MS. KYLA T SEMMENDINGER (Orcid ID : 0000-0002-0976-0140)

Article type : Technical Paper

A Probabilistic, Parcel-Level Inundation Prediction Tool for Medium-Range Flood Forecasting in Large Lake Systems

Kyla Semmendinger, Jillian Foley, Scott Steinschneider

Department of Biological and Environmental Engineering, Cornell University, Ithaca, New York, USA (Correspondence to Steinschneider: ss3378@cornell.edu)

Research Impact Statement: Incorporating uncertainty into inundation predictions provides a conservative forecast of flood risk for shoreline property owners on large lakes.

Abstract: This study contributes a bathtub-style inundation prediction model with abstractions of coastal processes (i.e. storm surge and wave runup) for flood forecasting at medium-range (weekly to monthly) timescales along the coastline of large lakes. Uncertainty from multiple data sources are propagated through the model to establish probabilistic bounds of inundation, providing a conservative measure of risk. The model is developed in a case study of the New York Lake Ontario shoreline, which has experienced two record-setting floods over the course of three years (2017-2019). Predictions are developed at a parcel-level and are validated using inundation accounts from an online survey and flyover imagery taken during the recent flood events. Model predictions are compared against a baseline, deterministic model that accounts for the same processes but does not propagate forward data uncertainties. Results suggest that a probabilistic approach helps capture observed instances of inundation that would otherwise be missed by a deterministic inundation model. However, downward biases are still present in

probabilistic predictions, especially for parcels impacted by wave runup. The goal of the tool is to provide community planners and property owners with a conservative, parcel-level assessment of flood risk to help inform short-term emergency response and better prepare for future flood events.

(KEYWORDS: *probabilistic predictions; Great Lakes; inundation verification; flood hazard)*

INTRODUCTION

Damage from coastal flooding is growing rapidly around the world (Jongman, *et al*., 2012; Paprotny *et al*., 2018). Along ocean coasts, flood frequency is projected to more than double in certain regions with sea level rise and increased storm activity (Vitousek *et al*., 2017). Along the coastline of large inland lakes, the situation is complicated by fluctuations in climate and hydrology that alter water level variability in different ways and over multiple timescales (Gronewold *et al*., 2013; Woolway *et al*., 2020). In the Laurentian Great Lakes, coastal communities have experienced record-setting high water levels over the last several years, leading to inundation of near-shore homes and businesses and flash floods during storm surge and high wave events (IJC-LOSLR Board, 2018; Gronewold and Rood, 2019). These record-setting floods highlight the need for information that can help communities reduce coastal flood impacts, especially information tailored for the unique hazards present in large lake systems.

Coastal communities make decisions regarding flood-risk reduction on various time scales, including short-term emergency response and long-term risk management. Short-term emergency response actions include alerting residents in low-lying regions (including seasonal residents), moving belongings to higher elevation, sealing low-elevation storm drains to avoid backflow, securing pumps to remove ponded water, and sandbagging homes and key infrastructure to reduce flood damages. These activities often require days or even weeks to implement. Long-term risk management actions include investing in shoreline protection and stabilization structures (e.g. vertical walls, revetments), elevating structures, retrofitting mechanical systems to operate under submerged conditions, installing sewer systems to avoid septic system failure, and retrofitting existing storm sewer outlets with control valves to avoid backflow. The necessary information and tools needed to make informed flood-risk decisions vary depending on the relevant time scales. In this study, we focus on developing a probabilistic inundation model that is designed to support medium-range (i.e. weekly to monthly) flood forecasting and short-term emergency response along lake coastlines but can also be adapted for screening-level assessments of long-term flood risk. This model is developed in a case study of the New York coastline of Lake Ontario, the 13th largest inland lake in the world and the last of the five Laurentian Great Lakes.

There are several commonly employed techniques to model inundation that range in complexity, from simplified conceptual models to 2-D and 3-D hydrodynamic models (Teng *et al.*, 2017). Deterministic, single-value water surface models (or "bathtub" models) predict inundation by comparing land elevation and static water level (NOAA-CSC, 2010). Because of their simplicity, bathtub models forego coastal process calculations such as storm surge and wave runup. However, on large lakes, wind fields can propagate significant wave runup and storm surge during periods of increased storm activity, and flood events can be induced by high static water levels, storm activity, or a combination of the two (Kreutzwiser *et al.*, 1992; Angel, 1995). Therefore, bathtub models often miss important factors that contribute to inundation.

Higher dimensional hydrodynamic models can accurately capture coastal processes at a fine temporal and spatial resolution using the governing laws of hydraulics and fluid motion (Bates *et al.*, 2010; Favaretto *et al.*, 2019). Given the importance of storm-related activities in inundation prediction (Spaulding *et al.*, 2017), a fine resolution hydrodynamic model is being used as the basis to update FEMA flood insurance rate maps for the entire Great Lakes shoreline (FEMA, 2014). While hydrodynamic models can capture accurately the coastal processes that contribute to inundation, they require granular meteorological inputs that are often unavailable when forecasting inundation at medium-range lead times (e.g., winds fields are typically forecasted out only a few days; Chu *et al.*, 2011). This complicates the direct use of predictions from hydrodynamic models in month-ahead flood risk estimates. However, we argue that these models still provide valuable information for medium-range forecasting. In this work, we forward an approach that develops statistical summaries of storm surge and wave runup from hindcasts of hydrodynamic model output, and then adds those components into bathtub models to provide a better characterization of flood risk at extended lead times.

Beyond accounting for multiple inundation processes, components of model error should be quantified and propagated to provide estimates of uncertainty around inundation predictions. It is important to account for data and model uncertainties to prevent decision-makers from relying on "precise, but potentially inaccurate" (Alfonso *et al*., 2016) data. This is particularly true during extreme high-water events, when relatively small uncertainties in water levels and elevation data can result in significantly different flooding impacts. This is demonstrated in Figure 1, which shows the range of properties along the shoreline of Lake Ontario that are inundated when water level and elevation uncertainties are considered and water levels are high (as during recent floods in 2017 and 2019).

Several uncertainties confound inundation predictions. For instance, there is underlying uncertainty in the vertical accuracy of elevation data, which is often derived from Light Detection and Ranging (LiDAR) data or digital elevation models (DEMs). Elevation data that overestimate the true land elevation would result in an underestimation of flood risk (Van de Sande *et al*., 2012). Similarly, inaccuracies in water levels from gage measurement error, interpolation to ungaged sites, datum conversion, or forecasts of hydrologic inputs can impact the accuracy of inundation predictions. Propagation and interactions in meteorologically induced surges, coastal seiches, and wave set-up, set-down, and runup further increase the uncertainty in total water levels that can induce coastal flooding (Mazas *et al.,* 2014). Probabilistic approaches for predicting inundation seek to account for these uncertainties by reporting the likelihood of inundation, rather than a deterministic, binary estimate of inundation occurrence (Leon *et al*., 2014). These approaches have grown in popularity over the past decade (Gesch, 2009; NOAA-CSC, 2010; Gesch, 2013; Schmid *et al*., 2014; Kane *et al*., 2015; Alfonso *et al*., 2016; Gesch, 2018; Kovanen *et al*., 2018; West et al., 2018), in part because they relay the reliability of inundation predictions and better communicate flood risk to shoreline communities (Di Baldassarre *et al*., 2010; Moser, 2014). For instance, the National Oceanic and Atmospheric Administration (NOAA) developed the Lake Level Viewer (coast.noaa.gov/llv) for each of the Great Lakes to incorporate uncertainty into bathtub-style predictions via the z-score method

(Schmid *et al*., 2014), albeit with limited resolution that could hinder its use for flood risk management at the local level (Komolafe *et al*., 2018).

In addition to uncertainty propagation, model validation is another critical step needed to ensure that decision-makers understand the accuracy of predicted flood risk information. Inundation models are often validated using observed water levels and streamflow (for riverine flooding) rather than accounts of property inundation due to data availability (Horritt, 2006). However, many recent studies have relied on crowdsourcing and citizen science to gather reports of flooding to use in the validation process (Horritt, 2006; Kutija *et al*., 2014; Blumberg *et al*., 2015; Yu *et al*., 2016; Loftis *et al*., 2017; Assumpção *et al*., 2018; Loftis *et al*., 2019), although this methodology is still not common practice (See, 2019). In addition, to the authors' knowledge, model validation has only been used to test deterministic inundation predictions; the verification of probabilistic inundation predictions using observed accounts of flooding is underexplored.

In order for inundation predictions to be utilized in coastal decision-making, such as emergency response actions, they must capture the underlying inundation-driving mechanisms while also quantifying uncertainty and stakeholder confidence in the predictions. Currently, medium-range inundation prediction techniques in lacustrine coastal regions do not take into account all of these factors. To address this gap, this study forwards a novel, probabilistic, and parcel-level inundation prediction and mapping tool that is used to address three underlying research questions: 1) How accurate are inundation predictions based on a deterministic bathtub model with abstractions of coastal processes (i.e., storm surge and wave runup) in large lake systems? 2) Under what conditions can predictions be improved by incorporating uncertainty? and 3) How does this accuracy vary depending on the mechanisms driving the inundation event?

The proposed model is adapted from an existing model (The Flood Tool) previously used for inundation predictions along the Great Lakes shoreline (Baird, 2005). The model estimates parcel-level inundation based on deterministic bathtub-style modeling with added modules to abstract storm surge and wave runup processes. This work provides three primary contributions over the original Flood Tool and to the broader literature. First, we provide an updated version of the model that can develop conservative probabilistic inundation predictions under retrospective and forecasted water level conditions while accounting for storm surge and wave runup processes. Second, we verify inundation predictions using flyover imagery and citizen-science reports of inundation via an online survey during recent flood events. As part of this work, we explore the spatial heterogeneity of prediction accuracy and its relation to the mechanisms that drive inundation along different areas of the shoreline. Finally, we demonstrate the use of the model for medium-range, probabilistic inundation forecasts along the New York Lake Ontario shoreline that can be updated with operational, multi-week forecasts of static water levels issued at sub-weekly timescales. The study concludes with a discussion of limitations of the proposed model, future research needs, and the potential of the model to be adapted for use in long-term planning efforts for lake level management.

DATA AND METHODS

The proposed model requires four components to probabilistically predict inundation at the parcel level: structure elevation, static water level, storm surge, and wave runup (Figure 2). Inundation predictions are based on a bathtub-style modeling framework, where the elevation of a structure on a parcel is compared against the total water level (i.e., the sum of static levels, storm surge, and wave runup) to estimate inundation. However, instead of a binary prediction of inundation, the model estimates the probability of inundation following the NOAA (or z-score) method (Schmid *et al*., 2014).

In this method, uncertainty is quantified for data associated with each component, which when taken together quantifies the cumulative uncertainty in the inundation calculation. This technique assumes that all data sources are unbiased, that the error in each data source is independent of errors for other data sources, and that the cumulative uncertainty can be approximated by a normal distribution. If these assumptions hold, then the cumulative uncertainty can be calculated by taking the root of the sum of the squares of the individual root mean square errors (RMSEs) for each data source. The cumulative RMSE is then used to calculate the z-score at a given structure using Equation 1.

$$Z \quad Score = \frac{Structure\ Elevation \quad (Static\ Water\ Level + Storm\ Surge + Wave\ Runup)}{RMSE_{cumulative}} \quad (1)$$

with

$$RMSE_{cumulative} = \sqrt{RMSE^{2}_{elev} + RMSE^{2}_{static.level} + RMSE^{2}_{surge} + RMSE^{2}_{runup}} \quad (2)$$

The z-score translates into a probability of inundation based on the cumulative distribution of a standard normal distribution evaluated at the z-score. Although some studies have shown not all errors are normally distributed, the assumption of normality of the NOAA method tends to more conservatively predict inundation (Gesch, 2009; Schmid *et al*., 2014). While over-predictions of flood risk might result in unnecessarily high flood protection costs, a conservative quantification of flood risk supports the risk-adverse nature of water managers and flood risk planners (O'Connor *et al*., 2005).

In the proposed model, the NOAA method and its components can be used to predict inundation for past events (verification mode) or future events (forecast mode). Some of the individual elements in Equation 1 will vary depending on the application (or mode) of inundation prediction. These elements, their data sources, and their associated uncertainties are described in more detail below. Geographic coverage for each data source is shown in Figure 3.

Structure Elevation

Structure elevation is defined as the elevation of the lakeward side of the structure of interest. For a conservative inundation prediction, the minimum elevation of the lakeward side of the structure is included in the z-score calculation. Elevation information is available in the form of digital elevation models (DEMs). There are seven New York counties with shorelines on Lake Ontario (Niagara, Orleans, Monroe, Wayne, Cayuga, Oswego, and Jefferson). All Lake Ontario shoreline counties, excluding Monroe County and Niagara County, are covered by a publicly available FEMA 1-meter DEM (FEMA, 2014. NYS Elevation Data. Accessed July 2018, https://gis.ny.gov/elevation). Monroe County is covered by a 1-foot DEM (Monroe County Department of Environmental Services, 2017. GIS Data. Accessed March 2019, https://www2.monroecounty.gov/gis-Data.php). Niagara County is covered by a publicly available 3-meter DEM (NOAA Office for Coastal Management, 2014. Coastal Digital Elevation Model: Lake Ontario. Accessed March 2019, https://inport.nmfs.noaa.gov/inport/item/48114).

Elevations were extracted for shoreline homes using tax parcel shapefiles in GIS software. Tax parcel information is publicly available for Niagara, Orleans, Wayne, Cayuga, and Jefferson counties. Tax parcel information for Monroe and Oswego counties were obtained from their respective GIS departments. The tax parcel shapefiles were used to identify the footprint of the structure of interest from the Microsoft Footprint Database (Microsoft, 2019. USBuildingFootprints. Accessed August 2019 - October 2019, https://github.com/microsoft/USBuildingFootprints), which was then used to extract the base elevation of the foundation of the structure from the compiled elevation dataset.

The uncertainty in the DEM elevations is assumed to be equal to the vertical error determined for the associated LiDAR data used to develop that DEM. These values are reported as a RMSE for each DEM dataset (Table 1).

Static Water Level

Static water level is defined as the still water level without any influence of storm related activities such as wave runup or storm surge. These data are input into the tool as either a lake-wide average of gage observations for a particular historic date (verification mode) or a forecasted static water level for a future date (forecast mode).

Verification Mode. In verification mode, gage observations on Lake Ontario are averaged to ensure there are no surge or seiche impacts when estimating the static water level. The six gages used in the calculation are located in both the United States and Canada (NOAA Great Lakes Environmental Research Laboratory, 2019. Great Lakes Water Levels Monitoring Network. Accessed August 2019, https://www.glerl.noaa.gov/data/wlevels/#monitoringNetwork) and include two long-term gages managed by NOAA at Rochester and Oswego, NY, as well as four long-term gages at Kingston, Cobourg, Toronto, and Port Weller, located in the Province of Ontario and managed by Fisheries and Oceans Canada. The static water level for a given date is calculated as the 5-day rolling average water level between the six gages in the monitoring network.

The uncertainty associated with static water levels in verification mode is the combination of gage measurement error (RMSE of 0.006 m (EPA, 2016)) and the error introduced by converting between datums. As static water level is calculated by averaging across six gages, the RMSE for the lake-wide average static level is approximately 0.002 m via the Central Limit Theorem. Error is introduced in datum conversion because all DEM-based elevations are reported with respect to the North American Vertical Datum of 1988 (NAVD88), while all water levels are reported with respect to the International Great Lakes Datum of 1985 (IGLD85). NOAA has calculated and reported the potential error associated with converting between these two datums (RMSE of 0.20 m (NOAA, 2016. VDatum. Accessed January 2019, https://vdatum.noaa.gov/docs/est_uncertainties.html#estTransform)).

Forecast Mode. Weekly water level forecasts are produced by the US Army Corps of Engineers – Detroit District (USACE) and Environment and Climate Change Canada (ECCC) and released every Friday (lre.usace.army.mil/Missions/Great-Lakes-Information/Great-Lakes-Water-Levels/Water-Level-Forecast/Weekly-Great-Lakes-Water-Levels). The forecasting system employs an ensemble of input hydroclimatic (e.g. precipitation, temperature, evaporation, runoff) forecasts at 1-4 week lead times and estimated inflows from the upper Great Lakes to Lake Ontario and from the Ottawa River to the St. Lawrence River. This ensemble of inputs is used to produce an ensemble of projected water levels on Lake Ontario.

In forecast mode, the ensemble mean is used as the static water level when predicting inundation. Uncertainty in the forecasted static level is quantified as the combination of datum conversion uncertainty and the 95% confidence interval of the ensemble, which is assumed to be +/- 2 standard deviations of a normal distribution centered around the mean forecast. From this confidence interval, we infer the standard deviation of the forecast and use it in the cumulative uncertainty term in Equation 1. The standard deviation will vary for each forecast issue, but at a 4-week lead time it is generally on the order of 0.10 m.

Storm Surge

Storm surge is defined as the increase in water level over the static mean level due to high wind activity or seiche events. In this study, hourly storm surge is taken from the Lake Ontario Operational Forecast System (LOOFS) that is managed by NOAA's National Ocean Service (NOS). The LOOFS is based on a gridded hydrodynamic model that uses atmospheric observations and weather prediction guidance to produce three dimensional predictions of water temperature and two-dimensional forecasts of water levels for Lake Ontario (Chu *et al*., 2011). The LOOFS also predicts deviations from the average lake level, i.e. seiche and storm surge events. The LOOFS provides two sources of data, short-term (1-48 hour) forecasts and nowcasts, the latter which is based on near real-time observations and provides a continuous estimate of present conditions across the lake. For any location along the shoreline, we utilize the nowcast data for hourly water level deviations from the lake level average (i.e., storm surge) at the grid cell nearest the location of interest. These gridded data are available along the entire coastline at a 5 km resolution from 2006 to present.

Verification Mode. When comparing model predictions of inundation to observed inundation events, we use the maximum LOOFS nowcast storm surge associated with the date of observed inundation. Because these are modeled data, we estimated their uncertainty by comparing nowcast surge values to surge values at hourly observations at gages across the shoreline, including those listed in Section 2.2.1 but also including additional gages managed by NOAA and the USGS. The observed hourly surge values were calculated by taking hourly observed water levels for each gage and subtracting from them a 3-day rolling average to estimate the water level deviation (i.e., surge) for each hour. The RMSE between the nowcast and observed surge was calculated for each gage for data between May 24, 2017 and July 19, 2019 (when all gages had available data), and then the RMSE values were averaged across gages to estimate an average RMSE for nowcast surge estimates that could be applied anywhere along the shoreline (RMSE of 0.026 m).

We also considered an alternative method to calculate storm surge at an arbitrary location along the shoreline based on the interpolation of hourly gaged observations to ungaged sites using an inverse distance weighting approach. This approach was compared against the LOOFS storm surge values under cross-validation. A determination was made to use the LOOFS surge data because it performed similarly to the interpolation approach and provided a longer dataset on which to base probabilistic estimates of surge. More detail is provided on this comparison in the Supplemental Material (Section S1, Figure S1).

Forecast Mode. When developing medium-range (weekly to monthly) forecasts of storm surge, there are not reliable weather forecasts of wind speed and direction at long lead times on which to base a forecasted surge event. Therefore, the user is provided with the option to select a scenario of wind speed and direction, and the tool then estimates the potential storm surge (with uncertainty) conditional on those wind parameters and the LOOFS nowcast data. This is accomplished for any grid cell along the shoreline using the following procedure:

1. Hourly wind speeds from the LOOFS nowcast data are categorized into bins ranging from 0 miles per hour (mph) to 100 mph by increments of 10 mph for a total of 10 wind speed bins.
2. Hourly wind directions are classified as the cardinal and intercardinal directions for a total of 8 wind direction bins.
3. Each combination of wind speed and wind direction is classified as a wind event (80 total wind events).
4. Each wind event is associated with some number *n* of hourly occurrences in the nowcast dataset, and each of those *n* occurrences has its own nowcast surge value at the grid cell of interest. This produces an empirical distribution of surge values for a given wind event and location. In addition, there is additional uncertainty in each individual nowcast surge value (as quantified in 2.3.1). We employ a mixture distribution (Figure 4) to compound the error in the modelled nowcast surge data with the uncertainty of potential surge values for any given wind event:

$$p(s|w) = \int p(s|\hat{s})\, p(\hat{s}|w)\, d\hat{s} \qquad (3)$$

Here, $p(s|w)$ is the distribution of the true surge value for a given wind event, $p(\hat{s}|w)$ is the distribution of modelled nowcast surge values for a given wind event, and $p(s|\hat{s})$ is the distribution of the true surge value around a particular nowcast surge value. We assume p $(s|w)$ can be approximated as a mixture of normal distributions, i.e., we assume normality in the nowcast surge values under any given wind event and in the errors of the nowcast surge values.

5. The mixture distribution allows us to determine the expected value of a surge event for any given wind event, as well as an estimate of its standard deviation. Here, the standard deviation quantifies uncertainty in both nowcast estimates of surge as well as the spread in surge events for a selected wind field. The expected surge value is input into the numerator of Equation 1 and the standard deviation is incorporated into the cumulative uncertainty in the denominator.

Wave Runup

Wave runup is defined as the water level increase resulting from near-shore wave breaking that propagates water up the shoreline. The method to calculate wave runup is adapted from the original formulation presented in the Flood Tool (Baird, 2005), which is based on the Mase method and is outlined by FEMA in the Guidelines and Specifications for Flood Hazard Mapping Partners (FEMA, 2009):

$$R = Mase(h_0) = 1.1 \times \xi^{0.7} \times h_0 \qquad (4)$$

Here, R is the wave height that exceeds the low bluff or vertical wall height, ξ is the surf similarity parameter, and h_0 is the offshore wave height. The surf similarity parameter is defined as:

$$\xi = \frac{tan\,\theta}{\sqrt{\frac{h_0}{L_0}}} \qquad (5)$$

where tan θ is the nearshore slope of the property and L_0 is the peak wave period. The vertical water depth added by wave runup on top of static water level and storm surge at a particular parcel is calculated using the shoreline profile slope and the vertical height, R.

The Mase equation in Eq. 4 is written as a function of h_0 to emphasize its dependence on offshore wave heights, which are assumed to be the primary source of uncertainty in this work.

Simulated offshore wave height and wave period data were collected from the US Army Corps of Engineers Wave Information Studies (WIS) dataset (United States Army Corps of Engineers (USACE), 2010. Wave Information Studies. Accessed February 2019, http://wis.usace.army.mil). The WIS uses discrete spectral wave models and input wind fields to provide estimates of wave height, period, and direction for gridded locations across the Lake Ontario shoreline (see Figure 3). These data, available from January 1, 1979 to December 31, 2014, were validated against a limited set of hourly observed wave heights measured at a buoy near Oswego, NY and were determined to be relatively unbiased, at least for that location (see Supplemental Material; Section S2, Figures S2-S3).

For inundation predictions in either verification or forecast mode, the average monthly wave height for the given date and WIS location nearest the parcel of interest is used as input into the wave runup calculation. The uncertainty of the monthly wave height (as quantified by the variance of WIS wave heights for that month and location, $\sigma_{h_0}^2$) is propagated into the wave runup calculation via the delta method:

$$\sigma_R^2 = \text{Mase}\,(h_0)^2 \sigma_{h_0}^2 \qquad (6)$$

Here, σ_R^2 is the variance of the vertical height and $\text{Mase}\,(h_0)$ is the derivative of the Mase equation with respect to h_0.

Shoreline profile information was retrieved from the Flood and Erosion Prediction System (FEPS) database (Baird, 2005) available for a large portion of the New York shoreline on Lake Ontario. This database includes parcel-level information for vertical wall/bluff height, distance from the structure to the vertical wall/bluff, and nearshore and backshore slope based on elevation data.

APPLICATION

Model Verification

Verification Data. Four datasets were used in model verification. The first was an online survey developed by Cornell University and New York Sea Grant (NYSG) and distributed to shoreline communities during the 2017 flood event that requested written and visual accounts of inundation (Steinschneider *et al.*, 2019). Responses were collected from approximately 500 participants. Survey responses were pre-screened to ensure that inundation occurred due to Lake Ontario water levels, rather than a connected waterway. Redundancy was purposefully included for key survey questions to ensure respondents fully understood the question and, to the best of their ability, answered it accurately. The survey included several questions about foundation inundation, which are used here as the basis to evaluate the inundation model. Some respondents provided the approximate date that foundation inundation began. In other cases, this field was left blank and the inundation event was associated with the date the survey was submitted. For each report of inundation, the maximum hourly water level at the associated tax parcel in the four weeks prior to and including the date of inundation was collected and used in the inundation prediction. We use this hourly water level, rather than the water level on the specific date of inundation, because there was often a lag between the inundation occurrence and survey reporting and some degree of uncertainty around the true date that foundation inundation began.

Imagery from three flyovers was collected and used as another source of observational data. One flyover was conducted on June 12, 2017 and was organized by the Eastern Lake Ontario Dune Coalition (hereafter abbreviated DUNE). This flyover included coverage of the entire southern shoreline of Lake Ontario. Another unmanned flyover on July 12, 2017 was conducted by the USGS. This flyover focused primarily along the coast of Wayne County, NY near the village of Sodus Point. The third flyover was completed June 15, 2019 by the non-profit group Save Our Sodus Inc. (SOS) with coverage of the entire New York shoreline. These images were provided to the research team through personal communication and can be made available upon request to SOS.

All of the flyover images were used to visually identify cases of foundation inundation along the shoreline. For each flyover product, images were scanned to identify properties with primary structures that could clearly be identified as having their foundation inundated or not inundated. These properties were assigned binary indicators (0,1) to record the inundation state of the structure foundation. A total of 63, 13, and 77 observations were collected from the DUNE, USGS, and SOS flyovers, respectively. Sample images with positive instances of foundation inundation from each of these flyovers can be found in the Supplemental Material (Section S3). Using Google Earth, the addresses of these properties were identified, and then they were aligned with the corresponding tax parcel data and building footprint using GIS software. The minimum elevation of the foundation of the home was then collected as well as the latitude and longitude. The date of inundation of the structure was input as the date of the flyover. For each image of inundation, the maximum hourly water level at the associated tax parcel on the date of inundation was collected and used in the inundation prediction.

Due to data availability, 458 of the 687 verification accounts (across both the survey and flyover images) were able to be fully validated and included in the results.

Verification Procedure. Model predictions of inundation were developed for all parcels where observations of inundation were available (i.e., parcels from the Cornell-NYSG survey, DUNE flyover, USGS flyover, and SOS flyover). It is highly unlikely that the flyover images capture waves in the still imagery; therefore, the wave runup calculation was not included in inundation predictions for these accounts. The wave runup calculation is included in predictions made for the survey accounts of inundation.

The tool produced probabilistic predictions of inundation for each property on the given date of inundation. These probabilistic predictions were also converted into deterministic predictions based on the sign of the numerator in Eq. 1 (e.g., a positive (negative) numerator in Eq. 1 was associated with a positive (negative) deterministic prediction of inundation).

We used these deterministic and probabilistic predictions to answer the first two research questions of this work: 1) How accurate are inundation predictions based on a deterministic bathtub model? and 2) Under what conditions can predictions be improved by incorporating uncertainty? The deterministic and probabilistic predictions were compared with the observed cases of inundation using a Brier Score (BS):

$$BS = \frac{1}{N}\sum_{i=1}^{N}(p_i \quad o_i)^2 \tag{7}$$

Here, p_i is a value between zero and one that reflects the predicted probability of inundation for the i^{th} parcel (e.g., $p_i = 0.57$ implies a predicted probability of inundation of 57%). We note that a deterministic prediction of inundation and no inundation is equivalent to a probabilistic prediction of 100% and 0% chance of inundation, respectively. The binary indicator o_i reflects the observed state of inundation for the i^{th} parcel (0 implies no inundation, 1 implies inundation). A BS of zero is ideal as it means that the predictions and the observed data are the same, whereas a BS closer to one indicates that the model is often either underpredicting or overpredicting cases of inundation.

The BS for the probabilistic and deterministic predictions were compared using a Brier Skill Score (BSS):

$$BSS = 1 - \frac{BS_{prob}}{BS_{det}} \tag{8}$$

The BSS quantifies the degree to which probabilistic predictions outperform deterministic predictions of inundation. A score of zero suggests that the probabilistic and deterministic predictions perform equally well, whereas a score above or below zero indicates that the probabilistic predictions are outperforming or underperforming the deterministic predictions, respectively. We note that if there is a large difference between structure elevation and total water levels for most parcels, the deterministic predictions will very often be correct and the deterministic BS score will likely be lower than the score for the probabilistic predictions (i.e., BSS will be negative). However, if in many cases the total water level and structure elevation are sufficiently close so that data uncertainty will significantly impact the inundation prediction, it is likely that the BSS will be positive and probabilistic predictions are needed to accurately assess and communicate inundation risk.

The BSS was calculated separately for each of the different observational datasets, as well as for observations associated with different shoreline types (i.e. open coastline or embayment). In addition, spatial patterns in the error between probabilistic predictions and observed inundation were analyzed to better determine regions along the shoreline where the model was more or less accurate. This analysis provided insight into which inundation-driving mechanisms were accurately being captured by the tool. Wave runup calculations solely impact open coastline properties, whereas static water level and storm surge calculations impact both open coastline and embayment properties. Therefore, geographic clustering of accurate or inaccurate model predictions in particular locations is important in diagnosing model performance.

Demonstration of Medium-Range Inundation Forecasts

To demonstrate inundation predictions under a medium-range water level forecast, we use the May 16, 2019 issue date for a June 14, 2019 forecast date. This water level forecast was chosen since it nearly overlaps with the SOS flyover, conducted on June 15, 2019. This allows us to compare known cases of inundation with month-ahead inundation predictions. The tool is publicly available online as a Google Earth Engine web application (https://kts48.users.earthengine.app/view/lake-ontario-flood-mapper).

There are two options for displaying inundation predictions based on forecasted water levels: a continuous map based on DEM grid cells or a parcel-specific inundation map based on structure polygons. Shoreline information required for wave runup calculations is only available for parcels in the FEPS database, not at the DEM grid cell resolution, and therefore wave runup is not included in continuous forecast maps at the DEM grid cell level. For each water level forecast, inundation predictions are made using the follow steps (shown here for DEM-based inundation maps):

1. The tool automatically retrieves the point forecast and uncertainty range (i.e. RMSE) of the USACE/ECCC issued static water level forecast at a month lead time.
2. The user selects a wind speed and wind direction (i.e. a wind event).
3. For the entire shoreline, that wind speed and direction is used to develop a storm surge distribution (via Eq. 3), providing a mean value and an RMSE.
4. These terms are included in Eqs. 1-2 to estimate the probability of inundation for every DEM grid cell.

A similar procedure is followed for parcel-specific inundation maps, but with an added step to include wave runup based on offshore wave heights for the month of interest.

RESULTS

Verification of Inundation Predictions

The reference model performance is based on deterministic inundation predictions, shown in Table 2. When all of the datasets are pooled together (Table 2a), the deterministic model accurately predicts no inundation and inundation for approximately 37% and 29% of the observations, respectively. However, the deterministic model underpredicts and overpredicts inundation for 21% and 13% verification accounts, respectively. If a conservative approach to inundation modeling is preferred, the underprediction rate of observed inundation events (approximately one in five) is concerning.

Model performance varies significantly across the different observational products. When compared to the survey verification product, the model shows a large percentage of underpredictions (Table 2b). It is worthwhile to reiterate that the survey is likely the only observational set that captures inundation events caused by wave processes in addition to static levels and storm surge. When compared to the USGS flyover, the model also has a similarly high rate of underpredictions, albeit based on a much smaller sample size (Table 2e). The model rarely underpredicts observations based on the SOS flyover, but overpredictions are common (the model predicts inundation when none is observed in 29% of all observations; Table 2d). No such overpredictions are reported for the DUNE flyover, and the rate of underpredictions is also relatively low (13%; Table 2c).

The probabilistic model provides a way to embed some degree of conservatism into the inundation predictions. Table 3 displays probabilistic inundation predictions (split into 10% increments) along with the reported inundation state for all observations pooled across verification products. Table 3 shows that the majority of observations that did not experience inundation were predicted to have a very low probability of being inundated (0-10% chance). The 98 observations that did experience inundation but were underpredicted by the deterministic model (see Table 2a) fall into the 0 – 50% range for the probabilistic predictions. Approximately 40% of these underpredicted properties have non-trivial flood-risk (i.e., 40 of the 98 underpredicted observations have a 10-50% chance of inundation). This demonstrates the need to propagate uncertainty into inundation predictions, so as to better represent flood risk for properties that may be underestimated using a deterministic approach. Still, 58 observations (~13%) that reported inundation were predicted to have a low likelihood of inundation (0-10% chance), suggesting some downward bias in the probabilistic predictions.

While Tables 2 and 3 highlight the potential value of probabilistic predictions as compared to a deterministic approach, this value is quantified using the BSS. Figure 5 shows the BSS for all verification accounts and the breakdown of skill between verification product (survey reports and SOS, DUNE, and USGS flyovers) and shoreline type (open coast versus embayment). Almost all products show an increase of prediction skill in probabilistic over deterministic predictions (i.e., a positive BSS). The largest BSS score is associated with the USGS product for embayments, but this result is only based on one observation. For categories with more observations, positive BSS scores range from 0.06 to 0.39. Embayment properties associated with the DUNE flyover are the only exception where deterministic predictions outperform probabilistic predictions (negative BSS). In this case, the deterministic predictions are very accurate; there are only 12 underpredictions and no overpredictions (Table 2c). Therefore, the incorporation of uncertainty is often not needed to explain discrepancies in the deterministic predictions.

When considering the full set of observations, the BSS score is higher for observations in embayments versus those on the open coast. This result is almost entirely driven by the survey verification product, which also shows better probabilistic predictions (as compared to a deterministic baseline) for embayment properties versus those on open coastline. Conversely, for the DUNE flyover product, probabilistic predictions tend to provide a larger improvement over a deterministic baseline for open coastline properties. For the SOS flyover, the BSS scores do not different significantly between open coastline and embayments, and little can be said for the USGS flyover because the sample size of embayment properties is too small.

The tendency for probabilistic predictions to more strongly outperform deterministic predictions in embayments for the survey, but not the flyovers, may reflect the processes captured by the different verification products. The flyovers provide still imagery taken on clear days with low wind activity to best capture images along the shoreline, and therefore do not likely capture inundation associated with wave activity. Conversely, the written survey reports integrate respondents' observations of flooding over a period of time prior to the survey date, and therefore can account for wave-related inundation. These wave-related flood events would be mostly limited to the open coast, since barrier beaches often protect properties within embayments from wave activity. Therefore, a lower BSS along the open coastline compared to embayments for the survey reports, but not for the flyover products, suggests that the tool may be systematically underestimating wave-related flooding.

To assess the spatial distribution of model prediction skill, Figure 6a shows the difference between the predicted probability of inundation and the binary inundation observation (hereafter the Inundation Probability – Observation (IPO) score) for all observations mapped across the Lake Ontario shoreline. Highly accurate and precise predictions are associated with IPO scores near zero, whereas IPO scores greater than and less than zero are indicative of probabilistic overpredictions and underpredictions, respectively. Figures 5b-d show the spatial distribution of IPO scores within specific regions of the shoreline in more detail. The distribution of IPO scores are shown for the full verification dataset and for each verification product in Figure 6e.

The probabilistic model is performing well along most of the shoreline, with IPO scores most often within a narrow range (-0.3, 0.3) around zero. While the median IPO score for most datasets (except the USGS) is approximately zero, the full dataset, survey, DUNE, and USGS products have IPO scores that tend to range below zero, indicating more underpredictions in these products. In contrast, most of the IPO scores for the SOS verification product are positive, indicating more overpredictions.

Model predictions along the eastern shoreline, such as regions around North Pond (Figure 6b) and Jefferson County, tend to be the most accurate. While large underpredictions and overpredictions (|IPO| > 0.9) do occur, they are more infrequent compared to other areas of the shoreline.

The largest concentration of underpredictions is located along the shoreline of Monroe County (Figure 6c) and are predominately from the survey verification product. Along the shoreline in this region, houses are located within close proximity to the waterfront with relatively unprotected shorelines. Survey accounts for these properties also reported significant wave activity as a major contributor to inundation, suggesting that model underpredictions in this area are linked to poorly characterized wave runup processes.

In Sodus Point (Figure 6d), there is a clustering of overpredictions along the peninsula (IPO > 0.9). All of the overpredictions in this region (n = 7) are from the SOS flyover verification product. In Wayne County, there is only 1 underprediction (IPO < - 0.9) for the SOS product, with the model doing a relatively accurate job of representing flood risk (- 0.9 < IPO < 0.9) for 118 verification accounts (94%). The modelled static water level used in the verification procedure for these accounts is 75.91 meters, which is the all-time high daily water level on Lake Ontario. In addition, the nowcast identified a positive surge anomaly in this area on the SOS flyover date. However, the observed water level at a USGS gage near Sodus Bay was 75.85 meters, suggesting the total water level used for verification was higher than that observed. This again illustrates that at high water levels, small discrepancies in measured data can result in significant uncertainty in predicted inundation.

Demonstration of Medium Range Inundation Forecasts

Figure 7 shows a medium range forecast of inundation for June 14, 2019, issued approximately a month beforehand (May 16, 2019). These inundation predictions are associated with a mean static water level forecast of 75.87 meters. The map displays inundation risk as low-, moderate-, and high-risk, which corresponds to probabilistic inundation predictions of 1-5% (yellow), 5-50% (orange), and 50-100% (red), respectively. The ranges for these categories were based on stakeholder feedback.

Under the static water level forecast, key locations in Wayne County NY were at moderate to high risk of inundation. The probabilistic inundation predictions for the SOS flyover observations on June 15, 2019 are shown in Table 4. At this lead time, the uncertainty in the static water level forecast is a major driver of the uncertainty in inundation predictions. There is only one property being significantly underpredicted (inundation probability < 10%). The model is overpredicting inundation significantly (inundation probability > 90%) for seven properties.

The probabilistic model has a BS of 0.20, which indicates the probabilistic model fits the observed data relatively well, even when based on the month-ahead forecasted static water level. The BSS for these observations is 0.35, which further demonstrates the benefits of the probabilistic predictions over a deterministic approach for medium range forecasts.

DISCUSSION

Model Limitations and Future Research Needs

The proposed model attempts to improve risk characterization over a deterministic approach by propagating known vertical errors into inundation predictions. However, some key uncertainties were not included in the model, particularly those related to structural uncertainties and biases in certain datasets and models. For instance, while our model accounts for an abstraction of wave runup processes via the Mase equation (Eq. 4), Melby *et al*. (2012) showed that these empirical models parameterized by deep-water wave conditions "will yield significant uncertainty in application to shallow water conditions with varied bathymetry." Improved results could be possible using different empirical approaches (Stockdon or EuroTop formula), or using hydrodynamic models capable of capturing wave transformation into shallow water regions, but these approaches were not considered. The abstractions of wave runup used here may be responsible for some regions of key underpredictions (e.g. Monroe County) and should be taken as a caution against using this model for precise inundation predictions in key areas susceptible to coastal processes that are difficult to characterize. Future work is needed to identify which components of the model lead to systematic biases and whether computationally efficient alternatives can be found. The model does not account for the duration of wind speeds on storm surge heights, nor the correlation of storm surge and waves. To more accurate capture flood risk, future work should determine how flood risk varies with the persistence and co-occurrence of extreme coastal events. Additionally, a variety of non-lake level processes (e.g. riverine flooding, ponding in nearshore areas, etc.) are excluded from the tool and are therefore not included in potential inundation impacts. Stakeholder communication is critical to ensure communities understand how this tool can (and cannot) be used to forecast their flood risk.

In an effort to present a robust verification of the model, multiple observational products were used to assess model predictions. However, these observational products all have measurement error, which should be considered when interpreting the results. For instance, while the flyover products are associated with precise dates, the survey information has more uncertainty in the date of the actual inundation event being reported. Survey responses could have been submitted at any time following an inundation event, and not all survey responses reported the date of flooding. There is also the possibility of human error in the survey product, as homeowners may have reported inundation for a different part of their property (e.g. utility shed, detached garage, etc.) when asked about foundation inundation for their primary residential structure. Further, while flyover images were carefully screened to ensure observations of inundation (or lack thereof) were accurate, there is the possibility that certain observations were incorrectly categorized.

When used in forecast mode, medium-range forecasts of static lake levels contribute a large portion of the uncertainty to inundation predictions. Any potential to improve the accuracy and precision of these forecasts could have significant value to coastal communities. Recent work has sought to improve water level forecasts in the Great Lakes region (Durnford *et al*., 2018). This effort requires that data and models be seamlessly integrated across the international border of the US and Canada (Gronewold *et al*., 2018), as demonstrated in the expansion of the National Water Model across the Great Lakes region (Mason *et al*., 2019). Forecasting efforts could also benefit from the assimilation of state-of-the-art measurements of antecedent conditions (e.g. snowpack (Arslan *et al.*, 2019), soil moisture (Entekhabi *et al*., 2010)), as well as additional runoff model intercomparisons (Gaborit *et al*., 2017) and improvements in models of open-water evapotranspiration (Charusombat *et al*., 2018). Medium-range water level forecasts would further improve with increased skill in precipitation and temperature forecasts at subseasonal to seasonal lead times (Vitart *et al*., 2017); recent efforts in the Great Lakes region have focused on developing a suite of seasonal forecast tools for this purpose (Bolinger *et al*., 2017). One of the strongest signals is related to the El Niño-Southern Oscillation (ENSO), which provides forecast information at the end of the fall season for winter and early spring water supplies. This coincides with the timing needed to prepare for and potentially reduce flood risk. While ENSO forecasts can be noisy over the Great Lakes, recent work suggests that non-linearity in the underlying teleconnections could be used to improve forecast skill (Carter *et al*., 2018; Fu *et al*., 2019).

Implications for Lake Level Management

Municipal-level decision-making for flood risk mitigation on Lake Ontario is complicated by water level management that influences flooding on the lake. Since the late 1950's, water levels on Lake Ontario have been regulated by the International Joint Commission (IJC) at the Moses-Saunders Power Dam, located downstream of Lake Ontario on the St. Lawrence River. The dam has been used to stabilize water levels on the lake for a variety of stakeholder interests, including domestic water use, navigation, hydropower, riparian protection, and recreational boating. However, water level stabilization caused significant stress to coastal wetlands and other ecosystems (Wang *et al*., 2015; Wilcox *et al*., 2018), leading the IJC to introduce a new water level management plan (Plan 2014) on January 1, 2017 that reintroduced some of the natural variability in water levels that had been reduced under the previous plan (IJC, 2014). A few months after Plan 2014 was implemented, Lake Ontario experienced the 2017 flood, and then the 2019 flood occurred two years later. These floods have caused significant public backlash against Plan 2014, threatening the environmental benefits promised under the plan. This has sparked a review of Plan 2014 and considerations of whether an alternative management regime could improve the tradeoff between riparian flood risk and environmental restoration.

To effectively quantify risk and expected benefits to multiple stakeholder interests, any candidate management plan needs to be assessed under an ensemble of plausible water supply scenarios. The proposed inundation prediction tool lends itself to aid in quantifying flood damages under a large ensemble of water supply scenarios due to its low computational cost. In addition, uncertainty propagation would lead to a conservative estimate of riparian flood risk, which would help address stakeholder concerns in the tense political environment. This is particularly important given the non-linear damage curve (see Figure 1), where small changes in peak water levels can lead to large changes in potential impacts. Therefore, the probabilistic inundation tool is well suited to assist in flood impact quantification in future lake level management studies (both on Lake Ontario and the upper Great Lakes). However, more work is needed to support this effort, e.g., regional extreme value models of storm surge and offshore wave heights anywhere along the shoreline. This effort is left for future work.

CONCLUSION

This study contributes a novel, probabilistic, and parcel-level inundation prediction and mapping tool that combines multiple flood-related processes (static water levels, storm surge, wave run-up) relevant to large lake systems while also accounting for and propagating uncertainty in each through to inundation predictions. The model acts as a computational efficient complement to other inundation prediction tools in the Great Lakes that is well adapted for repeated and conservative inundation prediction, as is needed for frequently issued flood forecasts during extreme high water events or in planning studies with large ensembles of water supply scenarios.

The tool was validated in a case study along the New York Lake Ontario shoreline with accounts of inundation from four separate observational products covering the record floods of 2017 and 2019. Validation efforts showed that the probabilistic tool provided more accurate inundation predictions than deterministic predictions. The probabilistic tool had areas of concentrated underpredictions, which were attributed to deficiencies in capturing wave runup and overtopping of shoreline protection structures (e.g. vertical walls). However, in most locations the probabilistic nature of the tool allowed for conservative inundation estimates that helps to avoid major underestimation of inundation risk.

The fully validated model will be made available to stakeholders as an online forecasting tool with the goal of supporting proactive risk management and accelerating community response to potential inundation. This tool joins a larger suite of models emerging to help communities mitigate heightened flood risk along the Great Lakes shoreline.

SUPPORTING INFORMATION

Additional supporting information may be found online under the Supporting Information tab for this article: a comparison of modelled and spatially-interpolated, observation storm surge data, an assessment of modelled wave height data, and sample images used for model validation.

AKNOWLEDGEMENTS

This work was prepared under project R/CHD-12 funded under award NA18OAR4170096 from the National Sea Grant College Program of the U.S. Department of Commerce's National Oceanic and Atmospheric Administration, to the Research Foundation for State University of New York on behalf of New York Sea Grant. The statements, findings, conclusions, views and recommendations are those of the author(s) and do not necessarily reflect the views of any of those organizations.

LITERATURE CITED

Alfonso, L., Mukolwe, M.M., and G. Di Baldassarre. 2016. "Probabilistic Flood Maps to Support Decision-Making: Mapping the Value of Information." *Water Resources Research* 52 (2): 1026-1043. https://doi.org/10.1002/2015WR017378

Angel, J.R. 1995. "Large-scale Storm Damage on the U.S. Shores of the Great Lakes." *Journal of Great Lakes Research* 21 (3): 287-293. https://doi.org/10.1016/S0380-1330(95)71039-5

Arslan, A.N. and Z. Akyürek. 2019. "Special Issue on Remote Sending of Snow and Its Applications." *Geosciences* 9 (6): 277. https://doi.org/10.3390/geosciences9060277

Assumpção, T. H., Popescu, I., Jonoski, A., and D.P. Solomatine. 2018. "Citizen Observations Contributing to Flood Modelling: Opportunities and Challenges." *Hydrology and Earth System Sciences* 22 (2): 1473-1489. https://doi.org/10.5194/hess-22-1473-2018

Baird, W.F. & Associates Coastal Engineers LTD. 2005. Flooding Performance Indicator: Methodology and Shared Vision Model Application. http://loslr.ijc.org/en/library/Coastal/FloodingPI%20MethodsReport.pdf

Bates, P.D., Horritt, M.S., and T.J. Fewtrell. 2010. "A Simple Inertial Formulation of the Shallow Water Equations for Efficient Two-Dimensional Flood Inundation Modelling." *Journal of Hydrology* 387 (1-2): 33-45. https://doi.org/10.1016/j.jhydrol.2010.03.027

Blumberg, A. F., Georgas, N., Yin, L., Herrington, T. O., and P.M. Orton. 2015. "Street-scale Modeling of Storm Surge Inundation along the New Jersey Hudson River Waterfront." *Journal of Atmospheric and Oceanic Technology* 32: 1486-1497. https://doi.org/10.1175/JTECH-D-14-00213.1

Bolinger, R.A., Gronewold, A.D., Kompoltowicz, K., and L.M. Fry. 2017. "Application of the NMME in the Development of a New Regional Seasonal Climate Forecast Tool."

Bulletin of the American Meteorological Society 555-564. https://doi.org/10.1175/BAMS-D-15-00107.1

Carter, E. and S. Steinschneider. 2018. “Hydroclimatological Drivers of Extreme Floods on Lake Ontario.” *Water Resources Research* 54 (7): 4461-4478. https://doi.org/10.1029/2018WR022908

Charusombat, U., Fujisaki-Manome, A., Gronewold, A.D., Lofgren, B.M., Anderson, E.J., Blanken, P.D., Spence, C., Lenters, J.D., Xiao, C., Fitzpatrick, L.E., and G. Cutrell. 2018. “Evaluating and Improving Modeled Turbulent Heat Fluxes Across the North American Great Lakes.” *Hydrology and Earth System Sciences* 22: 5559-5578. https://doi.org/10.5194/hess-22-5559-2018

Chu, P.Y., Kelley, J.G.W., Mott, G. V., Zhang, A., and G.A. Lang. 2011. “Development, Implementation, and Skill Assessment of the NOAA/NOS Great Lakes Operational Forecast System.” *Ocean Dynamics* 61: 1305–1316. https://doi.org/10.1007/s10236-011-0424-5

Di Baldassarre, G., Schumann, G., Bates, P.D., Freer, J.E., and K.J. Beven. 2010. “Flood-plain Mapping: A Critical Discussion of Deterministic and Probabilistic Approaches.” *Hydrological Sciences Journal* 55 (3): 364-376. https://doi.org/10.1080/02626661003683389

Durnford, D., Fortin, V., Smith, G.C., Archambault, B., Deacu, D., Dupont, F., Dyck, S., Martinez, Y., Klyszejko, E., Mackay, M., Liu, L., Pellerin, P., Pietroniro, A., Roy, F., Vu, V., Winter, B., Yu, W., Spence, C., Bruxer, J., and J. Dickhout. 2018. “Toward an Operational Water Cycle Prediction System for the Great Lakes and St. Lawrence River.” *Bulletin of the American Meteorological Society* 521-546. https://doi.org/10.1175/BAMS-D-16-0155.1

Entekhabi, D., Njoku, E.G., O’Neill, P.E., Kellogg, K.H., Crow, W.T., Edelstein, W.N., Entin, J.K., Goodman, S.D., Jackson, T.J., Johnson, J., Kimball, J., Piepmeier, J.R., Koster, R.D., Martin, N., McDonald, K.C., Moghaddam, M., Moran, S., Reichle, R., Shi, J.C., Spencer, M.W., Thurman, S.W., Tsang, L., and J. Van Zyl. 2010. “The Soil Moisture Active Passive (SMAP) Mission.” In: *Proceedings of the Institute of Electrical and Electronics Engineers,* edited by Robert J. Trew, 704-716. Piscataway, NJ: Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/JPROC.2010.2043918

Environmental Protection Agency (EPA). 2016. Technical Documentation: Great Lakes Water Levels and Temperatures. https://www.epa.gov/sites/production/files/2016-08/documents/great-lakes_documentation.pdf

Favaretto, C., Martinelli, L., and P. Ruol. 2019. "Coastal Flooding Hazard Due to Overflow Using a Level II Method: Application to the Venetian Littoral." *Water (Switzerland)* 11 (1): 1-16. https://doi.org/10.3390/w11010134

Federal Emergency Management Agency (FEMA). 2009. Guidelines and Specifications for Flood Hazard Mapping Partners. https://www.fema.gov/media-library-data/1537888452599-e36f6cb435d25c48ceff0ce520543aad/Guidelines_and_Specifications_for_Flood_Hazard_Mapping_Partners_Appendix_C_Guidance_for_Riverine_Flooding_Analyses_and_Mapping_(Nov_2009).pdf

Federal Emergency Management Agency (FEMA). 2014. FEMA Great Lakes Coastal Guidelines, Appendix D.3 Update. http://greatlakescoast.org/pubs/reports/Great_Lakes_Coastal_Guidelines_Update_Jan2014.pdf

Fu, W. and S. Steinschneider. 2019. "A Diagnostic-predictive Assessment of Winter Precipitation Over the Laurentian Great Lakes: Effects of ENSO and Other Teleconnections." *Journal of Hydrometeorology* 20 (1): 117-137. https://doi.org/10.1175/JHM-D-18-0128.1

Gaborit, É., Fortin, V., Tolson, B., Fry, L., Hunter, T., and A.D. Gronewold. 2017. "Great Lakes Runoff Inter-comparison Project, Phase 2: Lake Ontario (GRIP-O)." *Journal of Great Lakes Research* 43 (2): 217-227. https://doi.org/10.1016/j.jglr.2016.10.004

Gesch, D.B. 2009. "Analysis of Lidar Elevation Data for Improved Identification and Delineation of Lands Vulnerable to Sea-Level Rise." *Journal of Coastal Research* SI (53): 49-58. https://doi.org/10.2112/si53-006.1

Gesch, D.B. 2013. "Consideration of Vertical Uncertainty in Elevation-Based Sea-Level Rise Assessments: Mobile Bay, Alabama Case Study." *Journal of Coastal Research* SI (63): 197-210. https://doi.org/10.2112/si63-016.1

Gesch, D.B. 2018. "Best Practices for Elevation-based Assessments of Sea-level Rise and Coastal Flooding Exposure." *Frontiers in Earth Science* 6: 230. https://doi.org/10.3389/feart.2018.00230

Gronewold, A.D., Fortin, V., Lofgren, B., Clites, A., Stow, C.A., and F. Quinn. 2013. "Coasts, Water Levels, and Climate Change: A Great Lakes Perspective." *Climatic Change* 120: 697-711. https://doi.org/10.1007/s10584-013-0840-2

Gronewold, A.D., Fortin, V., Caldwell, R., and J. Noel. 2018. "Resolving Hydrometeorological Data Discontinuities Along an International Border." *Bulletin of the American Meteorological Society* 899-910. https://doi.org/10.1175/BAMS-D-16-0060.1

Gronewold, A.D. and R.B. Rood. 2018. "Recent Water Level Changes Across Earth's Largest Lake System and Implications for Future Variability." *Journal of Great Lakes Research* 45(1): 1-3. https://doi.org/10.1016/j.jglr.2018.10.012.

Horritt, M. S. 2006. "A Methodology for the Validation of Uncertain Flood Inundation Models." *Journal of Hydrology* 326 (1-4): 153-165. https://doi.org/10.1016/j.jhydrol.2005.10.027

International Joint Commission (IJC). 2014. Lake Ontario St. Lawrence River Plan 2014: Protecting against extreme water levels, restoring wetlands, and preparing for climate change. http://www.ijc.org/files/tinymce/uploaded/LOSLR/IJC_LOSR_EN_Web.pdf

International Joint Commission (IJC) Lake Ontario - St. Lawrence River (LOSLR) Board. 2018. Observed Conditions & Regulated Outflows in 2017. https://ijc.org/sites/default/files/2018-08/ILOSLRB_FloodReport2017.pdf

Jongman, B., Ward, P. J., and J.C.J.H. Aerts. 2012. "Global Exposure to River and Coastal Flooding: Long Term Trends and Changes." *Global Environmental Change* 22 (4): 823–835. https://doi.org/10.1016/j.gloenvcha.2012.07.004

Kane, H.H., Fletcher, C.H., Frazer, L.N., and M.M. Barbee. 2015. "Critical Elevation Levels for Flooding Due to Sea-level Rise in Hawai'i." *Regional Environmental Change* 15: 1679-1687. https://doi.org/10.1007/s10113-014-0725-6

Komolafe, A.A, Herath, S. and R. Avtar. 2018. "Sensitivity of Flood Damage Estimation to Spatial Resolution." *Journal of Flood Risk Management* 11 (S1): S370-S381. https://doi.org/10.1111/jfr3.12224

Kovanen, J., Oksanen, J., and T. Sarjakoski. 2018. "Near Real-Time Coastal Flood Inundation Simulation with Uncertainty Analysis and GPU Acceleration in a Web Environment." *Computers & Geosciences* 119: 39-48. https://doi.org/10.1016/j.cageo.2018.06.007

Kreutzwiser, R.D. and A.O. Gabriel. 1992. "Ontario's Great Lakes Flood History." *Journal of Great Lakes Research* 18 (1): 194-198. https://doi.org/10.1016/S0380-1330(92)71285-4

Kutija, V., Bertsch, R., Glenis, V., Alderson, D., Parkin, G., Walsh, C.L., Robinson, J., and C. Kilsby. 2014. "Model Validation Using Crowd-Sourced Data From a Large Pluvial Flood." In: *Proceedings of 11th International Conference on Hydroinformatics*, edited by Michael Piasecki. New York, NY: CUNY Academic Works. https://academicworks.cuny.edu/cc_conf_hic/415/

Leon, J.X., Heuvelink, G.B.M., and S.R. Phinn. 2014. "Incorporating DEM Uncertainty in Coastal Inundation Mapping." *PLoS One* 9 (9): e108727. https://doi.org/10.1371/journal.pone.0108727

Loftis, J.D., Wang, H., Forrest, D., Rhee, S., and C. Nguyen. 2017. "Emerging Flood Model Validation Frameworks for Street-level Inundation Modeling with Storm Sense." In: *Proceedings of the 2nd International Workshop on Science of Smart City Operations and Platforms Engineering,* edited by Rasit Eskicioglu, 13-18. Pittsburg, PA: Global City Teams Challenge. https://doi.org/10.1145/3063386.3063764

Loftis, J. D., Mitchell, M., Schatt, D., Forrest, D. R., Wang, H. V., Mayfield, D., and A.W Stiles. 2019. "Validating an Operational Flood Forecast Model Using Citizen Science in Hampton Roads, VA, USA." *Journal of Marine Science and Engineering* 7 (8): 242. https://doi.org/10.3390/jmse7080242

Mason, L.A., Gronewold, A.D., Laitta, M., Gochis, D., Sampson, K., Read, L., Klyszejko, E., Kwan, J., Fry, L., Jones, K., Steeves, P., Pietroniro, A., and M. Major. 2019. "New Transboundary Hydrographic Data Set for Advancing Regional Hydrological Modeling and Water Resources Management." *Journal of Water Resources Planning and Management* 145 (6): 06019004. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001073

Mazas, F., Kergadallan, X., Garat, P., and L. Hamm. 2014. "Applying POT Methods to the Revised Joint Probability Method for Determining Extreme Sea Levels." *Coastal Engineering* 91:140-150. https://doi.org/10.1016/j.coastaleng.2014.05.006

Melby, J., Caraballo-Nadal, N., and N. Kobayashi. 2012. “Wave Runup Prediction for Flood Mapping.” In: *Proceedings of the 33rd International Conference on Coastal Engineering,* edited by Patrick Lynett and Jane McKee Smith. Santander, Spain. https://doi.org/10.9753/icce.v33.management.79

Moser, S.C. 2014. “Communicating Adaptation to Climate Change: The Art and Science of Public Engagement When Climate Change Comes Home.” *WIREs Climate Change* 5 (3): 337-358. https://doi.org/10.1002/wcc.276

National Oceanic and Atmospheric Administration (NOAA) Coastal Services Center (CSC). 2010. Mapping Inundation Uncertainty. https://coast.noaa.gov/data/digitalcoast/pdf/mapping-inundation-uncertainty.pdf

O’Connor, R.E., Yarnal, B., Dow, K., Jocoy, C.L., and G.J. Carbonne. 2005. “Feeling at Risk Matters: Water Managers and the Decision to Use Forecasts.” *Risk Analysis* 25 (5):1265-1275. https://doi.org/10.1111/j.1539-6924.2005.00675.x

Paprotny, D., Sebastian, A., Morales-Nápoles, O., and S.N. Jonkman. 2018. “Trends in Flood Losses in Europe over the Past 150 Years. *Nature Communications* 9 (1985): 1-12. https://doi.org/10.1038/s41467-018-04253-1

Schmid, K., Hadley, B., and K. Waters. 2014. “Mapping and Portraying Inundation Uncertainty of Bathtub-Type Models.” *Journal of Coastal Research* 30 (3): 548-561. https://doi.org/10.2112/JCOASTRES-D-13-00118.1

See, L. 2019. “A Review of Citizen Science and Crowdsourcing in Applications of Pluvial Flooding.” *Frontiers in Earth Science* 7: 44. https://doi.org/10.3389/feart.2019.00044

Spaulding, M.L., Grilli, A., Damon, C., Fugate, G., Oakley, B.A., Isaji, T., and L. Schambach. 2017. “Application of State of Art Modeling Techniques to Predict Flooding and Waves for an Exposed Coastal Area.” *Journal of Marine Science and Engineering* 5 (1): 1-23. https://doi.org/10.3390/jmse5010010

Steinschneider, S., Styler, A., Stedman, R., and M. Austerman. 2019. “A Rapid Response Survey to Characterize the Impacts of the 2017 High Water Event on Lake Ontario.” *Journal of the American Water Resources Association* 55 (4): 1065-1079. https://doi.org/10.1111/1752-1688.12750

Teng, J., Jakeman, A.J., Vaze, J., Croke, B.F.W., Dutta, D., and S. Kim. 2017. “Flood Inundation Modelling: A Review of Methods, Recent Advances and Uncertainty Analysis.”

Environmental Modelling & Software 90: 201-216. https://doi.org/10.1016/j.envsoft.2017.01.006

Van de Sande, B., Lansen, J., and C. Hoyng. 2012. "Sensitivity of Coastal Flood Risk Assessments to Digital Elevation Models." *Water* 4 (3): 568-579. https://doi.org/10.3390/w4030568

Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C., Déqué, M., Ferranti, L., Fucile, E., Fuentes, M., Hendon, H., Hodgson, J., Kang, H.S., Kumar, A., Lin, H., Liu, G., Liu, X., Malguzzi, P., Mallas, I., Manoussakis, M., Mastrangelo, D., MacLachlan, C., McLean, P., Minami, A., Mladek, R., Nakazawa, T., Najm, S., Nie, Y., Rixen, M., Robertson, A.W., Ruti, P., Sun, C., Takaya, Y., Tolstykh, M., Venuti, F., Waliser, D., Woolnough, S., Wu, T., Won, D.J., Xiao, H., Zaripov, R., and L. Zhang. 2017. "The Subseasonal to Seasonal (S2S) Prediction Project Database." *Bulletin of the American Meteorological Society* 163-173. https://doi.org/10.1175/BAMS-D-16-0017.1

Vitousek, S., Barnard, P., Fletcher, C., Frazer, N., Erikson, L., and C.D. Storlazzi. 2017. "Doubling of Coastal Flooding Frequency within Decades due to Sea-Level Rise." *Scientific Reports* 7 (1399): 1-9. https://doi.org/10.1038/s41598-017-01362-7

Wang, L., Riseng, C.M., Mason, L.A., Wehrly, K.E., Rutherford, E.S., McKenna, J.E., Castiglione, C., Johnson, L.B., Infante, D.M., Sowa, S., Robertson, M., Schaeffer, J., Khoury, M., Gaiot, J., Hollenhorst, T., Brooks, C., and M. Coscarelli. 2015. "A Spatial Classification and Database for Management, Research, and Policy Making: The Great Lakes Aquatic Habitat Framework." *Journal of Great Lakes Research* 41 (2): 584-596. https://doi.org/10.1016/j.jglr.2015.03.017

West, H., Horswell, M., and N. Quinn. 2018. "Exploring the Sensitivity of Coastal Inundation Modelling to DEM Vertical Error." *International Journal of Geographical Information Science* 32 (6): 1172-1193. https://doi.org/10.1080/13658816.2018.1444165

Wilcox, D.A. and J.A. Bateman. 2018. "Photointerpretation Analysis of Plant Communities in Lake Ontario Wetlands Following 65 years of Lake-level Regulation." *Journal of Great Lakes Research* 44 (6): 1306-1313. https://doi.org/10.1016/j.jglr.2018.08.007

Woolway, R.I., Kraemer, B.M., Lenters, J.D., Merchant, C.J., O'Reilly, C.M., and S. Sharma. 2020. "Global Lake Responses to Climate Change." *Nature Reviews Earth & Environment* 1: 388-403. https://doi.org/10.1038/s43017-020-0067-5

Yu, D., Yin, J., and M. Liu. 2016. “Validating City-scale Surface Water Flood Modelling Using Crowd-sourced Data.” *Environmental Research Letters* 11: 124011. https://doi.org/10.1088/1748-9326/11/12/124011

FIGURE LEGENDS

Figure 1. The distribution of home elevations along the Lake Ontario shoreline (see Data). Under a high water level scenario (75.9 m, the average peak level between the floods of 2017 and 2019), 15% of homes on the Lake Ontario shoreline fall below that water level and are projected to be inundated. However, after propagating vertical errors in the mean water level and home elevations (see Methods), up to 43% of homes are at risk to be impacted.

Figure 2. Flowchart of probabilistic inundation predictions.

Figure 3. Map of study region and data sources.

Figure 4. Mixture distribution for storm surge at a particular location and for a particular wind event (e.g., 10 - 20 mph winds from the northeast). In this example, there are $n = 9$ hourly occurrences associated with this wind event, each with a different nowcast surge value. The distribution of true surge around any particular nowcast surge value is shown with dotted blue lines. The final mixture distribution of storm surge is shown in yellow.

Figure 5. The BSS for each combination of verification product and shoreline type. The BSS of all verification properties (orange line), open coastline properties (red line), and embayment properties (blue line) is displayed as the reference BSS for all product and shoreline combinations (shown numerically).

Figure 6. IPO scores mapped along the Lake Ontario shoreline (a). Negative values correspond to model underpredictions, while positive values correspond to model overpredictions. IPO scores are shown for Sandy Creek, Oswego County (b), Greece, Monroe County (c), and Sodus Point, Wayne County (d). The boxplot of IPO scores broken down by verification product is shown in (e).

Figure 7. The Google Earth Engine user interface of the online inundation prediction tool. Inundation predictions are shown for the June 2019 forecasted water level in Wayne County, New York.

TABLE CAPTIONS

Table 1. The reported RMSE for each DEM covering the Lake Ontario shoreline.

Source	County Coverage	RMSE (m)
FEMA 1-meter DEM	Orleans, Wayne, Cayuga, Oswego, Jefferson	0.127
Monroe County 1-foot DEM	Monroe	0.106
NOAA 3-meter DEM	Niagara	0.2

Table 2. Contingency tables for deterministic inundation predictions versus reported inundation accounts, shown by verification product (percentages of total sample size shown in parentheses).

a)

Full Data (n = 458)		**Reported** No	**Reported** Yes
Predicted	No	171 (37%)	98 (21%)
Predicted	Yes	57 (13%)	132 (29%)

b)

Survey (n = 226)		**Reported** No	**Reported** Yes
Predicted	No	55 (24%)	76 (34%)
Predicted	Yes	20 (9%)	75 (33%)

c)

DUNE (n = 89)		**Reported** No	**Reported** Yes
Predicted	No	53 (60%)	12 (13%)
Predicted	Yes	0 (0%)	24 (27%)

d)

SOS (n = 126)		**Reported** No	**Reported** Yes
Predicted	No	63 (50%)	4 (3%)
Predicted	Yes	37 (29%)	22 (18%)

e)

USGS (n = 17)		Reported	
		No	Yes
Predicted	No	0 (0%)	6 (35%)
	Yes	0 (0%)	11 (65%)

Table 3. Probabilistic inundation predictions for data from all verification products (n = 458).

Probability of Inundation (%)	Reported	
	No	Yes
0 - 10	131 (29%)	60 (13%)
10 - 20	12 (3%)	14 (3%)
20 - 30	10 (2%)	9 (2%)
30 - 40	9 (2%)	9 (2%)
40 - 50	9 (2%)	6 (1%)
50 - 60	10 (2%)	14 (3%)
60 - 70	8 (2%)	17 (4%)
70 - 80	7 (1%)	6 (1%)
80 - 90	12 (3%)	19 (4%)
90 - 100	20 (4%)	76 (17%)

Table 4. Probabilistic inundation predictions based on the month-ahead water level forecast for the SOS flyover verification product (n = 126).

Probability of Inundation (%)	**Reported**	
	No	Yes
0 - 10	41 (32%)	1 (1%)
10 - 20	7 (6%)	0 (0%)
20 - 30	8 (6%)	0 (0%)
30 - 40	5 (4%)	3 (2%)
40 - 50	8 (6%)	3 (2%)
50 - 60	7 (6%)	3 (2%)
60 - 70	6 (5%)	1 (1%)
70 - 80	8 (6%)	2 (2%)
80 - 90	3 (2%)	2 (2%)
90 - 100	7 (6%)	11 (9%)