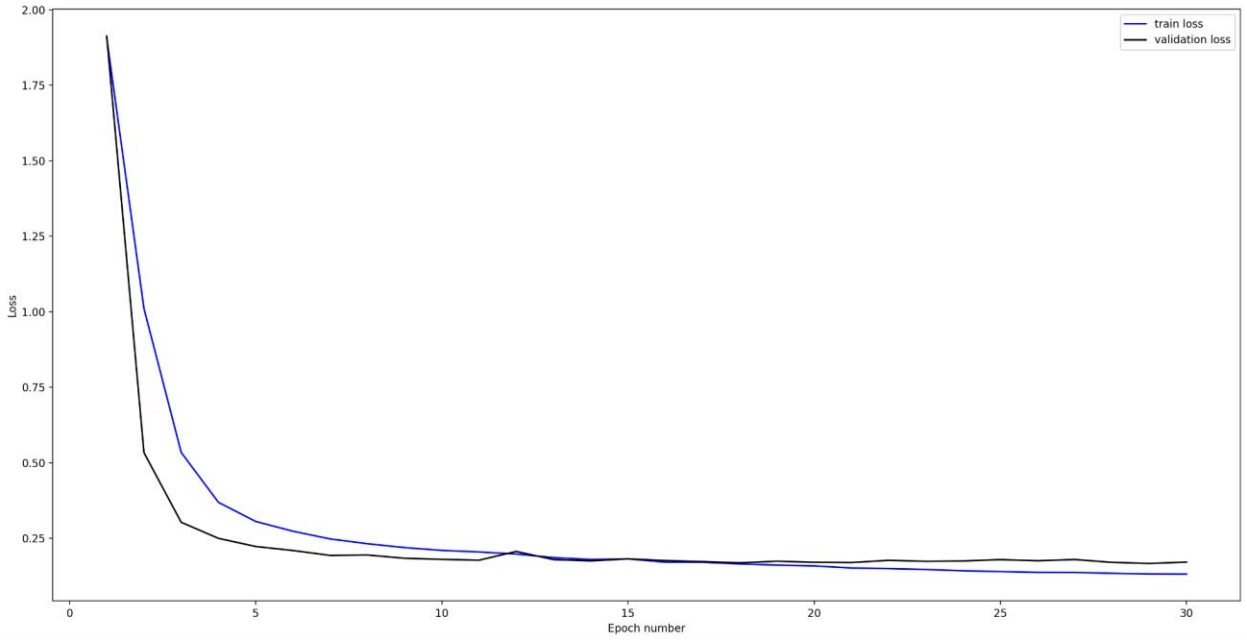


Figure 23. Training loss curve for the segmentation model over 30 training epochs with training loss noted in blue and validation loss noted in black.

An example model output can be shown below. These results are promising and demonstrate the model can capture roadway flooding from SuDs Imagery. Individual models made for each of the three sites (Down East, Carolina Beach, and Beaufort) show comparable results.

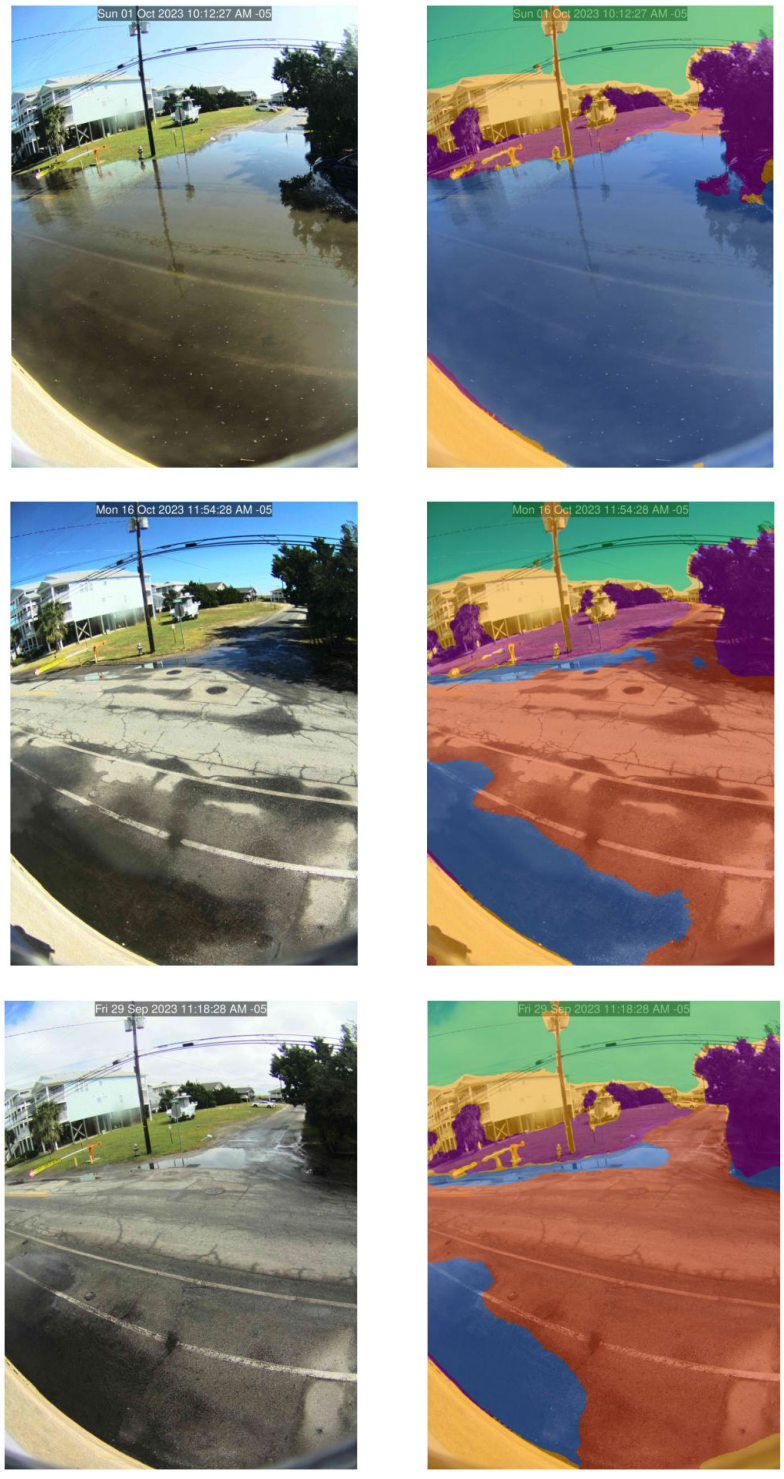


Figure 24. Example segmentation model results from Carolina Beach SuDS camera (left) and the segmentation label overlay created by the ML model (right): red is road, blue is water, yellow is building, purple is vegetation, and green is sky.

5.2. Spatial extent of flooding

Quantification of flood spatial extent of imagery is shown below using two methods. The first is a simple pixel-space estimate of flooding from segmented imagery while the second is a calculation from the geo-rectified imagery produced during the depth mapping method described in Section 2.4. The pixel-based estimate counts the number of pixels identified in the segmented image as water on the roadway. One drawback to the pixel-based approach is that the amount of area represented by each pixel changes as you move further away from the camera, meaning that pixels far away from the camera when flooded indicate a greater flooding extent than the same number of pixels close to the camera. So, this method can provide a misleading representation of flooding extent. For example, in the result provided below, the smaller cross-street extending away from the camera can be impassable with only about 30% of all roadway pixels in the image predicted to be water, whereas the main street running across the field of view is entirely flooded at about 90% of roadway pixels predicted to be water. This result could be further refined for specific locations by creating transects of pixels on portions of the roadway image to track the inundation of those roadway transects as percentages.

The second method calculates the spatial extent of the flood by counting the number of grid cells identified as water on the grid of a known resolution created as part of the geo-rectification process. This is a result of real-world coordinates with a defined area quantity (e.g. square meters) rather than the simple percentage obtained through the pixel-space approach.

Figure 25. An original image from Carolina Beach (left) and the segmentation overlay (right) with the total number of pixels identified as water and the percentage of roadway pixels predicted to be water.

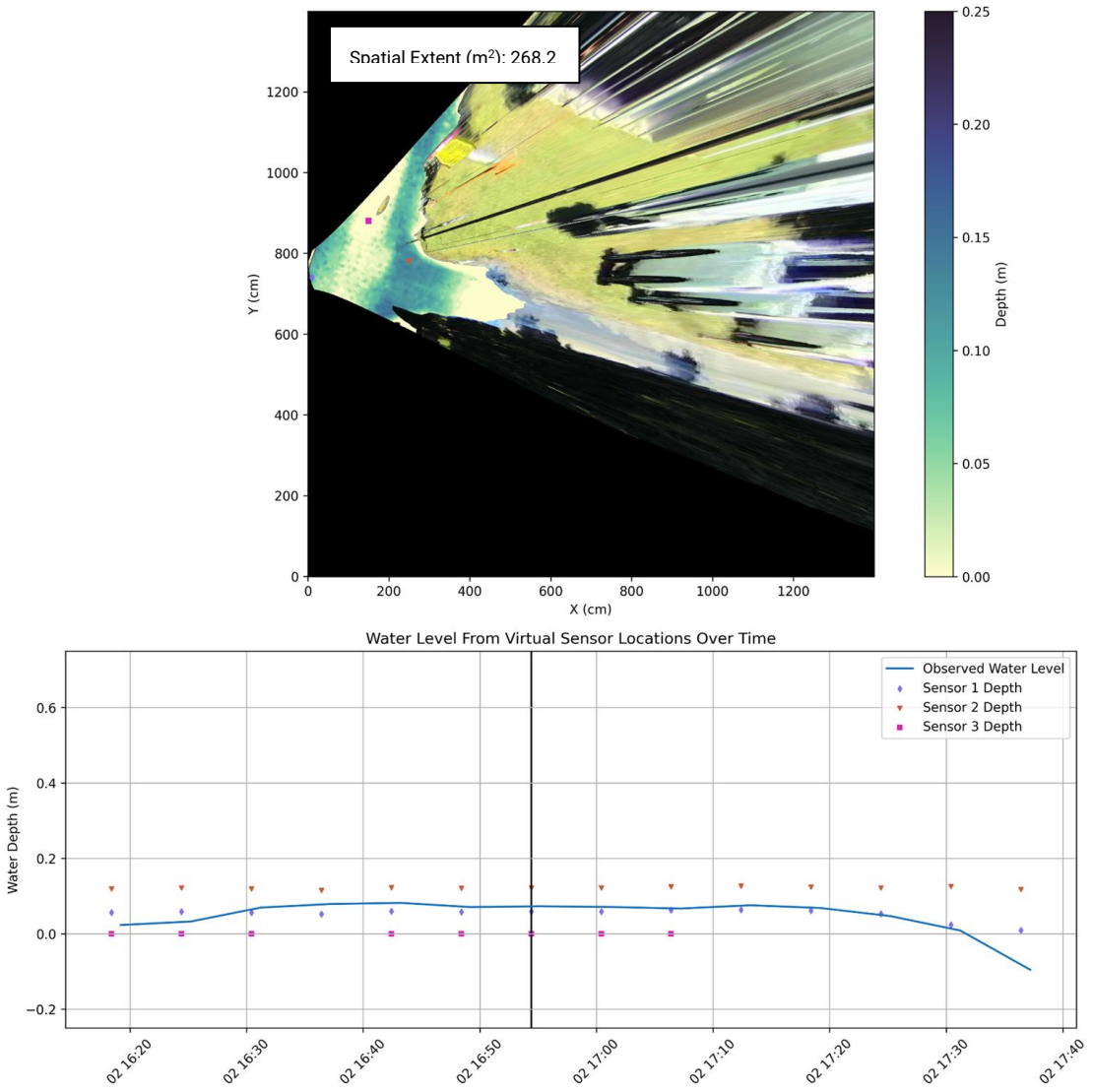


Figure 26. (a) Depth map of flooding in real-world coordinates on a 5 cm resolution grid. Spatial extent is denoted in square meters in the upper left and the color bar indicates depth in meters. Virtual sensor locations are shown with unique symbols. **(b)** The plot of the water level time series with the solid blue line indicates observations from the SuDS. Virtual sensor water depths are plotted using their unique symbols from the map in **(a)**. The purple diamond is the virtual sensor stationed as close to the SuDS as possible in the field of view.

5.3. Mapping flood depths spatially

Results for the flood depth mapping approach are shown for Carolina Beach above. As discussed above, the spatial extent of flooding can be calculated for each image and is displayed in the upper left corner of each output. The flood depth map generated following the methodology discussed in Section 2 is also shown. An additional utility of this method is the ability to track water levels at "virtual sensors" which is a matter of tracking the water depth at a particular grid location throughout of the imagery. In the example result below, we show how these virtual sensors can be located at any point of interest in the image (e.g. in the center of the roadway). This provides a similar water level time series as other traditional approaches to observing roadway inundation and can be used to track critical points of interest within the field of view.

6. Outcomes

In this project we used network science to identify key intersections in coastal roadway networks that if inundated, result in fracturing of the network. Through this work, we provided NCDOT with a methodology that can be used in planning efforts to identify nodes/intersections in any road network – including inland areas – that enhance vulnerability of the entire network based on a single metric (here, elevation). Additionally, we provided new land-based measures of roadway floods with in situ sensors and cameras, and new methods that automate detection of flooding on roadways from images using computer vision and machine learning. The in situ measures of flooding provide real-time information of flooding on land in coastal communities across the state (sunnydayflooding.com). These land-based sensors reveal that it is flooding on roadways ~25-125 days per year in coastal communities, which is far less than what is inferred based on tide gauge data (~1-10 days per year). This means that roadway accessibility is impaired much more frequently than previously thought.

Products that emerged from this work include novel software and models, maps, data and monitoring networks, peer-reviewed publications and manuscripts, as well as professional development. Specifically, the outcomes include:

1. Three open-source models:
    - Road networks model: https://github.com/NCRoadNetworks/NC_barrier_roadnetworks
    - TinyCamML binary classification model (flood/no flood): https://github.com/TinyCamML
    - Image segmentation model for roadway floods (% flooded): https://github.com/Doodleverse/segmentation_gym
2. Maps and rankings of vulnerable network intersections based on our analysis
3. New camera feeds and water level sensors
    - Water-level data is integrated into FIMAN
    - New website, with input from NCDOT personnel: sunnydayflooding.com
4. Two chapters in PhD theses, 5 undergraduate research projects
5. Leveraged funding

The following outcomes are still in development:

1. Depth map algorithm: https://github.com/rtmccune/depth_mapping
    - Description of limitations
    - Utility at other locations: Sea Level & New Bern

7. References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., … Zheng, X. (2016). TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems (Version 2). arXiv. https://doi.org/10.48550/ARXIV.1603.04467

Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.

Bruder, B. L., & Brodie, K. L. (2020). CIRN Quantitative Coastal Imaging Toolbox. *SoftwareX*, *12*. https://doi.org/10.1016/j.softx.2020.100582

Buscombe, D., Goldstein, E. B., Sherwood, C. R., Bodine, C., Brown, J. A., Favela, J., Fitzpatrick, S., Kranenburg, C. J., Over, J. R., Ritchie, A. C., Warrick, J. A., & Wernette, P. (2022). Human-in-the-Loop Segmentation of Earth Surface Imagery. Earth and Space Science, 9(3), e2021EA002085. https://doi.org/10.1029/2021EA002085

Buscombe, D., & Goldstein, E. B. (2022). A Reproducible and Reusable Pipeline for Segmentation of Geoscientific Imagery. *Earth and Space Science*, *9*(9), e2022EA002332. https://doi.org/10.1029/2022EA002332

Dahl, K. A., Spanger-Siegfried, E., Caldas, A., & Udvardy, S. (2017). Effective inundation of continental United States communities with 21st century sea level rise. *Elementa Science of the Anthropocene*, *5*(37). https://doi.org/10.1525/elementa.234

David, R., Duke, J., Jain, A., Reddi, V. J., Jeffries, N., Li, J., Kreeger, N., Nappier, I., Natraj, M., Wang, T., Warden, P., & Rhodes, R. (2021). TensorFlow Lite Micro: Embedded Machine Learning for TinyML Systems. Machine Learning and Systems, 3. https://proceedings.mlsys.org/paper_files/paper/2021/hash/6c44dc73014d66ba49b28d483a8f8b0d-Abstract.html

Farquhar, E., Goldstein, E., Bresnahan, P. J., Settin-Pitarra, B., Stasiewicz, J., Anarde, K. Detection of inland coastal flooding with TinyCamML, a low-cost, privacy-aware camera with onboard ML and cellular connectivity. (in prep for Earth & Space Science).

Hauer, M. E., Mueller, V., & Sheriff, G. (2023). Sea level rise already delays coastal commuters. *Environmental Research: Climate*, *2*(4), 045004. https://doi.org/10.1088/2752-5295/acf4b5

Hino, M., Belanger, S. T., Field, C. B., Davies, A. R., & Mach, K. J. (2019). High-tide flooding disrupts local economic activity SI. *Science Advances*, *5*(2), eaau2736. https://doi.org/10.1126/sciadv.aau2736

Hino, M., Anarde, K., Fridell, T., McCune, R., Thelen, T., Farquhar, E., Woodard, P., & Whipple, A., in press. Land-based sensors reveal high frequency of coastal flooding. Nature Communications Earth and Environment.

Gold, A., Anarde, K., Grimley, L., Neve, R., Srebnik, E. R., Thelen, T., Whipple, A., & Hino, M. (2023). Data From the Drain: A Sensor Framework That Captures Multiple Drivers of Chronic Coastal Floods. Water Resources Research, 59(4), e2022WR032392. https://doi.org/10.1029/2022WR032392

Goldstein, E. B., Buscombe, D., Lazarus, E. D., Mohanty, S. D., Rafique, S. N., Anarde, K. A., Ashton, A. D., Beuzen, T., Castagno, K. A., Cohn, N., Conlin, M. P., Ellenson, A., Gillen, M., Hovenga, P. A., Over, J. R., Palermo, R. V., Ratliff, K. M., Reeves, I. R. B., Sanborn, L. H., … Williams, H. E. (2021). Labeling Poststorm Coastal Imagery for Machine Learning: Measurement of Interrater Agreement. Earth and Space Science, 8(9), e2021EA001896. https://doi.org/10.1029/2021EA001896

McCune, R., Anarde, K., Goldstein, E., & Baker, C. Quantification of chronic coastal flooding: a machine-learning driven approach to water level extraction. (in prep for Water Resources Research).

Kang, B., Feagin, R. A., Huff, T., & Durán Vinent, O. (2024). Stochastic properties of coastal flooding events – Part 1: Convolutional-neural-network-based semantic segmentation for water detection. Earth Surface Dynamics, 12(1), 1–10. https://doi.org/10.5194/esurf-12-1-2024

Krippendorff, K. (1970). Estimating the Reliability, Systematic Error and Random Error of Interval Data. Educational and Psychological Measurement, 30(1), 61–70. https://doi.org/10.1177/001316447003000105

Kulp, S. A., & Strauss, B. H. (2019). New elevation data triple estimates of global vulnerability to sea-level rise and coastal flooding. Nature Communications, 10(1). https://doi.org/10.1038/s41467-019-12808-z

Liu, Y., Qi, N., Zhu, Q., & Li, W. (2019). CR-U-Net: Cascaded U-Net with Residual Mapping for Liver Segmentation in CT Images. 2019 IEEE Visual Communications and Image Processing (VCIP), 1–4. https://doi.org/10.1109/VCIP47243.2019.8966072

Li, S., Wahl, T., Barroso, A., Coats, S., Dangendorf, S., Piecuch, C., Sun, Q., Thompson, P., & Liu, L. (2022). Contributions of Different Sea-Level Processes to High-Tide Flooding Along the U.S. Coastline. *Journal of Geophysical Research: Oceans*, *127*(7), e2021JC018276. https://doi.org/10.1029/2021JC018276

Mahmoudi, S., Moftakhari, H., Muñoz, D. F., Sweet, W., & Moradkhani, H. (2024). Establishing flood thresholds for sea level rise impact communication. Nature Communications, 15(1), 4251. https://doi.org/10.1038/s41467-024-48545-1

Mydlarz, C., Sai Venkat Challagonda, P., Steers, B., Rucker, J., Brain, T., Branco, B., Burnett, H. E., Kaur, A., Fischman, R., Graziano, K., Krueger, K., Hénaff, E., Ignace, V., Jozwiak, E., Palchuri, J., Pierone, P., Rothman, P., Toledo-Crow, R., & Silverman, A. I. (2024). FloodNet: Low-Cost Ultrasonic Sensors for Real-Time Measurement of Hyperlocal, Street-Level Floods in New York City. *Water Resources Research*, *60*(5), e2023WR036806. https://doi.org/10.1029/2023WR036806

National Oceanic and Atmospheric Administration. (2024). *Water levels* [Dataset]. https://tidesandcurrents.noaa.gov/waterlevels.html?id=8656483

National Weather Service. (2024a). *Advanced Hydrologic Prediction Services*. https://water.noaa.gov/?wfo=ARX

National Weather Service. (2024b). *Significant Event Archive for Southeast NC*. https://www.weather.gov/ilm/SignificantLocalEvents

O'Donnell, K. L., Tomiczek, T., Higgins, A., Munoz, S., & Scyphers, S. (2024). Stakeholder Driven Sensor Deployments to Characterize Chronic Coastal Flooding in Key West Florida. *Earth's Future*, *12*(7), e2023EF003631. https://doi.org/10.1029/2023EF003631

Rinaldo, T., Ramakrishnan, K. A., Rodriguez-Iturbe, I., & Durán Vinent, O. (2021). Probabilistic structure of events controlling the after-storm recovery of coastal dunes. Proceedings of the National Academy of Sciences, 118(1), e2013254118. https://doi.org/10.1073/pnas.2013254118

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 (pp. 234–241). Springer International Publishing. https://doi.org/10.1007/978-3-319-24574-4_28

Rose, L., Widlansky, M. J., Feng, X., Thompson, P., Asher, T. G., Dusek, G., Blanton, B., Luettich, R. A., Callahan, J., Brooks, W., Keeney, A., Haddad, J., Sweet, W., Genz, A., Hovenga, P., Marra, J., & Tilson, J. (2024). Assessment of water levels from 43 years of NOAA's Coastal Ocean Reanalysis (CORA) for the Gulf of Mexico and East Coasts. Frontiers in Marine Science, 11, 1381228. https://doi.org/10.3389/fmars.2024.1381228

Sadler, J. M., Goodall, J. L., Behl, M., Bowes, B. D., & Morsy, M. M. (2020). Exploring real-time control of stormwater systems for mitigating flood risk due to sea level rise. *Journal of Hydrology*, *583*(October 2019). https://doi.org/10.1016/j.jhydrol.2020.124571

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. CVPR 2018 Open Access. https://openaccess.thecvf.com/content_cvpr_2018/html/Sandler_MobileNetV2_Inverted_Residuals_CVPR_2018_paper.html

Schmidt, V., Luccioni, A. S., Teng, M., Zhang, T., Reynaud, A., Raghupathi, S., Cosne, G., Juraver, A., Vardanyan, V., Hernandez-Garcia, A., & Bengio, Y. (2021). ClimateGAN: Raising Climate Change Awareness by Generating Images of Floods. https://doi.org/10.48550/ARXIV.2110.02871

Serafin, K. A., & Ruggiero, P. (2014). Simulating extreme total water levels using a time dependent, extreme value approach. Journal of Geophysical Research: Oceans, 119(9), 6305–6329. https://doi.org/10.1002/2014JC010093

Sweet, W., Dusek, G., Obeysekera, J., & Marra, J. J. (2018). *Patterns and Projections of High Tide Flooding Along the U.S. Coastline Using a Common Impact Threshold* (CO-OPS 086; NOAA Technical Report, p. 56).

Sweet, W. V., Hamlington, B. D., Kopp, R. E., Weaver, C. P., Barnard, P. L., Bekaert, D., Brooks, W., Craghan, M., Dusek, G., Frederikse, T., Garner, G., Genz, A. S., Krasting, J. P., Larour, E., Marcy, D., Marra, J. J., Obeysekera, J., Osler, M., Pendleton, M., … Zuzak, C. (2022). *Global and Regional Sea Level Rise Scenarios for the United States: Updated Mean Projections and Extreme Water Level Probabilities Along U.S. Coastlines* (p. 111). National Oceanic and Atmospheric Administration, National Ocean Service.

Sweet, W. V., Simon, S., Dusek, G., Marcy, D., Brooks, W., Pendleton, M., & Marra, J. J. (2021). *2021 State of High Tide Flooding and Annual Outlook*. National Oceanic and Atmospheric Administration.

Thelen, T., Anarde, K., Dietrich, J. C., & Hino, M. (2024). Wind and rain compound with tides to cause frequent and unexpected coastal floods. Water Research, 122339.

Walt, S. van der, Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., & Yu, T. (2014). scikit-image: Image processing in Python. *PeerJ*, *2*, e453. https://doi.org/10.7717/peerj.453

Zhang, H., Hong, X., Zhou, S., & Wang, Q. (2019). Infrared Image Segmentation for Photovoltaic Panels Based on Res-UNet. In Z. Lin, L. Wang, J. Yang, G. Shi, T. Tan, N. Zheng, X. Chen, & Y. Zhang (Eds.), Pattern Recognition and Computer Vision (pp. 611–622). Springer International Publishing. https://doi.org/10.1007/978-3-030-31654-9_52