Earth's Future

COMMENTARY
10.1029/2026EF008857

Reimagining How Flood Warnings Can Inform Decision-Making and Community Actions

Vinh Ngoc Tran[1,2], Xun Huan[3], Anindya Das Antar[4], Nikola Banovic[4], Jeff H. Bednar[5], Shannon Marie Bergt[6], Chen Cheng[3], Francina Dominguez[7], Simone Fatichi[8], Richard Gonzalez[9], Kevin Gray[7], Brian Jewett[7], Jongho Kim[10], Phong V.V. Le[11], Dan Lu[2], Snehal Prabhudesai[4], Deffi Putri[7], Sudhansu Rath[7], Khachik Sargsyan[12], Sarah H. Whitaker[13], Daniel B. Wright[14], Donghui Xu[15], John P. Ziker[13], and Valeriy Y. Ivanov[1]

[1]Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, MI, USA, [2]Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA, [3]Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI, USA, [4]Department of Computer Science and Engineering, University of Michigan, Ann Arbor, MI, USA, [5]Macomb County Public Works Commissioner Candice S. Miller, Clinton Township, MI, USA, [6]Department of Defense, US Army Garrison Detroit Arsenal Energy Manager, Detroit, MI, USA, [7]Department of Atmospheric Sciences, University of Illinois at Urbana-Champaign, Urbana, IL, USA, [8]Department of Civil and Environmental Engineering, National University of Singapore, Singapore, Singapore, [9]Department of Psychology, University of Michigan, Ann Arbor, MI, USA, [10]School of Civil and Environmental Engineering, University of Ulsan, Ulsan, South Korea, [11]Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA, [12]Sandia National Laboratories, Computational Physical Science and Uncertainty Quantification, Livermore, CA, USA, [13]Department of Anthropology, Boise State University, Boise, ID, USA, [14]Department of Civil and Environmental Engineering, University of Wisconsin-Madison, Madison, WI, USA, [15]Atmospheric, Climate, & Earth Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA

Key Points:

- FEWS-DM framework integrates flood prediction with human behavior and decision science to deliver actionable information
- Uncertainty quantification and communication are essential challenges that must account for user interpretation and cognitive constraints
- AI enables prediction and learning in flood systems but requires human-centered design and community feedback to build trust

Correspondence to:

V. N. Tran, V. Y. Ivanov and X. Huan,
vinhtn@ornl.gov;
ivanov@umich.edu;
xhuan@umich.edu

Citation:

Tran, V. N., Huan, X., Antar, A. D., Banovic, N., Bednar, J. H., Bergt, S. M., et al. (2026). Reimagining how flood warnings can inform decision-making and community actions. *Earth's Future*, *14*, e2026EF008857. https://doi.org/10.1029/2026EF008857

Received 29 APR 2026
Accepted 11 JUN 2026

Abstract

Society faces increasingly severe flood hazards, intensifying demand for flood early warning systems (FEWS) that deliver accurate and actionable information. However, most existing FEWS remain prediction-centric, treating decision-making as a downstream consumer of hazard forecasts while offering limited support for uncertainty interpretation, risk communication, and real-world response. This Perspective presents a vision and blueprint for a novel inland FEWS-decision-making (FEWS-DM) framework that repositions decision-making as an equal partner in the forecasting process—not a passive recipient of its outputs. The framework is built on three tightly coupled, co-evolving thrusts: Physical Science (T1), which advances flood prediction with quantified uncertainty informed by decision relevance; Human Science (T2), which incorporates psychology, behavior, and cultural and institutional context; and Decision Science (T3), which unifies physical predictions and human factors through principled, utility-based decision support with end-to-end uncertainty management. Rather than treating T1 as a solved problem, FEWS-DM recognizes that forecast development itself must be shaped by decision needs through continuous bidirectional feedback. We identify key scientific, behavioral, and operational challenges limiting such integration and discuss the enabling role of AI, while emphasizing human-centered design and community feedback as essential for building trust and improving flood risk management.

Plain Language Summary

Floods are becoming more severe, creating urgent demand for early warning systems that deliver useful, actionable information. Most current systems focus on predicting water levels but provide limited support for interpreting uncertainty or guiding real-world decisions. This paper introduces the FEWS-DM framework, which integrates flood forecasting with human behavior and decision science to transform warnings into decision-support tools tailored to user needs. We identify key challenges including improving forecast accuracy, communicating uncertainty effectively, supporting decision-making under multiple sources of uncertainty, and understanding how psychological, cultural, and institutional factors shape warning responses. We discuss how artificial intelligence can enhance prediction and communication capabilities while emphasizing that human-centered design and community feedback are essential for building trust and improving flood risk management.

1. Introduction

Flooding represents the most frequent and damaging natural disaster globally, in both human lives and economic losses. Approximately 1.8 billion people, constituting 23% of the global population, are directly exposed to 100-year flood events with inundation depths greater than 0.15 m (Rentschler et al., 2022). Global data show that more than 688,000 people lost their lives due to flooding between 1985 and 2021, with fatalities trending upward (Brakenridge, 2025). Since the 1990s, record-breaking rainfall events that drive extreme floods have deviated substantially from historical patterns and are projected to occur with increased frequency (Robinson et al., 2021), while remaining poorly captured by conventional extreme value analysis (Chen et al., 2025). Flood-related economic damages are estimated to intensify more than 20-fold by the end of this century (Hirabayashi et al., 2013; Paprotny et al., 2018; Winsemius et al., 2016).

Flood Early Warning Systems (FEWS) are among the most cost-effective disaster risk reduction tools available. FEWS are estimated to have reduced global flood-related fatalities by up to 43% (WHO, 2014; WMO, 2013) and economic costs by 35%–50% (Rogers & Tsirkunov, 2011). Cost–benefit analyses of European flood early warnings estimate economic returns of 400:1 (Pappenberger et al., 2015). In response, the UN Secretary-General launched the Early Warnings for All (EW4All) initiative, setting a target of universal FEWS coverage by 2027, building on the Sendai Framework's commitment to multi-hazard early warning systems by 2030 (Pearson & Pelling, 2015; WMO, 2022). The United Nations Office for Disaster Risk Reduction (UNDRR) reports measurable progress: 101 countries now have some form of multi-hazard early warning system (EWS), up substantially from a decade earlier, demonstrating that integrating scientific forecasting with community-level communication can reduce disaster losses (UNDRR, 2023). However, despite advances in monitoring, numerical weather prediction, and flood forecasting (Arduino et al., 2005; Cools et al., 2016; Hapuarachchi et al., 2011; Ivanov et al., 2021; Krzhizhanovskaya et al., 2011; Tran et al., 2023; Wu et al., 2020), substantial socioeconomic and human losses continue to occur during catastrophic events. For example, the July 2021 Western European floods killed over 220 people despite operational FEWS coverage (Koks et al., 2022). The April–May 2024 Rio Grande do Sul floods in Brazil caused 183 fatalities and displaced hundreds of thousands (Simoes-Sousa et al., 2025). The July 2025 Texas or the August 2025 Beijing flash floods exposed the same extreme vulnerabilities. The persistence of these tragedies warrants a systematic accounting of where current systems succeed, where they fail, and what structural innovations are needed.

Most operational FEWS predict rainfall and resulting streamflow that can be translated into approximate inundation levels near forecast locations (Cosgrove et al., 2024; Emerton et al., 2018; Emerton et al., 2016; Najafi et al., 2024; Reichstein et al., 2025). While valuable for emergency management authorities, such predictions are largely one-directional and detached from the realities of flood mitigation decision-making. Forecast users frequently struggle to translate hazard-oriented information into timely and appropriate responses, particularly at spatial scales of practical relevance (e.g., human-body, infrastructure, or buildings) and across different exposure contexts (e.g., static vs. mobile assets) (Reichstein et al., 2025; Thieken et al., 2023). The need for decision-relevant information is underscored by evidence that flood-related fatalities primarily occur among people traveling outdoors, rather than those sheltering indoors (Arrighi et al., 2015, 2017; Ivanov et al., 2021; Sanders et al., 2020). Yet few, if any, operational FEWS provide timely information directly usable by mobile users or decision-makers involved in route safety or real-time mobility decisions. Another major limitation of current FEWS is the absence of systematic approaches to characterize and communicate uncertainty across multiple sources, including observations, meteorological inputs, and flood-generating physical processes. This impacts confidence in warning systems, creating additional barriers to their effective implementation and user adoption. Furthermore, in addition to the aforementioned limitations, flood-related decisions are frequently constrained by behavioral, cognitive, and socio-economic aspects (Breznitz, 1984; Henrich et al., 2010; LeClerc & Joslyn, 2015; Ziker, 2007). When all of the above factors combine, they can lead to a "*flood preparedness impasse*", that is, the state in which decision makers become severely limited in their capacity to receive decision-relevant information, appropriately process, query, and interpret it, and undertake effective preventive measures.

The flood preparedness impasse continues to be common (Morss et al., 2016; Perera et al., 2020; Reichstein et al., 2025; Scolobig et al., 2012). Existing literature and our cross-disciplinary discussions with communities, residents, and stakeholders indicate that the linkage between FEWS and decision-making remains predominantly unidirectional: information flows from forecast systems to users, with little or no feedback capabilities to adapt forecasts to decision needs (Cools et al., 2016; Krzhizhanovskaya et al., 2011; Liu et al., 2018; Najafi et al., 2024;

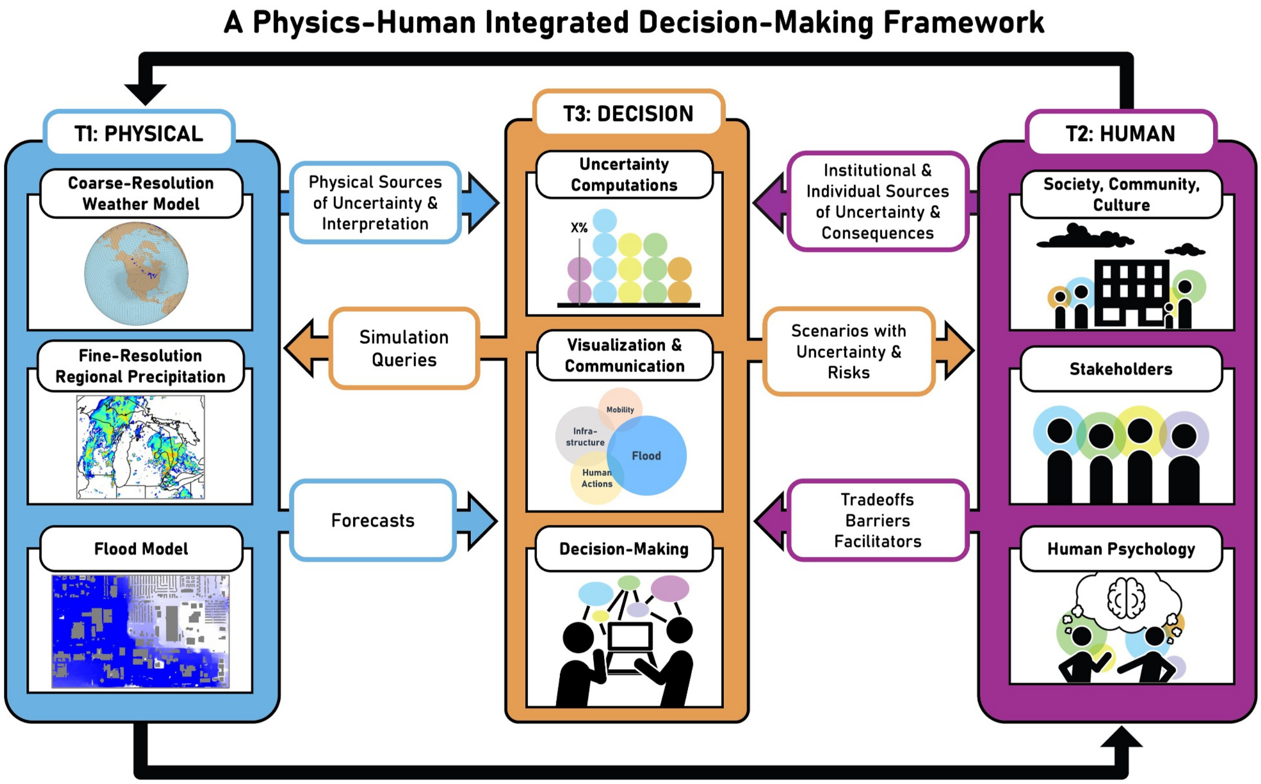

Figure 1. Description of the flood early warning framework, from prediction to decision. The framework contains three key Thrusts: **T1**–Physical sciences integrating climate, hydrometeorology, and flood predictions; **T2**–Human and social sciences relating interactions between individual psychology and the behavior of stakeholders, their community and culture; **T3**–Decision science integrating physical model predictions and human psychology and cultures with uncertainty quantification, management, and visualization. Black arrows denote feedback loops where T2 inform T1's refinement of predictions based on stakeholder needs and observed forecast errors, enhancing both accuracy and decision-relevance.

Perera et al., 2020; Zang et al., 2022). Decision effectiveness depends strongly on the spatiotemporal characteristics of forecasts, particularly lead time, that is, the period of time prior to a flood that is available for preparedness actions. A priori, however, well-structured flood risk mitigation should offer decision-makers the ability to query FEWS for information tailored to lead times, locations, and actions they anticipate. This highlights the need for a *bi-directional coupling between FEWS and decision-making* (i.e., FEWS-DM).

Drawing on a growing body of disaster post-mortems, behavioral research, and FEWS design evaluations, we contend that the largest gaps in flood early warning systems are in warning dissemination, uncertainty communication, and decision-making related response capability, and not in hazard monitoring and prediction per se (Cools et al., 2016; De Perez et al., 2022; Golding, 2022; Painter et al., 2025; Potter et al., 2025). No country has yet fully achieved all four WMO (World Meteorological Organization) EWS pillars (i.e., disaster risk knowledge, detection/forecasting, warning dissemination, and preparedness/response), and very few have fully functioning end-to-end systems (Painter et al., 2025; UNDRR, 2023).

What distinguishes the framework proposed here is a shift in the focus from a gradual incremental improvement in traditional forecast skill, to a structural redefinition of what should constitute a flood warning. Specifically, rather than treating decision-making as a downstream consumer of flood forecast information, we posit that decision objectives, behavioral constraints, and uncertainty tolerance need to be embedded within the forecasting loop itself. This shift, from a hazard-oriented prediction problem to a decision-conditioned prediction framework, requires continuous bilateral feedback between forecast generation and flood information use by decision-making.

Here, we outline a state-of-the-science real-time inland flood-forecasting framework that connects the current state of the science of prediction with flood response decision-making aiming to minimize flood risks. We propose a *Physics-Human Integrated Decision-Making Framework* (Figure 1) as the basis for FEWS-DM, built upon three tightly coupled thrusts: (**T1**) Physical Science, focused on simulating weather, extreme precipitation, and flooding with quantified uncertainty; (**T2**) Human Science, incorporating psychology, behavior, information perception, and stakeholder constraints within local, regional, and institutional contexts; and (**T3**) Decision Science, which unifies physical predictions and human factors through a principled, utility-based decision-making platform with end-to-end uncertainty management and visual communication.

2. The FEWS-DM Vision

2.1. What Current FEWS Have Achieved

The past three decades have witnessed remarkable advancements in operational flood forecasting or *hazard-based FEWS* systems that effectively provide real-time flood predictions. At continental and global scales, the Global Flood Awareness System (GloFAS) and European Flood Awareness System (EFAS) now provide ensemble probabilistic forecasts for all major river systems, with EFAS fully operational since 2012 (Emerton et al., 2016, 2018). In the United States, NOAA's (National Oceanic and Atmospheric Administration) National Water Model (NWM) produces streamflow forecasts for over 2.7 million river reaches, a step-change in spatial coverage relative to the historical ∼4,000-gauge network (Cosgrove et al., 2024). The Flanders (Belgium) EWS integrates ensemble numerical weather prediction (NWP) with radar composites to produce 48-hr and 10-day flood forecasts accessible via a public web portal, representing a model of multi-data-source FEWS integration (Fernández-Nóvoa et al., 2024).

Publicly accessible platforms have brought flood information closer to communities. For example, NWM, operated by NOAA, and River Forecast Centers (RFCs) together provide streamflow forecasts and flood stage guidance across the contiguous United States at multiple timescales: short-range (18-hr deterministic), medium-range (240-hr ensemble), and long-range (up to 30 days) (Cosgrove et al., 2024). The Iowa Flood Information System provides real-time sensor data, forecasts, and flood maps to community members and emergency managers over the State of Iowa, USA (Demir & Krajewski, 2013; Krajewski et al., 2017). Google Flood Hub provides AI-powered 7-day flood forecasts for thousands of locations worldwide (Nearing et al., 2024). Additionally, several other flood forecast systems can be mentioned, such as GloFAS (Emerton et al., 2016), EFAS (Smith et al., 2016), Delft-FEWS (Werner et al., 2013), FLASH (Gourley et al., 2017), FLOODIS (Rossi et al., 2015), or ANYWHERE (Sempere-Torres & Berenguer, 2024).

The transition to *Impact-Oriented FEWS* (IO-FEWS) represents a conceptual breakthrough beyond hazard prediction: rather than forecasting what the flood will be, IO-FEWS forecast what it will do—damage to infrastructure, displacement of populations, disruption of services (Merz et al., 2020; Najafi et al., 2024; Paul Zischg, 2024; Sempere-Torres & Berenguer, 2024). The principles of IO-FEWS are now codified in the WMO and implemented across member agencies worldwide (WMO, 2015). By extending FEWS outputs from physical hazard descriptors to consequence-oriented products, IO-FEWS begin to close the gap between hazard prediction and the information needs of decision-makers. Yet this shift, while significant, remains largely unidirectional: impact forecasts are delivered to users without systematic mechanisms for incorporating decision objectives, uncertainty tolerances, or behavioral constraints back into the forecasting process.

2.2. The Persistent Decision Gap

Despite achievements, a stark pattern emerges from a systematic analysis of high-casualty flood disasters: there is a persistent disconnect between the flood information generated, its communication to decision-makers, and the actions required. De Perez et al. (2022) analyzed the six deadliest and costliest hydro-meteorological disasters of the 21st century (Myanmar 2008, Russia 2010, Europe 2003, Somalia 2010, Philippines 2013, and India 2013) and found a consistent pattern across all cases: forecast systems performed adequately or fairly well, but warning dissemination and response capability failed. For example, in Typhoon Haiyan (Philippines, 2013; 7,000+ deaths), forecasts predicted landfall accurately 2 days in advance, but many residents did not understand the term "storm surge", and governance gaps prevented effective evacuation. In the 2021 Germany/Belgium floods, which resulted in over 220 deaths, Thieken et al. (2023) documented that warnings were issued but many residents did not receive them in a timely manner via official channels; rather, community members and environmental cues served as primary information sources for those who evacuated. Brazil's CEMADEN's (National Center for Monitoring and Alerts for Natural Disasters) post-event analysis of the 2024 Rio Grande do Sul floods identified "*challenges in risk communication and last-mile delivery*" as the primary failure mode despite adequate forecasting (Simoes-Sousa et al., 2025).

A comprehensive review of FEWS by Painter et al. (2025) found that no country has fully achieved all four WMO EWS pillars, and that the least-developed pillars are warning dissemination and response capability. Their analysis of the FloodNet NYC (New York City) system, the SIATA system (Sistema de Alertas Tempranas) in Colombia, and community-based EWS in Nepal all point to participatory co-design as a key differentiator of system effectiveness: communities need to have agency in sensor placement, communication protocol design, and warning dissemination. Broader global assessments reinforce this picture: Perera et al. (2020) identified information accessibility, last-mile communication, community engagement, and institutional coordination as the most critical unresolved barriers, while Potter et al. (2025) synthesized findings from 350+ WMO workshop participants to identify persistent research gaps including uncertainty communication, behavioral response quantification, and integration of modern technology with local emergency communication needs. Collectively, these studies converge on a shared conclusion: "*the potential benefits of scientifically sound forecasting systems will not materialize if a warning is not understood, is not useable or does not result in emergency response action*" (Cools et al., 2016).

This body of evidence motivates the core premise of FEWS-DM: addressing the continued needs for technical improvements of flood forecast lead time, accuracy, and space-time coverage will not close the decision gap on its own. Three structural innovations are required to create FEWS-DM framework: (a) bidirectional coupling between user decision contexts and forecast generation; (b) formal integration of human behavioral science and decision theory as co-equal framework components; and (c) continuous learning from decision outcomes for past events to improve forecast generation and communication.

2.3. FEWS-DM Architecture

The proposed framework involves highly interconnected collaboration among the three thrusts across both event and inter-event timescales. During flood events, T3 interacts with T1 by asking "*what will happen and when?*" for precipitation and flood events, and T1 responds by providing detailed predictions and forecasts; T2 informs T3 by addressing "*what is important?*" for the various stakeholder categories, describing key tradeoffs, facilitators, and barriers for pre-event and real-time decision-making; T3 communicates with T2, responding to "*how to understand flood forecasts,*" offering predictions with their uncertainties, consequences, and risks. Between events, T1's predictive capabilities are iteratively improved through feedback from T2 and T3: T2 provides insights into which forecast features and uncertainties most critically impact decision-making across different stakeholder groups, while T3 identifies systematic forecast errors, communication gaps, and user comprehension challenges that emerged during past events. This long-term iteration enables T1 to refine model parameterizations improving prediction accuracy, adjust uncertainty quantification approaches, and prioritize improvements in forecast attributes that matter most for effective decision-making, creating a continuous learning cycle that enhances both forecast skill and practical utility. This framework surpasses conventional FEWS by providing not only flood information, such as where and when flooding will occur, but also guidance on what end-users should do in culturally relevant terms, integrating uncertainties from both physical modeling and human behavioral and psychological factors.

The foundation of T1 is scientific understanding of flood-generating physical processes and their uncertainties, integrating observational and forecast data, while generating new information tailored to specific decision needs. This forecasting system is designed to be rapidly queried with performance improving over time as feedback accumulates. T2 describes the diversity of decision-makers and users of flood forecast information, capturing behavioral, cognitive, cultural, and socio-economic factors that shape how forecast information is perceived, trusted, and acted upon. T3 operationalizes decision-making by placing uncertainty assessment, management, visualization, and communication at the center of the forecasting–decision interface, explicitly linking probabilistic predictions to time-bounded action choices.

The decision-makers, or end-users of the FEWS-DM framework extend beyond government authorities and agencies with privileged access to forecast data. Specifically, the envisioned framework serves all individuals and organizations who need forecast information and must make decisions under imminent flooding risk, including disaster risk managers, administrators of residential areas, economically or strategically critical infrastructure operators, and even individual citizens. Importantly, the framework accommodates heterogeneous decision contexts, recognizing that optimal actions, acceptable risk levels, and required lead times vary across users and situations.

Artificial Intelligence (AI) capabilities can be leveraged throughout the framework to support these interactions. AI's ability to process large data volumes and accelerate forecasting is critical for enabling rapid, high-resolution predictions, while its use of surrogate modeling, uncertainty estimation, and information translation can enhance decision support. Within FEWS-DM, AI is envisioned as an enabling and augmentative component, rather than a replacement for physical understanding or human judgment, supporting automated decision assistance and adaptive communication with diverse user groups while operating within transparent, uncertainty-aware, and human-in-the-loop system designs (Tran et al., 2026).

To the best of our knowledge, FEWS-DM represents the first framework to integrate all four of the following structural features simultaneously. First, query-driven forecast generation: the physical modeling system (T1) does not simply broadcast generic products but responds to decision-specific queries—specifying location, variables of interest, lead time, spatial and temporal scales, and uncertainty format—that are formulated by the decision science component (T3) based on each user's declared objectives and constraints. No existing hazard-based FEWS implements this mechanism, including GloFAS, EFAS, NOAA NWM, and IFIS, as well as IO-FEWS (Najafi et al., 2024) or ANYWHERE (Sempere-Torres & Berenguer, 2024), they all deliver pre-defined forecast products to users, with no pathway for user decision contexts to modify what the physical model computes or prioritizes (Cosgrove et al., 2024; Emerton et al., 2018; Merz et al., 2020; Painter et al., 2025). Second, behavioral science as a co-equal model input: FEWS-DM formally encodes psychological mechanisms—risk aversion, loss aversion, anchoring, experiential processing, “cry-wolf” dynamics (Breznitz, 1984; LeClerc & Joslyn, 2015; Tversky & Kahneman, 1992)—as explicit inputs to T3's uncertainty communication design, not as post-hoc considerations. Existing community-based EWS and participatory platforms (Demir & Krajewski, 2013; Painter et al., 2025) introduce community input into warning dissemination but have not so far coupled behavioral models to the physical forecasting or uncertainty quantification engine. Third, end-to-end uncertainty propagation to decision outputs: FEWS-DM propagates uncertainty from meteorological inputs through hydrological modeling through the decision-theoretic interface, producing not a probabilistic forecast but a probabilistic recommendation—quantifying the expected utility of each available protective action under the full joint uncertainty distribution. This differs fundamentally from displaying ensemble spread or probability-of-exceedance curves to decision-makers and leaving their translation and interpretation to the user (Knotters et al., 2024; Prabhudesai et al., 2023). Such end-to-end propagation remains a substantial open challenge, requiring at the very minimum consistent uncertainty representation across various model components and computationally feasible inference within real-time constraints; mapping modeled probabilities to action success likelihood is another major frontier. Fourth, a closed inter-event learning loop: T2 and T3 systematically extract information from observed decision outcomes, communication failures, and forecast verification after each event, and feed this back to T1 to reprioritize forecast improvements—creating a compounding return on the system's practical utility that no existing operational FEWS achieves (Cools et al., 2016; De Perez et al., 2022; Potter et al., 2025). Each of these four features has been studied independently in prior research: decision-theoretic forecast valuation (Van Houtven, 2024), behavioral flood risk communication (Morss et al., 2016), impact forecasting (Merz et al., 2020), and adaptive FEWS design (Dale et al., 2014). What FEWS-DM proposes for the first time is their structural integration into a single, bidirectionally coupled system in which the forecasting loop and the decision loop are co-designed to condition each other. This is the specific and novel contribution of the framework outlined in this Perspective.

3. Challenges to Flood Forecasting (T1)

A foundational challenge of the framework is the need for forecasting information *prior* to the occurrence of flood events (T1 in Figure 1), with sufficient lead time for preparedness measures (Najafi et al., 2024). To effectively support decision-making and appropriate actions, FEWS should ideally provide the following forecast products: (a) precipitation forecast that realistically captures the intensity and duration of precipitation in the region; (b) spatially explicit inundation results at the “human action” scale resolution (finer than 100 m^2, e.g., the inset inundation map in Figure 2); (c) flood evolution timing with high temporal resolution (hourly or sub-hourly); (d) sufficient lead time for decision-making and implementation; and (e) probabilistic forecast results that explicitly include uncertainty. However, methodologies that *simultaneously* satisfy these requirements—using

10.1029/2026EF008857

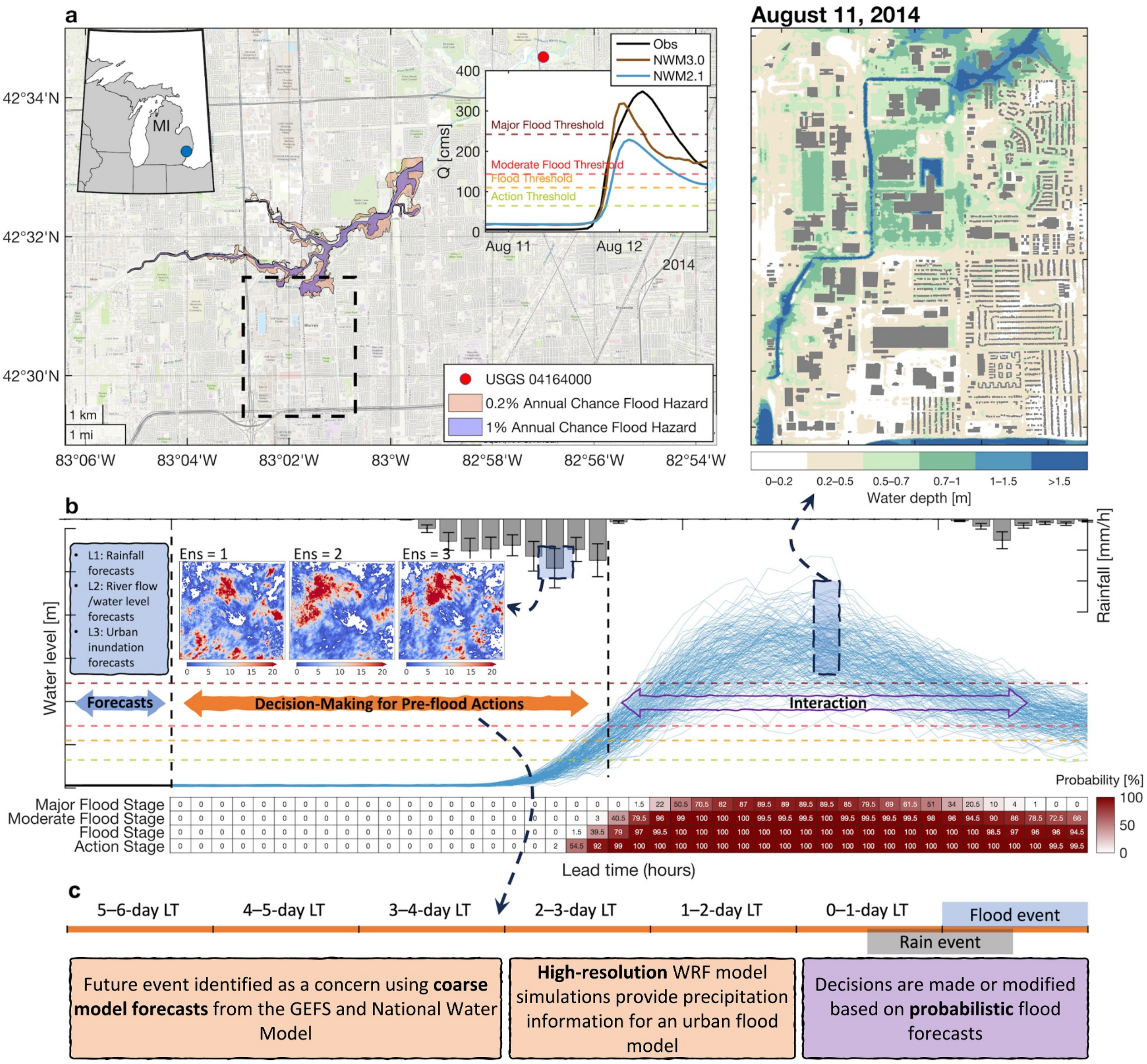

Figure 2. Illustration of a prototype of FEWS-DM system in operation using an exemplary location in Southeast Michigan, USA. Map (a) shows FEMA flood hazard maps with 0.2% and 1% annual exceedance probabilities, corresponding to 500- and 100-year return periods, respectively. The inset plot displays streamflow at the USGS (United States Geological Survey) gauge 04164000 during 11–12 August, 2014 flood event, showing the observed data and simulations from the NWM versions 2.1 and 3.0. Subplot (b) illustrates an example (synthetic) of real-time ensemble of flood (river water level) forecasts with ensemble rainfall forecasts, providing probabilistic information for flood occurrence based on flood warning levels provided by NOAA across different lead times. The inset with a map (top right) presents maximum inundation depths for the same flood event simulated using an urban flood model (Kim et al., 2012; Tran et al., 2024). In (b), the forecast timeline is divided into three phases: the **forecast** period occurs before flood onset (when water levels exceed flood thresholds) to ensure sufficient time for decision-making and implementing necessary pre-flood actions; the **decision-making** period encompasses the time for decisions and protective actions; and the **interaction** period represents the flood event itself, requiring real-time interaction between the FEWS decision-support system and users to minimize risk. (c) Decision-making phases could be split into the following: 3–6 days LT (lead time): decision makers notified of potential storm and provided coarse model forecasts. This includes dissemination of information to all potential end-users in jurisdictions that will be affected. Decision makers use information to come up with action plans and initiate communication across jurisdictions about how to manage the storm event. 2–3 days LT: Decision makers from relevant jurisdictions meet together to identify action plans for predicted events as well as alternative plans, should the storm behave unexpectedly. 0–1-day LT: DMs update action plan, ensure channels of communication are open so that rapid decisions can be made during the flood event.

first principles models, ingesting the required geospatial data, and enabling rapid pre-event and real-time queries—remain largely unavailable in operational settings. We identify a set of fundamental challenges that limit the implementation of these capabilities.

To the best of our knowledge, current operational FEWS do not provide predictions *simultaneously* satisfying all of the above requirements, despite progress in delivering basic "when and where" flood information (items 1 and 2 listed above). While timing and location are fundamental FEWS characteristics, even in developed countries, for example, the United States or Europe, such systems remain severely limited in their spatial resolution, temporal detail, or uncertainty characterization. Specifically, flood forecasting information in the United States is generally conveyed as streamflow or water level predictions for main channels and selected locations through systems operated by NOAA River Forecast Centers (Figure 2a). Spatial assessments of flood extent are largely limited to static FEMA (Federal Emergency Management Agency) hazard maps (Figure 2a), which delineate floodplains near open channels based on historical probability but do not predict when or where flooding will occur during specific events. Urban flooding (pluvial flooding) frequently occurs at locations far from rivers (Figure 2b) (Ivanov et al., 2024; Nelson-Mercer et al., 2025) yet remains largely unaddressed by operational forecasting systems. As a result, existing FEWS often fail to provide both temporal and spatial inundation information required for effective, real-time decision-making.

A second fundamental challenge for flood forecasting and decision support is the accuracy of simulated precipitation using NWP models and the modeling of water fluxes over and through the land surface using flood or overland flow models (OFM). Despite steady progress in the NWP forecasting technology (Bauer et al., 2015), substantial uncertainties persist in precipitation forecasts, particularly for extreme events (Boelee et al., 2019) and at fine spatiotemporal scales (Berne et al., 2009; Wright et al., 2013) associated with convective storms (Yano et al., 2018). These prediction challenges are critical because accurate precipitation forecasting at neighborhood scales, that is, minutes to hours in time and hundreds of meters in space, is essential for capturing flood dynamics in urban environments (Berne et al., 2004; Bracken et al., 2013; Pregnolato et al., 2017; Tran et al., 2024; Wainwright et al., 2011).

Decades of OFM research have produced models capable of fine-resolution simulations (up to 1 m resolution) that accurately describe the flow phenomena in complex areas such as urban environments (Balaian et al., 2024; Guidolin et al., 2016; Tran et al., 2024), and model structure itself is often not the main source of uncertainty. Instead, current challenges mostly stem from data-related issues, including representation of boundary conditions of the flood environment (topography, buildings, and infrastructure such as drainage systems, roads, ponds, dikes, or dams), initial conditions (e.g., soil moisture, water level in reservoirs, rivers, or underground conduits), and human interventions in water flows (flow regulation, reservoir and drainage operations, or construction that alters topography) (Galelli et al., 2025; Ivanov et al., 2021). At fine spatial resolutions, flood dynamics become increasingly sensitive to these factors, with complex spatial and temporal connectivity (Tran et al., 2024). As a result, data uncertainties can significantly degrade forecast accuracy, creating the need for systematic uncertainty quantification and reduction strategies (Bodoque et al., 2023).

A further obstacle is the computational burden associated with uncertainty quantification in precipitation inputs and flood modeling outputs. Uncertainty estimation typically requires running prediction models under multiple realizations of inputs, initial and boundary conditions, physics parameterizations, or human influence scenarios (Ivanov et al., 2021; Najafi et al., 2024; Tran et al., 2020). The number of required simulations increases dramatically when additional sources of uncertainty are included, making high-fidelity ensemble forecasting computationally prohibitive for real-time applications, with simulation times potentially exceeding the available decision window (Ivanov et al., 2021). This computational challenge remains one of the most common and daunting barriers to operational FEWS. For example, the NWM addresses some forecast uncertainty in the medium-term (up to 10 days) by providing six ensemble members, but the short-term forecast (up to 18 hr) is deterministic (i.e., only one ensemble member) (Cosgrove et al., 2024). This system produces channel water level and streamflow forecasts. For inundation mapping, rather than using dynamic inundation modeling, NOAA has experimented with the Height Above Nearest Drainage method (Nobre et al., 2011) to derive inundation extent and depth, but with significant limitations in accuracy, spatial detail, and temporal resolution. Together, these challenges present fundamental roadblocks to providing timely, uncertainty-aware flood predictions suitable for operational decision support.

There is a trend toward using AI and machine learning (ML) for weather and flood forecasting, motivated by their potential to improve computational efficiency and predictive skill, especially for streamflow predictions (Bao et al., 2025; Bi et al., 2023; Lam et al., 2023; Tran et al., 2025). However, many AI/ML studies in weather forecasting do not explicitly evaluate precipitation forecast quality for convective storms and extreme events (Bi et al., 2023; Lam et al., 2023). Although ML shows promise for certain regimes, current approaches often underestimate extreme events (Charlton-Perez et al., 2024; Pasche et al., 2025; Sun et al., 2025; Zhang et al., 2025), largely due to limited training data for rare events. Some studies have proposed training ML using satellite imagery (Frame et al., 2024); however, both spatial and temporal resolution of data are insufficient for the necessary level of detail, especially in urban areas.

More fundamentally, the pursuit of fully automated AI-driven forecasting systems overlooks a critical lesson from operational practice: the most reliable prediction systems integrate human expert judgment with automated capabilities. The River Forecast Centers in the United States, for example, consistently outperform purely automated systems such as the National Water Model because experienced meteorologists and hydrologists continuously monitor model outputs, observations, and forecasts, making manual adjustments to compensate for model limitations and uncertainty in input data such as numerical weather predictions (Cosgrove et al., 2024; Emerton et al., 2016; Tran et al., 2025). This "forecaster-in-the-loop" approach—where human expertise guides and corrects automated systems—has proven essential across high-stakes applications, from aviation (pilots plus autopilot plus ground control) to weather forecasting (Stuart et al., 2022; Tran et al., 2026). Accordingly, AI/ML approaches in FEWS should be designed to *augment*, rather than replace, human expertise, enabling experts to focus on critical judgment tasks while automation handles routine computation and monitoring.

4. Challenges to Decision-Making (T2 and T3)

Even where flood forecasts are accurate and timely, converting information into effective protective action by diverse users remains the most consequential unsolved challenge of operational FEWS.

Receiving forecast information from predictive models and transforming it into decisions for flood mitigation has always been challenging (Cools et al., 2016; De Perez et al., 2022; Painter et al., 2025). The process is influenced by numerous interacting factors (T3 in Figure 1). Fundamentally, the key questions are: who makes the decision, for whom is the decision made, and what information, constraints, and uncertainties shape that decision-making? Here we contend that flood forecasting information, beyond its traditional focus on flood management authorities (e.g., FEMA), should include the capability for user-specific decision-making. Millions of people affected by floods should have access to useful forecast data products to inform their decision-making based on their specific needs. While emergency authorities retain responsibility for issuing mandatory evacuation orders and other protective directives when public safety requires centralized coordination, individuals or specific organizations can benefit from accessible forecast information to make informed decisions within the bounds of official guidance. This necessitates a decision support system that allows diverse end-users to access and interact with flood forecast information. This represents a novel evolution of FEWS-aided decision-making systems. However, significant challenges remain.

First, there is no existing platform that allows the public to directly interact and communicate with FEWS forecast results. Personalized decision support systems—such as allowing users to specify their location, exposure context, or constraints, and receiving tailored warnings or safe-route guidance with clear explanations—remain largely absent from integrated operational frameworks. Some platforms, such as the US-NWS (National Weather Service) Alert and the EU-Alert cell-broadcast system, deliver location-aware alerts to mobile devices. However, none of these systems allows users to specify decision-specific constraints, nor do they integrate user feedback into the forecasting process. One significant challenge is that a diverse decision-making "ecosystem" includes stakeholders or end-users whose expertise is not in mathematics, computer science, or computational modeling (T2 in Figure 1). Current weather and flood model outputs and uncertainty information are not sufficiently accessible or interpretable by most decision-makers. Although approaches exist to achieve interpretability (Gilpin et al., 2018; Ras et al., 2018); these are mostly designed for technically trained users (e.g., model developers or analysts (Bhatt et al., 2020)), rather than for decision-makers who need to adequately understand forecast outputs, assess relevant measures of uncertainty, and act appropriately under time pressure. As a result, many existing methods that translate scientific information remain ineffective for non-scientist users (Kaur et al., 2020; Zhang et al., 2020) as they are poorly interpretable (Kaur et al., 2024) and misaligned with users' mental models (Kaur et al., 2022; Wang et al., 2019). Furthermore, evaluation of existing uncertainty communication mechanisms that aim to aid with interpretation of model results and transparency (Slack et al., 2021; Wang et al., 2021) has largely focused on how much decisionmakers agree with model predictions in the presence of uncertainty. However, studies of decision processes (e.g., how users make decisions under different uncertainty communication mechanisms) and the effect of uncertainty (e.g., how different uncertainty communication mechanisms affect decision quality) have shown that lay decisionmakers generally struggle to interpret mathematical and statistical visualizations of uncertainty (Prabhudesai et al., 2023).

Effective support systems must also enable interactions between human decision-makers and physical models, fostering trust by allowing users to actively explore "what-if" scenarios rather than passively viewing pre-selected outputs. However, such interactive capabilities remain largely absent from current flood decision-making systems. How these tools should be designed to align with the actual decision-making process across heterogeneous end-users remains to be explored. Furthermore, by allowing multiple users to interact with the same system, additional uncertainties can arise from competing priorities and decision-making approaches among diverse stakeholders, as well as from differences in how information is gathered and processed by various modeling procedures.

Serious gaming represents an underutilized avenue for building decision competence and generating behavioral data for FEWS design. Serious games place stakeholders in realistic flood scenarios, allowing practice of decision-making under uncertainty without real-world consequences (Garcia et al., 2022; Solinska-Nowak et al., 2018). This approach serves multiple functions within FEWS-DM: they reveal systematic cognitive biases and decision heuristics; they build decision competence and reduce the surprise of actual events; they enable empirical testing of different warning formats and uncertainty representations against decision quality outcomes; and they provide community engagement mechanisms that build trust and system legitimacy. A study using virtual reality simulation to test residents' wildfire evacuation intentions (Molan et al., 2022) found significant improvements in evacuation intention relative to traditional risk communication—suggesting that immersive simulation holds promise for flood contexts.

The interplay of various uncertainties in decision-making is further complicated by institutional, cultural, and individual factors. A number of studies demonstrate how cultural variability affects decision-making and assessment of risk and uncertainty (Atran et al., 2005; Bier & Lin, 2013; Bland & Schaefer, 2012; Boholm, 2003; Settembre-Blundo et al., 2021; Weber & Hsee, 2000; Yoe, 2019). There is variation in social and cultural systems at multiple levels: regional, intra-community, organizational, and even within different branches of organizations. This variation impacts individual and collective decision-making, especially when decisions are perceived as high stakes or require cross-jurisdictional collaboration (Lunstrum & Havice, 2025). Existing practices and logics of organizations can lead to the persistence of traditional ways of making decisions and limit innovation in the face of new hazards and risks (Aylett, 2013; Bohman et al., 2020; Bremer et al., 2021). A key challenge is the translation of advancements into practice: innovative knowledge tends to stay with those who use it, rather than percolate through institutions (Lane et al., 2019). The presence of multiple jurisdictions and entities responsible for urban flood resilience, and mismatches in social and cultural contexts of decision making in these institutions, limits coordination and execution of infrastructure resilience planning. Uncertainties originating from institutional, cultural, and individual variations generate constraints on adoption of new products, such as FEWS, because of concerns about credibility, security and privacy, as well as requirements made on end-users by organizations.

Human psychology can render even the most sophisticated forecasting systems or intelligent/optimized decision-making ineffective (Alvarado-Valencia & Barrero, 2014). Beyond uncertainties from technical elements in predictive systems, psychological factors—such as risk perception, trust in forecasts, and decision-making biases—introduce additional uncertainty that affects how people respond to warnings (Breznitz, 1984; Lawrence et al., 2006; LeClerc & Joslyn, 2015; Tversky & Kahneman, 1974, 1981). We contend that an effective system must not only provide reliable numerical information but must also possess the ability to capture the psychology of human decision-making. A key benefit of this approach is the ability to distinguish between different cognitive and emotional processes that shape flood response decisions. For instance, "risk aversion," that is, the tendency to prefer certain outcomes over uncertain ones even when the uncertain option has higher expected value, operates through different psychological mechanisms than "loss aversion," where people weigh potential losses more heavily than equivalent gains (Tversky & Kahneman, 1992). Other critical biases include "anchoring," when initial information disproportionately influences subsequent judgments, and "experiential processing," where people rely on vivid memories of past events rather than statistical information (Bosma et al., 2020). This decoupling is important because it points to direct interventions to address deficiencies on the human decision-making side. For example, a suboptimal decision due to risk aversion needs to be addressed differently than a suboptimal decision due to loss aversion; this affects modeling of how a decision-maker will respond to changes in uncertainty of flood water depth and velocity at particular locations because risk, potential loss, and their associated probabilities have separate underlying psychological, social, and cultural mechanisms. However, such criteria have never appeared in any decision-making system related to natural disasters, particularly flooding.

5. Holistic AI/ML-Empowered Prediction Modeling and Decision-Making

We contend that a holistic approach to FEWS-DM systems that integrates AI/ML offers a promising path forward. Recent research has highlighted the potential for AI/ML (Zhao et al., 2025) to enhance early warning systems (Nevo et al., 2022; Tran et al., 2025; Weber de Melo et al., 2025). ML applications can improve computational efficiency, enhance forecast accuracy, support transition from hazard-focused to impact-based predictions, democratize global access, and improve communication effectiveness (Reichstein et al., 2025). Within the FEWS-DM framework proposed here, AI/ML is envisioned as an enabling technology that supports prediction, uncertainty management, and decision assistance, rather than as a replacement for physical understanding of flood processes or human judgment.

Within this framework, AI/ML can contribute to FEWS-DM in the following areas:

1. *Higher resolution weather forecast products across space and time scales with next-generation computing:* Advanced AI-based weather prediction will fully leverage modern and evolving hardware architectures to enable simulations at unprecedented spatial and temporal resolutions with enhanced accuracy and reduced uncertainty (Kochkov et al., 2024). This involves reliance on high-performance computing resources, GPU acceleration, and cloud computing infrastructure to overcome current computational limitations in high-resolution atmospheric modeling. However, current weather ML models are predominantly trained and validated at resolutions of 0.1–1° (Bi et al., 2023; Lam et al., 2023), which are too coarse to capture the detailed dynamics of convective storms, particularly in urban areas where local flood response is heavily influenced by precipitation spatial distribution patterns (Berne et al., 2004; Ivanov et al., 2024; Meierdiercks et al., 2010; Peleg et al., 2017; Tran et al., 2024). Additionally, a comparison of AI weather prediction versus NWP for 1–7-day precipitation forecasting found NWP retains the advantage for extreme events (Radford et al., 2025); with traditional process-based models that outperform AI for record-breaking extremes (Zhang et al., 2025); and AI models consistently underestimate extreme precipitation (Charlton-Perez et al., 2024; Pasche et al., 2025). However, the field is advancing rapidly with new AI-based weather products emerging continuously. The key principle—resolution increases without commensurate accuracy improvements of the rainfall extremes do not serve FEWS needs—must guide AI weather model development for flood applications.

Post-processing and downscaling methods offer a complementary path (Wang et al., 2025). Statistical and deep learning downscaling trained on high-resolution observational products can improve extreme precipitation predictions (Wang et al., 2025). Importantly, this approach operates on different training data than direct prediction models, with quality constrained by observational benchmark accuracy. Emerging high-resolution ML models, such as HRRRcast (a data-driven emulator of the High-Resolution Rapid Refresh forecast in the U.S. (Abdi et al., 2025)), offer potential improvements for representing convection-driven extreme rainfall. HRRRcast is expected to operate at 3-km resolution, which remains coarser than the dominant spatial scales of rainfall variability (hundreds of meters to 1 km). When combined with downscaling postprocessing (Wang et al., 2025), it may enhance the accuracy of precipitation inputs for flood predictions.

2. *ML-based surrogate modeling for real-time flood forecasts:* Given the computational demands of high-resolution flood simulations, ML-based surrogate models offer a viable pathway to emulate high-fidelity physics-based models with dramatically reduced computational costs (Ivanov et al., 2021; Jones et al., 2023; Tran et al., 2023). When properly trained and validated, such surrogates can enable rapid ensemble forecasting and uncertainty quantification at operational timescales, supporting real-time decision-making within FEWS-DM.
3. *Enhanced observational data integration:* AI/ML methods can help assimilate and fuse the growing wealth of available data to continuously improve flood models (Ambika et al., 2025; Konapala et al., 2020; Lu et al., 2021; Tayal et al., 2024; Tran et al., 2025). This includes satellite observations, ground-based sensors, crowdsourced data, and Internet-of-Things (IoT) devices to improve boundary conditions, model validation, and real-time updating of flood forecasts (Dai et al., 2021; Fang et al., 2015). Effective use of these data

sources can reduce uncertainty and improve forecast relevance, particularly in data-sparse or rapidly evolving flood situations.

4. *Psychology-informed decision learning:* AI-based approaches, including reinforcement learning, provide opportunities to learn hidden objectives, constraints, and behavioral tendencies from historical decision data or structured "what-if" scenarios (Ziebart et al., 2010). Within FEWS-DM, such methods could support learning of user-specific or group-specific response patterns under uncertainty, enabling decision assistance that reflects known cognitive biases and behavioral heterogeneity.
5. *AI agent-based communication interface:* Large language models (LLMs) and AI agents provide user-friendly interfaces for interacting with complex forecast and decision-support systems (Gao et al., 2024). Specifically, conversational AI agents can serve as natural language interfaces that allow non-expert users to query probabilistic forecast outputs in plain language (e.g., "What is my flood risk at this location in the next 6 hr"? or "Which evacuation routes are safe given my vehicle type and my community's access constraints"?) and receive context-sensitive, uncertainty-aware responses. This mechanism does not require users to interpret mathematical probability distributions or hydrological model outputs directly. Recent demonstrations in related domains show LLM-based agents can effectively translate complex probabilistic outputs into actionable guidance, while maintaining transparency about uncertainty (Gao et al., 2024; Reichstein et al., 2025). The LLM component in FEWS-DM is envisioned as an interface layer, not a forecasting tool, that operates on outputs from verified physical and statistical models. Such agents can translate complex model outputs into accessible explanations, engage users in natural dialog, and provide personalized recommendations based on individual risk tolerance and objectives (Reichstein et al., 2025).

All components above can benefit from AI integration, particularly in supporting scale, consistency, and information flow. AI systems can help promote representation and procedural fairness in decision support, whereas human-driven decisions are susceptible to cognitive biases, emotional responses, and limited information processing capacity that intensify under stress or time pressure (Breznitz, 1984; Maule & Svenson, 2013). During emergencies, AI agents can assist by adhering to established protocols, integrating extensive data and transferring relevant knowledge across locations, while accounting for local context (Reichstein et al., 2025).

The beneficiaries of early warning systems should have meaningful input into system design, as this can enhance warning quality by incorporating local conditions and increase trustworthiness—an outcome linked to both system reliability and users' psychological identification with the technology (Bostrom et al., 2024). However, practical mechanisms for achieving such participation and determining the appropriate decision-making level (individual, community, district, national) remain unclear. In addition, users may still fear application of such systems depending on their digital fluency (Boyer et al., 2023; Ziker et al., 2025). Looking ahead, personalized text-based warnings could significantly improve localization and customization of alerts. When privacy and ethical concerns are adequately addressed, selective incorporation of personal data such as socioeconomic context, health information, and available infrastructure could enhance warning contextualization and user trust. This approach would make FEWS conceptually similar to existing personalized medicine systems for disease risk assessment. We envision that AI agents would support the adoption of FATES (Fairness, Accountability, Transparency, Ethics, and Sustainability) principles as essential foundations for equitable and trustworthy FEWS (Bernard & Balog, 2025; Cheong, 2024; Memarian & Doleck, 2023; Shaban-Nejad et al., 2021).

6. On the Operation of the Actionable Framework

Beyond the scientific, technical, and algorithmic obstacles outlined above, implementing effective early warning systems for diverse user-types presents significant operational complexities. These operational challenges correspond to the four foundational elements that the WMO identifies as essential for effective early warning systems: understanding risk, conducting monitoring and prediction activities, disseminating information, and building response capacity (WMO, 2022). Resolving these functional complexities can achieve equitable FEWS access to enhance risk understanding and monitoring capabilities across diverse societal contexts, while successful information delivery and responsive system design are crucial for enabling effective flood responses. Several key operational dimensions must be addressed to ensure system effectiveness and user adoption.

- *Credibility:* System reliability can promote user trust, which can be severely compromised by false positives or the lack of transparency. The "cry wolf" effect occurs when users become inert to warnings due to repeated false positives, leading them to ignore future alerts even when genuine threats emerge (LeClerc &

Joslyn, 2015). This erosion of trust is particularly problematic for FEWS, where the stakes are high and public safety depends on immediate response measures. Building and maintaining credibility requires consistent accuracy, transparent communication about uncertainties, and clear explanations of system limitations and capabilities.

- *Information security and privacy:* As personalized decision-making systems require users to submit sensitive information including personal data, real-time geolocation, travel patterns, and vulnerability assessments, robust privacy safeguards become critical. Current privacy protection frameworks may be inadequate for handling the volume and sensitivity of data required for effective personalized flood warnings. This includes concerns about data storage, transmission security, third-party access, and potential misuse of location and behavioral data. Regulatory compliance with data protection laws adds additional complexity to system design and operation.
- *Responsibility and liability:* Legal ambiguity surrounding accountability for missed warnings, false alarms, or incorrect forecasts presents significant operational challenges. Questions arise about who bears responsibility when automated systems fail: the technology developers, government agencies, or institutional operators? This uncertainty may lead to risk-averse behavior where agencies prefer conservative warnings, that is, issuing more frequent and broader alerts to avoid missing actual flood events, even at the cost of increased false positives that may reduce public trust and system effectiveness over time. Clear legal frameworks and liability structures are essential for encouraging innovation, while maintaining accountability and public safety standards.
- *Ownership, funding, and sustainability:* Long-term prediction system viability faces persistent challenges related to sustainable funding for continuous model improvement, new feature development, system upgrades, maintenance, and human resources. Many early warning systems suffer from inadequate investment in ongoing operations (Allaj & Sanfelici, 2023; Basher, 2006; Meckawy et al., 2022). Additionally, institutional fragmentation between meteorological and hydrological agencies, and emergency response organizations can create coordination challenges, duplicated efforts, and gaps in system coverage. Establishing clear ownership structures and sustainable funding mechanisms is crucial for system longevity.
- *Decision-making based on multiple systems:* The proliferation of technology inevitably leads to commercialization of flood warning and forecasting systems, creating a landscape similar to current weather prediction frameworks, in which multiple NWP models provide different forecasts (Brotzge et al., 2023; Stephens et al., 2012). This multiplication of platforms creates significant challenges when multiple real-time decision systems exist with potentially conflicting recommendations. Users may receive contradictory guidance from different platforms, leading to decision paralysis, reduced confidence in all systems, or dangerous choices based on preferred but potentially less accurate sources. This fragmentation can undermine the effectiveness of flood early warning efforts and highlights the need for standardization, quality control, and coordination mechanisms across different service providers to ensure consistent and reliable guidance for end-users.

7. Conclusions

Existing flood early warning systems have achieved remarkable progress across all components of the WMO EWS framework. Operational forecasting platforms now cover unprecedented geographic scales; IO-FEWS have shifted forecast products toward societal impacts; AI/ML has dramatically improved computational efficiency; and publicly accessible platforms have brought flood information closer to communities. These achievements are real and consequential. However, the persistence of high-casualty flood disasters in both technology-sophisticated and developing countries reflects a structural gap that forecast accuracy improvements alone cannot close: the bilateral coupling between FEWS outputs and the diverse decision needs of end-users is fundamentally incomplete. The failure modes of warning dissemination, uncertainty communication, and inadequate response implementation have been confirmed not only for floods, but also across multiple disasters, multiple continents, and multiple analytical frameworks.

We envision an *integrated physics-human flood early warning decision-making (FEWS-DM) framework* capable of interactive end-user engagement and rapid, uncertainty-aware decision support that can improve flood risk reduction relative to current FEWS. Specifically, the framework is built on three tightly coupled, co-evolving thrusts: Physical Science, advancing flood prediction with quantified uncertainty informed by decision relevance; Human Science, incorporating psychology, behavior, and cultural and institutional context into how warnings are received and acted upon; and Decision Science, unifying physical predictions and human factors through principled, utility-based decision support with end-to-end uncertainty management. Rather than treating any of these thrusts as solved or secondary, FEWS-DM positions them as equal partners in a bidirectionally coupled system. We have outlined the primary challenges in developing such systems, ranging from flood forecasting accuracy, uncertainty quantification, to decision-making under multiple uncertainty sources, from prediction modeling components to human psychology, cultural factors, and trust. Understanding these early warning constraints can inform academic research and operational innovation, helping to reduce system limitations and mitigate delayed or erroneous decisions. We also highlight the enabling role of AI in FEWS, noting its potential to augment prediction, learning, and communication when deployed within transparent, validated, and human-centered system designs.

Achieving this vision presents substantial challenges. Progress will require coordinated, cross-disciplinary efforts to address functional, operational, security, and accountability considerations, particularly given the life-or-death consequences associated with flood-related decisions. Despite the multitude of challenges, the convergence of advances in physical sciences, data availability, decision science, and computational capabilities suggests that the development of flood early warning systems tightly integrated with decision-making is increasingly within reach.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Availability Statement

The data used in Figure 2 include flood hazard map data obtained from FEMA (https://www.fema.gov/flood-maps), observed streamflow data from the U.S. Geological Survey (https://waterdata.usgs.gov/monitoring-location/04164000/), and retrospective simulation data sets from the NWM versions 2.1 and 3.0 (https://registry.opendata.aws/nwm-archive/).

Acknowledgments
Valeriy Y. Ivanov, Xun Huan, Nikola Banovic, Jeff H. Bednar, Chen Cheng, Anindya Das Antar, Francina Dominguez, Kevin Gray, Richard Gonzalez, Brian Jewett, Snehal Prabhudesai, Deffi Putri, Sudhansu Rath, Sarah H. Whitaker, and John P. Ziker acknowledge the support of the U.S. Department of Defense, Department of the Navy, the Office of Naval Research award #N00014-23-1-2735. Dan Lu and Vinh Ngoc Tran acknowledge the support of Dan Lu's Early Career Research Project, funded by the DOE Office of Biological and Environmental Research. Khachik Sargsyan acknowledges the support of the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research Earth Systems Model Development Program area of Earth and Environmental System Modeling. J. Kim was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korea government (MSIT) (RS-2022-NR070280). Phong V.V. Le was supported by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the U. S. Department of Energy. We are grateful to Dr. Emily L. Shroyer for offering valuable insights throughout the different stages of this research project.

References

Abdi, D., Jankov, I., Madden, P., Vargas, V., Smith, T. A., Frolov, S., et al. (2025). HRRRCast: A data-driven emulator for regional weather forecasting at convection allowing scales. *arXiv Preprint arXiv:2507.05658*. https://doi.org/10.1175/AIES-D-25-0061.1

Allaj, E., & Sanfelici, S. (2023). Early warning systems for identifying financial instability. *International Journal of Forecasting*, *39*(4), 1777–1803. https://doi.org/10.1016/j.ijforecast.2022.08.004

Alvarado-Valencia, J. A., & Barrero, L. H. (2014). Reliance, trust and heuristics in judgmental forecasting. *Computers in Human Behavior*, *36*, 102–113. https://doi.org/10.1016/j.chb.2014.03.047

Ambika, A. K., Tayal, K., Mishra, V., & Lu, D. (2025). Novel deep learning transformer model for short to sub-seasonal streamflow forecast. *Geophysical Research Letters*, *52*(14), e2025GL116707. https://doi.org/10.1029/2025gl116707

Arduino, G., Reggiani, P., & Todini, E. (2005). Recent advances in flood forecasting and flood risk assessment. *Hydrology and Earth System Sciences*, *9*(4), 280–284. https://doi.org/10.5194/hess-9-280-2005

Arrighi, C., Alcèrreca-Huerta, J., Oumeraci, H., & Castelli, F. (2015). Drag and lift contribution to the incipient motion of partly submerged flooded vehicles. *Journal of Fluids and Structures*, *57*, 170–184. https://doi.org/10.1016/j.jfluidstructs.2015.06.010

Arrighi, C., Oumeraci, H., & Castelli, F. (2017). Hydrodynamics of pedestrians' instability in floodwaters. *Hydrology and Earth System Sciences*, *21*(1), 515–531. https://doi.org/10.5194/hess-21-515-2017

Atran, S., Medin, D. L., & Ross, N. O. (2005). The cultural mind: Environmental decision making and cultural modeling within and across populations. *Psychological Review*, *112*(4), 744–776. https://doi.org/10.1037/0033-295x.112.4.744

Aylett, A. (2013). The socio-institutional dynamics of urban climate governance: A comparative analysis of innovation and change in Durban (KZN, South Africa) and Portland (OR, USA). *Urban Studies*, *50*(7), 1386–1402. https://doi.org/10.1177/0042098013480968

Balaian, S. K., Sanders, B. F., & Abdolhosseini Qomi, M. J. (2024). How urban form impacts flooding. *Nature Communications*, *15*(1), 6911. https://doi.org/10.1038/s41467-024-50347-4

Bao, D., George Xue, Z., Hiatt, M., Xu, K., Harris, C. K., & Trepanier, J. C. (2025). A machine learning-based prediction-to-map framework for rapid and accurate spatial flood prediction. *npj Natural Hazards*, *2*(1), 71. https://doi.org/10.1038/s44304-025-00122-2

Basher, R. (2006). Global early warning systems for natural hazards: Systematic and people-centred. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, *364*(1845), 2167–2182. https://doi.org/10.1098/rsta.2006.1819

Bauer, P., Thorpe, A., & Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature*, *525*(7567), 47–55. https://doi.org/10.1038/nature14956

Bernard, N., & Balog, K. (2025). A systematic review of fairness, accountability, transparency, and ethics in information retrieval. *ACM Computing Surveys*, *57*(6), 1–29. https://doi.org/10.1145/3637211

Berne, A., Delrieu, G., & Boudevillain, B. (2009). Variability of the spatial structure of intense Mediterranean precipitation. *Advances in Water Resources*, *32*(7), 1031–1042. https://doi.org/10.1016/j.advwatres.2008.11.008

Berne, A., Delrieu, G., Creutin, J.-D., & Obled, C. (2004). Temporal and spatial resolution of rainfall measurements required for urban hydrology. *Journal of Hydrology*, *299*(3–4), 166–179. https://doi.org/10.1016/s0022-1694(04)00363-4

Bhatt, U., Xiang, A., Sharma, S., Weller, A., Taly, A., Jia, Y., et al. (2020). Explainable machine learning in deployment, In *Proceedings of the 2020 Conference on fairness, accountability, and transparency*, (pp. 648–657). Association for Computing Machinery. https://doi.org/10.1145/3351095.3375624

Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., & Tian, Q. (2023). Accurate medium-range global weather forecasting with 3D neural networks. *Nature*, *619*(7970), 533–538. https://doi.org/10.1038/s41586-023-06185-3

Bier, V. M., & Lin, S. W. (2013). On the treatment of uncertainty and variability in making decisions about risk. *Risk Analysis*, *33*(10), 1899–1907. https://doi.org/10.1111/risa.12071

Bland, A. R., & Schaefer, A. (2012). Different varieties of uncertainty in human decision-making. *Frontiers in Neuroscience*, *6*, 85. https://doi.org/10.3389/fnins.2012.00085

Bodoque, J. M., Esteban-Muñoz, Á., & Ballesteros-Cánovas, J. A. (2023). Overlooking probabilistic mapping renders urban flood risk management inequitable. *Communications Earth & Environment*, *4*(1), 279. https://doi.org/10.1038/s43247-023-00940-0

Boelee, L., Lumbroso, D. M., Samuels, P. G., & Cloke, H. L. (2019). Estimation of uncertainty in flood Forecasts—A comparison of methods. *Journal of Flood Risk Management*, *12*(S1), e12516. https://doi.org/10.1111/jfr3.12516

Bohman, A., Glaas, E., & Karlson, M. (2020). Integrating sustainable stormwater management in urban planning: Ways forward towards institutional change and collaborative action. *Water*, *12*(1), 203. https://doi.org/10.3390/w12010203

Boholm, Å. (2003). The cultural nature of risk: Can there be an anthropology of uncertainty? *Ethnos*, *68*(2), 159–178. https://doi.org/10.1080/0014184032000097722

Bosma, C. D., Wright, D. B., Nguyen, P., Kossin, J. P., Herndon, D. C., & Shepherd, J. M. (2020). An intuitive metric to quantify and communicate tropical cyclone rainfall hazard. *Bulletin of the American Meteorological Society*, *101*(2), E206–E220. https://doi.org/10.1175/bams-d-19-0075.1

Bostrom, A., Demuth, J. L., Wirz, C. D., Cains, M. G., Schumacher, A., Madlambayan, D., et al. (2024). Trust and trustworthy artificial intelligence: A research agenda for AI in the environmental sciences. *Risk Analysis*, *44*(6), 1498–1513. https://doi.org/10.1111/risa.14245

Boyer, J. M. S., Wendell, J. A., Fails, K., House, & Ziker, J. (2023). Evolutionary perspectives on novel digital environments: Parental strategies in the ecology of fear. In *Paper presented at proceedings of the 22nd annual ACM Interaction Design and Children Conference*.

Bracken, L., Wainwright, J., Ali, G., Tetzlaff, D., Smith, M., Reaney, S., & Roy, A. (2013). Concepts of hydrological connectivity: Research approaches, pathways and future agendas. *Earth-Science Reviews*, *119*, 17–34. https://doi.org/10.1016/j.earscirev.2013.02.001

Brakenridge, G. (2025). Global active archive of large flood events. In *DFO–flood observatory*. University of Colorado.

Bremer, S., Glavovic, B., Meisch, S., Schneider, P., & Wardekker, A. (2021). Beyond rules: How institutional cultures and climate governance interact. *Wiley Interdisciplinary Reviews: Climate Change*, *12*(6), e739. https://doi.org/10.1002/wcc.739

Breznitz, S. (1984). Cry wolf: The psychology of false alarms.

Brotzge, J. A., Berchoff, D., Carlis, D. L., Carr, F. H., Carr, R. H., Gerth, J. J., et al. (2023). Challenges and opportunities in numerical weather prediction. *Bulletin of the American Meteorological Society*, *104*(3), E698–E705. https://doi.org/10.1175/bams-d-22-0172.1

Charlton-Perez, A. J., Dacre, H. F., Driscoll, S., Gray, S. L., Harvey, B., Harvey, N. J., et al. (2024). Do AI models produce better weather forecasts than physics-based models? A quantitative evaluation case study of storm Ciarán. *npj Climate and Atmospheric Science*, *7*(1), 93. https://doi.org/10.1038/s41612-024-00638-w

Chen, M., Mahto, S. S., He, X., Jun, C., Paschalis, A., Peleg, N., et al. (2025). Record-breaking rainfall: A stochastic approach for its prediction. *npj Natural Hazards*, *2*(1), 98. https://doi.org/10.1038/s44304-025-00148-6

Cheong, B. C. (2024). Transparency and accountability in AI systems: Safeguarding wellbeing in the age of algorithmic decision-making. *Frontiers in Human Dynamics*, *6*, 1421273. https://doi.org/10.3389/fhumd.2024.1421273

Cools, J., Innocenti, D., & O'Brien, S. (2016). Lessons from flood early warning systems. *Environmental Science & Policy*, *58*, 117–122. https://doi.org/10.1016/j.envsci.2016.01.006

Cosgrove, B., Gochis, D., Flowers, T., Dugger, A., Ogden, F., Graziano, T., et al. (2024). NOAA's national water model: Advancing operational hydrology through continental-scale modeling. *JAWRA Journal of the American Water Resources Association*, *60*(2), 247–272. https://doi.org/10.1111/1752-1688.13184

Dai, W., Tang, Y., Zhang, Z., & Cai, Z. (2021). Ensemble learning technology for coastal flood forecasting in internet-of-things-enabled smart city. *International Journal of Computational Intelligence Systems*, *14*(1), 166. https://doi.org/10.1007/s44196-021-00023-y

Dale, M., Wicks, J., Mylne, K., Pappenberger, F., Laeger, S., & Taylor, S. (2014). Probabilistic flood forecasting and decision-making: An innovative risk-based approach. *Natural Hazards*, *70*(1), 159–172. https://doi.org/10.1007/s11069-012-0483-z

Demir, I., & Krajewski, W. F. (2013). Towards an integrated flood information system: Centralized data access, analysis, and visualization. *Environmental Modelling and Software*, *50*, 77–84. https://doi.org/10.1016/j.envsoft.2013.08.009

De Perez, E. C., Berse, K. B., Depante, L. A. C., Easton-Calabria, E., Evidente, E. P. R., Ezike, T., et al. (2022). Learning from the past in moving to the future: Invest in communication and response to weather early warnings to reduce death and damage. *Climate Risk Management*, *38*, 100461. https://doi.org/10.1016/j.crm.2022.100461

Emerton, R., Zsoter, E., Arnal, L., Cloke, H. L., Muraro, D., Prudhomme, C., et al. (2018). Developing a global operational seasonal hydro-meteorological forecasting system: GloFAS-Seasonal v1. 0. *Geoscientific Model Development*, *11*(8), 3327–3346. https://doi.org/10.5194/gmd-11-3327-2018

Emerton, R. E., Stephens, E. M., Pappenberger, F., Pagano, T. C., Weerts, A. H., Wood, A. W., et al. (2016). Continental and global scale flood forecasting systems. *Wiley Interdisciplinary Reviews: Water*, *3*(3), 391–418. https://doi.org/10.1002/wat2.1137

Fang, S., Xu, L., Zhu, Y., Liu, Y., Liu, Z., Pei, H., et al. (2015). An integrated information system for snowmelt flood early-warning based on internet of things. *Information Systems Frontiers*, *17*(2), 321–335. https://doi.org/10.1007/s10796-013-9466-1

Fernández-Nóvoa, D., González-Cao, J., & García-Feal, O. (2024). Enhancing flood risk management: A comprehensive review on flood early warning systems with emphasis on numerical modeling. *Water*, *16*(10), 1408. https://doi.org/10.3390/w16101408

Frame, J. M., Nair, T., Sunkara, V., Popien, P., Chakrabarti, S., Anderson, T., et al. (2024). Rapid inundation mapping using the US national water model, satellite observations, and a convolutional neural network. *Geophysical Research Letters*, *51*(17), e2024GL109424. https://doi.org/10.1029/2024GL109424

Galelli, S., Turner, S. W. D., Pokhrel, Y., Yi Ng, J., Castelletti, A., Bierkens, M. F. P., et al. (2025). Advancing the representation of human actions in large-scale hydrological models: Challenges and future research directions. *Water Resources Research*, *61*(7), e2024WR039486. https://doi.org/10.1029/2024WR039486

Gao, C., Lan, X., Li, N., Yuan, Y., Ding, J., Zhou, Z., et al. (2024). Large language models empowered agent-based modeling and simulation: A survey and perspectives. *Humanities and Social Sciences Communications*, *11*(1), 1259. https://doi.org/10.1057/s41599-024-03611-3

Garcia, C. A., Savilaakso, S., Verburg, R. W., Stoudmann, N., Fernbach, P., Sloman, S. A., et al. (2022). Strategy games to improve environmental policymaking. *Nature Sustainability*, *5*(6), 464–471. https://doi.org/10.1038/s41893-022-00881-0

Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., & Kagal, L. (2018). Explaining explanations: An overview of interpretability of machine learning. In *Paper presented at 2018 IEEE 5th International Conference on data science and advanced analytics (DSAA)*.

Golding, B. (2022). *Towards the "perfect" weather warning: Bridging disciplinary gaps through partnership and communication*. Springer.

Gourley, J. J., Flamig, Z. L., Vergara, H., Kirstetter, P.-E., Clark, R. A., III., Argyle, E., et al. (2017). The FLASH project: Improving the tools for flash flood monitoring and prediction across the United States. *Bulletin of the American Meteorological Society*, *98*(2), 361–372. https://doi.org/10.1175/bams-d-15-00247.1

Guidolin, M., Chen, A. S., Ghimire, B., Keedwell, E. C., Djordjević, S., & Savić, D. A. (2016). A weighted cellular automata 2D inundation model for rapid flood analysis. *Environmental Modelling & Software*, *84*, 378–394. https://doi.org/10.1016/j.envsoft.2016.07.008

Hapuarachchi, H., Wang, Q., & Pagano, T. (2011). A review of advances in flash flood forecasting. *Hydrological Processes*, *25*(18), 2771–2784. https://doi.org/10.1002/hyp.8040

Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., et al. (2010). Markets, religion, community size, and the evolution of fairness and punishment. *Science*, *327*(5972), 1480–1484. https://doi.org/10.1126/science.1182238

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., et al. (2013). Global flood risk under climate change. *Nature Climate Change*, *3*(9), 816–821. https://doi.org/10.1038/nclimate1911

Ivanov, V. Y., Tran, V. N., Huang, W., Murphy, K., Daneshvar, F., Bednar, J. H., et al. (2024). Urban flooding is intensified by outdated design guidelines and a lack of a systems approach. *Nature Cities*, *1*(10), 626–627. https://doi.org/10.1038/s44284-024-00128-3

Ivanov, V. Y., Xu, D., Dwelle, M. C., Sargsyan, K., Wright, D. B., Katopodes, N., et al. (2021). Breaking down the computational barriers to real-time urban flood forecasting. *Geophysical Research Letters*, *48*(20), e2021GL093585. https://doi.org/10.1029/2021gl093585

Jones, A., Kuehnert, J., Fraccaro, P., Meuriot, O., Ishikawa, T., Edwards, B., et al. (2023). AI for climate impacts: Applications in flood risk. *npj Climate and Atmospheric Science*, *6*(1), 63. https://doi.org/10.1038/s41612-023-00388-1

Kaur, H., Adar, E., Gilbert, E., & Lampe, C. (2022). Sensible AI: Re-imagining interpretability and explainability using sensemaking theory. In *Paper presented at proceedings of the 2022 ACM Conference on fairness, accountability, and transparency*.

Kaur, H., Conrad, M. R., Rule, D., Lampe, C., & Gilbert, E. (2024). Interpretability gone bad: The role of bounded rationality in how practitioners understand machine learning. In *Proceedings of the ACM on human-computer interaction* (Vol. 8, pp. 1–34). https://doi.org/10.1145/3637354

Kaur, H., Nori, H., Jenkins, S., Caruana, R., Wallach, H., & Wortman Vaughan, J. (2020). Interpreting interpretability: Understanding data scientists' use of interpretability tools for machine learning. In *Paper presented at proceedings of the 2020 CHI Conference on human factors in computing systems*.

Kim, J., Warnock, A., Ivanov, V. Y., & Katopodes, N. D. (2012). Coupled modeling of hydrologic and hydrodynamic processes including overland and channel flow. *Advances in Water Resources*, *37*, 104–126. https://doi.org/10.1016/j.advwatres.2011.11.009

Knotters, M., Bokhove, O., Lamb, R., & Poortvliet, P. (2024). How to cope with uncertainty monsters in flood risk management? *Cambridge Prisms: Water*, *2*, e6. https://doi.org/10.1017/wat.2024.4

Kochkov, D., Yuval, J., Langmore, I., Norgaard, P., Smith, J., Mooers, G., et al. (2024). Neural general circulation models for weather and climate. *Nature*, *632*(8027), 1060–1066. https://doi.org/10.1038/s41586-024-07744-y

Koks, E. E., Van Ginkel, K. C., Van Marle, M. J., & Lemnitzer, A. (2022). Brief communication: Critical infrastructure impacts of the 2021 mid-July western European flood event. *Natural Hazards and Earth System Sciences*, *22*(12), 3831–3838. https://doi.org/10.5194/nhess-22-3831-2022

Konapala, G., Kao, S.-C., Painter, S. L., & Lu, D. (2020). Machine learning assisted hybrid models can improve streamflow simulation in diverse catchments across the conterminous US. *Environmental Research Letters*, *15*(10), 104022. https://doi.org/10.1088/1748-9326/aba927

Krajewski, W. F., Ceynar, D., Demir, I., Goska, R., Kruger, A., Langel, C., et al. (2017). Real-time flood forecasting and information system for the state of Iowa. *Bulletin of the American Meteorological Society*, *98*(3), 539–554. https://doi.org/10.1175/bams-d-15-00243.1

Krzhizhanovskaya, V. V., Shirshov, G., Melnikova, N. B., Belleman, R. G., Rusadi, F., Broekhuijsen, B., et al. (2011). Flood early warning system: Design, implementation and computational modules. *Procedia Computer Science*, *4*, 106–115. https://doi.org/10.1016/j.procs.2011.04.012

Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet, F., et al. (2023). Learning skillful medium-range global weather forecasting. *Science*, *382*(6677), 1416–1421. https://doi.org/10.1126/science.adi2336

Lane, A. K., Skvoretz, J., Ziker, J. P., Couch, B. A., Earl, B., Lewis, J. E., et al. (2019). Investigating how faculty social networks and peer influence relate to knowledge and use of evidence-based teaching practices. *International Journal of STEM Education*, *6*(1), 28. https://doi.org/10.1186/s40594-019-0182-3

Lawrence, M., Goodwin, P., O'Connor, M., & Önkal, D. (2006). Judgmental forecasting: A review of progress over the last 25 years. *International Journal of Forecasting*, *22*(3), 493–518. https://doi.org/10.1016/j.ijforecast.2006.03.007

LeClerc, J., & Joslyn, S. (2015). The cry wolf effect and weather-related decision making. *Risk Analysis*, *35*(3), 385–395. https://doi.org/10.1111/risa.12336

Liu, C., Guo, L., Ye, L., Zhang, S., Zhao, Y., & Song, T. (2018). A review of advances in China's flash flood early-warning system. *Natural Hazards*, *92*(2), 619–634. https://doi.org/10.1007/s11069-018-3173-7

Lu, D., Konapala, G., Painter, S. L., Kao, S.-C., & Gangrade, S. (2021). Streamflow simulation in data-scarce basins using Bayesian and physics-informed machine learning models. *Journal of Hydrometeorology*, *22*(6), 1421–1438. https://doi.org/10.1175/jhm-d-20-0082.1

Lunstrum, E., & Havice, E. (2025). Introducing jurisdiction. *Annals of the Association of American Geographers*, *115*(5), 1005–1028. https://doi.org/10.1080/24694452.2025.2467326

Maule, A. J., & Svenson, O. (2013). *Time pressure and stress in human judgment and decision making*. Springer Science & Business Media.

Meckawy, R., Stuckler, D., Mehta, A., Al-Ahdal, T., & Doebbeling, B. N. (2022). Effectiveness of early warning systems in the detection of infectious diseases outbreaks: A systematic review. *BMC Public Health*, *22*(1), 2216. https://doi.org/10.1186/s12889-022-14625-4

Meierdiercks, K. L., Smith, J. A., Baeck, M. L., & Miller, A. J. (2010). Heterogeneity of Hydrologic response in urban watersheds 1. *JAWRA Journal of the American Water Resources Association*, *46*(6), 1221–1237. https://doi.org/10.1111/j.1752-1688.2010.00487.x

Memarian, B., & Doleck, T. (2023). Fairness, accountability, transparency, and ethics (FATE) in artificial intelligence (AI) and higher education: A systematic review. *Computers and Education: Artificial Intelligence*, *5*, 100152. https://doi.org/10.1016/j.caeai.2023.100152

Merz, B., Kuhlicke, C., Kunz, M., Pittore, M., Babeyko, A., Bresch, D. N., et al. (2020). Impact forecasting to support emergency management of natural hazards. *Reviews of Geophysics*, *58*(4), e2020RG000704. https://doi.org/10.1029/2020rg000704

Molan, S., Weber, D., & Kor, M. (2022). Why can't they just evacuate early'–a study of the effect of a virtual reality experience on residents' intentions related to a wildfire risk. *International Journal of Disaster Risk Reduction*, *81*, 103268. https://doi.org/10.1016/j.ijdrr.2022.103268

Morss, R. E., Mulder, K. J., Lazo, J. K., & Demuth, J. L. (2016). How do people perceive, understand, and anticipate responding to flash flood risks and warnings? Results from a public survey in Boulder, Colorado, USA. *Journal of Hydrology*, *541*, 649–664. https://doi.org/10.1016/j.jhydrol.2015.11.047

Najafi, H., Shrestha, P. K., Rakovec, O., Apel, H., Vorogushyn, S., Kumar, R., et al. (2024). High-resolution impact-based early warning system for riverine flooding. *Nature Communications*, *15*(1), 3726. https://doi.org/10.1038/s41467-024-48065-y

Nearing, G., Cohen, D., Dube, V., Gauch, M., Gilon, O., Harrigan, S., et al. (2024). Global prediction of extreme floods in ungauged watersheds. *Nature*, *627*(8004), 559–563. https://doi.org/10.1038/s41586-024-07145-1

Nelson-Mercer, B., Kim, T., Tran, V. N., & Ivanov, V. (2025). Pluvial flood impacts and policyholder responses throughout the United States. *npj Natural Hazards*, *2*(1), 8. https://doi.org/10.1038/s44304-025-00058-7

Nevo, S., Morin, E., Gerzi Rosenthal, A., Metzger, A., Barshai, C., Weitzner, D., et al. (2022). Flood forecasting with machine learning models in an operational framework. *Hydrology and Earth System Sciences*, *26*(15), 4013–4032. https://doi.org/10.5194/hess-26-4013-2022

Nobre, A. D., Cuartas, L. A., Hodnett, M., Rennó, C. D., Rodrigues, G., Silveira, A., & Saleska, S. (2011). Height above the nearest drainage–a hydrologically relevant new terrain model. *Journal of Hydrology*, *404*(1–2), 13–29. https://doi.org/10.1016/j.jhydrol.2011.03.051

Painter, P., Semmens, K., Maxfield, K., Cattoën, C., & Carr, R. H. (2025). Designing effective flood early warning systems: A review of barriers, best practices, and key characteristics. *Journal of Flood Risk Management*, *18*(4), e70145. https://doi.org/10.1111/jfr3.70145

Pappenberger, F., Cloke, H. L., Parker, D. J., Wetterhall, F., Richardson, D. S., & Thielen, J. (2015). The monetary benefit of early flood warnings in Europe. *Environmental Science & Policy*, *51*, 278–291. https://doi.org/10.1016/j.envsci.2015.04.016

Paprotny, D., Sebastian, A., Morales-Nápoles, O., & Jonkman, S. N. (2018). Trends in flood losses in Europe over the past 150 years. *Nature Communications*, *9*(1), 1985. https://doi.org/10.1038/s41467-018-04253-1

Pasche, O. C., Wider, J., Zhang, Z., Zscheischler, J., & Engelke, S. (2025). Validating deep learning weather forecast models on recent high-impact extreme events. *Artificial Intelligence for the Earth Systems*, *4*(1), e240033. https://doi.org/10.1175/aies-d-24-0033.1

Paul Zischg, A. (2024). From flood impact modelling to flood impact forecasts. *Journal of Flood Risk Management*, *17*(2), e12995. https://doi.org/10.1111/jfr3.12995

Pearson, L., & Pelling, M. (2015). The UN Sendai framework for disaster risk reduction 2015–2030: Negotiation process and prospects for science and practice. *Journal of Extreme Events*, *2*(1), 1571001. https://doi.org/10.1142/s2345737615710013

Peleg, N., Blumensaat, F., Molnar, P., Fatichi, S., & Burlando, P. (2017). Partitioning the impacts of spatial and climatological rainfall variability in urban drainage modeling. *Hydrology and Earth System Sciences*, *21*(3), 1559–1572. https://doi.org/10.5194/hess-21-1559-2017

Perera, D., Agnihotri, J., Seidou, O., & Djalante, R. (2020). Identifying societal challenges in flood early warning systems. *International Journal of Disaster Risk Reduction*, *51*, 101794. https://doi.org/10.1016/j.ijdrr.2020.101794

Potter, S. H., Kox, T., Mills, B., Taylor, A., Robbins, J., Cerrudo, C., et al. (2025). Research gaps and challenges for impact-based forecasts and warnings: Results of international workshops for high impact weather in 2022. *International Journal of Disaster Risk Reduction*, *118*, 105234. https://doi.org/10.1016/j.ijdrr.2025.105234

Prabhudesai, S., Yang, L., Asthana, S., Huan, X., Liao, Q. V., & Banovic, N. (2023). Understanding uncertainty: How lay decision-makers perceive and interpret uncertainty in human-AI decision making. In *Paper presented at proceedings of the 28th International Conference on intelligent user interfaces*.

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

Radford, J. T., Ebert-Uphoff, I., & Stewart, J. Q. (2025). A comparison of AI weather prediction and numerical weather prediction models for 1–7-day precipitation forecasts. *Weather and Forecasting*, *40*(4), 561–575. https://doi.org/10.1175/waf-d-24-0081.1

Ras, G., van Gerven, M., & Haselager, P. (2018). Explanation methods in deep learning: Users, values, concerns and challenges. In H. J. Escalante, S. Escalera, I. Guyon, X. Baró, Y. Güçlütürk, U. Güçlü, et al. (Eds.), *Explainable and interpretable models in computer vision and machine learning* (pp. 19–36). Springer International Publishing. https://doi.org/10.1007/978-3-319-98131-4_2

Reichstein, M., Benson, V., Blunk, J., Camps-Valls, G., Creutzig, F., Fearnley, C. J., et al. (2025). Early warning of complex climate risk with integrated artificial intelligence. *Nature Communications*, *16*(1), 2564. https://doi.org/10.1038/s41467-025-57640-w

Rentschler, J., Salhab, M., & Jafino, B. A. (2022). Flood exposure and poverty in 188 countries. *Nature Communications*, *13*(1), 3527. https://doi.org/10.1038/s41467-022-30727-4

Robinson, A., Lehmann, J., Barriopedro, D., Rahmstorf, S., & Coumou, D. (2021). Increasing heat and rainfall extremes now far outside the historical climate. *npj Climate and Atmospheric Science*, *4*(1), 45. https://doi.org/10.1038/s41612-021-00202-w

Rogers, D., & Tsirkunov, V. (2011). Costs and benefits of early warning systems. *Global Assessment Rep*. http://documents.worldbank.org/curated/en/609951468330279598

Rossi, C., Favenza, A., Scullino, F., Macchia, V., Spoto, G. L., & Dominci, F. (2015). Evaluating FLOODIS: Mobile sensing for a flood emergency service in the cloud. In *Paper presented at 2015 International Conference on cloud technologies and applications (CloudTech)*. IEEE.

Sanders, B. F., Schubert, J. E., Goodrich, K. A., Houston, D., Feldman, D. L., Basolo, V., et al. (2020). Collaborative modeling with fine-resolution data enhances flood awareness, minimizes differences in flood perception, and produces actionable flood maps. *Earth's Future*, *8*(1), e2019EF001391. https://doi.org/10.1029/2019ef001391

Scolobig, A., De Marchi, B., & Borga, M. (2012). The missing link between flood risk awareness and preparedness: Findings from case studies in an Alpine region. *Natural Hazards*, *63*(2), 499–520. https://doi.org/10.1007/s11069-012-0161-1

Sempere-Torres, D., & Berenguer, M. (2024). The ANYWHERE paradigm shift in responding to weather and climate emergencies: Impact forecasting, dynamic vulnerability and the need for citizen's involvement. *Responding to Extreme Weather Events*, 1–22.

Settembre-Blundo, D., González-Sánchez, R., Medina-Salgado, S., & García-Muiña, F. E. (2021). Flexibility and resilience in corporate decision making: A new sustainability-based risk management system in uncertain times. *Global Journal of Flexible Systems Management*, *22*(Suppl 2), 107–132. https://doi.org/10.1007/s40171-021-00277-7

Shaban-Nejad, A., Michalowski, M., Brownstein, J. S., & Buckeridge, D. L. (2021). Guest editorial explainable AI: Towards fairness, accountability, transparency and trust in healthcare. *IEEE Journal of Biomedical and Health Informatics*, *25*(7), 2374–2375. https://doi.org/10.1109/jbhi.2021.3088832

Simoes-Sousa, I. T., Camargo, C. M., Tavora, J., Piffer-Braga, A., Farrar, J. T., & Pavelsky, T. M. (2025). The May 2024 flood disaster in southern Brazil: Causes, impacts, and SWOT-based volume estimation. *Geophysical Research Letters*, *52*(4), e2024GL112442. https://doi.org/10.1029/2024gl112442

Slack, D., Hilgard, S., Singh, S., & Lakkaraju, H. (2021). Reliable post hoc explanations: Modeling uncertainty in explainability, In *Proceedings of the 35th International Conference on neural information processing systems*, (p. 719). Curran Associates Inc.

Smith, P., Pappenberger, F., Wetterhall, F., Del Pozo, J. T., Krzeminski, B., Salamon, P., et al. (2016). On the operational implementation of the European flood awareness system (EFAS), In *Flood forecasting*, (pp. 313–348). Elsevier.

Solinska-Nowak, A., Magnuszewski, P., Curl, M., French, A., Keating, A., Mochizuki, J., et al. (2018). An overview of serious games for disaster risk management–prospects and limitations for informing actions to arrest increasing risk. *International Journal of Disaster Risk Reduction*, *31*, 1013–1029. https://doi.org/10.1016/j.ijdrr.2018.09.001

Stephens, E. M., Edwards, T. L., & Demeritt, D. (2012). Communicating probabilistic information from climate model ensembles—lessons from numerical weather prediction. *WIREs Climate Change*, *3*(5), 409–426. https://doi.org/10.1002/wcc.187

Stuart, N. A., Hartfield, G., Schultz, D. M., Wilson, K., West, G., Hoffman, R., et al. (2022). The evolving role of humans in weather prediction and communication. *Bulletin of the American Meteorological Society*, *103*(8), E1720–E1746. https://doi.org/10.1175/bams-d-20-0326.1

Sun, Y. Q., Hassanzadeh, P., Zand, M., Chattopadhyay, A., Weare, J., & Abbot, D. S. (2025). Can AI weather models predict out-of-distribution gray swan tropical cyclones? *Proceedings of the National Academy of Sciences*, *122*(21), e2420914122. https://doi.org/10.1073/pnas.2420914122

Tayal, K., Renganathan, A., & Lu, D. (2024). Improving streamflow predictions across CONUS by integrating advanced machine learning models and diverse data. *Environmental Research Letters*, *19*(10), 104009. https://doi.org/10.1088/1748-9326/ad6fb7

Thieken, A. H., Bubeck, P., Heidenreich, A., Von Keyserlingk, J., Dillenardt, L., & Otto, A. (2023). Performance of the flood warning system in Germany in July 2021–insights from affected residents. *Natural Hazards and Earth System Sciences*, *23*(2), 973–990. https://doi.org/10.5194/nhess-23-973-2023

Tran, V. N., Dwelle, M. C., Sargsyan, K., Ivanov, V. Y., & Kim, J. (2020). A novel modeling framework for computationally efficient and accurate real-time ensemble flood forecasting with uncertainty quantification. *Water Resources Research*, *56*(3), e2019WR025727. https://doi.org/10.1029/2019wr025727

Tran, V. N., Ivanov, V. Y., Huang, W., Murphy, K., Daneshvar, F., Bednar, J. H., et al. (2024). Connectivity in urbanscapes can cause unintended flood impacts from stormwater systems. *Nature Cities*, *1*(10), 654–664. https://doi.org/10.1038/s44284-024-00116-7

Tran, V. N., Ivanov, V. Y., Xu, D., & Kim, J. (2023). Closing in on hydrologic predictive accuracy: Combining the strengths of high-fidelity and physics-agnostic models. *Geophysical Research Letters*, *50*(17), e2023GL104464. https://doi.org/10.1029/2023GL104464

Tran, V. N., Kim, T., Xu, D., Tran, H., Le, M., Tran, T., et al. (2025). AI improves the accuracy, reliability, and economic value of continental-scale flood predictions. *AGU Advances*, *6*(3), e2025AV001678. https://doi.org/10.1029/2025AV001678

Tran, V. N., Xu, D., Le, P., Kim, J., Nguyen, T. V., Nguyen, G. T., et al. (2026). The value of forecasters-in-the-loop in real-time flood forecasting in the age of machine learning. *Geophysical Research Letters*, *53*(8), e2025GL118317. https://doi.org/10.1029/2025GL118317

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458. https://doi.org/10.1126/science.7455683

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323. https://doi.org/10.1007/bf00122574

UNDRR. (2023). *Global status of multi-hazard early warning systems 2023*. Stylus Publishing, LLC.

Van Houtven, G. (2024). Economic value of flood forecasts and early warning systems: A review. *Natural Hazards Review*, *25*(4), 03124002. https://doi.org/10.1061/nhrefo.nheng-2094

Wainwright, J., Turnbull, L., Ibrahim, T. G., Lexartza-Artza, I., Thornton, S. F., & Brazier, R. E. (2011). Linking environmental régimes, space and time: Interpretations of structural and functional connectivity. *Geomorphology*, *126*(3–4), 387–404. https://doi.org/10.1016/j.geomorph.2010.07.027

Wang, D., Zhang, W., & Lim, B. Y. (2021). Show or suppress? Managing input uncertainty in machine learning model explanations. *Artificial Intelligence*, *294*, 103456. https://doi.org/10.1016/j.artint.2021.103456

Wang, X., Choi, J.-Y., Kurihaya, T., Lyngaas, I., Yoon, H.-J., Xiao, X., et al. (2025). ORBIT-2: Scaling exascale vision foundation models for weather and climate downscaling. In *Paper presented at proceedings of the International Conference for high performance computing, networking, storage and analysis*.

Wang, Z., Huan, X., & Garikipati, K. (2019). Variational system identification of the partial differential equations governing the physics of pattern-formation: Inference under varying fidelity and noise. *Computer Methods in Applied Mechanics and Engineering*, *356*, 44–74. https://doi.org/10.1016/j.cma.2019.07.007

Weber, E., & Hsee, C. (2000). Culture and individual judgment and decision making. *Applied Psychology*, *49*(1), 32–61. https://doi.org/10.1111/1464-0597.00005

Weber de Melo, W., Iglesias, I., & Pinho, J. (2025). Early warning system for floods at estuarine areas: Combining artificial intelligence with process-based models. *Natural Hazards*, *121*(4), 4615–4638. https://doi.org/10.1007/s11069-024-06957-8

Werner, M., Schellekens, J., Gijsbers, P., van Dijk, M., van den Akker, O., & Heynert, K. (2013). The Delft-FEWS flow forecasting system. *Environmental Modelling & Software*, *40*, 65–77. https://doi.org/10.1016/j.envsoft.2012.07.010

WHO. (2014). Global report on drowning: Preventing a leading killer. In *Global report on drowning: Preventing a leading killer*.

Winsemius, H. C., Aerts, J. C., Van Beek, L. P., Bierkens, M. F., Bouwman, A., Jongman, B., et al. (2016). Global drivers of future river flood risk. *Nature Climate Change*, *6*(4), 381–385. https://doi.org/10.1038/nclimate2893

WMO. (2013). *Climate 2001–2010: A decade of climate Extremes—Summary report*. World Meteorological Organization. WMO no. 1119.

WMO. (2015). *WMO guidelines on multi-hazard impact-based forecast and warning services* (Vol. 1150, p. 34). WMO Doc.

WMO. (2022). *Early warnings for all: Executive action plan 2023–2027*Rep. World Meteorological Organization.

Wright, D. B., Smith, J. A., Villarini, G., & Baeck, M. L. (2013). Estimating the frequency of extreme rainfall using weather radar and stochastic storm transposition. *Journal of Hydrology*, *488*, 150–165. https://doi.org/10.1016/j.jhydrol.2013.03.003

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

Yano, J.-I., Ziemiański, M. Z., Cullen, M., Termonia, P., Onvlee, J., Bengtsson, L., et al. (2018). Scientific challenges of convective-scale numerical weather prediction. *Bulletin of the American Meteorological Society*, *99*(4), 699–710. https://doi.org/10.1175/bams-d-17-0125.1

Yoe, C. (2019). *Principles of risk analysis: Decision making under uncertainty*. CRC Press.

Zang, Y., Meng, Y., Guan, X., Lv, H., & Yan, D. (2022). Study on urban flood early warning system considering flood loss. *International Journal of Disaster Risk Reduction*, *77*, 103042. https://doi.org/10.1016/j.ijdrr.2022.103042

Zhang, Y., Liao, Q. V., & Bellamy, R. K. E. (2020). Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making, In *Proceedings of the 2020 Conference on fairness, accountability, and transparency*, (pp. 295–305). Association for Computing Machinery. https://doi.org/10.1145/3351095.3372852

Zhang, Z., Fischer, E., Zscheischler, J., & Engelke, S. (2025). Numerical models outperform AI weather forecasts of record-breaking extremes. *arXiv Preprint arXiv:2508.15724*. https://doi.org/10.48550/arXiv.2508.15724

Zhao, Y., Wu, X., Zhang, W., Lan, P., Qin, G., Li, X., & Li, H. (2025). A deep learning-based probabilistic approach to flash flood warnings in mountainous catchments. *Journal of Hydrology*, *652*, 132677. https://doi.org/10.1016/j.jhydrol.2025.132677

10.1029/2026EF008857

Ziebart, B. D., Bagnell, J. A., & Dey, A. K. (2010). Modeling interaction via the principle of maximum causal entropy, In *Proceedings of the 27th International Conference on machine learning*, (pp. 1255–1262). Omnipress.

Ziker, J. P. (2007). Subsistence and food sharing in northern Siberia: Social and nutritional ecology of the Dolgan and the Nganasan. *Ecology of Food and Nutrition*, *46*(5–6), 445–467. https://doi.org/10.1080/03670240701486743

Ziker, J. P., Fails, J. A., House, K., Boyer, J., Wendell, M., Abele, H., et al. (2025). Parent–child adaptive responses for digital resilience. *Social Sciences*, *14*(4), 197. https://doi.org/10.3390/socsci14040197