AGU ADVANCING EARTH AND SPACE SCIENCES

Water Resources Research

RESEARCH ARTICLE
10.1029/2024WR038424

Real-Time Flood Inundation Modeling With Flow Resistance Parameter Learning

Alexander Young[1], John D. Albertson[1], Giovanni Moretti[2], and Stefano Orlandini[2]

[1]School of Civil and Environmental Engineering, Cornell University, Ithaca, NY, USA, [2]Dipartimento di Ingegneria Enzo Ferrari, Università degli Studi di Modena e Reggio Emilia, Modena, Italy

Key Points:

- Flow resistance parameters representing the effects of land surface topography unresolved by digital terrain model data are determined hydraulically
- A sensitivity-based flow depth sensor network design method is developed to capture in real-time the full range of flood dynamics
- Bayesian optimization is found to enable reliable and rapid flow resistance parameter learning from simulated and observed flow depths

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

A. Young,
ay434@cornell.edu

Citation:

Young, A., Albertson, J. D., Moretti, G., & Orlandini, S. (2025). Real-time flood inundation modeling with flow resistance parameter learning. *Water Resources Research*, *61*, e2024WR038424. https://doi.org/10.1029/2024WR038424

Received 22 JUL 2024
Accepted 17 DEC 2024

Abstract Emergency response to flood plain inundations requires real-time forecasts of flow depth, velocity, and arrival time. Detailed and rapid flood inundation forecasts can be obtained from numerical solution of 2D unsteady flow equations based on high-resolution topographic data and geomorphologically informed unstructured meshes. However, flow resistance parameters representing the effects of land surface topography unresolved by digital terrain model data remain uncertain. In the present study, flow resistance parameters representing the effects of roughness, vegetation, and buildings are determined hydraulically in real-time using flow depth observations. A detailed numerical reproduction of a real flood has been largely corroborated by observations and subsequently used as a surrogate of the ground truth target. In synthetic numerical experiments, flow depth observations are obtained from a network of in-situ flow depth sensors assigned to hydraulically relevant locations in the flood plain. Starting from a generic resistance parameter set, the capability of a tandem 2D surface flow model and Bayesian optimization technique to achieve convergence to the target resistance parameter set is tested. Convergence to the target resistance parameter set was obtained with 50 or fewer tandem flow + optimization iterations for each forecasting cycle in which the difference between simulated and observed flow depths is minimized. The flood arrival time errors across a 52 km^2 flood plain inundation area were reduced by 3.13 hr with respect to results obtained without optimization from a fixed range of flow resistance parameters. Performance metrics like critical success index and probability of detection reach values above 90% across the flood plain.

Plain Language Summary When a flood inundation occurs over a populated flood plain due, for instance, to a levee or dam failure, emergency response requires numerical solutions of 2D unsteady flow equations to provide real-time forecasts of flow depth, velocity, and arrival time. These forecasts are important for supporting decisions regarding how emergency personnel resources should be allocated, and how evacuation of citizens and animals should be managed. As high-resolution topographic data and terrain analysis methods have recently made possible detailed and efficient flood plain inundation models based on geomorphologically informed meshes, an effort can now be made to hydraulically determine flow resistance parameters from flow observations in real time. In this work, illustrated with a synthetic case study, flow observations are provided by an optimally designed network of in-situ sensors that record flow depth over time. 2D flow model resistance parameters are determined with a Bayesian optimization technique based on observations collected by sensors with sufficient lead time for real-time forecasting applications. With the proposed methods, flood inundations are numerically reproduced in real-time in a hydraulically meaningful manner, and this has the potential to provide reliable flood forecasts and guidance for post-event recovery.

1. Introduction

Riverine (or fluvial) flooding continues to pose a major hazard to communities across the globe (Bates et al., 2021; Merz et al., 2021; Schubert et al., 2022). This is especially true for regions that use levees as flood protection, as the resulting flood from a levee breach can be unpredictable and particularly devastating (Orlandini et al., 2015). Communities may be built closer to rivers due to the relative safety that a levee provides, but this can increase the resulting damage when a breach occurs (Di Baldassarre et al., 2015). Predicting the location of potential levee breaches is difficult because of the uncertainty around levee failure mechanisms such as slope stability (Rossi et al., 2021) or piping due to burrowing mammal activity (Balistrocchi et al., 2021; Ceccato & Simonini, 2023; Orlandini et al., 2015). Traditional flood management involves mapping potential floods based on a wide array of possible scenarios using physics-based flood inundation models (Association of State Floodplain Managers, 2020; Wu et al., 2020). While this approach gives a good idea of high-risk areas, it cannot

predict how a specific flood will unfold (Cooper et al., 2022). Accurate and efficient predictions of how an individual flood will unfold are especially needed for real-time management of a flood inundation event (Rajib et al., 2020). These predictions are important for supporting decisions regarding how emergency personnel resources should be allocated, and how evacuation of citizens and animals should be managed (Potočki et al., 2022). For instance, it is critical for emergency personnel to know what conditions they are going to find in the field during a rescue mission so that the most appropriate means and equipment is selected (Fire Brigades Union, 2015). In addition, since the location and severity of levee failures is generally unknown before a breach occurs, attempting to simulate and prepare for all possible breach scenarios quickly becomes an enormous task. Therefore, levee breach floods call for a real-time forecasting approach that can generate accurate forecasts with sufficient lead time to be actionable immediately after the breach is discovered. This study provides a proof of concept demonstrating that a flood inundation modeling system can use high-resolution topographic data, flow depth observations from a sensor network, and Bayesian optimization for flow resistance learning to enable timely, actionable forecasts after levee breaches.

In general, the accuracy of flood inundation forecasts based on the 2D unsteady flow equations (also known as the shallow-water equations) depends on knowledge of the flow forcing, the site geometry, the operation of flood infrastructure, and resistance parameters (Pappenberger et al., 2008; Teng et al., 2017). There is uncertainty in all of these factors, but for this study the focus is on the uncertainty in flow parameters specifically. Hence, a scenario is considered whereby the flow forcing, site topography and hydraulic infrastructure are all known, and the feasibility of adjusting flow resistance parameters on the fly (based on flood depth data collected from a sensor network) to improve forecast accuracy is explored. The case considered in the present study is the one in which real-time flood inundation forecasting commences upon the discovery of the levee breach, and thus the exact position in terms of geographical coordinates of a levee breach can be considered known. Under this assumption, the breach hydrograph can be modeled using a riverflow model with its own set of inputs (precipitation, initial and boundary flow conditions), parameterizations (channel roughness), and geometry (river cross sections, levee breach evolution, and breach repairs). In practice, the hydrograph is derived from upstream and downstream gauge readings of river flow as well as observations of breach evolution (Orlandini et al., 2015). The uncertainty due to flood plain topography can be addressed by increasing the resolution of the digital terrain model (DTM) used to represent the geometry of the land surface (Bates, 2004, 2022; Kahl et al., 2022; Sanders & Schubert, 2019) or by using geomorphologically informed unstructured meshes preserving the important lines following the lowest part of valleys or channels (thalwegs) and the important lines of intersection at the top between opposite slopes (ridges), and relaxing the level of detail where these topographic features are inessential (Moretti & Orlandini, 2023; Pizzileo et al., 2024). Land surface topography that is not resolved by DTM data, including roughness, is in any case, a major source of uncertainty in a 2D flood inundation model and is also difficult to estimate (Marks & Bates, 2000; Raudkivi, 1967). In fact, a more general distinction can be made between resolved land surface topography formed by natural and human-made structures that are described explicitly by DTM data, and unresolved land surface topography that is not captured by DTM data such as variations on a smaller scale than the DTM cell (e.g., human-made structures like vehicles and agricultural swales or small-scale geologic features like sand bars and rills) or topographic structures that are filtered out when DTM data are obtained from digital surface model (DSM) data (e.g., buildings and vegetation). Part of the unresolved land surface topography like buildings or vegetation has previously been reintroduced explicitly in flood inundation models as done, for instance, by Schubert and Sanders (2012) or has been described explicitly from DSM data as reported by Pizzileo et al. (2024). In the present study, no attempt has been made to reintroduce or describe explicitly unresolved land surface topography, and its effects on surface flow propagation are represented by flow resistance parameters.

While the levee breach hydrograph and flood plain topography can in principle be measured and modeled, a rigorous determination of resistance to flow due to unresolved land surface topography implies the use of observed flows, which are rare and unique in the case of flood plain inundations. Uncertainty in flow resistance parameters is, at least in part, ascribed to variability in space and time. This is especially true for flow resistance parameters describing vegetation, which change throughout the year due to agriculture, seasonality of foliage and herbaceous plants, and the impact of previous flood events, thus invalidating the use of previously calibrated parameters (Arcement & Schneider, 1989). Historically, resistance parameters were assigned to different land surface types (e.g., urban, agriculture, forest) based on land cover maps and corresponding values from empirical studies (Arcement & Schneider, 1989; Chow, 1959). More recently, improvements in the estimation of the spatial

distribution of vegetation have been made with aerial imagery (Hossain et al., 2009) and lidar surveys (Bates, 2022). Although there is also some disagreement about whether the optimal resistance parameterization of a flood model corresponds to the land cover distribution, this assumption is often the best that can be made prior to a flood or without significant observations and computational resources (Hostache et al., 2010). Even today, the resistance parameters in many 2D flood inundation models use either empirical resistance coefficients or modelers optimize them with a basic grid search (Ming et al., 2020; Ongdas et al., 2020; Quirogaa et al., 2016).

A variety of model calibration strategies have been explored in the literature, broadly classified into static and dynamic approaches. Static approaches find the optimal parameter set or parameter probability distribution given a set of measurements. Dynamic approaches such as data assimilation continuously update the states and parameters of the model as new measurements become available over time. Gallagher and Doherty (2007) explore the use of the PEST (Parameter Estimation & Uncertainty Analysis) software to calibrate a lumped hydrologic model with seven parameters requiring 7,500 model evaluations. Fabio et al. (2010) also make use of PEST to calibrate a 2-D flood model with up to five flow resistance parameters, requiring around 200 model iterations. Quantification of the joint parameter probability distribution has been explored by Vrugt et al. (2008) and Hall et al. (2011), both making use of the Markov chain Monte Carlo algorithm to estimate the joint posterior parameter distribution for hydrologic and flood model parameters. Data assimilation approaches have increasingly been used to estimate flood states and parameters. Moradkhani et al. (2005) present a dual-state method for the estimation of hydrologic model parameters. In this study, a clear distinction is made between models developed for understanding physical processes and models developed for predicting or forecasting events, and the focus is explicitly made on the latter aim. Many data assimilation studies have shown how flood inundations in river systems and flood plains are forecasted accurately when the system states, and in some cases also the boundary conditions, are updated (Annis et al., 2022; Hostache et al., 2018; Lai et al., 2014; Neal et al., 2007; Pensoneault et al., 2023; Pujol et al., 2022; Revilla-Romero et al., 2016; Van Wesemael et al., 2019). Data assimilation can also be used for estimating or tuning model parameters (Annan et al., 2005; Hendricks Franssen & Kinzelbach, 2008; Madsen & Skotner, 2005). Flow resistance parameters are, however, often updated together with systems states making it hard to distinguish the respective contributions of the two updates. In addition, Annan et al. (2005) highlight how the problem of parameter estimation in climate modeling commonly uses simpler models than the current state-of-the-art models to meet computational efficiency requirements. Pensoneault et al. (2023) present an inverse ensemble Kalman estimation of distributed surface flow parameters in a drainage basin without updating system states. Both static and dynamic parameter calibration often requires 100 or more model evaluations for convergence criteria to be met. This computational expense may limit the application in real-time flood inundation modeling applications based on the numerical solution of 2D unsteady flow equations. In addition to the poor efficiency of numerical models, the scarcity of (integrated) in-situ observations and adequate satellite data, especially in the early stage of the flood, is another limiting factor for understanding and predicting flood inundations (Barthélémy et al., 2018; García-Pintado et al., 2013; Nguyen et al., 2022). Annis et al. (2022) develop a data assimilation method for incorporating crowdsourced data into a flood inundation model. The poor efficiency of detailed numerical models and the scarcity of reliable observations remain, however, critical issues in understanding and predicting flood inundations.

Instead of preselecting parameters or using ensemble forecasting, the real-time flood inundation model described in the present study is conditioned on the latest observations of flood behavior to produce, at any given time, hydraulically meaningful flow resistance parameter estimates that provide the best agreement between simulated and observed flows up to that time, and hence more accurate flood predictions for future times (Kim et al., 2019). Two-dimensional flood models are notoriously expensive to evaluate, and even with increasing computational power, evaluation times can range from hours to days (Ming et al., 2020). However, a recent development in geomorphologically informed mesh generation techniques based on thalweg and ridge networks extracted from high-resolution topographic data using the LANDMARK algorithm has dropped the run time from the order of hours to the order of minutes while retaining a forecast resolution of 1-m or less where strictly needed (Moretti & Orlandini, 2023; Pizzileo et al., 2024). Even so, model evaluation times on the order of minutes begin to compound because each time a forecast is requested, the model must be recalibrated with any new observation data available, potentially requiring hundreds or thousands of evaluations. Consequently, the number of model evaluations to arrive at a calibrated flow resistance parameter set must be kept to a minimum. To tackle the challenge of real time flood forecasting with a high-fidelity model, this present study implements an efficient Bayesian optimization approach. The Bayesian optimization approach converges to the optimal parameter set

with as few model evaluations as possible by efficiently minimizing the difference between the simulated and observed flow depths (Frazier, 2018). Bayesian optimization is a powerful, derivative-free model calibration tool that was designed to find the optimal parameter set for computationally expensive models (Frazier, 2018). Though there are limited applications to hydrologic models found in the literature, Ma et al. (2022) were able to reduce the number of model evaluations required to calibrate the SWAT model from 1,500 to 150 using a Bayesian optimization approach. No applications of Bayesian optimization of high-fidelity flood inundation models were found during the literature search, and so this research presents a novel application of the algorithm to the area of flood forecasting. Reducing the number of model evaluations required to calibrate model parameters will help open the door to high-resolution, real-time flood inundation flood forecasting.

Real-time flood forecasting requires observation data of past and current flood states in order for the calibration strategy to produce accurate forecasts. Historically, flood models could only be calibrated with data acquired after the event, using eyewitness reports and high water marks as benchmarks (Arcement & Schneider, 1989; Barnes, 1967; Hicks & Mason, 1991). Recent developments in satellite remote sensing have allowed for high-resolution spatial observation of flood extents, thus providing a robust source of data to calibrate against. However, in real-time forecasting scenarios, satellite imagery is often too temporally sparse or obstructed by clouds to provide a reliable source of flood observations (Bates, 2022). In levee breach flooding scenarios, an alternative to a remote sensing approach is to use in-situ flow depth sensors in the flood plain to obtain a nearly continuous set of flood observations. Even in urban areas where sensors may seem more accessible, there is still a scarcity of readily available flood observation data (Gallien, 2016; Rosenzweig et al., 2021). Previous research has investigated techniques for locating flow depth sensors in the flood plain, finding that sensor locations are most informative when they observe the largest difference in flood behavior from previously placed sensors (Liggett & Chen, 1994; Van Wesemael et al., 2019). The idea that sensor location should be sensitive to model parameters is critical for accurate and rapid model calibration, however, due to the scarcity of in-situ flood plain sensing research, this approach has not yet been explored in this context. In the present study, flow resistance parameters of the Manning type are determined directly from the observation of the flow occurring in the specific flood plain and under the specific circumstances of interest. To maximize the feasibility of this strategy during real flood inundation events, preplacement of flow depth sensors or their housings in the flood plain is informed prior to a breach by an intensive simulation of multiple possible breach locations. Due to the inherent a priori uncertainty in flow resistance parameters, an approach anchored in hydraulics and sensitivity is the appropriate strategy for determining hydraulically equivalent resistance parameters with respect to those originally used in laboratory experiments under controlled flow conditions (Chow, 1959; Ligrani & Moffat, 1986; Nikuradse, 1950; Raudkivi, 1967; Shockling et al., 2006).

The contribution of the present research centers on the necessity for an accurate flood inundation modeling framework in effective real-time flood forecasting. Model detail and efficiency in the numerical solution of the 2D unsteady flow equations are achieved through the use of geomorphologically informed unstructured meshes that capture the topographic detail offered by lidar surveys. Based on the combined use of the LANDMARK algorithm and of the HEC-RAS 2D flood inundation model, a framework is presented that addresses the combined challenge of designing a distributed flood observation network and efficiently incorporating the observations into a continuous forecasting system. The methods are presented in Section 2, the applicability and impact of these methods are evaluated in Section 3, the results are discussed in Section 4, and the conclusions are finally given in Section 5.

2. Methods

The proposed real-time flood forecasting framework is presented in Section 2.1. The attention is then focused on the HEC-RAS 2D surface flow model in Section 2.2, on the characterization of the land surface topography in Section 2.3, on flow resistance parameters in Section 2.4, on the design of a sensitivity-based flow depth sensor network in Section 2.5, and on real-time calibration of flow resistance parameters with Bayesian optimization techniques in Section 2.6.

2.1. Real-Time Flood Forecasting Framework

The real-time flood forecasting framework presented in this paper is designed around three key objectives: flood simulation accuracy, spatial resolution, and evaluation efficiency. Simulation accuracy ensures that flood

forecasts have utility for guiding emergency planning during flood events. Spatial resolution is key for understanding the spatial impact of a flood event. Simulation evaluation efficiency ensures there is sufficient time to act on flood forecasts. Figure 1a is a schematic of the framework components that, when combined, address all three objectives. The framework consists of pre-flood model instantiation (shown by the black boxes) and a real-time resistance parameter calibration procedure that is performed during the flood event (shown by the red boxes). Before the flood, a 2D surface flow model is created to represent possible inundations of the flood plain domain. On the left side of Figure 1a, this involves generating an unstructured computational mesh that conforms to the flood plain terrain using the LANDMARK terrain analysis algorithm developed by Moretti and Orlandini (2023) and described further in Section 2.3. The black box on the right side of Figure 1a represents the delineation in the 2D surface flow model of unresolved land surface topography due, for instance, to roughness and vegetation features that are not captured by the numerical mesh based on DTM data.

Since the flow resistance parameterization is unique to each flood event and may even vary during the event, the parameter values in the model must be calibrated as the flood occurs. The procedure of continual updating of parameters is described in the red boxes and arrows in Figure 1a. The sensors collect time series of flood flow depth that are used as the observations in a real-time resistance parameter calibration algorithm which selects parameter values that result in the closest match between the simulated and observed flood up to that point in time. Figure 1b is an illustration of how the real-time forecasting framework would be implemented during a flood event. A levee breach occurs at time t_0, which begins the real-time forecasting process. Once flow depth sensors are deployed, measurements are collected by the network of distributed flow depth sensors throughout the flood plain during the flow depth sensing step. Finally, the calibration algorithm updates the resistance parameters representing the effects of land surface topography unresolved by DTM data and thus not contributing to the flood plain hydraulic geometry in the 2D surface flow model as more flood observation data becomes available. During the flood event, data is collected from the sensor network over a period of time Δt. This data is fed into the calibration algorithm, the model is evaluated until the flow resistance parameter set converges to the optimum, and a downstream forecast is generated from the calibrated model. After a second sensing period, the calibration cycle is repeated, and a second forecast is generated. The sensing and calibration process continues until forecasts are no longer necessary. It should be noted that in this framework, the model parameters are calibrated on the full time series of observed surface flow data from t_0, rather than only on the data collected since the previous forecast.

2.2. HEC-RAS 2D Surface Flow Model

The choice of the 2D surface flow model influences the forecast accuracy and run time. When forecasting floods in real time, it is desirable to use any surface flow model which captures all topographic features that are resolved by DTM data and affect flow propagation. Roads and buildings are often only a few meters wide, and therefore assessing whether inundation will occur at a given location requires a resolution finer than that of the subject of interest. More generally, providing finer-scale data about flood inundation events better informs all aspects of flood management. Fine-resolution regular grids capable of describing macrotopographic structures affecting surface flow propagation may require prohibitive computational costs that are especially unsuitable to the proposed real-time forecasting framework. Recent advances in terrain analysis, however, make it possible to preserve thalweg and ridge networks extracted automatically, and thus objectively, from unaltered high-resolution topographic data when an unstructured computational mesh is generated (Moretti & Orlandini, 2023; Pizzileo et al., 2024). The combined use of the Moretti and Orlandini's (2023) LANDMARK terrain analysis algorithm and the HEC-RAS 2D surface flow model make possible significantly more accurate and efficient simulations of flood inundations than more traditional methods based on regular grids (Hydrologic Engineering Center, 2016; Moretti & Orlandini, 2023). The HEC-RAS 2D surface flow model requires three components for instantiation: (a) a representation of the land surface topography, (b) a boundary condition (e.g., a levee breach hydrograph), and (c) a parameterization of the surface roughness, more generally of the land surface topography that is unresolved by DTM data. The model accepts a DTM as the representation of the flood plain topography. The flood domain is discretized into an unstructured numerical solution mesh with the above-mentioned mesh generation algorithm. In the case of a levee breach, the model takes a levee breach hydrograph as input and solves the 2D unsteady flow equations across the mesh to simulate the flood inundation.

The 2D unsteady flow equations are derived by assuming that the water density is constant in space and time. When the terrain slope is small, the mass balance equation can be written as

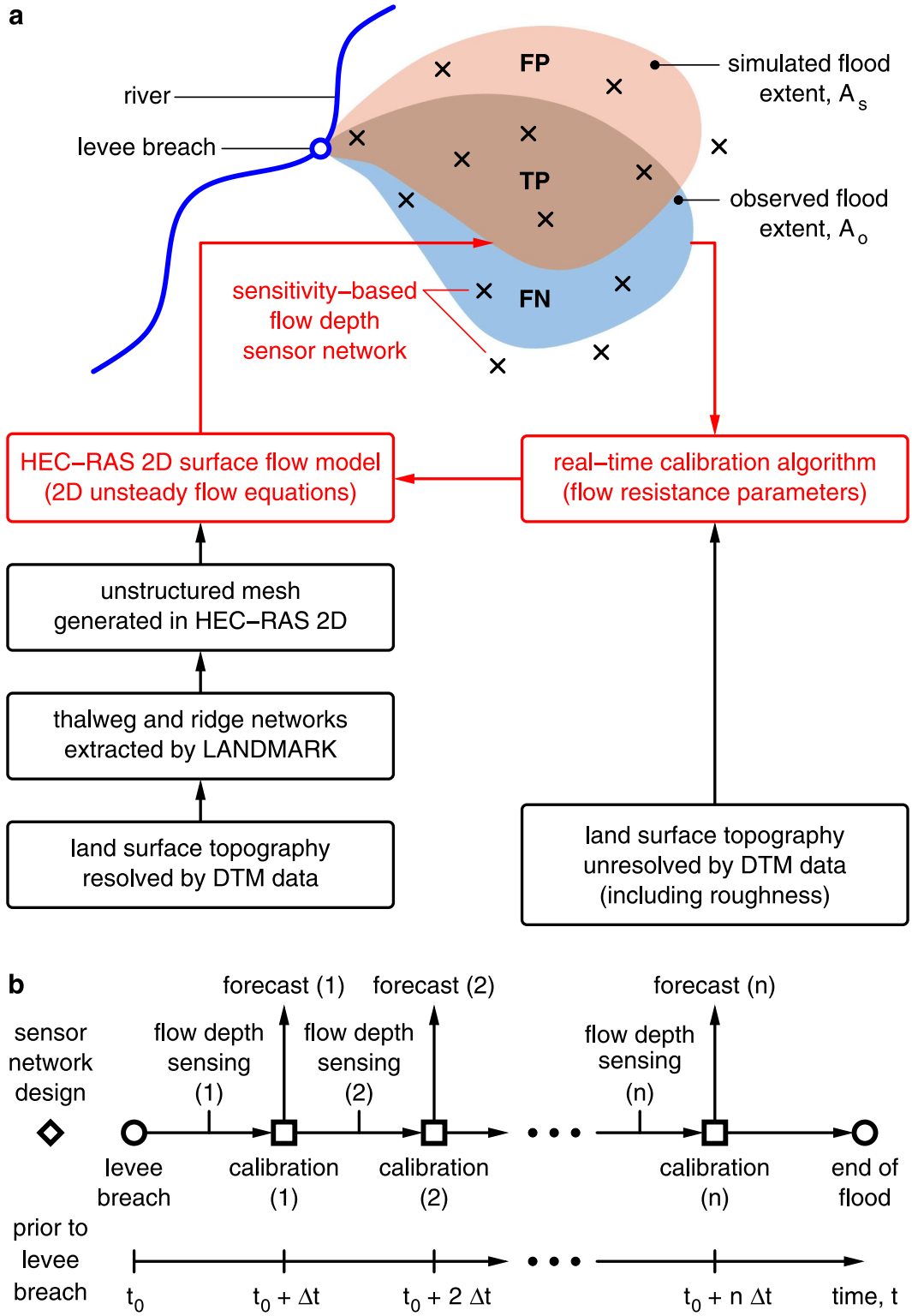

Figure 1.

$$\frac{\partial Z}{\partial t} + \frac{\partial (Zu)}{\partial x} + \frac{\partial (Zv)}{\partial y} = q, \tag{1}$$

where t is time, x and y are horizontal coordinates, Z (capital ζ) is the flow depth along the normal coordinate to the land surface, u is the component of flow velocity $\mathbf{v}$ along the x-direction, v is the component of flow velocity $\mathbf{v}$ along the y-direction, and q is the lateral inflow discharge per unit planar area. The projections of the momentum equation along the x and y directions can be written, respectively, as

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + g\frac{\partial h}{\partial x} - c_f u = 0, \tag{2}$$

and

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + g\frac{\partial h}{\partial y} - c_f v = 0, \tag{3}$$

where g is the acceleration due to gravity, $h = z + Z$ is the piezometric head, or hydraulic head, z is the land surface elevation, $c_f u = gn^2|\mathbf{v}|u/Z^{4/3}$ and $c_f v = gn^2|\mathbf{v}|v/Z^{4/3}$ are the friction slopes expressed by using the Manning's resistance coefficient n multiplied by g. In the HEC-RAS 2D flood model, the spatial distribution of Manning's n is encoded in a categorical raster file where each pixel value corresponds to a particular value of Manning's n. The wall-clock run time of each model evaluation is on the order of 10 minutes, and thus an efficient calibration algorithm is still required to enable real-time forecasting. As part of the real-time forecasting framework, preprocessing the land surface topography resolved by the DTM data in order to obtain a geomorphologically informed mesh in the flood inundation model is time intensive and is performed prior to the levee breach. A second preprocessing step is to make an initial guess of the resistance parameters that represent the effects of unresolved land surface topography based on land surface features. Upon discovery of the levee breach, the real-time model configuration steps are the characterization of the breach geometry including repair and parameterizations of flow resistance parameters representing unresolved land surface topography.

2.3. Land Surface Topography

A detailed and accurate description of the land surface topography is essential for ensuring a reliable prediction of surface flow propagation (Hydrologic Engineering Center, 2022, p. 14). In the present study, a distinction is made between macrotopography resolved by DTM data and used to determine the hydraulic geometry of the flood plain, and microtopography due to surface features, buildings, and vegetation that is not resolved by DTM data and is represented by flow resistance parameters. Resolved land surface topography can be described by a high-resolution (1-m or less) DTM capturing natural and human-made topographic structures such as ridges or levee, road, and railway embankments. Thalwegs and ridges are the essential topographic features that need to be extracted from observed topographic data and retained when terrain is partitioned. Thalwegs indicate main flow collectors and depression bottoms. Ridges indicate flow divides and barriers to flow. Since depression bottoms are thalwegs and depression borders are ridges, the preservation of depressions and the preservation of thalweg and ridge network structures are interrelated requirements. Using the LANDMARK terrain analysis algorithm described in Moretti and Orlandini (2023), natural and human-made thalwegs and ridges are extracted automatically, and thus objectively, from the original high-resolution (1 m or less) DTM without the need to apply coarsening or depression filling in the extraction procedure. The detail with which the extracted thalweg and ridge networks are represented is selected by setting a critical drainage area for thalweg representation and a critical dispersal area for ridge representation. The generation of an unstructured computational mesh that adapts to the

Figure 1. (a) Schematic of the flood forecasting framework showing how the LANDMARK terrain analysis and real-time resistance parameter algorithm fit together to generate accurate and efficient flood forecasts. The red arrows and boxes are the components presented in this paper. A sensitivity-based sensor network is overlaid on an illustration of the observed and simulated floods due to a levee breach. The true positive, false positive, and false negative areas are also shown. (b) Timeline of the continuous model calibration process after the levee breach. The breach occurs at time t_0 and forecasts are generated every Δt time step after model calibration with new observations.

extracted thalwegs and ridges makes it possible to describe accurately land surface topography and the related surface flows, as mesh element facets adapts to observed topographic features, while also increasing computational efficiency with respect to the use of regular grids, as a small number of mesh elements displaying variable size are used.

Extracted ridges are set as breaklines in HEC-RAS 2D so that an unstructured computational mesh adapting to these topographic features can be generated and the 2D unsteady flow equations can be accurately solved. It is specified that the 1-m DTM describing the land surface macrotopography is fully assimilated by HEC-RAS 2D along the facets of mesh elements and is only used to determine the distribution of surface water storage within the mesh elements (Hydrologic Engineering Center, 2022, p. 201). Ridges extracted by the LANDMARK algorithm act in HEC-RAS 2D as barriers to flow if they are set as breaklines and aligned to the mesh element facets. If extracted ridges are not aligned to the mesh element facets, they only affect the distribution of surface water storage within the mesh elements and are not seen as barriers to flow by the HEC-RAS 2D solver of the 2D unsteady flow equations. To ensure accurate and efficient surface flow simulations, it is therefore essential that ridges are identified and incorporated in HEC-RAS 2D as breaklines to obtain computational mesh that captures the topographic detail offered by high-resolution topographic data where needed, yet coarsens the mesh elements where terrain is regular. With respect to existing flood inundation models, the innovation in the combined use of LANDMARK and HEC-RAS 2D is not just a simple use of observed ridges as breaklines, which can even be performed manually, but rather the ability of LANDMARK to fully penetrate high-resolution topographic data for extracting ridges at any desired level of detail. Ridges and breaklines can in principle be identified and incorporated in HEC-RAS 2D independently from LANDMARK, but in practice this process becomes a prohibitive task without LANDMARK as the level of mesh refinement increases (Moretti & Orlandini, 2023; Pizzileo et al., 2024). The combined use of the LANDMARK algorithm and HEC-RAS 2D surface flow model makes it possible to obtain accurate and computationally efficient simulations of flood inundations as reported in Moretti and Orlandini (2023). The 1-m DTM used in the present study is obtained from the cloud of points acquired from a lidar survey by filtering vegetation and buildings. The effects of these filtered topographic structures and of the land surface microtopography that is not resolved by the DTM are represented in the HEC-RAS 2D surface flow model by seeking equivalent flow resistance parameters.

2.4. Flow Resistance Parameters

Land surface topography unresolved by DTM data is parameterized with flow resistance parameters of the Manning type. The resistance to flow in channels or flood plains of fixed geometry carrying clear water in a steady uniform flow can be predicted quite accurately (Nikuradse, 1950). But when the same channels or flood plains carry clear water in a nonuniform state of flow, the resistance problem becomes more complex. Past field studies have found that Manning's n for vegetation can range from 0.03 to 0.11 $m^{-1/3}$ s depending on the density and height of plants (Arcement & Schneider, 1989; Chow, 1959). As will be seen in Section 3, the impact of Manning's n on flood arrival time shows that changes as small as 0.02 $m^{-1/3}$ s can produce arrival time differences of more than 2 hr for downstream locations over a flood inundation with an area of 52 km^2 and a length of 28 km. Floods are highly sensitive to surface roughness, and accurate estimates of Manning's n are critical for ensuring that flood forecasts provide useful information. Flow resistance parameters are not the only source of uncertainty in a flood model, however, they are difficult to estimate prior to a flood and thus are the sole focus of the present study.

In the present study, flow resistance parameters are initially estimated on the basis of the best existing methods, and are then refined iteratively in real time by comparing forecasted and observed flows. Prior to a flood, the distribution of Manning's n is delineated based on assumptions about the land surface. This involves deciding on the number of land cover classes to include, each with its own Manning's n, and where to place those classes. Traditionally, land cover maps have been used to create the land cover classes used in hydrologic models (Arcement & Schneider, 1989). Given limited computational resources, optimization of the corresponding Manning's n parameters becomes computationally intensive when many different classes are delineated. Instead, it is preferable to discretize the land surface into only a few classes, thus preserving model efficiency and avoiding overparameterization. These parameters can then be refined using the real-time flood forecasting framework as new flood observations are collected and can be compared with the corresponding model forecasts. Due to the complex and integrated effect of flow resistance, refinement of a resistance parameter requires that the flood wave has interacted with the land area associated with that resistance parameter and then been measured by a flow depth

Table 1
Factorial Experiment Design for Three Roughness Classes of Manning's n

Treatment	n_1	n_2	n_3
1	+	+	+
2	+	+	−
3	+	−	+
4	+	−	−
5	−	+	+
6	−	+	−
7	−	−	+
8	−	−	−

sensor. Land cover classes that have not influenced the propagation of the flood are not able to have their flow resistance parameters calibrated, and thus best estimates from literature must be used.

2.5. Sensitivity-Based Flow Depth Sensor Network

Refinement of the flow resistance parameters requires observations of the flood state. Ideal observations are distributed in space and continuous in time. However, in practice there is usually a trade-off between spatial and temporal coverage. Satellites provide high spatial coverage but suffer from long return periods on the order of days and the image can be obscured by clouds (Bates, 2022). For this reason, this research examines using a network of in-situ flow depth sensors as the source of calibration data. In-situ sensors observe flow depths over time at a known geospatial location and operate continuously throughout the flood. They have the added benefit of being a flexible network where new sensors can be added and replaced as needed. Due to the uncertainty of a breach location and the rapid response needed after a breach occurs, the sensor locations must be preallocated based on an intensive simulation of possible breach locations as depicted in Figure 1b. Upon determination of a sensor network design that addresses multiple breach locations, the sensors or their housings can be permanently installed in the flood plain. If only the housings are installed, it is assumed that during emergency mobilization after the discovery of a breach, there is sufficient time to place sensors in the housings that are most appropriate for observing the resulting flood. The design of a flow depth sensor should be robust, cost-effective, and easy to install. A recent project in New York City called FloodNet has developed and installed ultrasonic flow depth sensors, which serves as an example of a flood observation network (Silverman et al., 2022). An important part of sensor network design is determining where sensors should be located to provide as much information as possible about the state of the flood. To achieve this, a sensitivity-based sensor placement algorithm was developed. The design of the flood plain sensor network in this paper focuses on its efficacy in enabling accurate calibration of the flow resistance parameter set (i.e., Manning's n). Locations in the flood plain that are sensitive to changes in upstream Manning's n are logically the best locations to place sensors. Conversely, if a location is not sensitive to changes in upstream Manning's n, it will not be useful for calibrating the resistance parameter set. The strategy for allocating sensors in the flood plain is to find those locations that are most sensitive to changes in all of the upstream Manning's n values that require calibration. For example, if the land surface has been divided into three roughness classes (vegetation, urban, and canal), the sensors should be in locations that see the greatest difference in flow depth as each Manning's n is varied. Disambiguating the sensitivity of specific locations to individual resistance parameters is complicated by the integrated effect of these parameters across the upstream contributing area. This requires a procedure that isolates each parameter's influence on the downstream area.

Sensitivity of locations to changes in Manning's n are assessed with a factorial experiment design. As an example, the land surface is partitioned into three resistance parameter classes and assigned Manning's n ranges of $\mathbf{n} = (n_1, n_2, n_3)$. Using values from literature, for example, from Chow (1959), the minimum and maximum realistic values of (n_1, n_2, n_3) are collected into two treatment arrays for "high," $\mathbf{n}^+ = \left(n_1^+, n_2^+, n_3^+\right)$, and "low," $\mathbf{n}^- = \left(n_1^-, n_2^-, n_3^-\right)$. The factorial experiment design is shown in Table 1 where "+" denotes the high Manning's n treatment for class f, n_f^+, and "−" denotes the low Manning's n treatment for class f, n_f^+. Each of the eight combinations of Manning's n are then used to run the flood model up to time T, generating eight corresponding depth time series for each location in the flood plain. The matrix of depth time series for the ith location is as follows

$$\mathbf{D}_i = \begin{bmatrix} d_{i,1}(t_1) & d_{i,1}(t_2) & \cdots & d_{i,1}(t_T) \\ d_{i,2}(t_1) & d_{i,2}(t_2) & \cdots & d_{i,2}(t_T) \\ \vdots & \vdots & \vdots & \vdots \\ d_{i,8}(t_1) & d_{i,8}(t_2) & \cdots & d_{i,8}(t_T) \end{bmatrix}, \quad (4)$$

where each row of $\mathbf{D}_i$ corresponds to the depth time series at location i for one of the eight factorial experiments in Table 1. The power of factorial experiments is that the sensitivity of a location to changes in each Manning's n can be quantified separately. Each depth time series in $\mathbf{D}_i$ can be assigned to a treatment group for each n_i based on whether the value of n_f was high or low. For example, rows 1−4 in $\mathbf{D}_i$ correspond to n_1^+ and rows 5−8 correspond to n_1^-. The sensitivity to changes in Manning's n is then quantified by the difference in means between the high group and the low group. When the group means are far apart, the location is relatively more sensitive to changes in Manning's n as to when the means are close together. Since the goal is to find the location that is most sensitive to changes in Manning's n, only the relative difference between locations matters.

The procedure is as follows:

1. Assign all depth time series in $\mathbf{D}_i$ to a factor's treatment group (i.e., n_f^+ and n_f^-).
2. Compute the ensemble mean vectors of each group as

$$\boldsymbol{\mu}_{i,f}^{+}(t) = \langle \mathbf{D}_{i,f}^{+}(t) \rangle \tag{5}$$

and

$$\boldsymbol{\mu}_{i,f}^{-}(t) = \langle \mathbf{D}_{i,f}^{-}(t) \rangle. \tag{6}$$

3. Compute the Euclidean distance between the two mean vectors, called the class sensitivity, K_f.

$$K_f = \left[\sum_{t=1}^{T} \left(\boldsymbol{\mu}_{i,f}^{+}(t) - \boldsymbol{\mu}_{i,f}^{-}(t) \right)^2 \right]^{1/2}. \tag{7}$$

Sensors are then assigned to the locations with the highest value of K_f. In practice, because the algorithm identifies numerical mesh cells that are sensitive to changes in Manning's n rather than point locations on the land surface, the last step is to identify the location within the cell that the sensor should be allocated. This can be done by visual inspection of inundated areas at the scale of the DEM or simply by choosing the location with the lowest elevation within the cell. Figure 2 shows the mean depth time series for the high and low treatment groups, $\boldsymbol{\mu}_{i,f}^{+}(t)$ and $\boldsymbol{\mu}_{i,f}^{-}(t)$, for two numerical mesh cells in the model used in the case study (Section 3.5). Cell 2 has a larger difference between the group means and thus is more sensitive to changes in Manning's n. The class sensitivities for cell 1 and cell 2 are $K_f = 37.7$ m and $K_f = 6.2$ m, respectively.

2.6. Real-Time Calibration Algorithm

Real-time flood inundation modeling places a strict time constraint on the calibration procedure. A forecast must be issued with sufficient lead time to allow for evacuation and other preparation. In this study, we first define an objective function that quantifies the modeled flow depths in relation to the flood observations. This is done by computing the root-mean-square error (RMSE) between the modeled and observed response at each sensor location, denoted RMSE_j for the jth location as given by

$$\mathrm{RMSE}_j = \sqrt{\frac{1}{K} \sum_{k=1}^{K} \left(\hat{d}(t_k) - d(t_k) \right)^2}, \tag{8}$$

where $\hat{d}(t_k)$ is the simulated flow depth at time t_k and time step k and $d(t_k)$ is the observed flow depth at time t_k and time step k. The mean of all the RMSE values from each sensor location, $\langle \mathrm{RMSE}(\mathbf{n}) \rangle$, is then the objective function value given a set of Manning's n, $\mathbf{n} = (n_1, n_2, \ldots)$ as given by

$$\langle \mathrm{RMSE}(\mathbf{n}) \rangle = \frac{1}{N} \sum_{j=1}^{N} \mathrm{RMSE}_j. \tag{9}$$

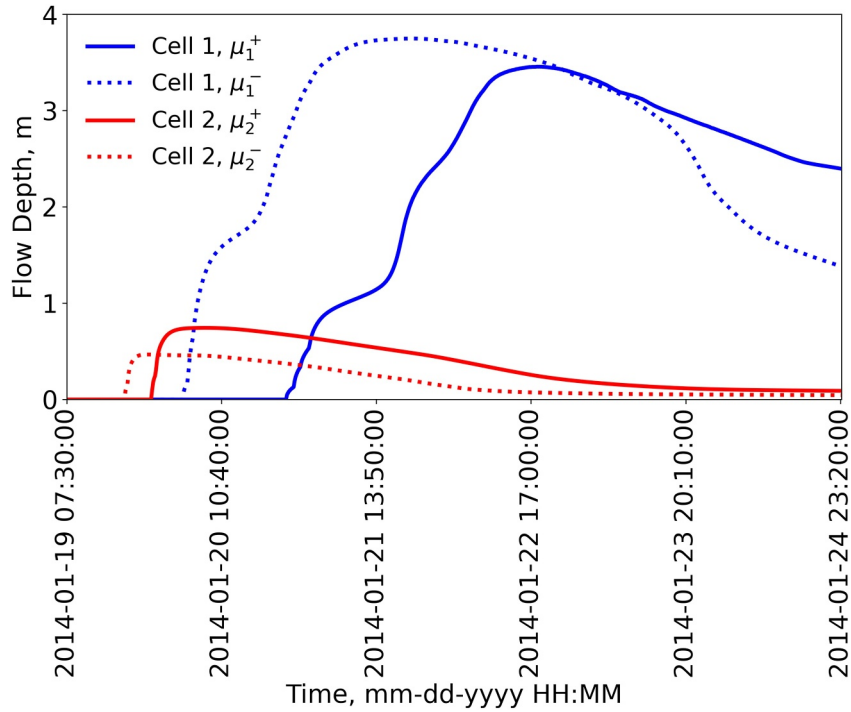

Figure 2. Mean depth time series from each treatment group for two cells. Cell 1 has a class sensitivity of $K_f = 37.7$ m and cell 2 has a class sensitivity of $K_f = 6.2$ m.

The goal of calibration is to find the set of Manning's n that minimizes $\langle \text{RMSE}(\mathbf{n}) \rangle$. Each time the calibration procedure is run, the model is calibrated on the full depth time series from each flow depth sensor starting from the levee breach initiation.

For calibration of the resistance parameters, a Bayesian optimization algorithm is implemented to converge on the parameter set that minimizes $\langle \text{RMSE}(\mathbf{n}) \rangle$. Bayesian optimization has been primarily used in the field of machine learning for finding optimal parameter sets for computationally expensive neural networks using the fewest model iterations. It is most useful for models that are expensive to evaluate and when an analytical solution is intractable (Frazier, 2018). HEC-RAS 2D fits this profile because evaluations require minutes of wall-clock time and the solution is computed numerically. Mathematically, the goal of Bayesian optimization is to efficiently solve the inverse problem as given by

$$\mathbf{n}^* = \arg\max_{\mathbf{n}} f(\mathbf{n}), \tag{10}$$

where $\mathbf{n}^*$ is the parameter set that produces the global optimum of the objective function $f(\mathbf{n})$. The Bayesian optimization procedure is comprised of two steps: a Gaussian process regression surrogate, which approximates the objective function and an acquisition function, which chooses the next parameter set with which to evaluate the model. The Gaussian process surrogate provides a flexible representation of the objective function surface and allows for new function observations to be incorporated while retaining its Gaussian properties (Rasmussen & Williams, 2006).

A Gaussian process can be thought of as a multivariate normal distribution over a set of functions that span the parameter space. The distribution is fully specified by a mean function, μ, that governs the objective function mean at any location $\mathbf{n} = (n_1, n_2, \dots)$ and a covariance matrix, $\mathbf{\Sigma}$, that encodes the covariance structure of the objective function. The mean function is often taken to be a constant mean of the observed values of the objective function. Some assumptions have to be made about the covariance structure of the objective function like, for instance, if it is isotropic and homogeneous. The covariance function between points, also known as the kernel, is then used to generate the covariance matrix. The Matérn kernel is commonly used because of its flexibility in

representing covariance structures (Frazier, 2018). The Gaussian process distribution is generally written as $f \sim N(\mu, \mathbf{\Sigma})$ where f is the objective function, μ is the mean function, and $\mathbf{\Sigma}$ is the covariance matrix. As new observations of the objective function are made at locations in parameter space, the Gaussian process prior distribution is updated to reflect the additional information and the posterior distribution remains Gaussian. The posterior distribution, following Rasmussen and Williams (2006), is computed with the conditional distribution of the multivariate normal distribution as given by

$$f_{\text{test}} | \mathbf{n}_{\text{test}}, \mathbf{n}_{\text{obs}}, f_{\text{obs}} \sim N(\mu, \mathbf{\Sigma}), \tag{11}$$

where f_{obs} are all the previous observations of the objective function at locations $\mathbf{n}_{\text{obs}}$ and f_{test} is the objective function value to estimate at location $\mathbf{n}_{\text{test}}$. The mean function, μ, is given by

$$\mu = K(\mathbf{n}_{\text{test}}, \mathbf{n}_{\text{obs}}) K(\mathbf{n}_{\text{obs}}, \mathbf{n}_{\text{obs}})^{-1} f_{\text{obs}}, \tag{12}$$

and the covariance matrix, $\mathbf{\Sigma}$ is given by

$$\mathbf{\Sigma} = K(\mathbf{n}_{\text{test}}, \mathbf{n}_{\text{test}}) - K(\mathbf{n}_{\text{test}}, \mathbf{n}_{\text{obs}}) K(\mathbf{n}_{\text{obs}}, \mathbf{n}_{\text{obs}})^{-1} K(\mathbf{n}_{\text{obs}}, \mathbf{n}_{\text{test}}), \tag{13}$$

where $K(\cdot, \cdot)$ is the covariance function or kernel between two locations in parameter space.

The acquisition function uses the properties of the Gaussian process surrogate to select new sets of parameters that have the highest probability of producing the global minimum or maximum of the objective function (Frazier, 2018). For example, the expected improvement (EI) acquisition function finds the location in parameter space that maximizes the expected value of the difference between the posterior distribution given a new sampling location in parameter space, $f(\mathbf{n})$, and the current maximum observed value, $f^*(\mathbf{n})$ (Jones et al., 1998). The EI is given by

$$\text{EI}(\mathbf{n}) \coloneqq E\left[(f(\mathbf{n}) - f^*(\mathbf{n}))^+\right], \tag{14}$$

where $a^+ = \max(a, 0)$ (Frazier, 2018). The parameter set $\mathbf{n}$ with the maximum EI($\mathbf{n}$) is then the new sampling location for the next model evaluation. Expected improvement is not the only possible acquisition function, and others have been developed for different use cases such as the knowledge gradient and entropy search among others (Frazier et al., 2009; Hennig & Schuler, 2012). The Bayesian optimization procedure is run serially where in each iteration a parameter set is chosen by the acquisition function, the model is evaluated, and the Gaussian process surrogate is updated with the new value of the objective function. An illustration of the Bayesian optimization procedure is shown in Figure 3. An initial parameter set is chosen based on the best estimate of Manning's n from land surface characteristics and the HEC-RAS 2D flood inundation model is evaluated. The simulated water depths are compared to the true water depths using a predetermined metric (e.g., $\langle \text{RMSE}(\mathbf{n}) \rangle$) that serves as the objective function. Given the parameter set and the objective function value, the Gaussian process surrogate is updated and the acquisition function is evaluated to determine the next parameter set. The process is repeated until a convergence criterion is met or the computational budget for evaluation time is reached.

2.7. Forecast Assessment

Once forecasts are generated from the calibrated parameter values for each experiment, the accuracy is assessed with two sets of metrics that are relevant to emergency planning and public confidence in flood prediction. The first metric addresses the timing of the flood wave at locations throughout the domain. The second set of metrics measures how well the forecast predicted inundation at all locations in the flood plain.

2.7.1. Flood Wave Arrival Time

During a flood, it is important to predict when water will inundate critical locations such as roads and infrastructure so that emergency resources can be allocated efficiently. To address this, a critical flow depth, d_c, is defined as the depth of flood water above which roads become impassable or there is damage to critical infrastructure. This depth could be different for different types of infrastructure and can be adjusted based on the

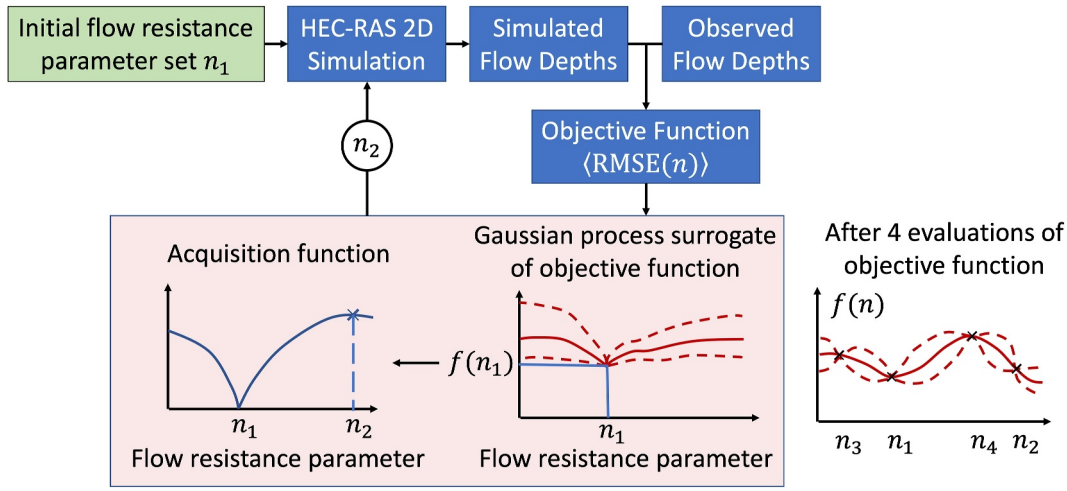

Figure 3. Illustration of the Bayesian optimization procedure as applied to the HEC-RAS model that is parameterized with a single value for Manning's n. Parameters n_1, n_2, n_3, n_4 are obtained from sequential iterations. After four iterations, n_4 is the best estimate of the optimal parameter set since it produces the current global maximum. The solid red line represents the mean function and the dotted lines represent the 95% confidence interval.

forecast goals. In this study, the critical flow depth was defined as 10 cm at each location in the flood plain. The critical flow depth of 10 cm was chosen because vehicle speed slows by 80% as flow depths reach that level, thus significantly slowing evacuation (Pregnolato et al., 2017). The time to the critical flow depth, t_c, is the elapsed time between the levee breach and when the water reaches d_c at a given location. The observed time to critical flow depth is denoted t_c, while the simulated or forecasted time to critical flow depth is denoted $\hat{t}_c$. The difference in time to critical flow depth between the observations and simulation is termed the critical flow depth time discrepancy and is denoted δ_{t_c} (Figure 4). A positive value of δ_{t_c} means that the simulated flood wave arrived later than the observed flood wave. A negative value of δ_{t_c} means that the simulated flood arrival time is earlier than what was observed. Negative values are therefore a more conservative case, while positive values mean that the model forecasted a later arrival time than in reality. Model performance is evaluated by computing the critical flow depth time discrepancy for all flooded cells over the duration of the flood. With a more accurate model calibration, the critical flow depth time discrepancy should move closer to zero for all inundated cells.

$$\delta_{t_c} = \hat{t}_c - t_c \tag{15}$$

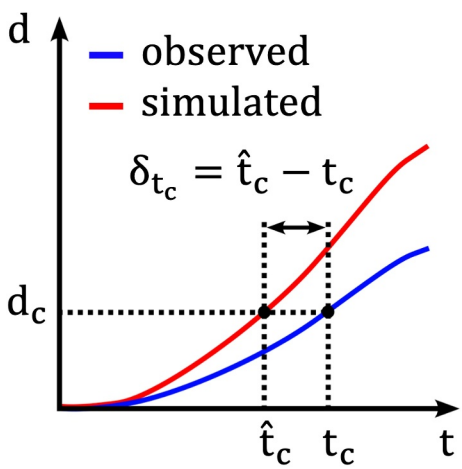

Figure 4. Illustration of the critical flow depth time discrepancy for a single location. In this example, the simulation predicted the flood wave would exceed the critical depth earlier than reality, thus $\delta_{t_c} < 0$.

2.7.2. Inundation Metrics

Inundation metrics quantify how well the model simulation captured the observed inundation extent. This is a binary classification where cells are either inundated or not inundated. Flood model performance is grouped into three categories: true positive (TP), false negative (FN), and false positive (FP). True positive is the area where the simulated and observed inundation overlap; FN is defined as the area where inundation was observed, but not simulated; and FP is defined as the area where inundation was simulated, but not observed. True negative is unbounded and thus not useful. In the equations below, TP is TP, FP is FP, and FN is FN. The diagram at the top of Figure 1a shows an illustration of the TP, FP, and FN areas along with the simulated flood extent, A_s, and the observed flood extent, A_o. Three metrics are calculated to quantify how the simulated inundation compares with the observed inundation. The false alarm ratio (FAR) is the ratio of the FP area, FP, to the simulated flood extent, $A_e = \text{TP} + \text{FP}$, that is

$$\text{FAR} = \frac{\text{FP}}{\text{TP} + \text{FP}}. \tag{16}$$

The FAR is zero when there are no false positives and the flood extent is equal to or entirely contained within the observed flood extent. The second metric is the critical success index (CSI) which is the ratio of the TP area to the extent of the union of the simulated and observed flood extents, that is

$$\mathrm{CSI} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP} + \mathrm{FN}}. \tag{17}$$

The CSI takes a value of unity when the simulated and observed flood extents overlap completely. The presence of either false positives or negatives will reduce the value toward zero. The third metric is the probability of detection (POD) which is the ratio of the TP area to the observed flood extent, that is

$$\mathrm{POD} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}. \tag{18}$$

The POD takes a value of unity when the TP area completely encompasses the observed inundation area. A value of unity also implies that the model produces no false negatives.

3. Case Study

The real-time forecasting framework is illustrated for a flood event that occurred in Northern Italy in January 2014 (Orlandini et al., 2015). After the flood, a detailed HEC-RAS 2D model of the scenario was developed for the purpose of understanding the flood dynamics. The aim of this case study is to demonstrate how the real-time forecasting framework can be implemented during a flood event and to evaluate its performance with respect to flood wave timing and the spatial extent of inundation. Due to the lack of in-situ flow depth sensors present during the flood event, synthetic "ground truth" data generated from a HEC-RAS 2D model of the event, which have been largely corroborated by observations, were used instead. Aerial imagery of the flood provided a source of inundation extent information to assess the influence of the vegetation and urban resistance parameters on the accuracy of the flood inundation model (Figures S1–S118 in Supporting Information S1).

3.1. Scenario

In Northern Italy, the Secchia and Panaro Rivers flow 150 km north from their origins in the Apennines Mountains to the Po River. They enclose a strip of land that contains a mix of agricultural lands and settlements that range in size from small villages to the city of Modena with a population of nearly 200,000 people. The land is protected from the rivers by a long system of levees that serve as the primary flood control measure. An unusually large rainstorm storm in January 2014 led to a collapse in a section of the levee along the Secchia River 6 km north of Modena that was found to be altered by mammal bioerosion (Orlandini et al., 2015). The breach, located at 44 °41′57.85″N and 10 °56′41.68″E, occurred at 6:30 in the morning on January 19th and rapidly widened to 80 m by the end of the day. The resulting flood traveled approximately 28 km, inundated 52 km^2 over the following week, and had a volume of approximately 36×10^6 m^3 and a peak flow of 434 $m^3\ s^{-1}$. This resulted in the displacement of people who lived in the downstream towns and caused one fatality (see Figure 1 in Orlandini et al. (2015) for graphical map of the flood). Views of the developed HEC-RAS 2D model and observations used for the corroboration of the obtained results are shown, for instance, in Figures 5 and 6, respectively (Figures S1–S118 in Supporting Information S1).

3.2. Surface Flow Model

Two HEC-RAS models were constructed to better understand the flood dynamics. The first model is a 1D dynamic wave model of the Secchia River that generates the flood hydrographs from the levee breach. This model takes as boundary conditions the inflow and outflow hydrographs derived from river gauges for the selected channel reach and the observed evolution of the levee breach. The resulting hydrograph from the levee breach gives an estimate of the water volume that escaped through the Secchia River levee breach (Orlandini et al., 2015). The breach hydrograph from this model is then used as the input to a HEC-RAS 2D flood plain model, the mechanics of which were described in Section 2. The DTM data used in the 2D model was derived from a 1-m lidar survey collected between 2008 and 2015 with a point cloud density of 1.5 points/m^2. The DTM exhibits a vertical accuracy not greater than 15 cm and a planimetric accuracy not greater than 30 cm (Di Martire

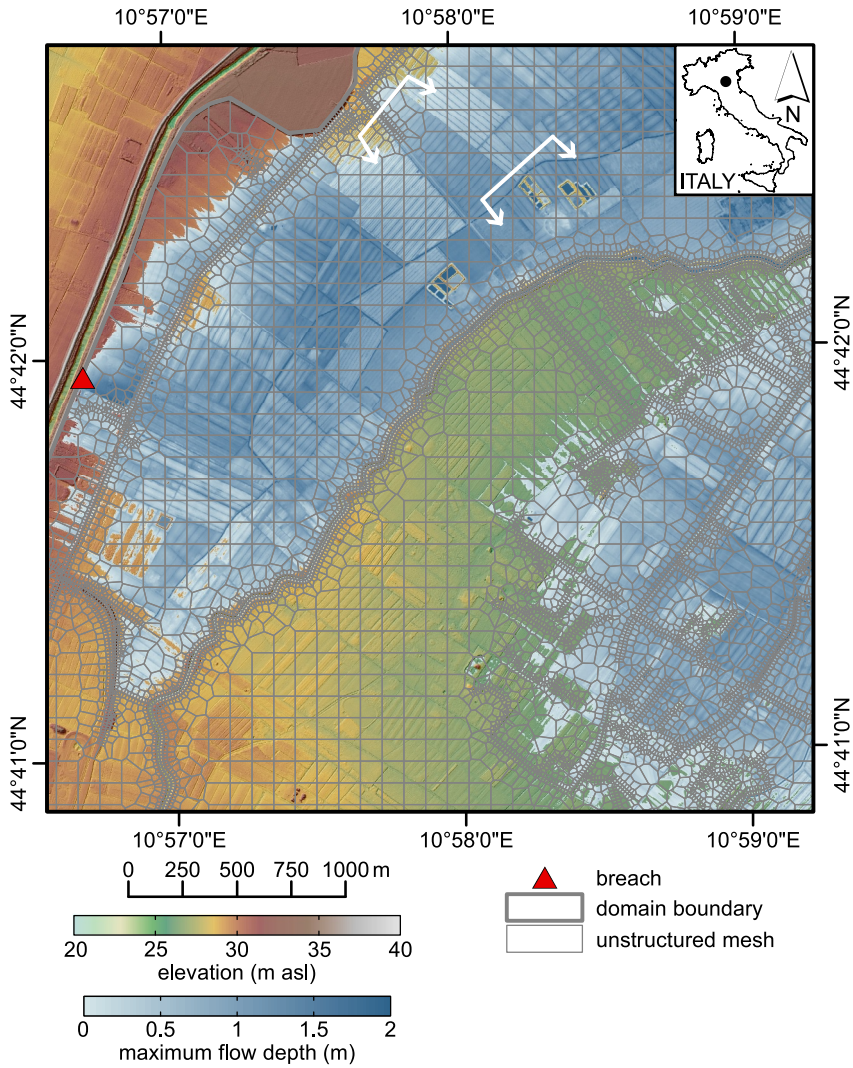

Figure 5. Terrain partitioning of a lowland terrain located between the Secchia and Panaro Rivers (Po River valley, Italy) obtained by using an unstructured mesh that adapts to the ridge network extracted through the LANDMARK algorithm by setting the critical value for the dispersal area $S_c = 10^4$ m^2 (Moretti & Orlandini, 2023). Shown in shades of blue is the related simulation of flood areas during the event of 19 January 2014.

et al., 2017). The LANDMARK terrain analysis algorithm (Moretti & Orlandini, 2023) is then run on the DTM and partitions the flood plain into a geomorphologically informed unstructured mesh consisting of 65,704 cells ranging from 0.03 to 1 ha in area. An example of the mesh is shown in Figure 5. The complete flood is modeled over an approximately 5-day interval starting on 19 January 2014 at 02:40 UTC + 01:00 and ending on 25 January 2014 at 00:00 UTC + 01:00. The computational time step is 1-min and the state of the flood at each mesh cell is saved in 10-min intervals. Conventional performance metrics such as FAR, CSI, and POD, as well as error functions such as mean error, mean absolute error, RMSE, volume conservation error e_V, type-1 error, and type-2 error are used to evaluate the flood inundation model (Pizzileo et al., 2024). The HEC-RAS 2D "full momentum" solver (an Eulerian-Lagrangian method) is found to provide accurate solutions when the 1-min time step is used in preference to smaller time steps for model efficiency requirements (Figures S119–S122 in Supporting Information S1).

3.3. Spatial Distribution of Flow Resistance Parameters

Flow resistance is parameterized in HEC-RAS 2D with a raster layer where each 10-m by 10-m pixel is assigned to user-delineated land cover classes, each with its own Manning's n. For the calibration model, the land surface

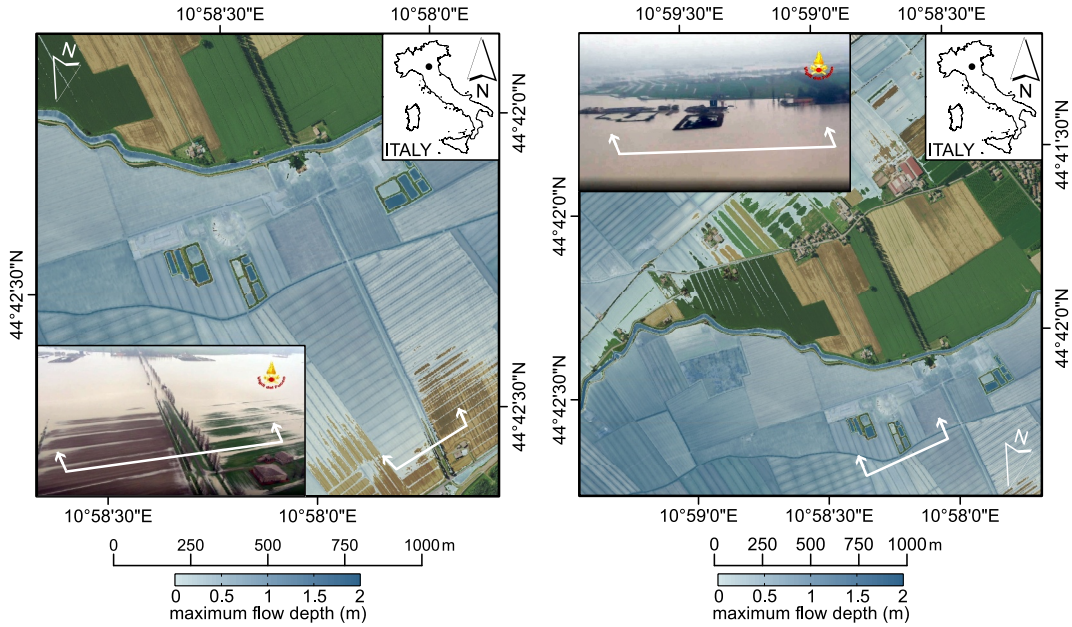

Figure 6. Comparison between simulated and observed flood areas in the terrain located between the Secchia and Panaro Rivers (Po River valley, Italy) during the event of 19 January 2014. The terrain partitioning obtained by using the automatically extracted ridge network reported in Figure 5 is essential to accurately describe human-made macrotopographic structures such as road embankments and the related hydraulic effects on surface flow propagation.

was divided into three land cover classes: vegetation, urban, and canal. The spatial locations of each land cover class were delineated based on land cover maps of the flood plain. The map of the land cover delineation used in the calibration model is shown in Figure 7a. Additional complexity was added to the ground truth model land cover classes by further dividing the three classes into subclasses (Figure 7b). This procedure is further described in Section 3.6.

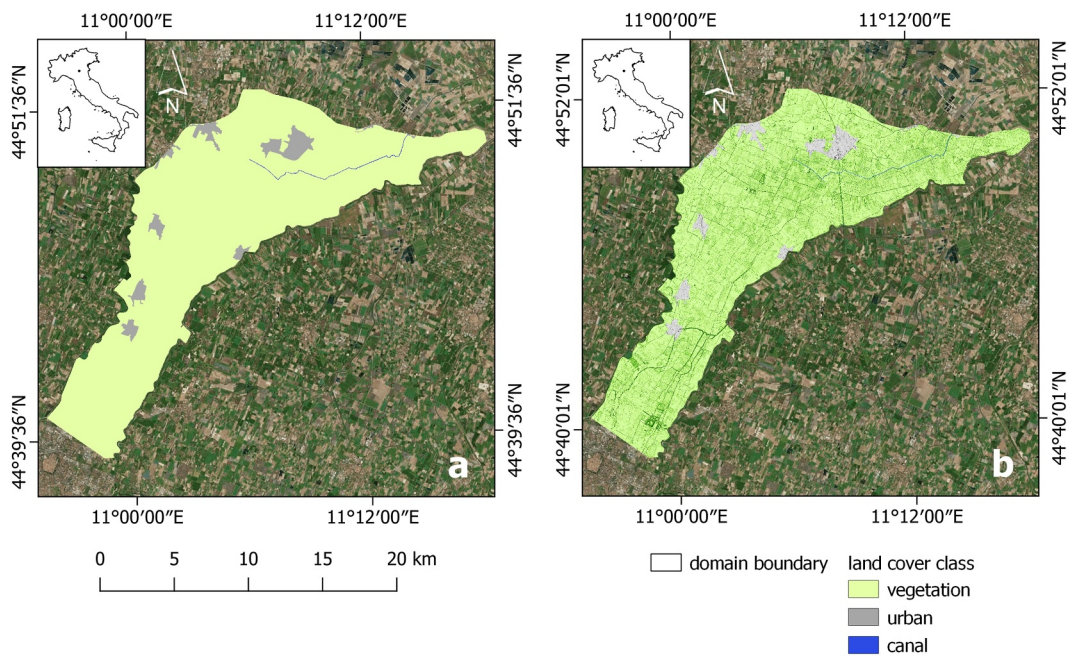

Figure 7. Land cover class distribution used to assign Manning's *n* values in panels (a) the simulation model and (b) the ground truth model with each land cover class further broken into subclasses.

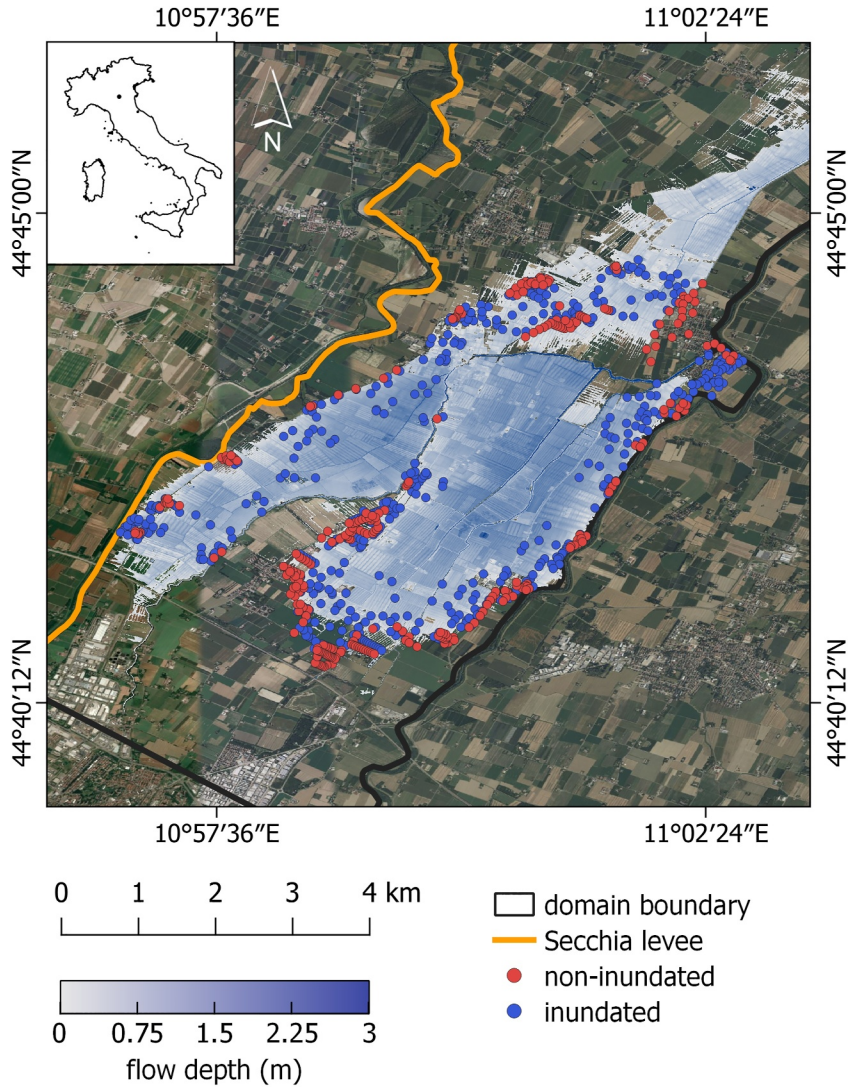

Figure 8. Observations of inundated and non-inundated locations obtained from aerial imagery at 11:40 a.m. on 20 January 2014. Simulated flood extent from the reference model is shown for comparison.

3.4. Accuracy Assessment of Surface Flow Model

To ensure that the HEC-RAS simulation is a faithful representation of the real-world flood processes, the performance of the surface flow model is compared to images of the flood event captured at 11:40 a.m. on 20 January 2014 by emergency personnel. Based on the flood extent shown in each image, certain geospatial locations were manually labeled as "inundated" or "non-inundated" (Figures S1–S118 in Supporting Information S1). Because images were captured by emergency personnel at an oblique angle from the ground, it was not feasible to automatically extract the inundated and non-inundated areas. Due to this constraint, care was taken to systematically choose locations in the images that were clearly inundated or non-inundated. In addition, points that were easy to map by using reference structures in their proximity were selected. Figure 8 shows the locations of the image-derived inundation samples along with the simulated flood extent at 11:40 a.m. on 20 January 2014 using the same model configuration as in Orlandini et al. (2015). The results from this analysis are also used to define the set of reference Manning's n values in the ground truth simulation model. The accuracy of the model with respect to the choice of Manning's n for vegetated and urban land cover classes was assessed with three metrics: FAR, CSI, and POD (Section 2.7.2). The canal land cover class was not yet inundated at the time the observations were made, and thus does not influence the propagation of the flood up to that point. Each metric was computed for a vegetation Manning's n range between 0.03 and 0.09 $m^{-1/3}$ s. For each of these values the urban Manning's n

10.1029/2024WR038424

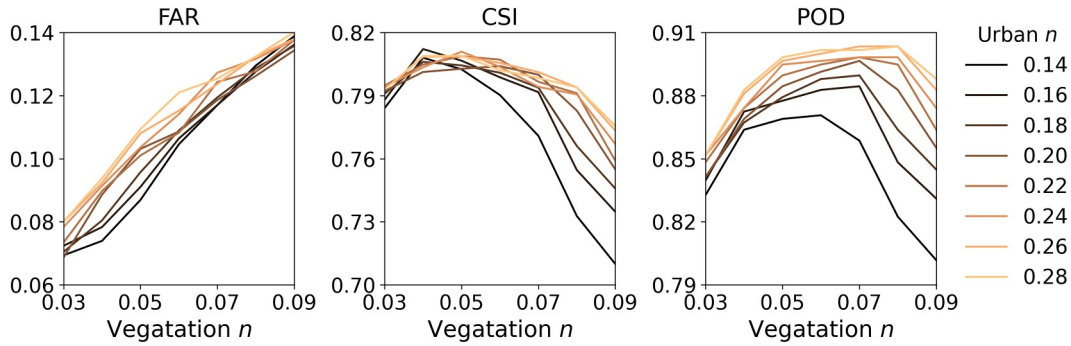

Figure 9. False alarm ratio, critical success index, and probability of detection over a range of values for vegetation and urban Manning's n.

was varied between 0.14 and 0.28 $m^{-1/3}$ s. Figure 9 shows the resulting behavior of each metric over the range of Manning's n values. The FAR shows a positive relationship with both the vegetation and urban Manning's n and a stronger sensitivity to vegetation. The CSI peaks between a vegetation n value of 0.04 and 0.05 $m^{-1/3}$ s and then drops off as n increases, however, some of this drop-off is attenuated by an increasing value of the urban Manning's n. The POD increases with an increasing urban Manning's n until $n = 0.26$.

3.5. Flow Depth Sensor Allocation

A feasibility requirement for a sensitivity-based flow depth sensor network is partial deployment before a levee breach. Optimal sensor locations can be identified based on potential flood hydrographs from various breach points, with some housings permanently equipped with sensors and others set up for quick installation upon breach detection. The case study examines an ideal scenario with known breach location and hydrograph, serving as a preliminary model for more extensive future sensor network designs. The factorial experiment design described in Section 2.5 assigned low and high Manning's n values to each of the three land cover classes totaling eight model runs with each combination: Manning's n values for vegetation, urban area, and canals were $n_1 = (0.030, 0.120)$ $m^{-1/3}$ s, $n_2 = (0.120, 0.200)$ $m^{-1/3}$ s, and $n_3 = (0.025, 0.060)$ $m^{-1/3}$ s, respectively. The model is run for the full 5-day duration of the flood with a computational time step of one second. Sensor locations are assigned at hourly intervals following the start of the breach to ensure coverage throughout the duration of the flood. Two processing steps were performed before the final sensor locations were chosen. The first was to filter out numerical mesh cells that exhibited numerical instabilities. Numerical instabilities in the flow depth time series were found to cause erroneously high values for K_f and thus sensors would be misallocated. Instabilities were identified when the depth time series of a cell exhibited more than four oscillations of 4 cm or more over the course a 3-hr moving window. This operation filtered out approximately 2,367 numerical mesh cells, or 3.6% of the 65,704 total cells. Spatial and temporal correlations were found at the most sensitive locations in the flood plain. In other words, locations sensitive to Manning's n at one time were likely to be sensitive to Manning's n at temporally proximal times. Highly sensitive areas were also spatially correlated in that they were often found within close spatial proximity to one another. As a final processing step, the DBSCAN algorithm with a neighborhood radius of 100 m was used to find spatial clusters of sensor locations (Ester et al., 1996). The final sensor location was chosen as the location in the cluster situated first temporally. For example, if there are two locations in a cluster, the first found at hour 3 and the second found at hour 4, the first location is used for sensor placement. A total of 34 sensor locations were found for the flood (11 for vegetation, 16 for urban, and 7 for canal) and the resulting map is shown in Figure 10.

A theoretical maximum sensor density network scenario treating all 23,180 inundated mesh cells in the ground truth model as "sensors" is presented for comparison (shown as white points in Figure 10). This case is clearly impractical, but is theoretically informative about the performance of the proposed learning methods when the availability of sensors is not a limiting factor. The presence of a low ridge in the topography results in tight packing of the maximum sensor locations near the breach; mesh cells are smaller near ridges and thalwegs as specified in the LANDMARK mesh generation algorithm (Moretti & Orlandini, 2023).

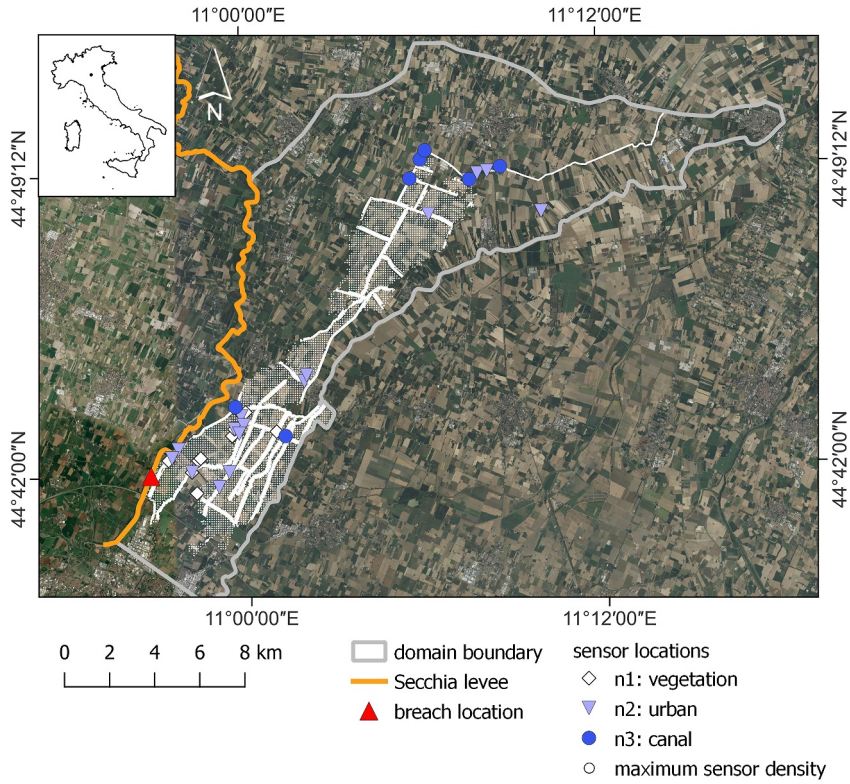

Figure 10. Flow depth sensor locations. Sensors are labeled according to which land cover class they are sensitive to. The white circular markers represent the maximum sensor density network scenario where all 10-cm inundated mesh cells are used as a sensor network.

3.6. Synthetic Ground Truth Model

Flow depth sensor information is not available for the January 2014 Secchia River flood inundation event, so synthetic observation data was generated using the HEC-RAS 2D model of the flood inundation. The synthetic ground truth model is run for the full 5-day duration of the flood and serves as both a source of calibration observations and forecast performance benchmark for the real-time forecasts. A major pitfall of using synthetically generated data is that it may not reflect the same complex dynamics as would be seen in the real world since any model is inherently a simplification. This is especially true when the same model is used for generating synthetic data and for testing a calibration procedure on that data. This is because a deterministic model can exactly reproduce previously generated outputs given the same set of parameters and initial conditions. Therefore, to create a representative study, complexity was added to the synthetic ground truth model by separating each of the three initial land cover classes into subclasses using a terrain roughness index. The terrain roughness index was generated for the flood plain by computing the standard deviation of a moving three-pixel by three-pixel window centered on each pixel of the DTM raster. The index was then binned into eight equal-interval subsets and used to create Manning's n subclasses, resulting in 6 subclasses for vegetation, 7 for urban, and 8 for canal. The differing number of subclasses is a result of the uneven spatial distribution of terrain roughness, where the relatively smooth vegetated area does not reach the two highest roughness classes found in the urban and canal classes. The subclass Manning's n values were then drawn from a normal distribution with variance equal to 0.020 $m^{-2/3}$ s^2 and centered on 0.060 $m^{-1/3}$ s for vegetation, 0.220 $m^{-1/3}$ s for urban, and 0.025 $m^{-1/3}$ s for canal. The resulting values of Manning's n were defined for vegetation as $n_1 = (0.0328, 0.0430, 0.0505, 0.0637, 0.0691, 0.0772)$ $m^{-1/3}$ s, for urban as $n_2 = (0.1780, 0.1930, 0.1990, 0.2110, 0.2130, 0.2210, 0.2250)$ $m^{-1/3}$ s, and for canal as $n_3 = (0.0189, 0.0206, 0.0211, 0.0255, 0.0260, 0.0343, 0.0415, 0.0555)$ $m^{-1/3}$ s. Figure 7b shows the final Manning's n map for the ground truth model. This procedure ensures there is no "true" Manning's n parameter set that can exactly replicate the ground truth model output, thus better approximating a real world scenario where there is also no exact parameter set. The flood simulation from obtained this model is then used as the calibration target.

Table 2
Calibrated Manning's n Values in $m^{-1/3}$ *s From the 10-hr and 43-hr Land Cover Class Experiments for Each Sensor Array and Maximum Sensor Density Network Cases*

Calibration run	n_1: Vegetation	n_2: Urban	n_3: Canal
10-hr	0.0576	0.203	–
10-hr maximum sensor density network	0.0576	0.208	–
43-hr	0.0576	0.175	0.0376
43-hr maximum sensor density network	0.0548	0.175	0.0307

3.7. Calibration and Forecasting

To demonstrate real-time forecasting, the HEC-RAS 2D model was calibrated at 10 hr after the breach and 43 hr after the breach using the sensor array and maximum sensor density network cases as shown in Figure 10 for a total of four calibration experiments. The two calibration times (10 and 43 hr) represent two calibration cycles out of many that could be conducted as the flood wave progresses downstream; refer to Figure 1 for the continuous calibration framework. At 10 hr after the breach, only the vegetation and urban land cover classes had been inundated, and at 43 hr, all three classes had been inundated, thereby demonstrating two- and three-parameter calibration cases. As discussed in Section 2.6, the Bayesian Optimization algorithm converges on the parameter set that produces the global minimum of the objective function. The algorithm was implemented in Python using the BoTorch package which automatically fits the Gaussian process surrogate and allows for different acquisition function options (Balandat et al., 2020). An Intel Xeon 2.40 GHz CPU was used to run the optimizations; each 10-hr simulation required 3–5 min of wall clock time and each 43-hr simulation required 30 min of wall clock time. For these experiments, the Gaussian process surrogate is fit on an objective function which is taken to be the ensemble root mean squared error (i.e., $\langle \mathrm{RMSE}(\mathbf{n}) \rangle$) between all the observed and simulated flow depth sensor time series. The acquisition function was chosen to be "Noisy EI," which is a variation on the commonly used "Expected Improvement" acquisition function as described in Section 2.6. Noisy EI is robust to any noise that may be present in the flow depth sensor readings and hence the objective function (Letham et al., 2017).

For each experiment, the parameter search space was seeded by running the model five or six times with Sobol-sampled values for Manning's n for the 10-hr and 43-hr cases, respectively (Sobol, 1967). The parameter search space was bounded using ranges derived from Chow (1959), thus the Manning's n bounds for vegetation, urban area, and canals were $n_1 = (0.02, 0.20)$ $m^{-1/3}$ s, $n_2 = (0.12, 0.26)$ $m^{-1/3}$ s, and $n_3 = (0.01, 0.06)$ $m^{-1/3}$ s, respectively. The Sobol-sampled points ensure there is an even distribution across the parameter space. Then, the Bayesian optimization algorithm was run serially for an additional 25 iterations for the 10-hr cases, and 44 iterations for the 43-hr cases. Since each iteration requires a single model evaluation, the model was evaluated 30 and 50 times for the 10-hr and 43-hr cases, respectively. The number of iterations was chosen based on prior observations of how many Bayesian optimization cycles were required to achieve parameter convergence. In both experiments, convergence to an optimal set of parameters was achieved after 10 iterations for the vegetation Manning's n and 20 for the urban and canal Manning's n (supporting information, Figure S123 in Supporting Information S1). Since the flood has progressed longer in the 43-hr case than the 10-hr case, there is greater interaction between the flood and the land cover, leading to different dynamics that show up as differences in the urban Manning's n value. For each of the four experiments, a forecast for the entire 5-day flood duration was generated using the optimal flow resistance parameter set obtained from calibration. The calibrated Manning's n values are reported in Table 2.

3.8. Impact of Resistance Parameters and Depth Sensors

The accuracy of the forecasts generated from the four experiments are assessed using the critical flow depth time discrepancy and the inundation metrics. Prior to a flood, the best estimate of flow resistance parameters is to use empirical studies and expert knowledge of the flood plain. For this reason, the forecast performance of the experiments is compared against empirical estimates of the vegetation flow resistance parameter obtained from literature while holding the urban and canal parameters constant at $n = 0.020$ $m^{-1/3}$ s and $n = 0.025$ $m^{-1/3}$ s, respectively (Arcement & Schneider, 1989; Chow, 1959). The accuracy of the experiment forecasts is depends on

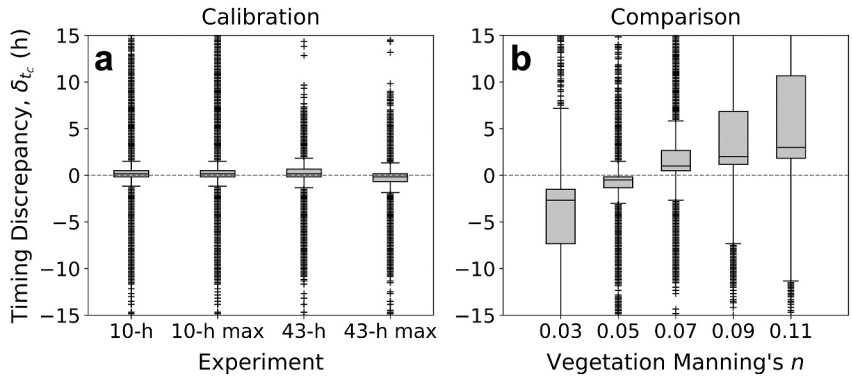

Figure 11. Arrival time discrepancy for 10-hr and 43-hr Bayesian optimization calibration experiments, and experiments using the maximum sensor density network are denoted by "max" (a). Comparison provided with varying value of vegetation land cover class Manning's n (b).

how well the Manning's n for each dynamic class captures the roughness effects as well as the amount of information about the flood state provided by the sensor network.

Figure 11 shows box plots of the time to critical flow depth discrepancy, δ_{t_c}, for all inundated mesh cells between the Bayesian optimization calibration forecast and the synthetic ground truth model over the entire flood duration. In this figure, a value of zero means that there was no difference between the true and predicted arrival time. In Figure 11a, the results from the four Bayesian optimization experiments show that calibrating resistance parameters using observations from flow depth sensors result in median arrival time discrepancies of 0.16 hr for the 10-hr, 10-hr max, and 43-hr experiments, and −0.16 hr for the 43-hr max experiment. The impact of increasing the number of sensors to the maximum sensor density network case does not have a marked impact on forecast performance as measured by arrival time discrepancy. The comparison box plots in Figure 11b are the results of five forecasts using a fixed range of Manning's n values for the vegetation land cover class in Figure 7a while holding the urban and channel Manning's n values constant at 0.2 $m^{-1/3}$ s and 0.025 $m^{-1/3}$ s, respectively. According to Chow (1959), short grass has a Manning's n value of 0.030 $m^{-1/3}$ s and dense brush has a Manning's n value of 0.110 $m^{-1/3}$ s, both of which are commonly present in agricultural areas and thus serve as a reasonable range for the vegetation Manning's n. In the case that there are no flow depth sensors installed in the floodplain, the flood modeler must estimate the vegetation flow resistance parameter from experience, and the comparison boxes represent the forecast performance from a potential selection. In the case that flow depth sensors have been installed in the floodplain, the five vegetation Manning's n values represent a simple, five-increment grid search approach to estimating the optimal parameter value, which would be chosen to be $n = 0.050$ $m^{-1/3}$ s, as it produces the smallest arrival time discrepancy. All Bayesian optimization calibration cases reduce the median arrival time discrepancy over all comparison cases and tighten the box plot whiskers (defined as 1.5 times the interquartile range) for all but the $n = 0.050$ $m^{-1/3}$ s comparison case. The magnitude of the reduction in forecast arrival time discrepancy between the calibration and comparison cases across the floodplain depends on the comparison case Manning's n. Across all inundated locations in the floodplain, the arrival time discrepancy is reduced, on average, by 4.13 hr when comparing the 10-hr and 43-hr calibration cases to the $n = 0.03$ $m^{-1/3}$ s case, but by 28.2 min when comparing the 10-hr and 43-hr calibration cases to the $n = 0.05$ $m^{-1/3}$ s case. The average reduction in arrival time discrepancy when comparing the 10-hr and 45-hr calibration cases to all comparison cases is 3.13 hr. The box plots are marked by a significant number of outliers, or points that fall outside of the whiskers, and are shown by the plus signs in Figure 11. The fractions of outliers as percentages of the total number of points included in the box plots are 18.4%, 18.8%, 11.9%, and 7.8% for the 10 hr, 10 hr max, 43 hr, and 43 hr max experiments (where max indicates the maximum sensor density network), respectively.

The inundation statistics FAR, CSI, and POD were computed for each of the four calibration experiments when the ground truth inundation was at its maximum areal extent; this corresponded to 21 January 2019 at 09:30 or 54 hr and 50 min after the breach. Inundation was defined as a numerical mesh cell having an average depth greater than 10 cm. Table 3 shows the values of the three metrics for each of the four experiments. The calibrated model for all four experiments has a high degree of accuracy with FAR remaining under 0.04, CSI remaining

Table 3
Inundation Metrics for the Four Calibration Experiments at the Time Step of Ground Truth Maximum Inundation Extent (21 January 2019 at 09:30 or 54 hr and 50 min After Breach)

Calibration run	FAR (–)	CSI (–)	POD (–)
10-hr	0.029	0.958	0.986
10-hr maximum sensor density network	0.029	0.955	0.983
43-hr	0.031	0.936	0.994
43-hr maximum sensor density network	0.026	0.968	0.993

above 0.93, and POD remaining above 0.098. The high POD values means that the calibrated model inundation nearly overlaps completely with the observed inundation with a low fraction of false negatives. The presence of false positives lowers the value of the CSI with respect to the POD, implying that the calibrated model is slightly over-predicting inundation. This is especially true for the 43-hr cases, where the addition of false positives in the CSI calculation reduces the POD from 0.994 to 0.936. As in Figure 11, the maximum sensor case does not show marked improvement over the 34-sensor array in any of the inundation metrics for the 10-hr experiments. For the 43-hr case, there is a slight improvement in the FAR and CSI when using the maximum sensor density network.

4. Discussion

This study investigates the potential for a flood inundation model based on the numerical solution of the 2D unsteady flow equations to be used in real-time flood forecasting in combination with a flow depth sensor network design method and a resistance parameter learning procedure (Figure 1). While detail and efficiency of the flood inundation model are ensured by the use of a geomorphologically informed unstructured mesh preserving thalwegs and ridges observed in high-resolution topographic data (Figures 5 and 6), reliability is sought in the present study by estimating flow resistance parameters based on observed flow depth (Figures 3 and 7). Flow resistance parameterization is an important source of uncertainty and is difficult to accurately quantify prior to a flood, not least because microtopography and vegetation roughness unresolved by DTM data vary widely from season to season. Even floods closely spaced in time will experience different flow resistance effects due to altering of vegetation by the previous flood. In fact, in river systems delimited by well-defined banks or levees as well as in flood plains inundated by levee failures, microtopography and vegetation effects on surface flow propagation are not time invariant. Rather, they depend on agricultural practices, vegetation season, soil moisture status and flow conditions (Arcement & Schneider, 1989). This suggests a generalizable forecasting approach where flow resistance parameterizations are updated based on observed flood behavior. To do this, a framework is developed that addresses the combined challenge of observing flow behavior and calibrating a computationally detailed flood inundation model. Flow depth sensors placed at strategic locations in the flood plain serve as the source of calibration data. The sensor locations are assigned with a sensor network algorithm that chooses locations most sensitive to flow resistance (Figures 2 and 10; Sections 2.5 and 3.5). Efficient calibration of the model is achieved with a Bayesian optimization approach, which reduces the number of model evaluations required to find the optimal parameter set to 50 or fewer (Figure 3, Sections 2.6 and 3.7). The forecasts from the framework must be made readily accessible to emergency planners throughout and after the flood. This could be accomplished through a cloud-based platform that distributes flood alerts and other information through an online dashboard (Deren et al., 2021). With these pieces in place, when a flood occurs, the observation-calibration-forecasting cycle seamlessly integrates with the management procedures already in place.

Binary inundation observations obtained from aerial videography provided a unique source of model validation data that served two purposes in the case study (Section 3.4, Figure 8; Figures S1–S118 in Supporting Information S1). The first purpose was to validate that the flood model was a realistic representation of the true flood dynamics and the second purpose was to help inform the choice of vegetation and urban Manning's n values in the synthetic ground truth model. Based on the FAR, CSI, and POD metrics plotted in Figure 9, the vegetation resistance parameter has a greater influence on flow propagation than the urban resistance parameter. Taken together, these three metrics paint a picture of how the model most accurately reproduces the observations at low values of vegetation Manning's n and higher values of urban Manning's n. However, as Manning's n for both classes increase past a certain point, the model begins to overpredict inundation (Figure 9). This analysis highlights the need for accurately estimating flow resistance parameters to generate actionable flood forecasts. Binary point inundation measurements derived from aerial imagery are an additional source of model validation data that have the potential to augment in-situ and satellite measurements. However, extra care must be taken in the interpretation of inundation metrics due to spatial clustering or biases in where the images are collected and the method by which they are digitized.

The placement process for the sensor network is outlined in Sections 2.5 and 3.5, and since it relies on simple hydraulic theory, it is sufficiently flexible to apply to any flood plain. The requirements are a 2D unsteady surface

flow model and time to evaluate the model for a range of land surface roughness parameterizations (eight times in the case study described in Section 2.5). The forecast prediction results presented in Section 3.8, Figure 11, and Table 3 demonstrate that 34 in-situ sensors result in nearly identical calibration performance as the maximum sensor density network, with only slight improvements in the CSI and FAR for the 43-hr, maximum sensor density network case. By discretizing the heterogeneous "true" land surface into three land cover classes (vegetation, urban, and canal), the assumption is that the aggregate influence of the land cover can be captured by three resistance parameters. Furthermore, the sensors are assigned based on simulated sensitivity to these three aggregate classes, rather than sensitivity to the "true," heterogeneous land cover due to the difficulty of accurately determining the latter. The aggregation of the true land cover into three aggregate classes may require further evaluation as demonstrated by the presence of outliers in Figure 11. In this study, representing the true spatial distribution of flow resistance with three aggregate land cover classes is the limiting factor in producing accurate forecasts, rather than the allocation of the sensor network. This study did not investigate the minimum number of sensors to achieve an accurate forecast, but it is expected that some redundancy, as demonstrated here, increases the probability that reliable flow depth observations will be available during a flood event.

The Bayesian optimization algorithm found parameter sets that outperformed most of the comparison cases with fewer than 50 model evaluations (Figure 11). In other words, the algorithm avoided any local optimum traps that would lead to a large misidentification of parameters and poor forecast performance. The comparison cases in Figure 11b represent the performance of either a selection of a single vegetation resistance parameter if flow depth sensor data is not available, or a simple grid search approach to estimating the best flow resistance parameter value if flow depth sensors have been installed in the flood plain. In the absence of flow depth sensors, the vegetation flow resistance parameter used to generate a forecast could diverge from the optimal value and lead to large discrepancies in forecasted flood arrival time as illustrated by the box plots in Figure 11b for Manning's n values of 0.09 $m^{-1/3}$ s or 0.11 $m^{-1/3}$ s. With the addition of flow depth sensors, the grid search approach to calibrating the vegetation resistance parameter could, in theory, be performed automatically instead of the Bayesian optimization procedure. However, this approach does not scale well with increasing parameter dimensionality. The relative improvement in median flood arrival time of the forecasts using Bayesian optimization to calibrate all three flow resistance parameters over simply using the best performing parameter from this grid search approach (0.05 $m^{-1/3}$ s) is at minimum 28.2 min across the flood inundation extent (Figure 11b). The relatively small difference between the two approaches may result from the large fraction of the floodplain that the vegetation resistance parameter represents, rendering the effects of urban and canal flow resistance parameters negligible (Figure 7). A more complex flow resistance parameter distribution would reduce the efficacy of a grid search approach due to the lack of a clear single representative flow resistance parameter, potentially leading to a larger relative improvement by the Bayesian optimization approach. The greatest benefit of the Bayesian optimization approach is when there exist multiple flow resistance parameters that must be calibrated automatically. With the five-increment grid search approach, 125 model evaluations would be needed to estimate vegetation, urban, and canal flow resistance parameters, 75 more than the Bayesian optimization approach.

The practical implementation of the real-time forecasting framework has inherent limitations, the nature of which are related to the fundamental uncertainties in flood forecasting and the computational expense required to run a high-resolution flood inundation model. The uncertainty due to flow resistance parameterization was addressed in this paper. Flow forcing, site geometry, and operation of flood infrastructure all constitute other sources of uncertainty, but were not addressed by the Bayesian optimization procedure. The presented flood forecasting framework could, in theory, be extended to these other sources of uncertainty as long as they are parameterized in the flood inundation model. Because flow resistance parameters are represented by a single constant Manning's n, they are particularly amenable to the presented approach and other sources of uncertainty may be better-addressed through other means: increased DTM resolution or a better understanding of the physics of levee breaches, for example, Prior to a flood, the exact position of the levee breach is unknown, and thus the determination of flow depth sensor locations should be based, in principle, on multiple breach locations and with varying breach hydrographs. Sensors or their housings would then be preallocated in the floodplain, and, upon discovery of a breach, activated or installed to match the optimal sensor layout for the most similar simulated breach scenario. In this study, this extensive analysis is not considered, and we have instead used a single breach location and hydrograph to illustrate the sensor placement procedure. This corresponds to the best-case scenario where the true breach location and hydrograph exactly match one that had been presimulated. Future work could investigate the number of breach locations and flow hydrographs that need to be presimulated to ensure the flow depth sensor

network is able to provide valuable observations for any breach. The real-time flood forecasting framework is generalizable to flood plains in any geographical area since the choice of sensors and calibration algorithm is independent of land surface characteristics. However, larger and flatter flood plains are more amenable to real-time forecasting than short and steep catchments due to the slower progression of the flood wave. In steep catchments, observation of a sufficiently long flow depth time series may significantly cut into forecast lead time. This lead time is greatly improved by the Bayesian optimization approach, but is still limited by the computational expense required to evaluate the flood model. In this study, each 10-hr flood simulation required 3 min for geometry preprocessing and around 3 min for simulation with commonly available workstations. This meant that with 25 optimization iterations, a forecast was available after 2.5 hr. Each 43-hr flood simulation required 3 min for geometry preprocessing and 25 min for simulation. With a total of 50 model iterations, the full calibration process required nearly 24 hr of wall clock time. For flood inundations of this scale (52 km^2 in terms of area, 28 km in terms of length), the computation time required by surface flow simulation make the proposed framework useful with commonly available workstations at early times in the flood, but may become too computationally expensive later in the flood (Moretti & Orlandini, 2023).

The promising results of the case study presented in this research outline a roadmap toward the development of a real-time flood forecasting system. A field trial should be conducted to elucidate some of the challenges that come with the transition from a synthetic study to a real-world application. Real-time flood forecasting is continuously evolving as new technologies (terrain analysis, computational fluid mechanics, satellites, sensors, computers) make detailed modeling and observation feasible. As technology advances, new conceptual improvements become possible. This study illustrates the added value of including continuous flow depth observation for real-time estimation of resistance parameters in a flood modeling framework. It is shown that real-time calibration of resistance parameters leads to significant improvement in our ability to predict inundations in terms of timing and inundation extent (Section 3.8). The availability of an accurate and efficient 2D flood inundation model based on geomorphologically informed unstructured terrain partitioning, effective sensor network design methods, and a robust and efficient Bayesian optimization technique are the essential requirements to enable real-time resistance parameter learning. Although the proposed flood modeling framework has not been tested during a real flood inundation event, the numerical experiments reported in this study are a sound proof-of-concept and support further research in the field with real sensors and direct application in a flood inundation emergency.

5. Conclusions

This study presents a physics-based real-time flood inundation modeling framework that produces flood forecasts with sufficient lead time for emergency planning and evacuation. This is achieved with an iterative cycle of flow depth sensing, calibration, and forecasting. The model accuracy and efficiency rely on the use of geomorphologically informed unstructured computational meshes (Section 2.3). The framework reliability relies on the use of fast model evaluations and rapid calibration of resistance parameters to provide the most up-to-date and accurate forecasts of flood extent possible (Figure 1). Flow depth sensing with in-situ sensors allows for the flow resistance parameterization of the flood inundation model to be hydraulically representative of the effects of unresolved land surface topography on the flood at hand (Section 2.5). The location of the depth sensors are chosen based on their sensitivity to the land surface roughness, thus enabling accurate calibration (Figures 2 and 10). Model calibration employs a Bayesian optimization algorithm, which was chosen because it converges rapidly to the optimal parameter set with as few model evaluations as possible, a key requirement for sufficient forecast lead time (Section 2.6 and Figure 3). A case study using synthetic flow depth data describes how the forecasting framework would be applied in a real flood scenario (Section 3 and Figure 6). A numerical reproduction of a real flood inundation event is created with a HEC-RAS 2D model as the ground truth target for calibration (Section 3.6). The case study analyzes the framework forecast performance at two times during the flood and for two sensor array layouts (Section 3.7). The resulting flood forecasts are compared against the ground truth simulation with two sets of metrics to assess the calibration performance (Section 3.8).

Flood inundation model calibration required fewer than 50 model evaluations to converge to the target resistance parameter set (Section 3.7). The computer workstation used to perform the experiments in this study was fast enough to produce forecasts 10-hr after the breach, but too slow to provide meaningful lead time for the 43-hr calibration (Section 3.7). Therefore it is essential to ensure that there is sufficient computational resources at hand during a flood event. The four experiments showed that with 34 in-situ flow depth sensors allocated in

hydraulically meaningful locations throughout the floodplain, the framework can produce downstream forecasts with accuracy comparable to those produced with a high-density sensor network as measured by flood arrival time and inundation extent (Figure 11 and Table 3). The median flood arrival time forecast errors across the flood plain are reduced by 1.67 hr or 86% on average (Section 3.8). After optimization, correct classification of inundation extent as measured by FAR is below 4% and above 90% for CSI and POD (Section 3.8 and Table 3). Future work is required to design a pilot-scale study to investigate the challenges that will be faced during real-world implementation. The transition from calibrating on synthetic data to real flow depth observations may reduce the efficacy of the calibration algorithm due to sensor noise, malfunctions, and poor placement within the flood plain. Locating flow depth sensors may be more complicated in topographically varied environments and larger networks will be required when levee breach location uncertainty is high. Extensive preliminary modeling must be carried out before the framework is made operational. Differently from other existing real-time flood inundation models, flow resistance parameters are calibrated using real-time flow depth observations to represent the effects of land surface topography unresolved by DTM data in a hydraulically meaningful manner and this makes the results of the proposed modeling framework specifically suitable to guide post-event recovery and prevention of flood inundations.

Data Availability Statement

Data used to generate the results reported in this manuscript are available on Zenodo at https://zenodo.org/record/11643497 (Young et al., 2024). The link allowing editors and reviewers to access the data during the review process can be found in the TXT file "zenodo-shared-link.txt" provided as part of the submission material. The HEC-RAS optimization code used for this study is available through the "rasopt" Python package at https://github.com/alex-l-young/rasopt.

Acknowledgments

The authors declare no conflicts of interest relevant to this study. The research reported in the present paper was supported by Fondazione Cassa di Risparmio di Modena through the Grant 2018-0093, by the University of Modena and Reggio Emilia through the Grant FAR 2020 Mission Oriented, by the European Union NextGenerationEU/NRRP, Mission 4 Component 2 Investment 1.5, Call DD 3277 (12/30/2021), Award 0001052 (06/23/2022), under the project ECS00000033 "Ecosystem for Sustainable Transition in Emilia-Romagna," Spoke 6 "Ecological Transition Based on HPC and Data Technology," CUP E93C2200110001, by the European Union NextGenerationEU/NRRP, Mission 4 Component 2 Investment 1.1, Call PRIN 2022 D.D. 104 (02/02/2022), under the project 2022M4798K "Rivers Affected by Mammal Bioerosion (RAMB)" CUP E53D23004420006, and by the European Union NextGenerationEU/NRRP, Mission 4 Component 2 Investment 1.3, Call DD 2356 (02/23/2024), under the project PE00000005 "Multi-Risk sciEnce for resilienT commUnities undeR a changiNg climate (RETURN)," Spoke VS1 "Water," subproject "Riduzione del Rischio di Collasso Arginale (R2CA)," CUP D43C22003030002. High-resolution topographic data were provided by Ministero dell'Ambiente, della Tutela del Territorio e del Mare (Rome, Italy), Agenzia Interregionale per il Fiume Po (Parma, Italy), and Regione Emilia-Romagna (Bologna, Italy). The observations of the flood event occurred along the Po Valley flood plain on 19 January 2014 were provided by the Civil Protection of the Regione Emilia-Romagna (Italy) and by the National Fire and Rescue Service (Modena, Italy). The authors thank the Editor Peter Troch, the Associate Editor Ryan Morrison, the Reviewer Brett Sanders, and three anonymous reviewers for comments that led to improvements in the manuscript.

References

Annan, J. D., Hargreaves, J. C., Edwards, N. R., & Marsh, R. (2005). Parameter estimation in an intermediate complexity earth system model using an ensemble Kalman filter. *Ocean Modelling*, *8*(1), 135–154. https://doi.org/10.1016/j.ocemod.2003.12.004

Annis, A., Nardi, F., & Castelli, F. (2022). Simultaneous assimilation of water levels from river gauges and satellite flood maps for near-real-time flood mapping. *Hydrology and Earth System Sciences*, *26*(4), 1019–1041. https://doi.org/10.5194/hess-26-1019-2022

Arcement, G. J., & Schneider, V. R. (1989). *Guide for selecting mannings roughness coefficients for natural channels and flood plains*. U.S. Geol. Surv. Water Supply. Paper 2339.

Association of State Floodplain Managers. (2020). Flood mapping for the nation: A cost analysis for completing and maintaining the Nation's NFIP flood map Inventory (Tech. Rep.). Retrieved from https://asfpm-library.s3-us-west-2.amazonaws.com/FSC/MapNation/ASFPM_MaptheNation_Report_2020.pdf

Balandat, M., Karrer, B., Jiang, D. R., Daulton, S., Letham, B., Wilson, A. G., & Bakshy, E. (2020). BoTorch: A framework for efficient Monte-Carlo Bayesian optimization. In *Advances in neural information processing systems* (Vol. 33). Retrieved from https://proceedings.neurips.cc/paper/2020/hash/f5b1b89d98b7286673128a5fb112cb9a-Abstract.html

Balistrocchi, M., Moretti, G., Ranzi, R., & Orlandini, S. (2021). Failure probability analysis of levees affected by mammal bioerosion. *Water Resources Research*, *57*(12), e2021WR030559. https://doi.org/10.1029/2021WR030559

Barnes, H. H. (1967). Roughness characteristics of natural channels. *U. S. Geological Survey Water-Supply Paper*, *1849*.

Barthélémy, S., Ricci, S., Morel, T., Goutal, N., Le Pape, E., & Zaoui, F. (2018). On operational flood forecasting system involving 1D/2D coupled hydraulic model and data assimilation. *Journal of Hydrology*, *562*, 623–634. https://doi.org/10.1016/j.jhydrol.2018.05.007

Bates, P. D. (2004). Remote sensing and flood inundation modelling. *Hydrological Processes*, *18*(13), 2593–2597. https://doi.org/10.1002/hyp.5649

Bates, P. D. (2022). Flood inundation prediction. *Annual Review of Fluid Mechanics*, *54*(1), 287–315. https://doi.org/10.1146/annurev-fluid-030121-113138

Bates, P. D., Quinn, N., Sampson, C., Smith, A., Wing, O., Sosa, J., et al. (2021). Combined modeling of us fluvial, pluvial, and coastal flood hazard under current and future climates. *Water Resources Research*, *57*(2), e2020WR028673. https://doi.org/10.1029/2020WR028673

Ceccato, F., & Simonini, P. (2023). The effect of heterogeneities and small cavities on levee failures: The case study of the Panaro levee breach (Italy) on 6 December 2020. *Journal of Flood Risk Management*, *16*(2). https://doi.org/10.1111/jfr3.12882

Chow, V. T. (1959). *Open-channel hydraulics*. McGraw-Hill.

Cooper, C. M., Sharma, S., Nicholas, R. E., & Keller, K. (2022). Toward more actionable flood-risk information. *Earth's Future*, *10*(11), e2022EF003093. https://doi.org/10.1029/2022EF003093

Deren, L., Wenbo, Y., & Zhenfeng, S. (2021). Smart city based on digital twins. *Computational Urban Science*, *1*(1), 4. https://doi.org/10.1007/s43762-021-00005-y

Di Baldassarre, G., Viglione, A., Carr, G., Kuil, L., Yan, K., Brandimarte, L., & Blöschl, G. (2015). Debates-perspectives on socio-hydrology: Capturing feedbacks between physical and social processes. *Water Resources Research*, *51*(6), 4770–4781. https://doi.org/10.1002/2014wr016416

Di Martire, D., Paci, M., Confuorto, P., Costabile, S., Guastaferro, F., Verta, A., & Calcaterra, D. (2017). A nation-wide system for landslide mapping and risk management in Italy: The second Not-ordinary Plan of Environmental Remote Sensing. *International Journal of Applied Earth Observation and Geoinformation*, *63*, 143–157. https://doi.org/10.1016/j.jag.2017.07.018

Ester, M., Kriegel, H. P., Sander, J., & Xiaowei, X. (1996). *A density-based algorithm for discovering clusters in large spatial databases with noise*. AAAI Press. Retrieved from https://www.osti.gov/biblio/421283

Fabio, P., Aronica, G. T., & Apel, H. (2010). Towards automatic calibration of 2-d flood propagation models. *Hydrology and Earth System Sciences*, *14*(6), 911–924. https://doi.org/10.5194/hess-14-911-2010

Fire Brigades Union. (2015). Inundated: The lessons of recent flooding for the fire and rescue service (Tech. Rep.). https://www.fbu.org.uk/publications

Frazier, P. I. (2018). A tutorial on Bayesian optimization (Tech. Rep. No. arXiv:1807.02811). https://doi.org/10.48550/arXiv.1807.02811

Frazier, P. I., Powell, W., & Dayanik, S. (2009). The knowledge-gradient policy for correlated normal beliefs. *INFORMS Journal on Computing*, *21*(4), 599–613. https://doi.org/10.1287/ijoc.1080.0314

Gallagher, M., & Doherty, J. (2007). Parameter estimation and uncertainty analysis for a watershed model. *Environmental Modelling & Software*, *22*(7), 1000–1020. https://doi.org/10.1016/j.envsoft.2006.06.007

Gallien, T. (2016). Validated coastal flood modeling at Imperial Beach, California: Comparing total water level, empirical and numerical overtopping methodologies. *Coastal Engineering*, *111*, 95–104. https://doi.org/10.1016/j.coastaleng.2016.01.014

García-Pintado, J., Neal, J. C., Mason, D. C., Dance, S. L., & Bates, P. D. (2013). Scheduling satellite-based SAR acquisition for sequential assimilation of water level observations into flood modelling. *Journal of Hydrology*, *495*, 252–266. https://doi.org/10.1016/j.jhydrol.2013.03.050

Hall, J. W., Manning, L. J., & Hankin, R. K. S. (2011). Bayesian calibration of a flood inundation model using spatial data. *Water Resources Research*, *47*(5). https://doi.org/10.1029/2009WR008541

Hendricks Franssen, H. J., & Kinzelbach, W. (2008). Real-time groundwater flow modeling with the Ensemble Kalman Filter: Joint estimation of states and parameters and the filter inbreeding problem. *Water Resources Research*, *44*(9). https://doi.org/10.1029/2007WR006505

Hennig, P., & Schuler, C. J. (2012). Entropy search for information-efficient global optimization. *Journal of Machine Learning Research*, *13*(null), 1809–1837.

Hicks, D. M., & Mason, P. (1991). *Roughness characteristics of New Zealand rivers*. Water Resources Survey, DSIR Marine and Freshwater.

Hossain, A., Jia, Y., & Chao, X. (2009). Estimation of Manning's roughness coefficient distribution for hydrodynamic model using remotely sensed land cover features. In *2009 17th International Conference on Geoinformatics*. IEEE. https://doi.org/10.1109/geoinformatics.2009.5293484

Hostache, R., Chini, M., Giustarini, L., Neal, J., Kavetski, D., Wood, M., et al. (2018). Near-real-time assimilation of SAR-derived flood maps for improving flood forecasts. *Water Resources Research*, *54*(8), 5516–5535. https://doi.org/10.1029/2017WR022205

Hostache, R., Lai, X., Monnier, J., & Puech, C. (2010). Assimilation of one satellite SAR image for flood simulations. Method and test case (Moser river). In *AOGS 2010—Asia oceania geosciences society* (Vol. 22, pp. 49–65). World Scientific. Retrieved from https://hal.archives-ouvertes.fr/hal-00908214

Hydrologic Engineering Center. (2016). *HEC-RAS: River analysis system (hydraulic reference manual)*. US Army Corps of Engineers Hydrologic Engineering Center.

Hydrologic Engineering Center. (2022). *Development of the 2D computational mesh (training document)*. US Army Corps of Engineers Hydrologic Engineering Center.

Jones, D. R., Schonlau, M., & Welch, W. J. (1998). Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, *13*(4), 455–492. https://doi.org/10.1023/a:1008306431147

Kahl, D. T., Schubert, J. E., Jong-Levinger, A., & Sanders, B. F. (2022). Grid edge classification method to enhance levee resolution in dual-grid flood inundation models. *Advances in Water Resources*, *168*, 104287. https://doi.org/10.1016/j.advwatres.2022.104287

Kim, B., Choi, S. Y., & Han, K.-Y. (2019). Integrated real-time flood forecasting and inundation analysis in small–medium streams. *Water*, *11*(5), 919. https://doi.org/10.3390/w11050919

Lai, X., Liang, Q., Yesou, H., & Daillet, S. (2014). Variational assimilation of remotely sensed flood extents using a 2-D flood model. *Hydrology and Earth System Sciences*, *18*(11), 4325–4339. https://doi.org/10.5194/hess-18-4325-2014

Letham, B., Karrer, B., Ottoni, G., & Bakshy, E. (2017). Constrained Bayesian optimization with noisy experiments. *arXiv*. https://doi.org/10.48550/ARXIV.1706.07094

Liggett, J. A., & Chen, L.-C. (1994). Inverse transient analysis in pipe networks. *Journal of Hydraulic Engineering*, *120*(8), 934–955. https://doi.org/10.1061/(ASCE)0733-9429(1994)120:8(934

Ligrani, P. M., & Moffat, R. J. (1986). Structure of transitionally rough and fully rough and fully rough turbulent boundary layers. *Journal of Fluid Mechanics*, *162*(162), 69–98. https://doi.org/10.1017/s0022112086001933

Ma, J., Zhang, J., Li, R., Zheng, H., & Li, W. (2022). Using Bayesian optimization to automate the calibration of complex hydrological models: Framework and application. *Environmental Modelling & Software*, *147*, 105235. https://doi.org/10.1016/j.envsoft.2021.105235

Madsen, H., & Skotner, C. (2005). Adaptive state updating in real-time river flow forecasting—A combined filtering and error forecasting procedure. *Journal of Hydrology*, *308*(1), 302–312. https://doi.org/10.1016/j.jhydrol.2004.10.030

Marks, K., & Bates, P. (2000). Integration of high-resolution topographic data with floodplain flow models. *Hydrological Processes*, *14*(11–12), 2109–2122. https://doi.org/10.1002/1099-1085(20000815/30)14:11/12⟨2109::AID-HYP58⟩3.0.CO;2-1

Merz, B., Blöschl, G., Vorogushyn, S., Dottori, F., Aerts, J. C. J. H., Bates, P., et al. (2021). Causes, impacts and patterns of disastrous river floods. *Nature Reviews Earth & Environment*, *2*(9), 592–609. https://doi.org/10.1038/s43017-021-00195-3

Ming, X., Liang, Q., Xia, X., Li, D., & Fowler, H. J. (2020). Real-time flood forecasting based on a high-performance 2-d hydrodynamic model and numerical weather predictions. *Water Resources Research*, *56*(7), e2019WR025583. https://doi.org/10.1029/2019WR025583

Moradkhani, H., Sorooshian, S., Gupta, H. V., & Houser, P. R. (2005). Dual state–parameter estimation of hydrological models using ensemble Kalman filter. *Advances in Water Resources*, *28*(2), 135–147. https://doi.org/10.1016/j.advwatres.2004.09.002

Moretti, G., & Orlandini, S. (2023). Thalweg and ridge network extraction from unaltered topographic data as a basis for terrain partitioning. *Journal of Geophysical Research: Earth Surface*, *128*(4), e2022JF006943. https://doi.org/10.1029/2022JF006943

Neal, J. C., Atkinson, P. M., & Hutton, C. W. (2007). Flood inundation model updating using an ensemble Kalman filter and spatially distributed measurements. *Journal of Hydrology*, *336*(3), 401–415. https://doi.org/10.1016/j.jhydrol.2007.01.012

Nguyen, T. H., Ricci, S., Fatras, C., Piacentini, A., Delmotte, A., Lavergne, E., & Kettig, P. (2022). Improvement of flood extent representation with remote sensing data and data assimilation. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1–22. https://doi.org/10.1109/TGRS.2022.3147429

Nikuradse, J. (1950). Laws of flow in rough pipes. NACA TM 1292 (English Traslation of "Str´omungsgesetze in rauhen Röhren," VDI Forschungsheft 361).

Ongdas, N., Akiyanova, F., Karakulov, Y., Muratbayeva, A., & Zinabdin, N. (2020). Application of HEC-RAS (2D) for flood hazard maps generation for Yesil (Ishim) river in Kazakhstan. *Water*, *12*(10), 2672. https://doi.org/10.3390/w12102672

Orlandini, S., Moretti, G., & Albertson, J. D. (2015). Evidence of an emerging levee failure mechanism causing disastrous floods in Italy. *Water Resources Research*, *51*(10), 7995–8011. https://doi.org/10.1002/2015wr017426

Pappenberger, F., Beven, K. J., Ratto, M., & Matgen, P. (2008). Multi-method global sensitivity analysis of flood inundation models. *Advances in Water Resources*, *31*(1), 1–14. https://doi.org/10.1016/j.advwatres.2007.04.009

Pensoneault, A., Krajewski, W. F., Velásquez, N., Zhu, X., & Mantilla, R. (2023). Ensemble Kalman Inversion for upstream parameter estimation and indirect streamflow correction: A simulation study. *Advances in Water Resources*, *181*, 104545. https://doi.org/10.1016/j.advwatres.2023.104545

Pizzileo, S., Moretti, G., & Orlandini, S. (2024). Flood plain inundation modeling with explicit description of land surface macrostructures. *Advances in Water Resources*, *188*(6), 104713. https://doi.org/10.1016/j.advwatres.2024.104713

Potočki, K., Hartmann, T., Slavíková, L., Collentine, D., Veidemane, K., Raška, P., et al. (2022). Land policy for flood risk management—Toward a new working paradigm. *Earth's Future*, *10*(4), e2021EF002491. https://doi.org/10.1029/2021EF002491

Pregnolato, M., Ford, A., Wilkinson, S. M., & Dawson, R. J. (2017). The impact of flooding on road transport: A depth-disruption function. *Transportation Research Part D: Transport and Environment*, *55*, 67–81. https://doi.org/10.1016/j.trd.2017.06.020

Pujol, L., Garambois, P.-A., & Monnier, J. (2022). Multi-dimensional hydrological–hydraulic model with variational data assimilation for river networks and floodplains. *Geoscientific Model Development*, *15*(15), 6085–6113. https://doi.org/10.5194/gmd-15-6085-2022

Quirogaa, V. M., Kurea, S., Udoa, K., & Manoa, A. (2016). Application of 2D numerical simulation for the analysis of the February 2014 Bolivian Amazonia flood: Application of the new HEC-RAS version 5. *Ribagua*, *3*(1), 25–33. https://doi.org/10.1016/j.riba.2015.12.001

Rajib, A., Liu, Z., Merwade, V., Tavakoly, A. A., & Follum, M. L. (2020). Towards a large-scale locally relevant flood inundation modeling framework using swat and lisflood-fp. *Journal of Hydrology*, *581*, 124406. https://doi.org/10.1016/j.jhydrol.2019.124406

Rasmussen, C., & Williams, C. K. I. (2006). *Gaussian processes for machine learning*. MIT Press.

Raudkivi, A. J. (1967). *Loose boundary hydraulics*. Pergamon Press.

Revilla-Romero, B., Wanders, N., Burek, P., Salamon, P., & de Roo, A. (2016). Integrating remotely sensed surface water extent into continental scale hydrology. *Journal of Hydrology*, *543*, 659–670. https://doi.org/10.1016/j.jhydrol.2016.10.041

Rosenzweig, B. R., Herreros Cantis, P., Kim, Y., Cohn, A., Grove, K., Brock, J., et al. (2021). The value of urban flood modeling. *Earth's Future*, *9*(1). https://doi.org/10.1029/2020ef001739

Rossi, N., Bačić, M., Kovačević, M. S., & Librić, L. (2021). Development of fragility curves for piping and slope stability of river levees. *Water*, *13*(5), 738. https://doi.org/10.3390/w13050738

Sanders, B. F., & Schubert, J. E. (2019). Primo: Parallel raster inundation model. *Advances in Water Resources*, *126*, 79–95. https://doi.org/10.1016/j.advwatres.2019.02.007

Schubert, J. E., Luke, A., AghaKouchak, A., & Sanders, B. F. (2022). A framework for mechanistic flood inundation forecasting at the metropolitan scale. *Water Resources Research*, *58*(10), e2021WR031279. https://doi.org/10.1029/2021WR031279

Schubert, J. E., & Sanders, B. F. (2012). Building treatments for urban flood inundation models and implications for predictive skill and modeling efficiency. *Advances in Water Resources*, *41*, 49–64. https://doi.org/10.1016/j.advwatres.2012.02.012

Shockling, M. A., Allen, J. J., & Smits, J. (2006). Roughness effects in turbulent pipe flow. *Journal of Fluid Mechanics*, *564*(564), 267–285. https://doi.org/10.1017/s0022112006001467

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

Sobol, I. (1967). On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Computational Mathematics and Mathematical Physics*, *7*(4), 86–112. https://doi.org/10.1016/0041-5553(67)90144-9

Teng, J., Jakeman, A., Vaze, J., Croke, B., Dutta, D., & Kim, S. (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 Wesemael, A., Landuyt, L., Lievens, H., & Verhoest, N. E. C. (2019). Improving flood inundation forecasts through the assimilation of in situ floodplain water level measurements based on alternative observation network configurations. *Advances in Water Resources*, *130*, 229–243. https://doi.org/10.1016/j.advwatres.2019.05.025

Vrugt, J. A., Ter Braak, C. J. F., Clark, M. P., Hyman, J. M., & Robinson, B. A. (2008). Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation. *Water Resources Research*, *44*(12). https://doi.org/10.1029/2007wr006720

Wu, W., Emerton, R., Duan, Q., Wood, A. W., Wetterhall, F., & Robertson, D. E. (2020). Ensemble flood forecasting: Current status and future opportunities. *WIREs Water*, *7*(3), e1432. https://doi.org/10.1002/wat2.1432

Young, A., Orlandini, S., Moretti, G., & Albertson, J. D. (2024). Real-time flood inundation modeling with flow resistance parameter learning [Dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.11643497