Uncertainty-Aware Pluvial Flood Forecasting and Decision Support in Data Scarce Urban Areas

Adele Therese Young

UNCERTAINTY-AWARE PLUVIAL FLOOD FORECASTING AND DECISION SUPPORT IN DATA SCARCE URBAN AREAS

Adele Therese Young

UNCERTAINTY-AWARE PLUVIAL FLOOD FORECASTING AND DECISION SUPPORT IN DATA SCARCE URBAN AREAS

DISSERTATION

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus Prof.dr.ir. H. Bijl,
chair of the Board for Doctorates
and
in fulfilment of the requirement of the Vice Rector of IHE Delft
Institute for Water Education, Prof.dr. G.P.W. Jewitt,
to be defended in public on
Thursday, 23 April 2026 at 15:00 hours

by

Adele Therese YOUNG

This dissertation has been approved by the (co)promotor.

Composition of the doctoral committee:

Rector Magnificus TU Delft	chairperson
Vice Rector IHE Delft	vice-chairperson
Em.Prof.dr.ir. C. Zevenbergen	TU Delft / IHE Delft, promotor
Dr. B. Bhattacharya	IHE Delft, copromotor
Independent members:	
Prof.dr.ir. R. Uijlenhoet	TU Delft
Prof.dr. W.A.A.I. El Barki	University of Alexandria, Egypt
Prof.dr.ir. B. Kolen	UvA
Dr. M. Ramos	INRAE / Sorbonne University, France
Prof.dr. I.I. Popescu	TU Delft / IHE Delft, reserve member

This research was conducted under the auspices of the Graduate School for Socio-Economic and Natural Sciences of the Environment (SENSE)

Cover design by Alicia Conner (colour alteration and digital painting). Original photo, flooding in Alexandria by ZUMA Press used with permission.

Published by IHE Delft Institute for Water Education
www.un-ihe.org
ISBN 978-90-73445-79-6

"*The child is not perfect, but the child is yours*".

To parents who support and love unconditionally.

"*If a (wo)man will begin with certainties, (s)he shall end in doubts; but if (s)he will be content to begin with doubts (s)he shall end in certainties*". - Sir Francis Bacon

ACKNOWLEDGEMENTS

"*Supervision is an opportunity to bring someone back to their own mind, to show them how good they can be*" — Nancy Kline

I would like to thank my supervisory team for the continued guidance through my PhD journey. Prof. dr. Chris Zevenbergen and Dr. Biswa Bhattacharya, we are a well-balanced team: the eternal optimist, the strategic optimist, and the rational optimist. Chris, thank you for creating a space for me to explore ideas while always keeping the big picture in focus. Throughout the PhD, you have also provided me with opportunities to explore activities beyond my research, which has opened new doors. Biswa, throughout the PhD and until the very end, you have supported and encouraged me to proclaim my successes and to be a "good harvester". A skill I am admittedly still working on. Without your encouragement, I ~~definitely~~ would have never pursued a PhD. I am grateful for all our discussions, your time and patience. Dr. Emma Daniëls, my external supervisor, I am extremely thankful for your guidance with the WRF modelling, continued enthusiasm, encouraging words and for being my "writing motivation".

"Start by doing what is necessary; then do what is possible; and suddenly you are doing the impossible"— *St. Francis of Assisi*

This research would not have been possible without the support of others. Thanks to Colis Allen, for my introduction to WRF modelling. I would also like to thank Dr. Gerald Corzo for facilitating my access to SURF's Snellius HPC and to Dr. Schalk Jan van Andel for assisting with access to data and taking the time to answer my "hallway" questions. I am extremely grateful to Dr. Mohamed Hasan Khalil, the staff at the Egyptian Meteorological Agency and the Alexandria Sanitary and Drainage Company, who provided invaluable insights. I would also like to thank the "Omars" for graciously answering all my questions related to Egypt and Alexandria. This research also would not have been possible without Faisal Mahmood whose MSc research contributed to this work.

They say it takes a village, in this case it took continents.

One of the greatest joys of doing a PhD has been meeting so many kindred spirits; my IHE family. A special thank you to Milk, Claudia, Thaine, Ana Maria, Alyssa and Yared who have been with me through professional and personal triumphs and lows. Thank you, my amazing friends and colleagues, Bich, Balbina and Adam (ABBA), your reflections and feedback have been invaluable. To other friends and colleagues, Eliana, Gaby, Enya, Celia, Omar, Sumi, Cristiane, Deborah, Eliza, Ingrid, Noor, Haris, Henry, Krishna, Vanida, Bota, Coti, Khin, Tanya, Mercy, Janvière, Dikman, Ha, Mila and Irene. All work and no play makes the PhD very boring. Thank you for all the memories and the laughs. I will truly miss all our lunch discussions and debates. A special thank you to the extended

IHE family, Alex, Juliette, Stefan, Thomas and Jhon. I will always be indebted to my EWSYP colleagues, my best forecasting friend (BFF) Erika, Lydia and Niko.

I would also like to thank the Hydro informatics staff (my adopted department) Leo, Andreja and Arlex who have also entertained my questions regarding the decision modelling and urban flood modelling. The CURR and SIDS Research group colleagues Assela, Lucy, Willam, Akosua and other IHE staff Alessandro, Micha, Francesco, Michael and also Amit, for my brief introduction to decolonising science. I am extremely grateful to other IHE staff, Floor, Jolanda and Anique, the IT department, Erwin, for taking the time to chat on evenings and so many others at IHE who have been along my journey in one way or another.

I would like to thank Fazir, for putting me on this professional path and supporting my decision to pursue the MSc and PhD. I would also like to thank Catalina, Liz, Renate, Irene, Mary Anne, Dorothy, Erin and others from the Red Cross Red Crescent Climate Centre, for giving me the space to grow and the opportunity to cultivate my interest outside of my PhD. Big up to my Caribbean sisters Erin, Danneille, Sharoma, Uwe, Oclaya and Kambiri for your friendship and support. We're one step closer to "Tryna take over the world". So many people have come into my life during this PhD and it would be impossible to name them all, so thank you thank you thank you!

"I sustain myself with the love of family." — Maya Angelou

The greatest sacrifice is being away from the ones you love. To Uncle Peter and Momma. COVID was a difficult period for many, but I am appreciative of the time during the PhD we spent together. Your prayers, love and support have always and continue to be my greatest encouragement. To my siblings, Alicia (and personal graphic designer), Andrea and My Jono and cousin Karla, we are spread across continents but thank you for your continuous support and words of encouragement. Thank you to Christopher, Jeff, Ally and Jason for your support. To Jadey, Eli, Philip, Charlie and Jax, you probably have no idea what Aunty Adele has been doing but thanks for bringing small joys to my life. To my forever friends, Danielle, Hannah and Asa for reminding me that friendship can stand the test of time. Thank you to all the Aunties, Uncles and cousins for your continued support. We finally have a doctor in the family, yes a "real" one.

"A good husband makes a good wife and a good wife makes a good husband and…."

Finally, to my life partner, Justin. There was life before you came to the Netherlands and life after you came to the Netherlands. I warned you about life with a final year(s) PhD, a challenge you openly accepted. As you have always said, "Once there is a will, there is a way". You too have made sacrifices to be here with me and I am beyond grateful to you and your family. I am extremely grateful for your patience, your relentless optimism, your hugs, and the wonderful meals. Thank you for entertaining all my questions. You have been here since the beginning of the PhD. Next, I look forward to life with you after the PhD.

SUMMARY

Pluvial flooding is on the rise as more cities are challenged by a changing climate and local drivers: increased urbanisation and inadequate sewer systems. Sustainable flood risk management requires a hybrid of structural and non-structural measures to ensure water-hazard resilient cities. In this regard, flood forecasting and early warning systems have been proposed as a "low regret" measure to reduce flooding and increase preparedness through forecast-based actions. However, many cities, especially in data-scarce regions, lack the capacity to produce high-quality rainfall forecasts and well-calibrated flood predictions (including timing, water levels, extent and impact). This limitation creates a cascading effect, hindering the ability to make reliable decisions due to uncertainties in the forecast or inaccuracies in the input data. For example, decisions in anticipatory flood management become problematic due to their dependencies on knowledge derived from uncertain data and the consequences of incorrect predictions and/or actions.

Previous studies have focused on improving the accuracy of predictions or models, but there has been comparatively less focus on addressing the complexities of flood forecasts and the decision support chain, particularly under conditions of limited data and uncertainty. Probabilistic forecast knowledge is beneficial for quantifying uncertainty and has been acknowledged to support decision-making, but there is no consensus on the most suitable and effective way to incorporate it for decision-making.

This thesis aims to explore how pluvial flood forecasting and decision support at local scales in data-scarce cities can be improved, with a focus on using available data, method suitability and incorporating uncertainty into the decision-making process for anticipatory flood action. This research was motivated by practitioners' need to take action ahead of time, despite limited and uncertain data, which is so often the case in cities in the Global South. The research was carried out in the coastal city of Alexandria, Egypt, which experiences annual flooding due to winter storms.

The research adopted an incremental and sequential approach, in which each research outcome is built upon in the next step and used as input for the next objective. Thereby integrating the data, models and knowledge. First, the research investigated how available data can be combined for flood forecasting using a rainfall threshold method, crowd-sourced information and global ensemble forecasts. Next, the study evaluated the suitability and limitations of using a downscaled high-resolution Weather Research and Forecasting model for deterministic rainfall forecasts at multiple spatial scales for different urban flood forecasting approaches. A high-resolution limited area ensemble rainfall forecasting model was then configured and simulated. The resulting probabilistic distributions were incorporated into a Bayesian decision framework to translate probabilistic rainfall forecasts into actionable information.

Findings show that, despite limited data, using rainfall forecasts alongside non-traditional data sources can effectively predict flooding and improve decision-making. Although this method introduces some complexity compared to widely used deterministic rainfall forecasts, it offers a balance between a simplified approach and forecast uncertainty when limited data exist.

Regarding the use of higher-resolution forecasts, the results showed that, despite the need for very high-resolution forecasts for urban flood forecasting, there are limits when using downscaled models, as outcomes vary depending on the model configuration, parameterisation and the flood forecast approach. These findings highlight the need to consider the interdependencies between meteorological forecasts and their application to flood forecasting.

Finally, the thesis demonstrated the value of using a Bayesian decision framework for incorporating knowledge of flood vulnerability and local flooding at the neighbourhood scale for triggering equitable flood early actions. The proposed probabilistic framework showed the potential benefits of a user-preference parametric loss function that assigns weights to different vulnerability classes. The Bayesian decision framework served as a valuable decision support tool which incorporates uncertainty, decision makers' risk-averse attitudes and prioritises equitable anticipatory flood actions to reduce flood damage.

The primary aim of this research was not to make forecasts more accurate but rather to develop a greater understanding of the usefulness of imperfect data, forecasts and models in decision-making for pluvial flood forecasting and local anticipatory flood management. The research raises themes of fit for purpose data and creating value from imperfect information by examining the use and limitations of data and the interdependencies of the flood forecasts and decision-making chain. This approach supports uncertainty-aware flood forecasting and decision-making instead of creating an illusion of certainty.

SAMENVATTING

Overstromingen door neerslag nemen toe nu steeds meer steden worden geconfronteerd met de gevolgen van klimaatverandering en lokale oorzaken, zoals toenemende verstedelijking en een inadequaat rioleringssysteem. Duurzaam overstromingsrisicobeheer vereist een combinatie van structurele en niet-structurele maatregelen om steden veerkrachtig te maken tegen watergerelateerde gevaren. In dit verband worden overstromingsvoorspellings- en waarschuwingssystemen beschouwd als een “low regret”-maatregel om overstromingen te verminderen en de paraatheid te vergroten door middel van op voorspellingen gebaseerde acties.

Veel steden, met name in data-arme regio’s, beschikken echter niet over de capaciteit om betrouwbare neerslagvoorspellingen en goed gekalibreerde overstromingsvoorspellingen (met betrekking tot timing, waterstanden, omvang en impact) te produceren. Deze beperking veroorzaakt een kettingreactie, waardoor het moeilijk wordt om betrouwbare beslissingen te nemen vanwege onzekerheden in de voorspellingen of onnauwkeurigheden in de invoergegevens. Besluitvorming binnen anticiperend overstromingsbeheer wordt hierdoor problematisch, aangezien zij afhankelijk is van kennis die gebaseerd is op onzekere data en de mogelijke gevolgen van foutieve voorspellingen en/of acties.

Eerder onderzoek richtte zich voornamelijk op het verbeteren van de nauwkeurigheid van voorspellingen of modellen, terwijl er relatief weinig aandacht is besteed aan de complexiteit van overstromingsvoorspellingen en de bijbehorende besluitvormingsketen — vooral onder omstandigheden van beperkte data en onzekerheid. Probabilistische (kans)voorspellingen zijn waardevol voor het kwantificeren van onzekerheid en worden erkend als steun bij besluitvorming, maar er bestaat geen consensus over de meest geschikte en effectieve manier om deze te integreren in het besluitvormingsproces.

Dit proefschrift heeft tot doel te onderzoeken hoe pluviale overstromingsvoorspellingen en besluitvorming op lokale schaal in data-arme steden kunnen worden verbeterd. Daarbij ligt de nadruk op het gebruik van beschikbare data, de geschiktheid van methoden, en het integreren van onzekerheid in het besluitvormingsproces voor anticiperende overstromingsacties. Het onderzoek werd gemotiveerd door de behoefte van praktijkdeskundigen om tijdig actie te ondernemen, ondanks beperkte en onzekere data — een situatie die vaak voorkomt in steden in het mondiale zuiden. De kuststad Alexandrië in Egypte diende als casestudy.

Het onderzoek volgde een incrementele en sequentiële aanpak waarbij elk onderzoeksresultaat voortbouwt op het voorgaande en dient als input voor de volgende stap. Zo werden data, modellen en kennis geïntegreerd. Eerst werd onderzocht hoe

beschikbare data kunnen worden gecombineerd voor overstromingsvoorspellingen met behulp van een neerslagdrempelmethode, informatie van burgers (crowdsourcing) en wereldwijde ensemblevoorspellingen. Vervolgens werd de geschiktheid en beperkingen geëvalueerd van het gebruik van een downscaled, hoge-resolutie Weather Research and Forecasting (WRF)-model voor deterministische neerslagvoorspellingen op verschillende ruimtelijke schalen, toegepast in verschillende stedelijke overstromingsvoorspellingsbenaderingen. Daarna werd een hoge-resolutie ensemble-neerslagvoorspellingsmodel ontwikkeld, waarvan de resulterende probabilistische verdelingen werden geïntegreerd in een Bayesiaans beslissingskader om probabilistische neerslagvoorspellingen te vertalen naar bruikbare informatie.

Het onderzoek, uitgevoerd in de kuststad Alexandrië (Egypte), waar overstromingen voornamelijk in de winter voorkomen, toonde aan dat ondanks beperkte data, het combineren van neerslagvoorspellingen met niet-traditionele databronnen effectief kan worden gebruikt om overstromingen te voorspellen en besluitvorming te verbeteren. Hoewel deze methode meer complexiteit introduceert dan de veelgebruikte deterministische voorspellingen, biedt zij een balans tussen een vereenvoudigde aanpak en het inschatten van onzekerheid wanneer data beperkt zijn.

Wat betreft het gebruik van hoge-resolutievoorspellingen toonden de resultaten aan dat, hoewel zeer gedetailleerde voorspellingen belangrijk zijn voor stedelijke overstromingsmodellen, er grenzen zijn aan het gebruik van downscaled modellen. De resultaten variëren bovendien afhankelijk van de configuratie van het neerslagvoorspellingsmodel, de parameterisatie van convectie op kleine schalen en de gekozen overstromingsvoorspellingsaanpak. Dit benadrukt de noodzaak om rekening te houden met de onderlinge afhankelijkheden tussen meteorologische voorspellingen en hun toepassing bij overstromingsvoorspellingen.

Tot slot toonde het proefschrift de waarde aan van het gebruik van een Bayesiaans beslissingskader waarin ensemblevoorspellingen van overstromingen, lokale kennis op buurtniveau en een gebruikersspecifieke parametrische verliesfunctie worden geïntegreerd, die verschillende kwetsbaarheidsklassen gewichten toekent. De voorgestelde methodologie fungeert als een waardevol besluitvormingsinstrument dat onzekerheid en risicomijdende houdingen van besluitvormers incorporeert, en dat rechtvaardige anticiperende acties tegen overstromingen prioriteert om schade te beperken.

Het primaire doel van dit onderzoek was niet om voorspellingen nauwkeuriger te maken, maar om beter te begrijpen hoe onvolmaakte data, voorspellingen en modellen bruikbaar kunnen zijn in besluitvorming rond pluviale overstromingsvoorspellingen en lokaal anticiperend overstromingsbeheer. Het onderzoek behandelt thema's als 'fit-for-purpose'-data en het creëren van waarde uit onvolmaakte informatie door het bestuderen

van de gebruiksmogelijkheden en beperkingen van data en de onderlinge afhankelijkheid tussen voorspellingen en besluitvorming. Deze aanpak bevordert onzekerheidsbewuste overstromingsvoorspellingen en besluitvorming, in plaats van het creëren van een illusie van zekerheid, met als doel betere beslissingen te nemen en de paraatheid te vergroten via op voorspellingen gebaseerde acties.

CONTENTS

1

INTRODUCTION

1.1 Background

Globally, weather-related disasters continue, with astounding economic damage, loss of life and social and economic disruption. The year 2024 alone saw 368 billion US dollars (USD) in economic losses from weather-related disasters, of which 348 billion USD was attributed to inland flooding (AON, 2025). While a changing climate continues to influence the frequency and intensity of extreme precipitation events, urbanisation and population growth also pose significant challenges to flood risk management. Higher populations pose a challenge due to increased vulnerabilities and exposed assets (Jha et al., 2012). As of 2024, 57.7% of the world's population lives in urban areas and this is expected to increase to 68% by 2050 (UNDESA, 2024).

Urban pluvial (rain-driven) flooding, commonly called surface water flooding, is characterised by high-intensity, short-duration rainfall that contributes to direct runoff, surcharging sewers, overtopping urban drains and inundating low-lying areas. The sixth Intergovernmental Panel on Climate Change Assessment Report, AR6 (2023), predicts that the increased frequency and intensity of heavy rainfall events translate into higher frequency and greater magnitude of pluvial and surface water floods. While large-scale events capture global headlines, many cities are plagued by these smaller, sometimes more frequent localised pluvial events which result in high cumulative losses, socio-economic impacts and disruption to critical infrastructure and everyday life (Acosta-Coll et al., 2018; Jiang et al., 2018; J. A.E. ten Veldhuis, 2011). With projected increases in urbanisation and impermeable surfaces and the future intensification and increased frequency of rainfall events from climate change, this type of flooding has emerged as a critical issue (Houston et al., 2011; Nicklin et al., 2019; Rosenzweig et al., 2018; Zevenbergen et al., 2008; Zhou et al., 2012). Pluvial and surface water floods account for one-third of all floods in the United Kingdom and have also increased in China's megacities: Beijing, Shanghai, Wuhan and Shenzhen (Jiang et al., 2018). Cities in the Middle Eastern and North African region, such as Alexandria, Casablanca and Tunis, have also experienced increased flooding over the last decade, accompanied by poor infrastructure and ad hoc informal development within the cities (Banerjee et al., 2014; Bigio, 2009) and by unprecedented rainfall in Dubai.

To cope, there has been a call to adapt and transition from purely structural measures towards more hybrid, multi-functional measures which are low regret, less expensive and more resilient and adaptable to future uncertainties (Difrancesco & Tullos, 2014). Early Warning Systems (EWS) and early action initiatives such as Anticipatory Flood Management (AFMA) are proactive approaches that consider weather predictions and future hazards and impacts, issuing warnings and explicitly taking action before an event. When disseminated and communicated sufficiently in advance of an event, warnings and the implementation of emergency measures can significantly reduce damage to assets and lives whilst mitigating residual risks (Viktor Rözer et al., 2016).

In cities and urban areas, there is a huge potential for anticipatory actions to assist emergency managers. Possible anticipatory flood mitigation measures consist of selecting decisions from a pre-defined categorical catalogue of actions, including issuing of warnings, redistribution of flow in a system to allow storage, dispatching of temporary flood defences and sandbags, control and monitoring of weirs and pumps, cleaning of drains and grates and allocation of resources. The success of such strategies requires accurate predictions of not only the occurrence but also the magnitude, location and timing of hazards and impacts to provide critical actionable information.

Advancements in technology and computing power have fostered progress in data collection and availability, the integration of models, improvements in forecast skill and the development of high-resolution global numerical weather prediction models that can predict several days ahead and ensemble forecasts with predictive uncertainty quantification. Supported by advances in hydrodynamic inundation models, surface delineation runoff and flood processes or AI methods have improved, which increases the speed and efficiency of computations.

However, at the city scale, many existing and emerging challenges lead to uncertainty along the pluvial flood forecast and decision-making chain in both initial developmental stages and persist into the operational phase. National weather and disaster agencies' forecasts are not always satisfactory for local-scale actions. For example, many weather and flood forecasting capabilities and warnings are done at the national level. High spatial and temporal resolution rainfall products are not always available to realistically represent structure and variability in the rainfall patterns. Local actions however, require high-resolution data and models to capture the dynamic and heterogeneous nature of urban catchments, as well as knowledge of the drainage system, terrain and initial conditions.

Data scarcity is characterised by insufficient input data, which, for rainfall data, is often inadequate or absent long-term records with high spatiotemporal resolution, posing a significant challenge for hydrological forecasting (Kidd et al., 2017). Data-scarce cities tend to lack the data to build complex drainage systems and calibrate models. The lack of data can influence the selection of models, the quality of the knowledge produced by models and the application of this knowledge in local-scale decision-making. Therefore, given the interdependencies and connectivity among these components, a closer look at each component is needed to identify opportunities for improvement. The following sections further describe each component of the pluvial forecasting, decision and response chain in the context of data-scarce cities.

1.2 FORECASTING - DECISION -RESPONSE FRAMEWORK FOR AFMA

The United Nations Disaster Risk Reduction Agency (UNDRR) defines an early warning system as "*An integrated system of hazard monitoring, forecasting and prediction*

disaster risk assessment, communication and preparedness activities, systems and processes that enable individuals, communities, governments, businesses and others to take timely action to reduce disaster risk in advance of hazardous events". It consists of 4 pillars: Disaster risk knowledge, Forecasting, monitoring & detection, Communication & dissemination and Preparedness & response (WMO, 2018).

Flood forecasting, decision-making and response systems are subsets of the overall EWS, consisting of several subsystems that can be decomposed into different elements. In AFMA, response is framed as an anticipatory response to forecast and pre-agreed actions, whereas in traditional humanitarian contexts, response refers to reactive measures taken during or after a disaster to manage its impacts.

In general, the design, implementation and operation of such flood forecasting require an understanding of the system's response to past and future events. It can be conceptualised (Figure 1.1) as using data that feeds into models to generate knowledge used to make decisions and take actions, such as whether to take an action, what action and when, where and how to implement it. There is a critical link between data, models, knowledge, decisions and responses; therefore, in data-scarce cities, the quality and quantity of data used and the knowledge produced also influence the ability to implement flood forecasting systems for anticipatory flood management. Henonin et al. (2013) compared the flood forecasting process to that of a house structure where each component's quality contributes to the overall system's integrity and durability. Consequently, limitations in one component will inevitably result in a trickle-down effect and the failure of the entire structure/system.

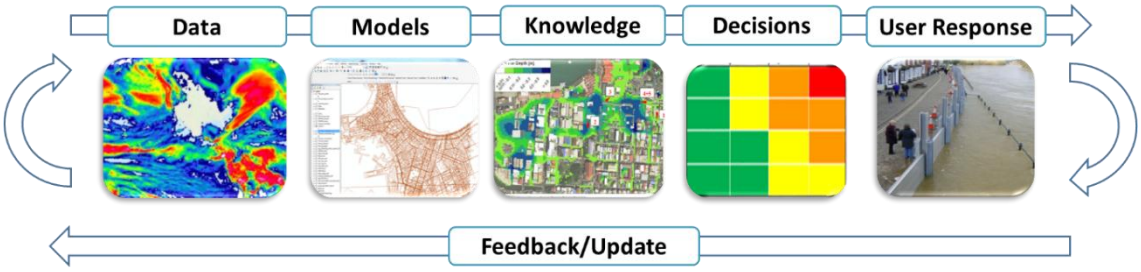


Figure 1.1 Connectivity of data, models, knowledge, decisions and response chain

In addition, this chain of components can be further decomposed into meteorological forecasting, hydrological/flood forecasting, decision-making subsystem and warning-response subsystem (Figure 1.2) (Verkade and Werner, 2011). The methods used in a flood forecasting system depend on the type or source of flooding, but in general, the framework consists of meteorological forecasts that provide information on the magnitude, timing and location of rainfall and other variables. Rainfall is fed into hydrological models, which provide inputs to hydrodynamic models, or river-level observations are used to forecast hydrological variables: flood depths, extent, populations

at risk and impacts. Based on the information, triggers or thresholds and other decision support tools are used to issue warnings or take actions. This process is supported by constant monitoring of forecasts and observations at different lead times.

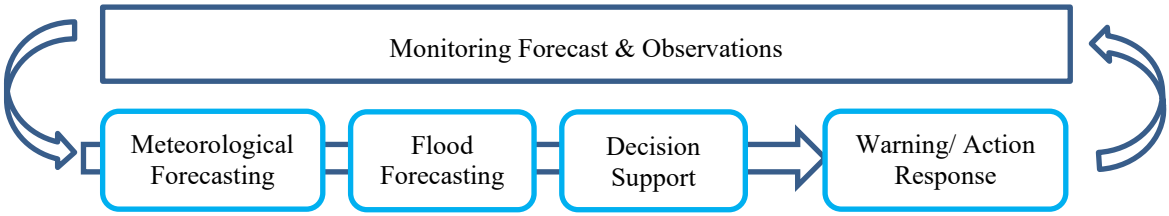


Figure 1.2 Flood forecasting- Decision- Warning- Response System. Adapted from Verkade and Werner 2011

In the decision subsystem, a decision-maker is a user of a warning but may also be a producer of a warning for different users. Issuing a warning can also be considered a decision/action for meteorological or flood forecasters vs end users who take protective action. For this thesis, decisions will refer to actions taken by emergency managers and vary with the information available at different lead times. It should be noted that the decision support system is separate from the actual decision, reinforcing that it is meant only to aid the decision process and does not dictate the final decision. Weather warnings have also been described as a value chain approach that represents the challenges of linking the interdisciplinary capabilities and outputs of observations, weather forecasting, hazard forecasting, impacts, warnings and decisions with their associated communities and the flood forecast chain (Golding et al., 2019). While advancements in the respective fields are shown by the height of the mountain, less progress has been made in impact forecasting and decision-making, which is supported by socio-economic modelling and behavioural science, respectively.

The shift towards a more proactive rather than reactive approach has gained traction and popularity, particularly through initiatives such as the WMO's Early Warning for All and other early warning and early action initiatives. The concept has evolved and is influenced by improvements in forecasting, risk and vulnerability assessment and the increased focus on preparedness and response before an event. Anticipatory Action mechanisms in particular have been adopted by humanitarian organisations where anticipatory actions refer to "*actions taken to reduce the humanitarian impacts of a forecast hazard before it occurs, or before its most acute impacts are felt and backed by pre-arranged financing*" (IFRC, 2020). The flood forecasting-decision-response framework for anticipatory flood actions is shown in Figure 1.3. Using real-time information and forecasts to take flood actions before an event is not new. Real-time control (RTC) of sewer systems in anticipation of high flows has been practised in flood management to control the operation of pumps, overflows and other hydraulic structures

(García et al., 2015). Using sensors and forecasts, RTC systems mitigate flooding by operating the hydraulic components of an urban sewer system to reduce the peak flood response via redistribution or storage before, during and after an event.

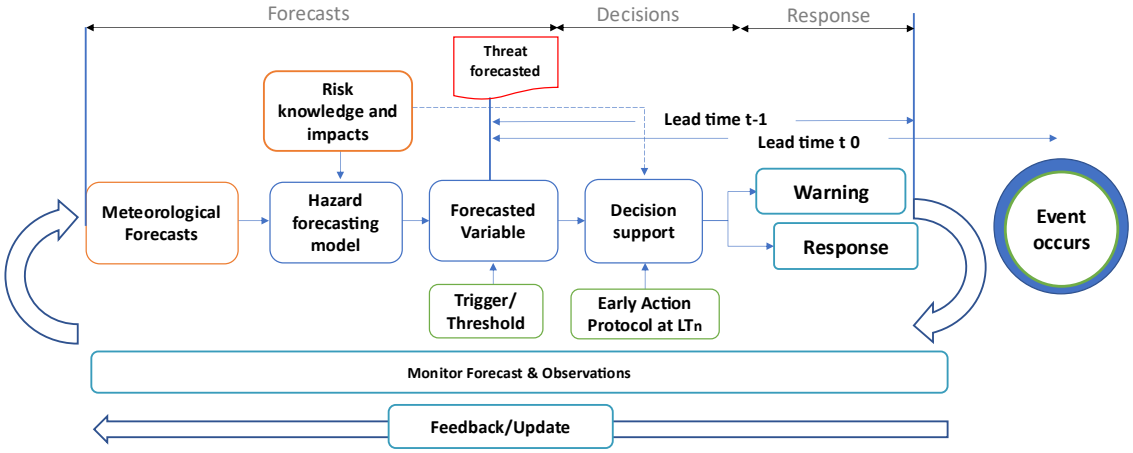


Figure 1.3 Flood forecasting-decision-response framework for anticipatory flood action

While there have been advances in the forecasting component for taking actions, there has also been a shift towards more approaches which emphasise the type of information produced (such as impact-based or probabilistic), how it is communicated and how useful it is to decision makers (Doyle et al., 2019; Jean et al., 2023; Kox et al., 2018; M.-H. Ramos et al., 2010; WMO, 2018). Although different frameworks have been used (as discussed above), they all strongly highlight the connectivity and dependencies between the various components. The next sections will describe the different components of the meteorological forecasting, flood forecasting and decision-making components.

1.3 METEOROLOGICAL FORECASTING

The advent of computational approaches meant more complex calculations using numerical weather prediction models based on the physics of the atmosphere and the development of more sophisticated hydrological models which enabled more accurate predictions of hydrological variables. The introduction of satellite remote sensing provided a new way to collect global hydrological and meteorological data in near real-time. As computing power increased, this facilitated high-resolution models and ensemble forecasts which run at shorter times and forecasting days, weeks and months in advance, along with the use of risk and impact information in the decision-making subsystem.

Flooding in urban areas can be caused by large-scale frontal precipitation systems or harder to predict small-scale convective type systems. The accuracy with which this rainfall is forecasted significantly influences the reliability of urban pluvial flood forecasting (B. W. Golding, 2009; Hapuarachchi et al., 2011). Rainfall variability is a significant source of uncertainty in the prediction of sewer flooding. Whereas high-resolution forecast data is needed to represent urban processes, timing, intensity and spatial variability. There are several precipitation data sources at varying qualities, resolutions and lead times; weather radars, quantitative precipitation estimates, reanalysis models and Quantitative Precipitation Forecast (QPFs) from Numerical Weather Prediction models (NWPs).

Quantitative Precipitation Forecasts (QPF) serve as forcing data into rainfall-runoff models. Forecasts drive rainfall-runoff and hydraulic models to forecast flows, water levels, flood extent, depths and impact. QPFs from NWP models use mathematical models to simulate the state of the atmosphere. Models are initialised using observations as initial conditions to model the evolution of the atmosphere. New observations are assimilated into the model to capture the most recent conditions. Global Ensemble NWPs are produced in many centres around the world on a global scale. Forecasts are uncertain due to the chaotic nature of the atmosphere and this uncertainty increases with lead time. Ensemble Prediction Systems (EPS) provides a way of estimating uncertainty in a weather forecast. Instead of running the NWP model once (a deterministic forecast), the model is run repeatedly from perturbed initial conditions to produce an "ensemble" of forecasts up to 15 days (Buizza *et al.*, 1999). Ensemble forecasts evaluate uncertainty while extending the forecast lead time (Cloke & Pappenberger, 2009; Duan et al., 2018; Gofa, 2010) and are advantageous in providing probabilistic information in the decision-making process (Dale et al., 2012). While identifying a future event, an acceptable lead time depends on the divergence of ensembles which, in turn, is generally related to the spatial and temporal resolution.

Downscaled regional or Limited Area Models increase resolution and accuracy at smaller scales. The Weather Research and Forecasting (WRF) model is a widely used dynamic downscaling model used in many environmental applications (Goodarzi et al., 2019; He et al., 2013; S. Ibrahim & Afandi, 2014b; Skamarock et al., 2008). Global models are used as lateral and boundary conditions confining nested domains at finer resolutions, allowing simulation of more localised phenomena. There has been improvement in resolving small-scale weather variability through convection-permitting prediction models (Clark *et al.,* 2016), but there are still limitations with NWPs at the city scale (Kumar & Venkatesha Prasad, 2018; Thorndahl et al., 2013, 2016) and over shorter lead-times.

To overcome the challenges with localised convective processes for pluvial flooding, high-resolution radar-based estimates have been used (Chen et al., 2017; Codo & Rico-

Ramirez, 2018; Duncan et al., 2011; Schellart et al., 2014). Forecasting applications have focused on how to improve accuracy, extend lead times by merging with radar advected patterns with QPFs (Schellart et al., 2011; Yoon, 2019) and reducing uncertainty propagation or radar rainfall and radar-based forecasts through the use of radar ensembles and probabilistic forecasting. Jasper-Tönnies *et al.* (2018) successfully combined ensemble radar nowcast with high-resolution QPFs for urban flood modelling. However, there are still challenges with the use of radar in forecasting rainfall in small catchments, which require accurate estimation of initial states (Abdalla et al., 2013). Nowcasting models which statistically extrapolate future rainfall distributions, provide high spatial and temporal resolution at short lead times but the accuracy is limited beyond 6 hours (B. W. Golding, 1998; He et al., 2013). Although radar estimates are capable of capturing intensity at smaller grids, there is a trade-off between input resolutions and forecast lead times, which is not ideal for actions that require longer implementation periods to be effective. Further, many countries are just not covered by weather radar observations (Saltikoff et al., 2019; WMO, 2020) or such observations are not used for nowcasting or hydrological forecasting.

Post-processing of forecast precipitation is another method used to improve the forecast's accuracy, sharpness and reliability by removing systematic errors (Crochemore, Ramos, & Pappenberger, 2016; Verkade et al., 2013). Although a correction that improves the performance of NWP's outputs is not always reflected as an improvement in the hydrological forecasts (Wetterhall & Smith, 2019). Additionally, post-processing of forecasts is another key challenge in data-scarce regions given its reliance on long-duration climatological data observation. Continued research is needed to address how to improve high-resolution forecasts in the absence of high-resolution observed data.

1.4 URBAN PLUVIAL FLOOD FORECASTING APPROACHES AND PROCESSES

Urban pluvial catchments are defined by smaller basins, short time of concentrations and a high sensitivity to interactions between paved areas and urban features. In urban areas, there are a high proportion of impervious surfaces and most of the rainfall is transformed into runoff. Flooding occurs when drainage capacity is exceeded and overflows onto the street or surface runoff cannot enter drains and ponding occurs in low areas. ten Veldhuis *et al.* (2011) found that component failure and blockages were significant contributors to flooding but receives minor attention compared to overloading of the sewer. Urban flood models consist of a rainfall-runoff and hydrodynamic component. Flow in sewers is characterised as unsteady non-uniform flow. Usually, free surface flow is assumed but during high-intensity events, the pipes may become surcharged, resulting in pressurised pipe flow.

Several studies have investigated the sensitivity of urban catchments to rainfall spatiotemporal variability (Cristiano et al., 2017; Ichiba et al., 2018; Ochoa-Rodriguez et al., 2015). Berne *et al.* (2004) found that urban catchments in the order of 10km^2 require a spatial and temporal resolution of 3 km and 5 mins respectively and values decrease with decreasing size of the catchment. However, in reality, this resolution is rarely attainable and even less so for forecasting purposes, primarily due to the incoherent spatio-temporal scale of atmospheric models and urban flood generation. The sensitivity to rainfall intensity and complex hydrological response at the urban scale coupled with the mitigation time required to implement actions means that each subsystem requires high-resolution data requirements and model discretisation which can capture spatially distributed processes while balancing the computational run time and required lead-times and mitigation times. Lead time refers to the minimum time required to implement effective preparatory actions. Figure 1.3 highlights that shorter lead times also coincides with the requirements of higher resolution data. This requirement for high-resolution data also influences the selection of the flood forecasting approach and the sophistication of the triggering mechanism and requisite forecast variables.

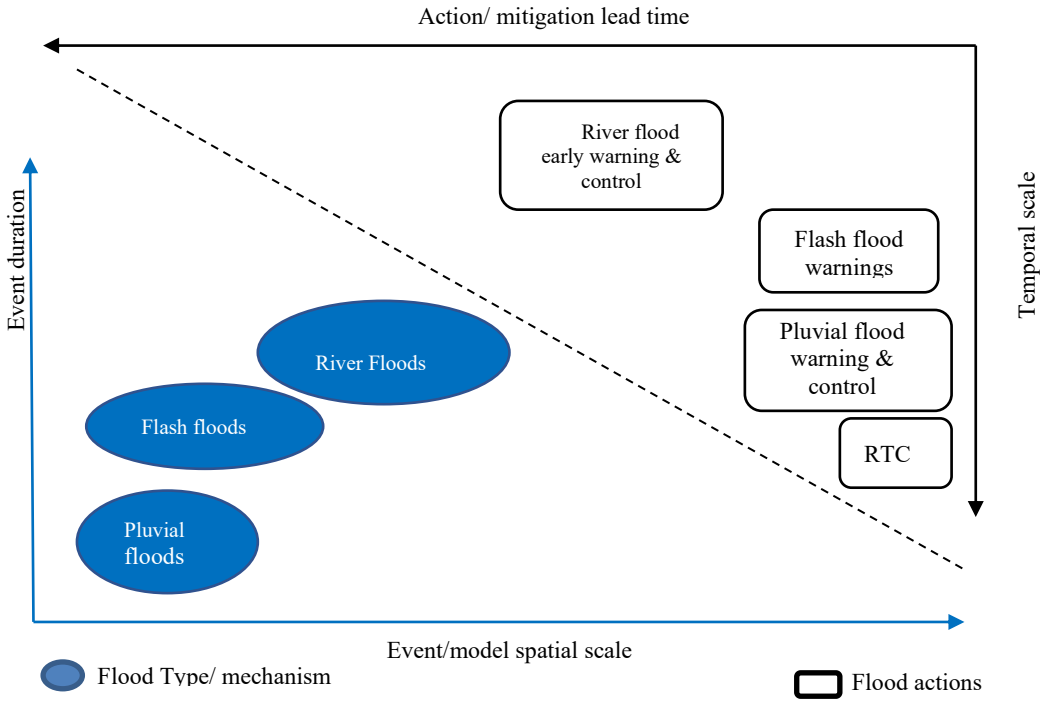


Figure 1.3 Comparison of flood type event duration scale and spatial scale and lead time/mitigation time for flood actions. Adapted from Merz 2020 and van Andel 2011

1.4.1 Hazard estimation methods

Henonin et al. (2013) classified flood forecasting hazard models as empirical, pre-simulated scenario-based (offline) or real-time simulations (online) (Figure 1.4). These

can be further classified by the type of data and forecast: deterministic vs probabilistic (representing a probability distribution of a future variable) or risk-based, or impact-based.

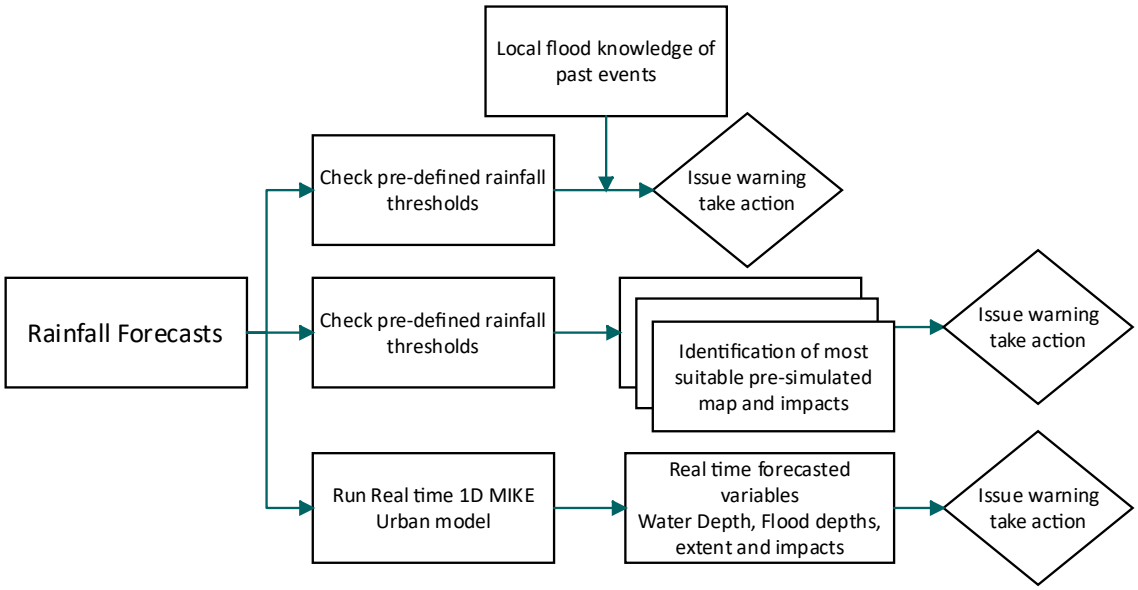


Figure 1.4 Different pluvial flood forecasting approaches

Empirical methods such as rainfall thresholds directly compare precipitation forecasts with historical or user-defined critical rainfall thresholds (Falconer et al., 2009; Georgakakos, 2006; Martina et al., 2006; Parker et al., 2011; S. J. Wu et al., 2015). Comparing rainfall forecasts from NWPs with critical reference rainfall thresholds derived by stochastic analysis of historical measurements (Lorenzo Alfieri et al., 2012). An example is the (Georgakakos, 2006) the Flash Flood Guidance System (FFGS) which uses forecast rainfall depths that are likely to result in discharges associated with floods and found they produce a high probability of detection of flash floods (Georgakakos, 2006; Z. Wu et al., 2023). These alerts are issued based on different user-defined rainfall thresholds or hydrological predictors such as soil moisture conditions. However, this method is sensitive to the trigger or warning thresholds and a major challenge is ensuring that decision thresholds are robust (de Perez et al., 2016). (S. J. Wu et al., 2015)compared methods to determine the performance of rainfall thresholds results depending on the size of the catchment and the quantity of data available. For a small urban catchment, Bouwens et al. (2018) and Tian et al. (2019) proposed deriving rainfall thresholds not with hydrodynamic models but by correlating over 70,000 citizens' reported flood incidents overflow pumping values with sub-daily rainfall data. Tian et al. (2019) further derived thresholds using decision tree models. Although there is some level of uncertainty, it can be useful in providing some insight in predicting floods but only when this type of data is available. This simplified Early Warning Systems (EWS) approach provides an immediate short-term solution for data-scarce regions, which may lack the technical capacity and resources to implement complex methods. New emerging methods are being

used to improve flood forecasting by integrating crowdsourced hydrological data, low-cost sensors, web technologies and monitoring from local stakeholders (Verbeiren et al., 2019).

Real-time simulations forecast with real-time hydrodynamic models forecast the water levels, flood extent, locations, damage and impact. Real-time models have the benefit of providing the most accurate simulation when updated with the most recently available observations and state conditions (J. R. René, 2014). Real-time forecasting examples for pluvial flooding are limited. However, this approach encounters the same demand of high-resolution spatial data and copious input parameters into complex models associated with uncertainty and long computational times, limiting its use for real-time operational use. The advantage of real-time approaches allows direct model spatial variability and impact which is beneficial to areas that do not have high-resolution observations (Hofmann & Schüttrumpf, 2019). Therefore, there must be a balance between model complexity and flood analysis quality (Henonin et al., 2013; Ochoa-Rodríguez et al., 2015). The different types of models used for real-time forecasting are further discussed in the next section.

However, to overcome the challenges of computational times, scenario interpolation methods or query-based approaches use pre-simulated scenarios to issue a warning based on a comparison between forecasted rainfall patterns and historical rainfall patterns. Pre-simulated scenario-based systems utilise a catalogue of previous hydrodynamic simulations and the accuracy is dependent on both model complexity and input data. An example of this system is the Scottish Environmental Protection Agency's Grid to Grid model. Scenario-based approaches have the advantage of fast computing over real-time approaches. However, it requires many runs and observations to capture the different range of possibilities, especially when synthetic rainfall hydrographs are used to develop scenarios. Previous research has proposed "quasi real-time simulations" which combine scenarios and real-time simulations. Hofmann and Schüttrumpf (2019) utilised a hydrodynamic simulation to understand more about the flood hazard and validate the model, but a pre-simulated library of flood inundation maps is used in real-time. Alternatively, Filianoti *et al.* (2020) used a 1D model to produce rough flood maps and proposed a method of matching 2D flood maps run with synthetic hyetographs.

1.4.2 Urban flood models

Real-time and pre-simulated scenarios both use urban flood models to forecast floods with real-time variables or with a library of different run scenarios. Extensive research has been done on the use of urban flood models, but fewer studies have been conducted on their specific application to pluvial flood forecasting. Urban flood models generally consist of a rainfall-runoff model combined with flow routing and hydraulic modelling. Complexity ranges from conceptual, physically based to black box. Models used for

design and scenario analysis differ from those required for real-time flood forecasting. In the former, the current trend is towards more efficient, higher spatial resolution models which include all key hydraulic features and accurately capture the hydrological and hydraulic processes. However, as previously mentioned, real-time flood forecasting needs to minimise run times and model instabilities with sufficient acceptable accuracy (Moore *et al*., 2015). Models can benefit from including existing conditions such as obstructions due to urban debris which is often excluded from flood forecasting models.

Urban models can be 1D, 2D, or 1D-1D and 1D-2D models, whereby overland surface flow runoff from rainfall is combined with subsurface sewer routing. The 1D model quantifies runoff from urban areas and models the flow through the pipe system. When coupled with a 2D surface model, it can be used for inundation mapping. While more accurate in representing surface flow conveyance, the resolution of models is linked to the 2D surface grid and its connection with the 1D model (Henonin et al., 2013). Flooding only occurs from manhole nodes once water has accumulated in the system and excludes pooling of water before reaching a drainage pipe or channel.

Several authors have attempted to increase model accuracy. Leandro *et al.* (2009) demonstrated within acceptable tolerance the agreement between the results of a 1D-1D and 1D-2D model. In large cities, attempts have been made to limit the simulation to smaller areas. Jahanbazi and Egger (2014) used a 2D surface model at areas susceptible to flooding only. Hénonin *et al.* (2015) proposed using a multi-grid approach which combines a fast coarse grid and finer grid to reduce computational time. (Ming et al., 2020) used a 2D new forecasting system by coupling a graphics processing unit (GPU) accelerated hydrodynamic model with NWP products to provide high-resolution, catchment-scale forecasting of rainfall-runoff and flooding processes induced by intense rainfall. Despite improvements in computational time, there was some overestimation of the floods compared to high-resolution grids due to differences in runoff volume. In general, errors are introduced as localised simulations and closed boundary conditions do not fully capture the complex connectivity of the city (Xing et al., 2019).

Data Driven Models

Alternatively, data-driven or machine learning models such as Artificial Neural Networks (ANN) have been proposed to reduce run times of computationally demanding physically based models (Berkhahn et al., 2019; Berkhahn & Neuweiler, 2024; Duncan et al., 2011) through statistical relationships derived from rainfall and flow data.

Although more commonly used in fluvial flood forecasting, there is the potential to use data-driven model approaches for Pluvial flood forecasting (Campolo et al., 2003; Jain et al., 2017). Mosavi *et al.* (2018) reviewed machine learning methods for flood prediction and overwhelmingly, this approach demonstrates a reduction in the run time. Berkhahn *et al.* (2019) found good agreement between an ANN model and the physically based

models in pluvial catchments. However, ANN's application to urban pluvial catchments is still limited due to data requirements and machine learning methods are unable to capture the changes in system.

Bermúdez *et al.* (2018) demonstrated comparable results and reductions in run time using two surrogate models (1) based on a 1D sewer network model and GIS overland flood network and (2) lumped conceptual hydraulic model using a combined data-driven and physically based modelling approach. However, there is a challenge to apply this method given its dependence on large quantities of data to reproduce a relationship between input and output variables. Alternatively, simplification of models by removing secondary network elements has also been considered (Leitão et al., 2010; Ngo et al., 2018). Although there was some agreement in the peak times, these models tend to predict higher flows and water depths. Calibration and validation is also required for the output to be useful. Real-time data assimilation is also difficult due to the requirement of collecting observed flows and short simulation time. In the absence of observed data, Leandro *et al.* (2011) demonstrated the applicability of a 1D/2D model to calibrate a 1D1D model. Recently, crowdsourced data has been used more frequently to validate models of urban flooding in the absence of observed data (Hsiang, 2019; See, 2019; Yu et al., 2016). (Hsiang, 2019; See, 2019; Yu et al., 2016)

Regardless of the flood model type, different challenges exist for an operational pluvial flood forecasting system. The main purpose of the model is to produce knowledge of required variables to guide decision-makers in selecting mitigation strategies and action. Therefore, the usefulness of the flood model becomes more important and dependent on the type and quality of the input and output and the flexibility and speed of the model (Leskens et al., 2014). The next section describes the progress that has been made with the inclusion of impacts, risk and probabilistic forecasts.

1.4.3 Impact-based and risk-based forecasts

Henonin *et al.* (2013) categorised approaches based on the hazard estimation method, but approaches can also be defined as impact-based and risk-based depending on the input type and resulting knowledge produced. Much of the discussion thus far has centred on forecasting hazard, however impact-based forecasts and warnings put focus on the inclusion of vulnerability, consequences and potential impacts to improve understanding and implementation of actions (Merz et al., 2020; WMO, 2015, 2021). The inclusion of impacts on flood forecasts considers factors such as the number of persons or infrastructure exposed and the consequences of the hazard to those elements. It provides quantitative assessments of impacts and damages based on the hazard and assets at risk (Schmidt et al., 2011). However, producing impact forecasts adds an extra level of complexity and requirement for actual observed impacts or vulnerabilities and can also

be dynamically changing as impacts unfold and as people take protective actions (Harrison et al., 2022).

Impacts are classified as direct tangible or indirect intangible. Direct tangible impacts refer to physical damage resulting from direct contact with flood waters, e.g. damage to buildings, property, vehicles and crops. Indirect intangible damages are a causation of direct damages for example, service or traffic disruptions, jobs losses, population displacement, contamination and psychological effects(Hammond et al., 2015). Previous research has been done on flood impact and damage assessments (Aldridge et al., 2020; Merz et al., 2020) but fewer studies have focused on urban damage and impact assessments. Hammond *et al. (*2015) completed a review on urban flood impact assessment and found that the difficulty lies in a lack of good quality flood impact data which results in problems with the validation and calibration of impacts. Real-time Assessment of Flash Flood Impacts (ReAFFIRM) is a real-time high-resolution approach for estimating flash flood hazard and corresponding impacts by combining hazard, flood map and impact assessment (Ritter et al., 2020). The impact-based forecast module included socio-economic exposure and vulnerability data based on different land use classes, critical infrastructure and population densities. Hofmann and Schüttrumpf (2019) utilised a risk-based approach defined by real-time hazard simulation maps and coupled flood predictions with pre-calculated urban damage potential. Rözer *et al. (*2021) went one step further and combined a contaminant transport model with neural-network based urban inundation model and a damage model for residential buildings.

1.4.4 Probabilistic flood forecasts

Probabilistic forecasts provide a range of possible outcomes using a probability distribution of input values and are useful in representing and estimating forecast uncertainty. Several sources of uncertainty exist in the flood forecast chain due to rainfall input variables (as discussed in section 1.3), model prediction uncertainty, uncertainty in consequences and damages for impact-based forecasts **(**Figure 1.5**)**. Uncertainty sources are defined as aleatory, due to natural variabilities and randomness or as epistemic uncertainty because of a lack of knowledge which can be reduced by improving the quantity and quality of data and the better representation of the meteorological and hydrological processes. Uncertainty in the modelling process arises from the mathematical representation and simplification of physical processes such as overland flow, storage, sub-catchment aggregation and estimation of model input parameters (Ghosh & Hellweger, 2012). Additionally, uncertainty in flood models also arises due to uncertainty in the system's initial maintenance conditions and an inability to calibrate and validate models from sparse and uncertain measurements and observations. As a result, uncertainty estimation in flood modelling is considered fundamental by a growing number of researchers. However, uncertainty is not always considered in practice and

accounting for the full range of uncertainty down the FF chain remains a challenge (W. Wu et al., 2020).

When driven by meteorological forecast, the term predictive uncertainty is used to describe the probability of a future event conditioned on the information available up to that time (prior knowledge). The importance of considering predictive uncertainty (uncertainty propagation) has been discussed by (R. Buizza et al., 1999; Krzysztofowicz, 1999; Todini, 2009; Van Steenbergen & Willems, 2014; Weerts et al., 2011). Predictive uncertainty can be represented by statistical post-processors or ensemble techniques (ensemble of initial conditions or parameterisation). Statistical post-processing methods use the historical performance of forecasts against observations to derive probabilistic distributions of future forecasts, whereas probabilistic hydrological forecasts are produced from hydrological simulations forced with the meteorological ensembles derived from numerical weather predictions (as discussed in section 1.3).

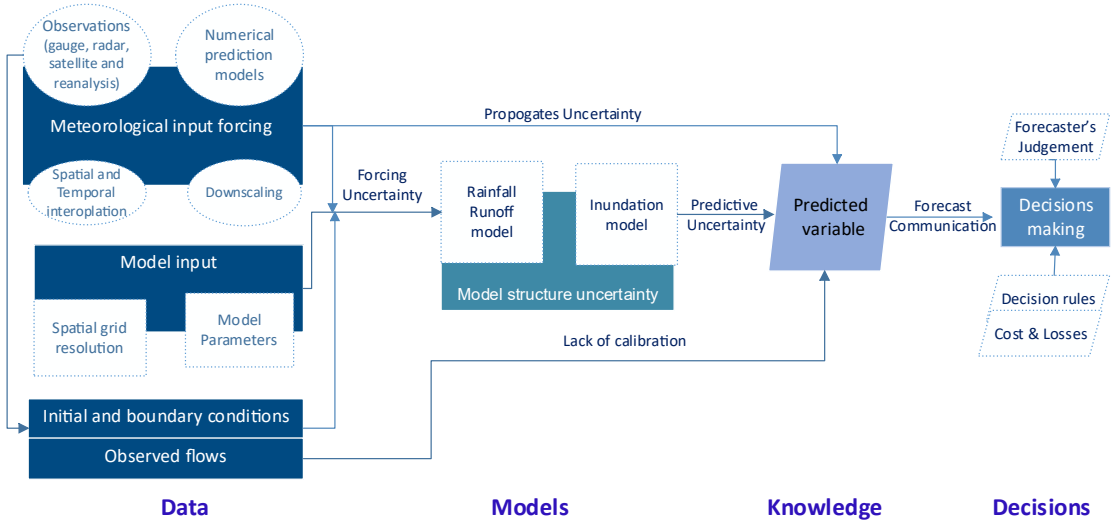


Figure 1.5 Sources of uncertainty in the input data, model, output variables and the decision-making component

Boelee et al. (2019) conducted an overview of publications using statistical, ensembles or a combined approach for predictive uncertainty, suggesting that statistical methods are best used when observed data is available, whereas ensemble methods can be used when targeting uncertainty from a particular source and in the absence of observed data.

In Section 1.4.2 the computational requirements flood models were discussed. Therefore, the idea of ensemble/ multiple run for flood inundation models further complicates the use of ensembles for pluvial flood forecasting (Wu et al., 2020). Despite this, using probabilistic forecasts over deterministic forecasts has been found to be a useful means to quantify uncertainty through probability distribution functions and supports the use of risk-based decision approaches (Davis et al., 2022; Iqbal, 2017; M. H. Ramos et al., 2013;

J. R. René, 2014; Zanchetta & Coulibaly, 2020). Therefore, further research is needed to determine the best way to utilise ensemble forecasts and ensure the added value of ensemble forecasts. A summary of the pluvial flood forecast approaches, data, triggers and outputs is shown in Table 1.1. Regardless of the approach, flood model type or data, different challenges in implementing an operational pluvial flood forecasting system. The main purpose of the model is to produce knowledge of required variables to guide decision-makers in selecting mitigation strategies and action. However, Leskens et al. (2014) found there can be discrepancies between what models offer and what decision makers require. Therefore, there is still a need for the flood forecast approach to incorporate the user's needs (Flack et al., 2019; L. J. Speight et al., 2021; A. L. Taylor et al., 2018; Q. Zhang et al., 2019). Therefore, the usefulness of the flood model becomes more important and dependent on the type and quality of the input and output, as well as the flexibility and speed of the model.

Table 1.1 Summary of data, model and trigger mechanisms of different pluvial FF approaches.

	Hazard estimation methods					
Approach	Rainfall threshold	Data driven	Pre-simulated Scenarios	Real-time Simulated Scenarios	Probabilistic/ Ensemble	Impact-based/ Risk based
Data	Historical rainfall Historical flood Quantitative precipitation forecast/estimates Observed rainfall Critical flood areas				Ensemble rainfall & depth Probability distribution of model input parameters	Hazard + Exposure + Vulnerability + impact + Damage data
		- Topography/DEM - Land Use/land cover - Network data		Discharge		
Tigger mechanism	Critical rainfall thresholds		Critical rainfall thresholds	Real-time catchment	Exceed probability threshold	
Model	Look up charts	Mathematical expression (ANN, SVM	Pre- run process based models	Simulation run in real-time with real		
			Conceptual models 1D, 1D-2D hydrodynamic, GPU			
Knowledge	Forecast or observed rainfall exceedance Dichotomous forecast Categorised warning levels	Forecasted water level or Discharge exceeds critical predetermined value	Warning levels per scenario corresponding action	Water levels Hydrographs Flood extent Flood evolution		
Limitations	Limited characterisation of flood processes	Large historical data requirements	Capture dynamic real-time Changes	Limited number of scenarios	High Computation Time	Requires frequent updating
Reference	Young et al. (2021), Yang et al. (2016) (Jang, 2015) Falconer et al. (2009)	Berkhahn et al. (2019) Duncan et al. (2011) Berkhahn et al. (2024)	Speight et al. (2018) Aldridge et al. (2020)	Thorndahl et al. (2016) Speight et al. (2018) Schellart et al. (2011)	Hofmann & Schüttrumpf, (2019) Thorndahl et al. (2016	Rozer et al (2021) Hofmann & Schüttrumpf, (2019))

	Hazard estimation methods					
Approach	Rainfall threshold	Data driven	Pre-simulated Scenarios	Real-time Simulated Scenarios	Probabilistic/ Ensemble	Impact-based/ Risk based
		Ke et al. (2020) Rozer et al (2021)		J.-R. René et al. (2013)	Davis et al. (2022) J.-R. René et al. (2013)	

1.5 DECISION SUPPORT & ANTICIPATORY ACTION

Decisions are a key aspect of the flood forecasting chain, but incorrect decisions are often made. In operational urban water management for flood mitigation, decisions include but are not limited to the combination of several actions such as issuing warnings, redistribution of flow in a system to allow storage, dispatching of temporary flood defences and sandbags, control and monitoring of weirs and pumps, cleaning of drains and grates and allocation of resources. Finally, decisions must also be made regarding when to cease all actions once the threat has passed.

Decision Support Systems (DSS) are defined as a collection of tools which integrate, visualise or analyse multiple sources of information or evaluate alternative decisions. Mirfenderesk *et al. (*2016) describes the elements of an ideal DSS should have timely warnings, comprehensiveness, accuracy, speed, flexibility, ease of construction, operation and maintenance, accessibility and the effectiveness of the system and communication to a wide range of audiences.

The sophistication of the decision-making approach depends on several factors such as the flood forecasting approach, predictability, the level of uncertainty, the spatial scale of impacts and time scales and the technical capabilities of decision-makers (Simpson et al., 2016). Approaches range from experience using formalised heuristic pre-defined set of control rules, sophisticated optimisation methodologies, classical decision analysis theories or graphical user interfaces combining models. GIS layers platforms such as Delft FEWS provide a system shell to integrate observed and forecast data, models and support decision-making (Werner et al., 2013). The Probabilistic Urban Flash flood Informative Nexis PUFFIN application is used to integrate quantitative precipitation forecast and real-time modelling to identify infrastructure components at risk (Brendel et al., 2020). These DSS platforms can integrate other types of information such as vulnerable areas, hot spot areas and critical infrastructure. However, too much information can also complicate the decision-making process and overwhelm users (Jean et al., 2023). They also seldom provide a means for decision analysis techniques to evaluate different alternatives and actions.

A formalised systematic approach to decision-making is beneficial because it allows decisions to be repeatable, justifiable, transparent and able to stand up to scrutiny and it

allows for the implementation of a pre-defined plan. Madruga De Brito and Evers (2016) reviewed multi-criteria decision-making frameworks for flood risk management and found most focus on ranking alternatives, the Analytical Hierarchy Process (AHP) method, TOPSIS or simple weighting. However, uncertainty analysis is not usually considered.

Multiple factors contribute to decision-making: regulatory, instructional, political, limited resources and others (Morss et al., 2005). Hence, it is necessary to differentiate between prescriptive and normative decision-making and descriptive decision-making. The former suggest what decision-makers "*should*" do (rational), whereas descriptive decision theories are more rooted in psychology and attempt to explain and predict how people make decisions (Hansson, 2011). The latter will not be addressed here although its importance is acknowledged. Prescriptive and normative analysis approaches present a systematic approach to decision-making from a set of alternatives, integrating variables such as cost-loss (potential impacts). These can be represented in a mathematical framework to represent the decision which should be taken and can incorporate uncertainty.

Decision theory has been applied in many fields: economics, statistics and philosophy to name a few. In theoretical terms (Figure 1.6), decisions can be represented by alternatives, states of nature, outcomes of the decision and their value which can be expressed as cost or profit. Alternatives refer to different actions/ measures which can be implemented. States of nature refer to the possible events for example, the future evolution of a weather event and associated flood levels which are beyond the control of the decision maker. The outcome is the consequences of choosing an alternative based on the state of nature which occurs and is thus used to explain the effect of choosing an alternative.

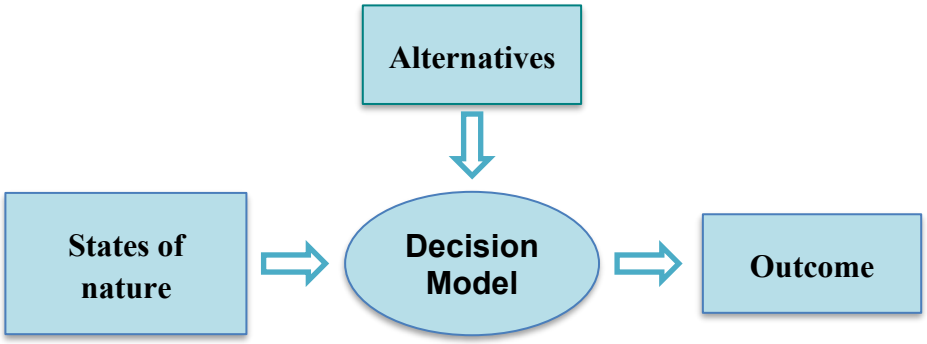


Figure 1.6 Theoretical decision model highlighting states of nature, alternatives to determine an outcome

Decision-making under uncertainty is a challenge in flood forecasting (Krzysztofowicz, 2001; Todini, 2017). Uncertainty quantification is particularly relevant for urban areas, as it informs decision-making processes (Ivanov et al., 2021). Probabilities derived from

probabilistic forecasts are used to express the likelihood of the future state of nature hence the attractiveness of using probabilistic forecasts to support decision-making (M. H. Ramos et al., 2013). The cost-loss ratio approach (Murphy, 1977) was proposed in flood warnings and has since undergone some iterations (Murphy, 1993; Richardson, 2000; Mylne, 2002; Roulin, 2007; Buizza, 2008; Verkade and Werner, 2011). In its simplest form, this method is based on making decisions when the probability of occurrence (a threshold), P exceeds the ratio of the cost, C of a measure/warning versus the losses L_a that would be avoided. If the ratio is greater than 1, then the cost of the warning is greater than the loss. However, the use of this method has been described as risk-neutral and its suitability in flood forecasting is debated (Addor et al., 2011; Matte et al., 2017). Dale *et al.* (2012) also proposed a benefit-cost decision framework which similarly accounts for uncertainty and level of impact. This method employs ensemble members, each of which is assigned a monetary value based on water level and impact.

In Bayesian decision approaches on choosing the alternative which maximises the expected utility (DeGroot, 2004; Savage, 1954) where utility (function) refers to the value of outcomes and can be represented as minimisation of expected losses or damages (Economou et al., 2016; Krzysztofowicz, 1993; Todini, 2017). Goodarzi et al. (2019) used an atmospheric ensemble forecast and a Bayesian network to estimate the flood peak and Fuzzy- TOPSIS model for flood warning decisions. Silvestro et al. (2019) proposed a warning decision process based on forecast impacts by computing direct economic damage and impacts on the population Bischiniotis et al. (2019) explored the use of sequential decision-making to account for limitations of the forecast quality as well as the decision-makers dilemma between acting upon limited-quality forecast information and taking less effective actions.

All authors found added benefits in using these approaches, but all require damage data or functions to estimate loss from a hazard and its impact. The reliance on economic and damage data is one of the main challenges in its application. Damage functions are derived from observed damage and impact, which are often unavailable, especially in already data-scarce regions. Thus, inaccurate damage data may introduce additional uncertainty into the decision-making process. Therefore, decision support systems should also be complemented by sensitivity analysis of different criteria and an understanding of uncertainties (Madruga De Brito & Evers, 2016).

In summary, DSS platforms visualise and communicate model and GIS outputs but often do not translate forecast information into actions or alternatives. Decision analysis frameworks, such as Bayesian decision-making approaches, can be built into DSSs to translate probabilistic forecasts into actionable information while accounting for user preferences. However, decision analysis approaches are not to be perceived as rigid frameworks but rather as recommended actions that may be overruled based on forecasters' or decision-makers' judgement.

Despite the implementation of decision support systems, questions remain about the optimal use of model outputs for decision-making, the use and communication of probabilistic forecasts to end users and the balance between data limitations and decision requirements. It is also acknowledged that there is the tendency for decision-makers to prioritise resources and institutional constraints over technical model information especially under pressure and uncertainty of the situation (Leskens et al., 2014; Morss et al., 2005). Ideally, any decision framework should integrate all these components so as the limit the introduction of biases into the decision-making process

1.6 PROBLEM STATEMENT AND RESEARCH GAPS

The flood forecasting chain can be conceptualised as data, models, decisions and responses or as several subsystems (Section 1.2). Each requires different data, models, approaches and involvement in disciplines. There are different sources of uncertainty in each component leading to a cascading effect across the forecast chain. Most explicitly demonstrated how data (in)accuracy and uncertainty affect model outputs (knowledge), how these outputs influence decisions and finally how the decisions and response will ultimately reduce damage and loss of life.

Previous studies have focused on improving each component separately with most of the research focusing on improving prediction aspects by improving global meteorological forecasts for local flood forecasting and quantifying uncertainty. Approaches vary in complexity, by the level of warning/service required (hazard, flood extent, impact-based etc.), actions to be taken (anticipatory, sequential or real-time), data available but in most circumstances they are dependent on the quantity and quality of available data and the technical expertise, capabilities and resources.

Within the context of operational flood forecasting in an urban data-scarce region, the requirement for high spatial and temporal observations and the inclusion of impacts is still a limiting factor. In many data-scarce countries, flood forecasting and warnings are done at a national level and not at the localised scale in which pluvial floods occur. Although there is the availability of satellite data, radar and numerical weather predictions, this is often associated with large uncertainties and is inadequate to accurately model flooding magnitude, timing and location at the city scale. Additionally, the lack of observations for uncertainty quantification is a challenge and affects the ability to make decisions (Figure 1.7).

Much of the scientific literature focuses on improving flood forecast accuracy through various methods such as ensemble approaches, data assimilation and more efficient models or incorporating impacts. These efforts are largely independent of the practical decision-making process and the actions that need to be taken based on these forecasts. The information required varies with each decision maker and specific actions to be taken.

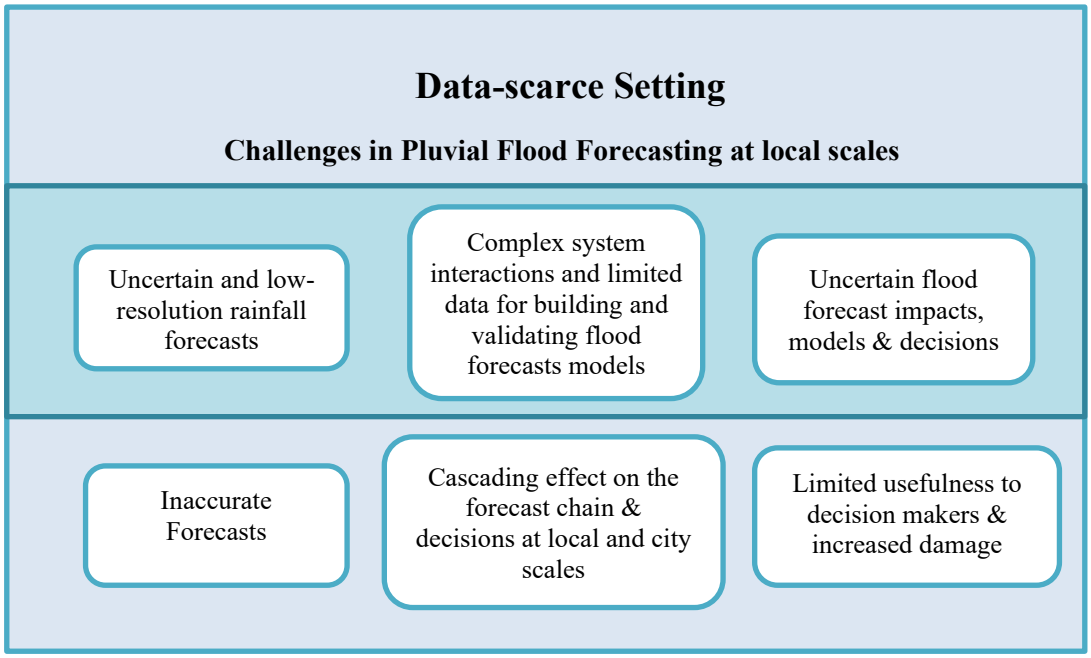


Figure 1.7 Summary of causes and effects of challenges in pluvial flood forecasting at local scales in data-scarce urban areas

1.6.1 Research questions and objectives

The overall aim of this research is *to investigate flood forecasting and decision support approaches tailored to data-scarce urban environments, with particular emphasis on assessing the suitability of available data and methods and on evaluating and incorporating uncertainty into anticipatory flood management*. This aims to achieve a greater understanding of how to align the available data with model and end-user requirements, while considering multiple dimensions that influence the Forecast-Decision-Response chain.

To achieve the research aim, four research questions have been proposed:

i. How can available data be utilised to improve flood forecasting in data-scarce urban areas?

ii. How suitable are using high-resolution rainfall forecasts at multiple spatial scales for urban flood forecasting methods?

iii. How well do high-resolution ensemble rainfall forecasts capture intensities compared to deterministic forecasts and what is the potential for its use in probabilistic decision frameworks***?***

iv. How can forecast uncertainty be incorporated into decision frameworks to support anticipatory flood management?

The above research questions will be answered by achieving the following research objectives:

i. Objective 1 - To investigate how available data can be combined to forecast floods in a limited urban data context.

ii. Objective 2 - To evaluate the suitability of high-resolution rainfall forecasts at different scales for use in different flood forecasting approaches in data-scarce cities

iii. Objective 3 - To assess the performance of ensemble rainfall forecasts and deterministic forecasts using categorical and probabilistic metrics for different events.

iv. Objective 4 - To develop and evaluate a Bayesian decision-making framework using uncertainty forecast information to support anticipatory flood action.

1.6.2 Practical value

The aim of this research is not to improve forecasts or predictions but instead to gain an understanding of the "usefulness" of imperfect data, forecasts and models. Forecasts and models are only of use if they can bring out benefits or reduce impacts by allowing decision-makers to take timely action (Mylne, 2002).

Characterising "fit for purpose" data and models is not based on how well they predict values, but rather on whether they enable the selection of an action that reduces flood losses and impacts. Regardless of the cascading effects of data and model uncertainty, decision-makers are expected to issue warnings or take actions to protect the public and property.

This research supports an integrated approach to flood forecasting and early warning rather than creating the illusion of certainty. Once accepted and incorporated, forecasters and decision-makers can collaborate to make better, more timely decisions and increase preparedness through forecast-based actions.

This research will adopt a methodology that recognises that effective flood forecasting and decision support in data-scarce urban contexts demand the adaptive integration of various data sources, modelling techniques and decision tools. The study will combine approaches evaluating, adapting and synthesising high-resolution rainfall forecasts, ensemble techniques and context-specific decision frameworks.

1.7 THESIS OUTLINE

This chapter introduced the concepts and reviewed the literature on pluvial flood forecasting and decision-making for anticipatory flood management, research gaps, research questions and objectives.

Chapter 2 introduces the study area and overall research methodology and briefly summarises the methods used in each component.

Chapter 3 presents a practical approach of incorporating critical rainfall thresholds, historical flood data (crowdsourced) and ensemble precipitation forecasts for forecasting extreme rainfall.

Chapter 4 addresses the concepts of whether increasing the resolution of WRF rainfall forecasts will lead to improved flood forecasting in urban pluvial catchments, a problem which has not been explored sufficiently, especially in data-scarce cities. Additionally, it explores how the difference in scale influences the different flood forecasting approaches.

Chapter 5 explores the skill of high-resolution ensemble precipitation forecasts to be used in a probabilistic risk-based decision-making framework (Chapter 6).

Chapter 6 presents a framework for using a probabilistic decision framework using the output of Chapter 5. Bayesian decision theory is used to trigger spatially varying anticipatory flood decisions using a rainfall threshold method.

Chapter 7 synthesises and reflects on the research findings and proposes areas for future research.

2

RESEARCH METHODOLOGICAL FRAMEWORK AND STUDY AREA

2.1 RESEARCH METHODOLOGY

This research aims to determine how flood forecasting and decision support for local flood-anticipatory actions can be improved by better aligning data quality and quantity, flood forecasting approaches and the incorporation of uncertainty. To answer the main research questions (**Chapter 1.6.1**), four research objectives have been proposed. The research adopts a sequential approach, in which each outcome serves as input to the next objective, thereby integrating data, methods and limitations.

Objective 1 - To investigate how available data can be combined to forecast floods in a limited urban data context. (**Chapter 3**)

Objective 2 - To evaluate the suitability of high-resolution rainfall forecasts at different scales for use in different flood forecasting approaches in data-scarce cities (**Chapter 4**)

Objective 3 -To assess the performance of ensemble rainfall forecasts and deterministic forecasts using categorical and probabilistic metrics for different events. (**Chapter 5**)

Objective 4 - To develop and evaluate a Bayesian decision-making framework using uncertainty forecast information to support anticipatory flood action. (**Chapter 6**)

An overview of the research methodology, which aligns with the research objectives, is summarised in Figure 2.1. A more detailed description of the methods for each research objective is given in their respective chapters; however, the same data and models overlap in some of the chapters. Therefore, the following sections provide an overview of the data and the theoretical background of the models in this research.

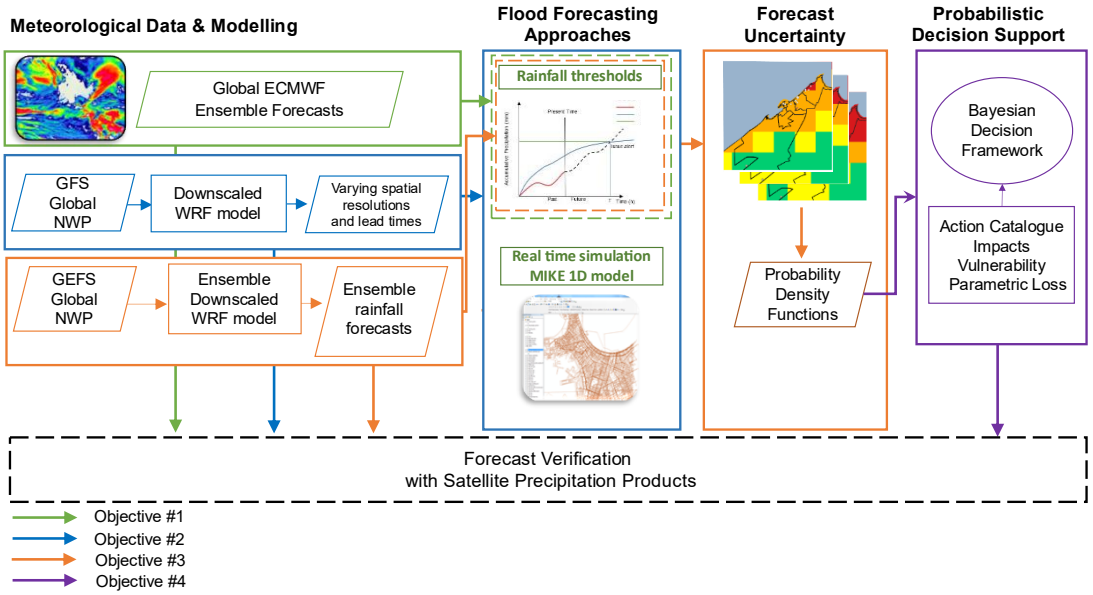


Figure 2.1 Proposed research methodology overview

2.2 STUDY AREA

Alexandria is located in the Alexandria Governorate in the Nile Delta, on the southern boundary of the Mediterranean Sea (Figure 2.2). A renowned tourist destination, the city is known for its cultural heritage and landmarks. It is the second-largest city in Egypt, spanning over 2,300 km2 and with a population of approximately 5.5 million (UN, 2023) and comprises smaller districts. Like many Middle Eastern and North African cities, it has experienced urban expansion over the years and informal settlements account for a third of Alexandria's total population (AASTMT and Egis BCEOM International, 2011).

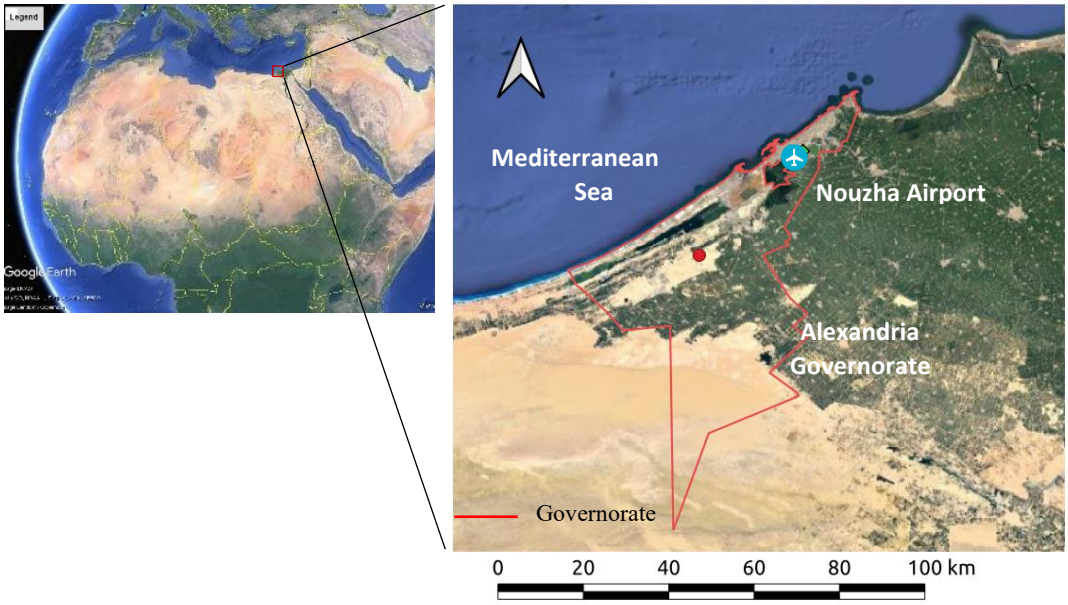


Figure 2.2 Geographical location of Alexandria Governorate, indicated by red lines.

Climate

The coastal city is characterised by an arid Mediterranean climate with rainy winters, occasional storms (from October to April) and long warm summer months (May to September) with no rain. There is high annual rainfall variability, with 360 mm (max) and 70 mm (min), averaging 195 mm/year. In the winter months, Alexandria is characterised by high temporal rainfall variability, a distinguishing characteristic of the Mediterranean climate (Hasanean, 2004). Winter storms, locally referred to as "Al-Nawat", are the result of migratory cyclones and fronts, upper-level troughs (low pressure) accompanied by strong winds, rain and storm surges, or strong winds only.

Alexandria City was selected as the case study, given the increased number of flood events in recent years and the challenges faced with developing a flood forecasting system with limited data and resources. Alexandria suffers from occasional pluvial flooding due

to runoff accumulation and surcharging of the city's dense sewer network and these risks are expected to increase with a changing climate (AASTMT et al., 2011).

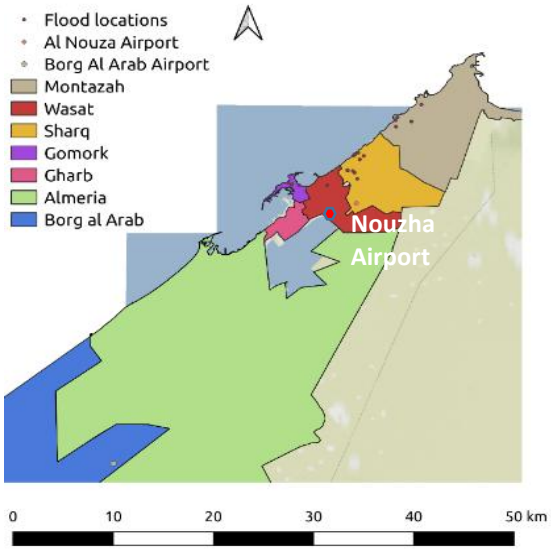


Figure 2.3 Geographical location of Alexandria City and districts. The El Gomork district is shown in purple.

Drainage and Sewer Infrastructure

The system receives domestic and industrial wastewater, as well as stormwater. Recent data from the Alexandria Sanitary Drainage Company (ASDCO) indicate approximately 93.4% of the urban area is connected to the sewer system and the system has a capacity of roughly 1.6 million m3/day (Zevenbergen et al., 2016). During the summer, the system's load can double due to tourists, but the summer does not coincide with the storm period. Due to the city's rapid urbanisation over the years, much of the aged drainage and sanitation infrastructure has become increasingly overwhelmed. The system comprises a complex interaction of sewer networks, catch basins, manholes, pumps, force mains, treatment plants and drainage channels before being discharged into Lake Maryut. The El Gomork district (Figure 2.3) experiences occasional flooding and traffic disruptions along the Corniche. Stormwater is collected via a combined sewer system and pumped to the West Treatment plant via lift stations and force main pipes. The sewer system is designed for an arid climate and has a 2-year drainage design of 26mm/day (AASTMT et al., 2011) and 10-year rainfall intensity of 9mm/h (ASDCO). Alexandria does not have an operational flood forecasting system and lacks reliable sub-daily rainfall gauge data. City officials currently use information from the National Meteorological Agency and historical knowledge of floods to issue warnings and take action. With the threat of

emerging risk, AFMA has been proposed as a viable solution to increase preparedness and reduce damage (Biswa Bhattacharya et al., 2018; Zevenbergen et al., 2016).

Land Cover and Land Use maps and elevation

Alexandria is considered very urbanised and densely populated city. Residential areas account for the majority of land use along with roads (Figure 2.4). Lake Maryut, used for navigation and aquaculture, also accounts for a significant part of the land cover and some agricultural areas are located on the city's outskirts. Zevenbergen et al. (2016) reported urban areas have increased by almost 40% (from semi-bare) over the past fifteen years, suggesting a significant reduction in open permeable areas.

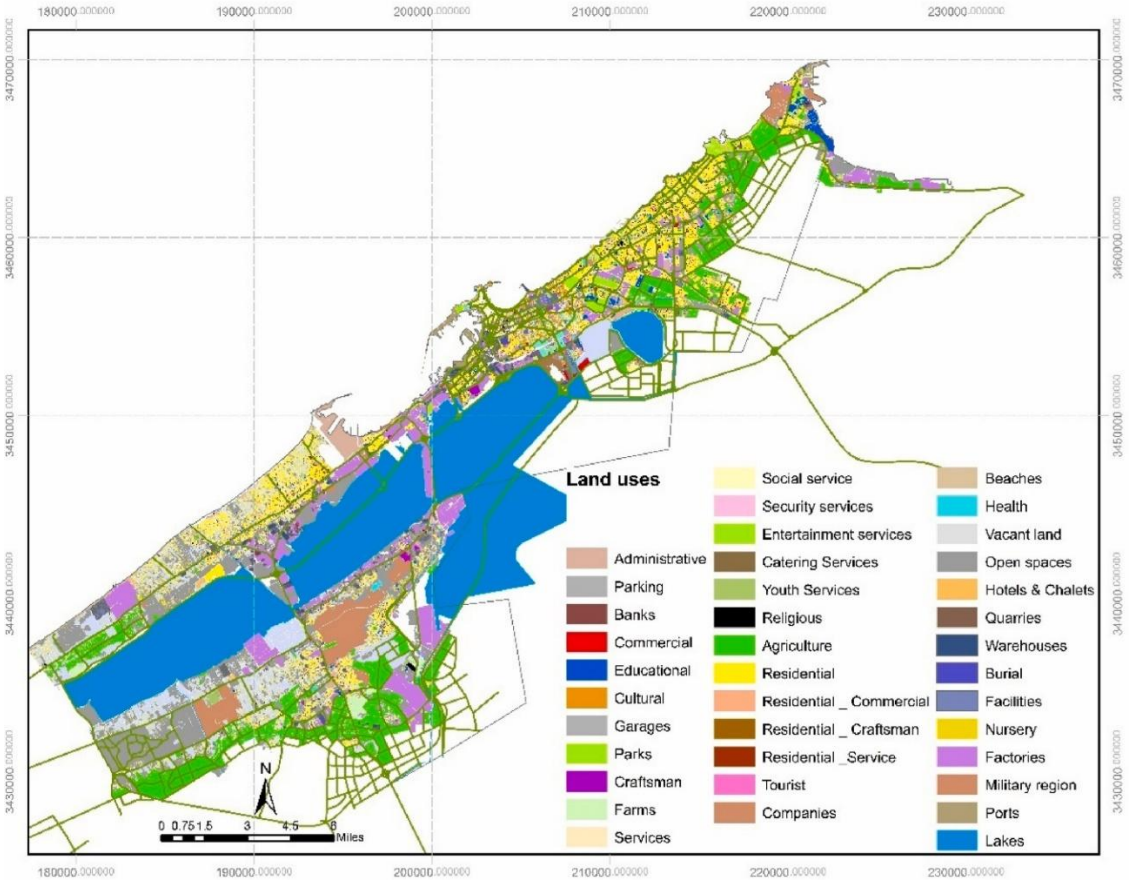


Figure 2.4 Land use classifications for Alexandria 2023 showing a high designation of residential areas in the city. Source: (Mabrouk et al., 2023)

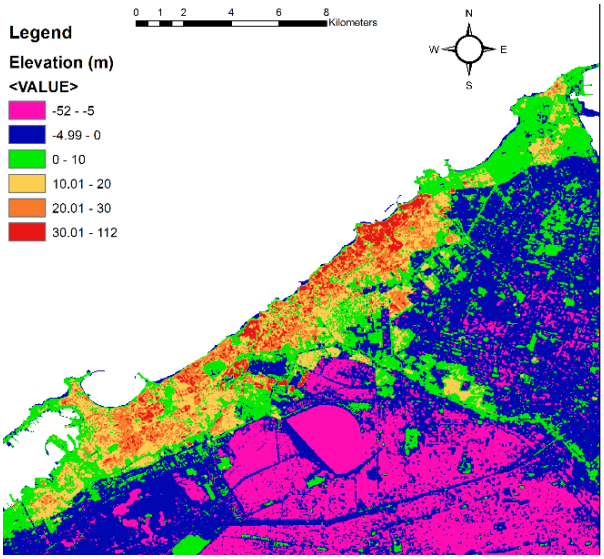


Figure 2.5 Digital elevation model of Alexandria City. Resolution 30m. Source: Japan Aerospace Exploration Agency (JAXA)

Alexandria is characterised by irregular hills in the southern parts, with elevations ranging from 0 to 40 meters above mean sea level (UNISDR, 2010) and higher elevations in some coastal areas. As the city moves farther from the coast, the topography is characterised by low-lying areas below mean sea level, with a significant portion 5m below sea level (Figure 2.5). There are no known open-source high-resolution Digital Elevation Models (DEMs) (less than 5km) available for Alexandria**.**

Flood impacts

Alexandria City was selected as a case study given the increased number of flood events in recent years and the challenges faced with developing a flood forecasting system with limited data and resources. Alexandria suffers from occasional pluvial flooding due to runoff accumulation and surcharging of the city's dense sewer network and these risks are expected to increase with a changing climate (AASTMT et al., 2011). For example, the El Gomork, El Montaza and Smouha districts (Figure 2.3) experience occasional flooding and traffic disruptions. Alexandria does not have an operational flood forecasting system and lacks reliable sub-daily gauge data. City officials currently use extreme weather warnings issued by the national meteorological agency and historical knowledge of floods to issue warnings and take action. With the threat of emerging risk, AFMA has been proposed as a viable solution to increase preparedness and reduce damage (B. Bhattacharya et al., 2018; Zevenbergen et al., 2016).

Current Forecasting Operations and Flow of Information

The Egyptian Meteorological Agency (EMA) is tasked with forecasting extreme weather events and issuing warnings for the entire country of Egypt, including the North Coast. They use several global numerical weather prediction models, such as ECMWF and their

own downscaled limited-area models, such as the Weather Research and Forecasting Model. The forecast resolution ranges from ~0.1 to ~0.5 °, depending on the product. They monitor several proxy variables (geopotential height, jet stream, atmospheric pressure) as indicators of severe weather along with the forecaster's experience and expertise. Forecasts are qualitative and descriptive, indicating the movement, severity of weather systems and possible flood impacts (based on past experience) without specifying rainfall magnitude. The Water Resources Research Institute (WRRI) also have a mandate for flash flood forecasting in Wadi systems under the National Water Research Centre (NWRC) of the Ministry of Water Resources and Irrigation (MWRI). WRRI also issues forecasts that are used by ASDCO, but as of recently, they are no longer issued publicly on social media.

Discussions were held with a representative of the EMA to understand what and how information is shared with decision-makers and its suitability for flood forecasting and decision-making in Alexandria. The authority monitors various weather elements through its stations in all governorates of the Republic, including surface monitoring stations, stations for monitoring the upper atmosphere, stations for monitoring air pollution, solar radiation and ozone, as well as receiving outputs from global and local numerical models and satellite images from different channels. The flow of information is structured so that the EMA disseminates forecasts via several public and private channels, including a WhatsApp group with Alexandria's Governor and ASDCO. This is considered the most convenient way of sharing information. Warnings are also issued to the Alexandria Port and marine vessels.

While the EMA issues rainfall likelihoods, this is based on experience and does not produce ensemble quantitative precipitation forecasts (QPF). They would like to use products in the future. The facility has recently acquired radar capabilities on the North Coast. The radar system is currently undergoing testing and plans are to implement nowcasting with a 1-2 hour lead time. While the current products are sufficient for national-scale forecasts, they remain relatively coarse for cities such as Alexandria.

Understanding Emergency Managers' Needs in Alexandria

In Alexandria, different stakeholders include the ASDCO, the Governorate, MWRI, road and traffic operators, electricity providers and local NGOs, who must work together to reduce the impacts of flooding, directly or indirectly. ASDCO and the Governorate receive forecasts simultaneously via a WhatsApp group from the EMA and MRRI. ASDCO is responsible for all actions related to the sewer network; the Governorate is responsible for declaring a state of emergency and school closures; while the Ministry of Irrigation monitors the surrounding drainage canals, lakes and pumping at El Max point. After the 2015 flood event, ASDCO closely monitors Satellite images with a typical latency of 3 hours and uses forecast information from both the EMA and MWRI. The

more extreme outlook is used to trigger an action. The MWRI maps show the 24-hour rainfall accumulation over the next three days.

Discussions with ASDCO representatives gave insight into the types of actions they take, the information they use and need to take actions and the local knowledge of past events they use to issue warnings. In Alexandria, the challenge continues to be how to translate national-scale warnings into local-scale actions. The representative of ASDCO has stated, *"Yes, we use the forecasts. The challenge is that sometimes the forecast is not very accurate and not very spatially representative of the study area. It does not give the intensity and the locations of rainfall. It is very general".* The information currently received is useful to an extent, as it gives an indication of when there will be no rainfall and the possibility of an impending extreme event, but it is not sufficient. ASDCO would like information on the intensity, location and flood depths, linked to GIS layers that can be used to prioritise where and which actions to take. The actions taken by ASDCO include the hazard-reducing measures shown in Table 2.1. Twenty-four to seventy-two (24-72) hours was also expressed as the ideal lead time and the representative responded positively when asked whether knowing the likelihood or probability of flooding and its impacts would help them make decisions.

From the discussion, the type of actions taken by ASDCO includes hazard-reducing measures shown in Table 2.1:

Table 2.1 Examples of emergency management actions taken by emergency managers before a rainfall event

	List of Actions		**List of Actions contd..**
1.	Staff and equipment preparedness / putting staff on stand-by, coordinating and increasing staff before an event	7.	Emptying lift station wells
2.	Issuing warnings to the public	8.	School cancellations
3.	Deployment of staff to respond	9.	Road closures
4.	Dispatching of suction trucks to known hotspot locations before event.	10.	Sourcing additional pump trucks.
5.	Checking and cleaning of known hotspot manholes and lift stations	11.	Issuing a state of emergency
6.	Deployment of temporary structures	12.	Lowering levels in irrigation drainage canals to receive additional stormwater

As of 2021, ASDCO has 121 suction trucks (Figure 2.6). In one instance, it was reported that over 180 trucks were dispatched to different locations as ASDCO can also request trucks from private companies, neighbouring Governorates and the Army in extreme circumstances. Figure 2.7 shows pump trucks stationed at hotspot locations.

Although the Governate continues to implement emergency management actions, they have also mentioned that emphasis is being placed on permanent solutions and strategies, including separating sewer and stormwater.

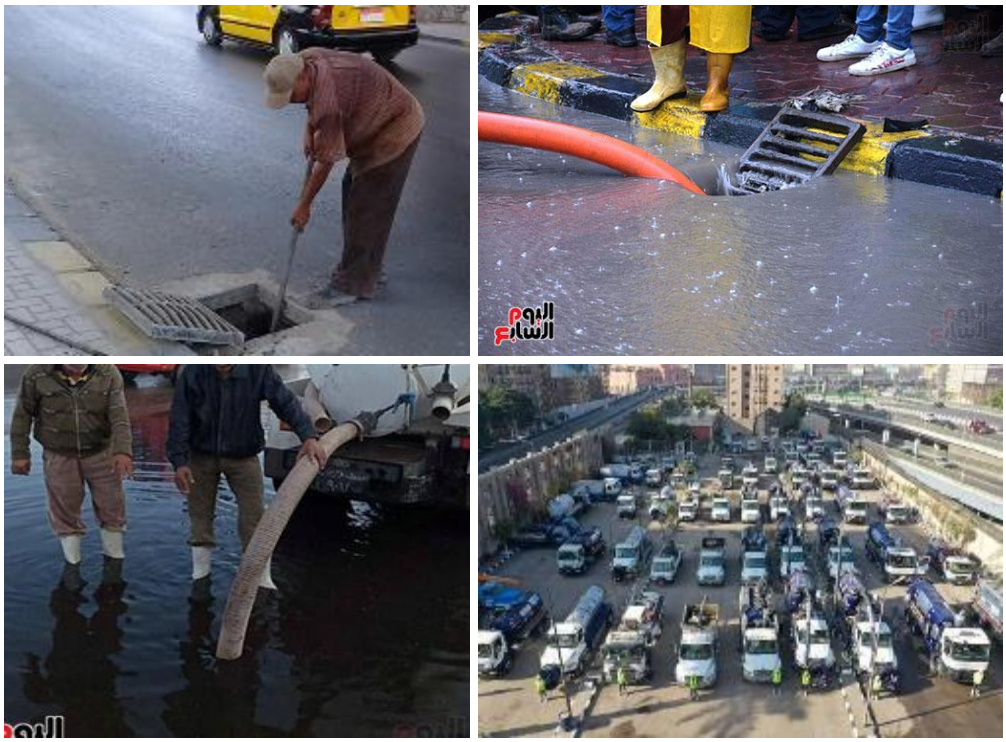


Figure 2.6 Cleaning of drainage inlets in preparation for the storm season (topleft) Removal of water from manholes during and after an event (topright and bottom left) the fleet of suction vehicles dispatched by the ASDCO (bottom right) Source: https://www.youm7.com/

Figure 2.7 Many hotspots are in tunnels (tunnel exit), Nov 20 2021 Source:(YouTube: Chave Weather-daily videos)(left) Hose at the same location. Tunnel entrance (right.

The Governorate chairs an emergency meeting with the Executive authorities. They issue states of emergency, including requests for citizens to stay home and limit movement and school closures to reduce traffic and allow suction trucks to remove water. The flow of information for warnings and emergency actions are shown in Figure 2.8. The preparatory actions taken by the Governorate and ASDCO are well documented in national newspapers. ASDCO and MWRI work together to take actions to lower water levels in drainage channels and to operate drainage and flow control structures. Many of the same actions are part of the standard emergency protocol and are taken once the EMA issues a severe weather warning. However, what varies is the number of trucks dispatched, the location, the amount of water being discharged, the level to which channels are reduced and the severity of the extreme weather warning.

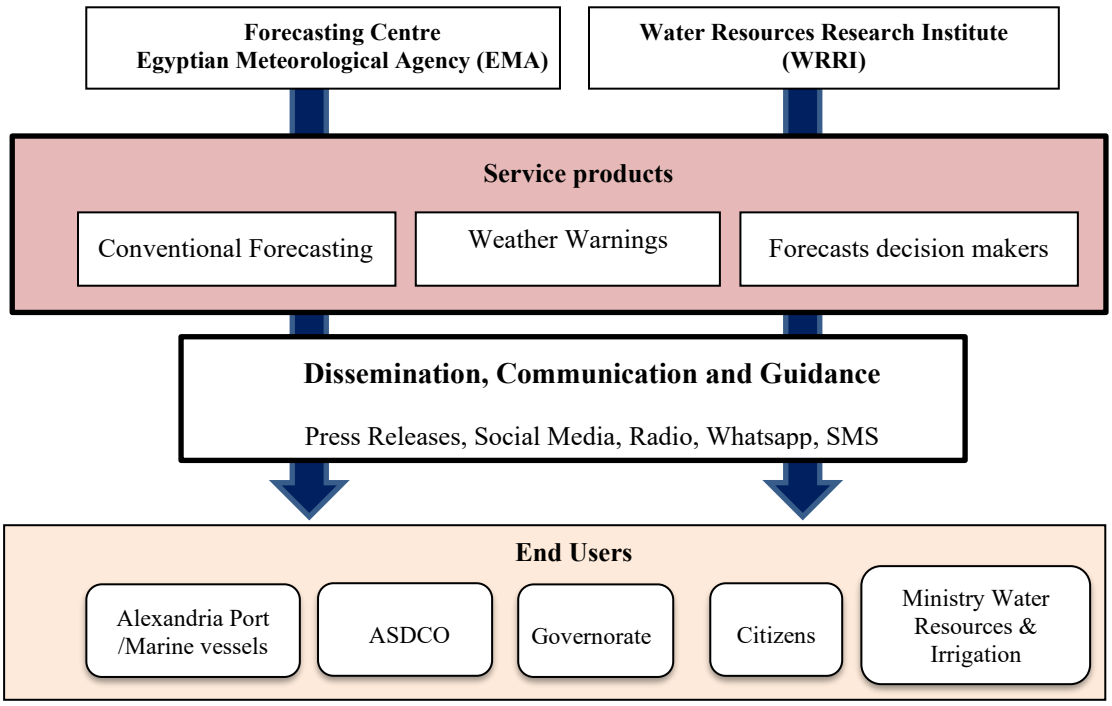


Figure 2.8 Summary of the flow of information for warnings and emergency actions in Alexandria city

2.3 DATA

Both historical observed gauge rainfall and satellite precipitation products were used in this research. Historical gauge rainfall records from Nouzha Airport were sourced from the open, online source Tutiempo and verified against the National Centre for Environmental Information, Global Historical Climatology Network daily (GHCNd). This data source has the advantage of a relatively long dataset from 1957 to present. Tutiempo also reports hourly descriptive rainfall data, which were used for dichotomous verification of rainfall and for identifying the timing of rainfall events.

Forecast rainfall is defined as an area-averaged grid, whereas observed gauge rainfall represents point accumulations and does not fully represent the rainfall spatial variability. In data-scarce areas, rainfall gauge data is often available only as daily totals and are sparsely distributed. Satellite-based gridded precipitation products have been used to represent observed precipitation in poorly or ungauged basins (Arias-Hidalgo et al., 2013). The resolution of satellite precipitation products has progressively improved with products such as Integrated Multi-satellitE Retrievals for GPM (IMERG) Final Precipitation Products V7. (Huffman et al., 2019) which was launched in 2014 and Multi-Source Weighted-Ensemble Precipitation MSWEP V2 which merges several gauge, satellite and reanalysis-based data (Beck et al., 2019). GPM is the continuation of

Tropical Rainfall Measuring Mission (TRMM) and the IMERG Final product has a latency of approximately 3.5 months. Both products have a spatial resolution of 0.1 degrees and the highest temporal resolutions of half-hour and three hours, respectively. Both datasets have been found suitable for detecting rainfall events in North Africa (Mekonnen et al., 2023; Nashwan et al., 2019). A summary of rainfall products used for this research are shown in Table 2.2**.**

Table 2.2 Summary of precipitation sources observed and forecast and different spatial and temporal resolutions used in this research

Data Source	Spatial resolution	Temporal resolution	Temporal Coverage	Type
Tutiempo GHCNd	Point	Daily	1957- Present	Quantitative and qualitative observations
Multi-Source Weighted-Ensemble Precipitation (MSWEP)	0.1 ~ 11km	3 hourly	1979- 2020	Reanalysis Gauge calibrated
The Integrated Multi-satellitE Retrievals for GPM GPM-IMERG V 7 Final	0.1 ~ 11km	30mins	2000-06-01 to Present (3.5 month latency)	Satellite Precipitation Gauge calibrated

2.4 HIGH-RESOLUTION METEOROLOGICAL FORECASTING MODEL

The non-hydrostatic dynamic-core Advanced Research Weather Research Forecasting (WRF-ARW) model, version 4.3, was used to downscale gridded global weather models for past events (hindcasting). WRF-ARW is an open-source mesoscale convection-permitting dynamic downscaling tool that integrates the compressible, non-hydrostatic Euler mass solver for atmospheric research and operational use. The ARW system uses an Eulerian solver to solve ordinary differential equations. Other features include initialisation capabilities, lateral boundary condition options, grid nesting, parameterisation physics options and data assimilation (Skamarock et al., 2008). As a numerical model, WRF uses the atmospheric state as initial conditions and iterates over timesteps to solve the dynamic weather equations that govern future weather states. The initial conditions are discretised onto a grid using a third-order Runge Kutta method for temporal separation and an Arakawa C-grid staggering scheme for spatial discretisation. (Skamarock et al., 2008).

In addition to atmospheric conditions, lateral boundary conditions and lower boundary conditions for terrain, land use and sea surface temperature, static geographical data were

used. Although referred to as the WRF model, the workflow consists of WRF Pre-processing systems (WPS) and other sub-processes to prepare the terrestrial data, terrain and global forecast (boundary conditions) to define the WRF grid. Geogrid.exe is used to interpolate static terrestrial data, Ungird.exe decodes met data from the forcing date and Metgrid.exe interpolates data horizontally onto the model domain. Real.exe interpolates data onto the vertical grid and the output is used as input to the wrf.exe model. WRF also has nudging capabilities using Four-Dimensional Data Assimilation (FDDA), which nudges the forecast toward observations; however, this was not explored for this analysis. A summary of the WPS and WRF workflow is shown in Figure 2.9.

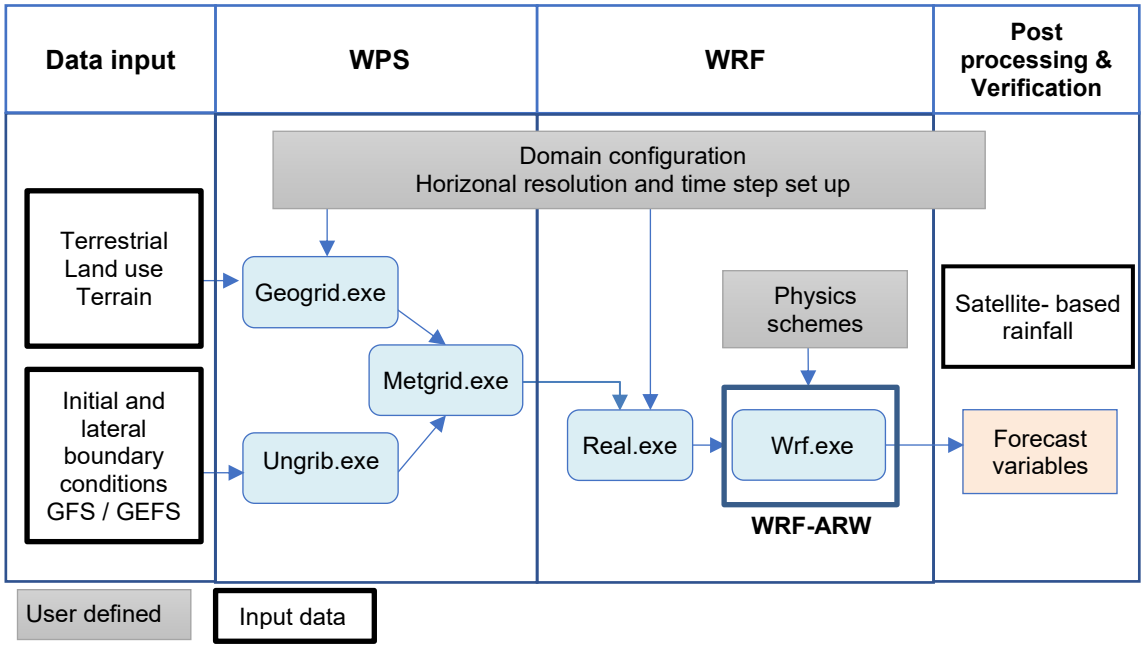


Figure 2.9 Summary of WPS and WRF workflow components

WRF Model Setup

The WRF model is optimised to accelerate computations by parallelising processes and threads. The model is configured with MPI (Message Passing Interface) capabilities, which enable distributed-memory parallel computing for large-scale computations. The national high-performance computing system Snellius, operated by SURF in the Netherlands, was used to perform both deterministic and ensemble simulations. The system runs on a Linux-based operating environment, which was used to configure the WPS and WRF models before execution on the SURF nodes. Model runs were distributed across multiple computing nodes, each equipped with dual AMD Epyc Rome processors (64 cores per processor, 128 cores per node) and 256 GB of RAM, enabling large-scale parallel computing for different parameterisation domain configurations and multiple runs of ensemble forecasts.

Model Parameterisation

Different physics schemes were configured for this analysis using grid parameterisations and physics schemes (Table 2.3). The selection of physics schemes was guided by Y. Liu et al. (2021), who optimised the selection for Egypt. The parameterisation process involves selecting physics schemes to resolve small-scale or complex features by representing them as variables at scales the model can resolve, simplifying the process. The microphysics represents grid-scale cloud and rainfall processes, while sub-grid scale "convective rain" is resolved using the Cumulus Parameterisation (CP) schemes. The radiation scheme provides atmospheric heating due to longwave and shortwave radiation. The Planetary Boundary Layer (PBL) scheme distributes surface fluxes through boundary-layer eddy fluxes and allows for PBL growth via entrainment (Janjić, 1994). The land surface model (LSM) provides heat and moisture fluxes over land to provide a lower boundary condition for vertical transport in the PBL scheme.

Table 2.3 Summary of parameterisation schemes to be used in the WRF ARW model

Scheme name	Model	Scheme#	Reference
Microphysics	Single-moment 6-classWSM6	6	(NOAA, 2001)
Longwave Radiation	RRTM	1	(Iacono et al., 2008)
Shortwave Radiation	RRTM	1	
Planetary Boundary Layer	Mellor–Yamada–Janjic	2	(Janjić, 1994)
Surface layers	Eta Similarity	2	(Jiménez et al., 2012)
Land surface model	Unified Noah Land Surface Model	2	(Tewari et al., 2004)
Cumulus	Grell Freitas Ensemble	3	(Grell and Freitas, 2014)

Forcing data and domain configuration

Analysis was carried out for both deterministic and ensemble forecasts. For deterministic forecasts, global atmospheric forcing data were sourced from the National Centres for Environmental Prediction Global Forecast System (GFS). Global Forecast Grids Historical Archive ds084.1 for lateral boundary conditions and MODIS15 land-use map. Data is available at 0.25° and 3-hourly intervals (Figure 2.10). All spatial domains had a vertical resolution of 50 layers. For ensemble runs, the WRF model was run 21 times using the Global Ensemble Forecast System (GEFS) 21-member ensemble as lateral boundary conditions. Ensemble data is archived at 1 ° and 0.5 ° from 2017 to present, with 6-hourly temporal intervals. All spatial domains had a vertical resolution of 45 layers. Models were run at different lead times,72, 48, 24 and 12 hrs and 1hr output. A 12-hour

spin-up period was used to balance the inconsistencies that may arise between the model physics forecasts and the initial and boundary conditions provided by the forcing data.

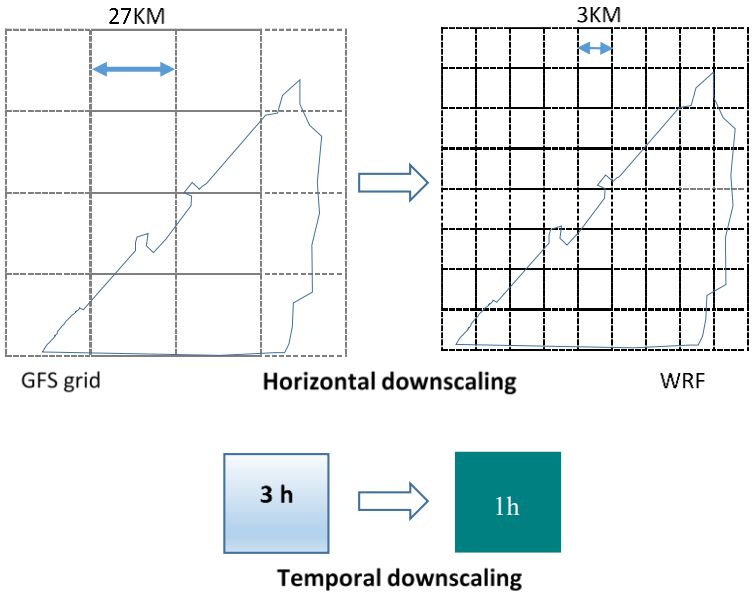


Figure 2.10 Example of horizontal and temporal downscaling

The deterministic and ensemble models were nested across different domain configurations and sizes, as specified by the forcing data. Further details on the specific model setup and domain configuration for the deterministic and ensemble forecasts are available in **Chapters 4 and 5**, respectively.

2.5 RAINFALL VERIFICATION METHODS

Several Verification methods are used,

In Chapters 3, 4, and 5, categorical scores are used to evaluate a dichotomous forecast, i.e. rain or no rain or flood or no flood or flood with different classes. Categorical verification measures such as Probability of Detection (POD), False Alarm Ratio (FAR), Critical Success Index (CSI) and Accuracy were used to discriminate between warning classes related to the observed event category. The Accuracy metric was used to measure both events and non-events. A 2x2 contingency table was used to quantify hits, misses, false alarms and correct negatives into a single value score.

Table 2.4 2x 2 Contingency table used to assess categorical scores

	Observed	Yes	No
Forecast	Yes	Hit	False Alarm
	No	Miss	Correct Negative

Categorical descriptions of a hit, miss and or false alarm are presented in Table 2.4. The Probability of Detection (POD), False Alarm Ratio (FAR) and Critical Success Index (CSI) and Accuracy were used as the performance indices and are defined as follows (Table 2.5):

Table 2.5: Summary of forecast verification metrics used

Hit Rate (Probability of Detection) =	$\frac{Hit}{Hit + Miss}$	*(2.1)*
False Alarm Ratio (FAR)=	$\frac{False\ Alarm}{Hit + False\ Alarm}$	*(2.2)*
Threat Score/ Critical Success Index (CSI) =	$\frac{Hit}{Hit + False\ Alarm + Miss}$	*(2.3)*
Accuracy =	$\frac{Hit\ + Correct\ Negative}{Hit + False\ Alarm + Miss\ + Correct\ Negative}$	*(2.4)*

Spatial Scores: Fraction Skill Score

Spatial Scores are introduced in Chapter 4 through the Fractional Skill Score (FSS). The FSS metric directly compares the fractional coverage of forecast and observed rainfall, which exceeds a specified threshold for a neighbourhood size. FSS is calculated for each neighbourhood size using Eq (2.5), where F*i* and O*i* are the forecast and observed rainfall fractions exceeding the specified thresholds, respectively and N is the number of spatial grids in the neighbourhood. An FSS of 1 is perfect, an FSS above 0.5 is considered useful and an FSS of 0 is not a skill.

$$\boldsymbol{FSS} = 1 - \frac{\frac{1}{N}\sum_{i=i}^{N}(F_i - O_i)^2}{\frac{1}{N}\left(\sum_{i=1}^{N}(F_i)^2 + \sum_{i=1}^{N}(O_i)^2\right)} \qquad (2.5)$$

Ensemble Rainfall Metrics

Continuous Ranked Probability Score (CRPS)

For this study, the reliability and the discrimination, the ability to differentiate between events and non-events were evaluated. The Continuous Ranked Probability Score (CRPS) is a measure of the difference between the cumulative distribution function (CDF) of the ensemble forecasts and the observation, also expressed as a single a score (Brown, 2016). The CRPS reduces to the mean absolute error for an ensemble forecast; therefore, it can be used directly to compare the accuracy of an ensemble forecast with accuracy of a deterministic forecast (Casati et al., 2008).

$$\mathrm{CRPS} = \int_{-\infty}^{\infty} \left(\mathrm{F_y(y)} \quad 1\{\mathrm{y} \geq x\}\right)^2 \mathrm{dy},$$

$$\overline{\mathrm{CRPS}} = {}^{1}/_{\mathrm{n}} \sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{CRPS_i} \qquad (2.6)$$

Where Fy (y) – Forecast variable, 1{y≥ x} is a step function that assumes probability of 1 for values ≥ to the observed and 0 otherwise, n- Number of forecast and observed pairs

Brier Score

The Brier Score was used to calculate the average squared error of the probability forecast for the event. It measures the average square difference between the forecast and observed probability, which is 1 if the event occurred and 0 otherwise (Brown, 2016). Usually, the sample is separated into 10 bins based on the forecast probability and the corresponding observation. A perfect Brier score will equal to zero. A smaller score is desirable which occurs when the forecast probability equals the observed frequency. The Brier score can be decomposed into three components: reliability, resolution and uncertainty, which are useful for exploring the dependence of probability forecasts on ensemble characteristics (Gofa, 2010).

$$BS = \frac{1}{N} \sum_{t=1}^{N} (f_t - o_t)^2 \qquad (2.7)$$

N = the number of observed and forecast pairs, ft = the forecast probability, ot - corresponding observation cumulative probability; (f_t -o_t) is 1 if f_t greater than the observed value o_t, otherwise the value is 0.

2.6 URBAN FLOOD FORECASTING MODEL

For research objective 2 (Chapter 4), the Danish Hydraulic Institute (DHI) MIKE Plus model was used to simulate flows and water levels in urban storm drainage systems. The 1D MIKE Plus model was developed for the El Gomork district (~5km^2) by Mahmood (2021). The downscaled WRF outputs at different resolutions were used to simulate flows over sub-catchments and routed through pipes or channel networks at various time steps.

It consists of a rainfall-runoff model that transforms rainfall time series to runoff hydrographs for input into the hydraulic network model (Figure 2.11). 1D model simulations were performed using only the deterministic rainfall forecast runs in Chapter 4.

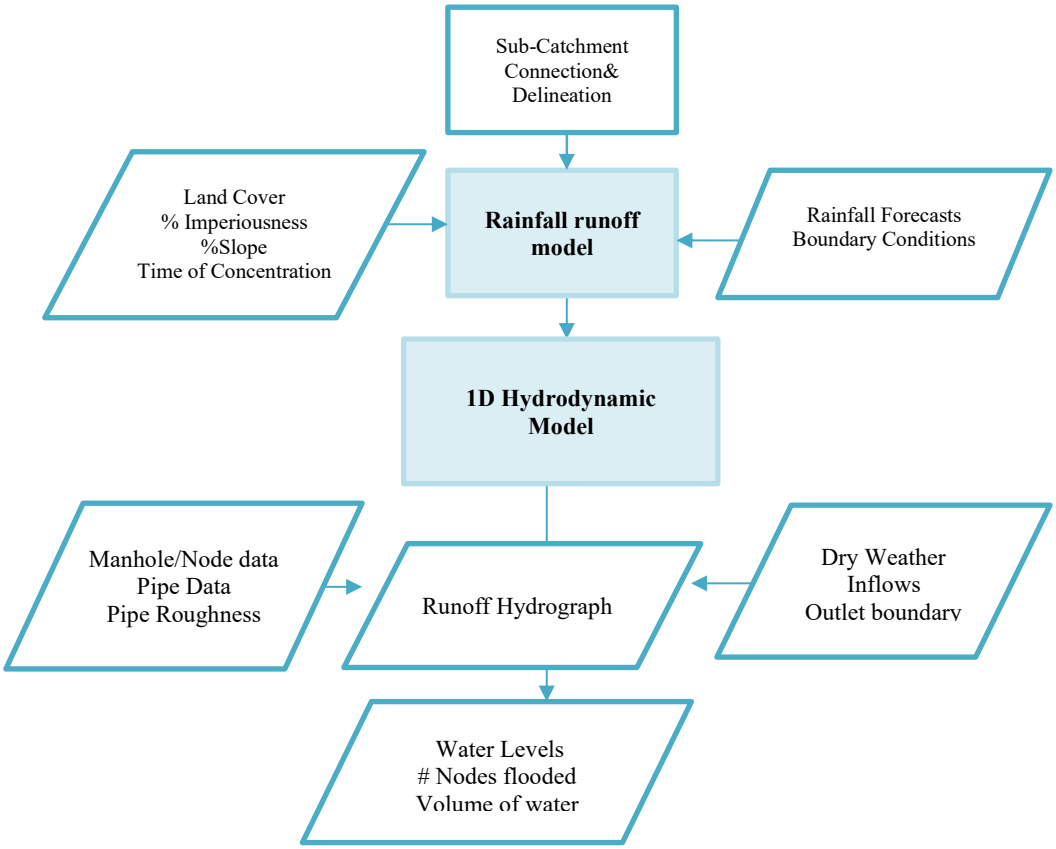


Figure 2.11 Summary of the Urban Flood Model Setup used in Chapter 4

The hydrodynamic model uses the MOUSE engine to solve the 1D St Venant equations, a coupled partial differential equation describing the continuity (Eq. 2.8) and momentum balance (Eq. 2.9). Under regular flow conditions, 1D pipe flow can be modelled as open channel/non-pressurised however the transitions from open flow to pressurised flow means the St Venant equations are no longer valid. Therefore, in order for the St Venant equations to remain valid, MIKE introduces a fictitious slot at the top of the pipe that generalises the St Venant equations (DHI, 2020). At the nodes, the water level is computed based on the water level from the previous time step and the flow contributions during the time step from each connected pipe and from externally connected flows, such as catchment runoff discharge. In the 1D model, flooding is assumed to occur when the simulated hydraulic head in the node exceeds the ground level. In that case, MIKE 1D introduces an artificial reservoir at the top of the node, with a surface area 1000 times that

of the node's surface (by default). The surcharged water is stored in the reservoir, to be returned back into the sewer. In that case, MIKE 1D introduces an artificial reservoir at the top of the node, with a surface area 1000 times that of the node's surface (by default). The surcharged water is stored in the reservoir, to be returned to the sewer (DHI, 2017). Additional details on the MIKE Plus model can be found in MIKE 1D Reference Manual for collection systems.

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = 0 \quad (2.8)$$

$$\frac{1}{A}\frac{\partial Q}{\partial t} + \frac{1}{A}\frac{\partial}{\partial x}\left(\frac{Q^2}{A}\right) + g\frac{\partial h}{\partial x} = g(S - J) \quad (2.9)$$

Where Q is the flow in the x-direction and A is the area of the channel cross section. For the momentum equation, h is the water depth, g is the gravitational acceleration, S is the bed slope and J is the friction slope.

1D Hydraulic model Set up and Drainage network

The Gomork network comprises 3482 local and collector manholes connected by sewer pipes ranging in diameter from 300 to 2000mm. Pipes flow via gravity to two sanitary lift stations: Outlet #1 (Kaytbay lift station) and Outlet #2 (El Meena). The combined sewage is then pumped from the lift stations and treated at the West Treatment plant before discharging into Lake Maryut. The model setup and the locations of the West Treatment Plant and Lake Maryut are shown in Figure 2.12. Tides and wave action are not considered to be contributing to flooding in the model setup.

Manholes covers were set as normal to allow water to spill if flooded. Pipes in Alexandria are vitrified clay. A Manning roughness value of 0.013 was used, which is consistent with closed conduits flowing partly full for sanitary sewers coated with sewage slime with bends and connections (US Army Corps et of Engineers, Hydrologic Engineering Centre, 2020).

Alexandria city is a very densely urbanised area with buildings and paved roadways. A land cover GIS dataset comprising buildings, roads and green infrastructure was used to derive imperviousness for the rainfall-runoff model. The model was calibrated by Mahmood (2021) using available drainage network data and satellite precipitation data sets MSWEP and GPM IMERG.

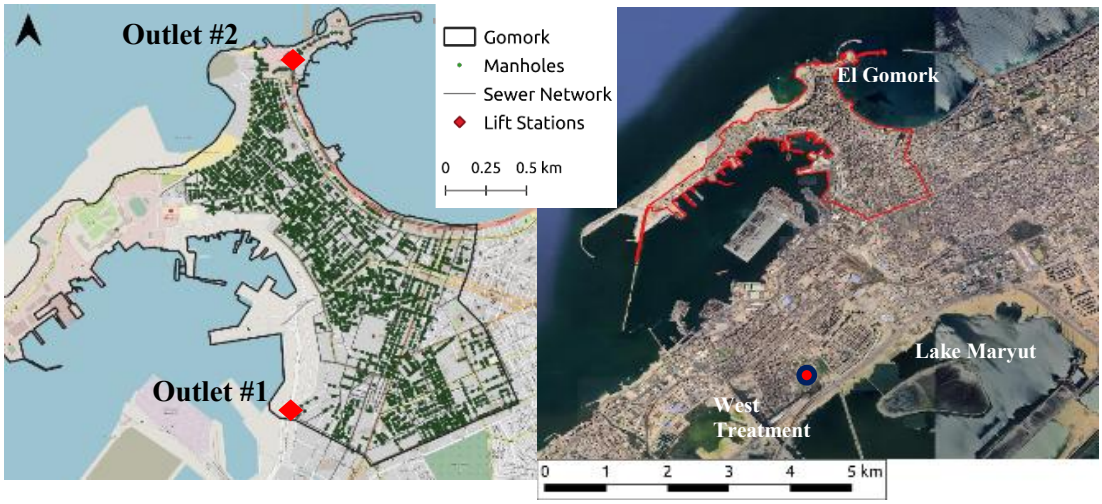


Figure 2.12 Mike Urban plus drainage model set up for El Gomork district showing manholes, sewer pipe layout and location outlet lift stations (left) which pumps combined stormwater and sewage to the West Treatment Plan

Catchment delineation and processing

Catchments were delineated using point layer (Thiessen polygons)-Voronoi partitioning and manhole nodes used as the points for creating the Thiessen polygons. For catchment processing, catchment connections were established from the centroid of each catchment to the nodes. Available GIS layers were used to calculate the total surface imperviousness of each catchment's corresponding land covers. For buildings impervious was assumed to be 95%, roads: 80% and green areas: 5% based on the Urban Storm Drainage Criteria Manual (Mile High Flood District (MHFD), 2024)A composite percentage impervious value for each catchment was determined based on the proportion of landcover in the catchment.

Kirpich's (1940) formula used for calculating Time of concentration of catchment the study area. In the absence of a high-resolution DEM file, percentage slope was calculated from a 1m DEM surface was created using the manhole ground level elevations. For the surface runoff model, the Time Area method was used for simplicity. MIKE Plus uses a formula that covers curve shapes in-between the three standard ones The curves were chosen for different catchments on the basis of calculating 'Area-50, as the part of the catchment that is within the distance corresponding to 50% of the time of concentration Mahmood (2021. The curves are specified by giving the Time Area coefficient directly instead of specifying a Time-Area curve (DHI, 2017). More details on the urban model is described in Mahmood (2021).

Initial condition, boundary conditions and model calibration

To establish initial conditions, dry weather flows were used based on the daily flows distributions and model calibrated to determine the average actual flow. Population densities were used to derive the daily flows (Table 2.6). El Gomork districts has a population of ~98,000 in 2023 and is further broken down into population densities which were used as an indicator of the dry weather flows. Typical Dry weather flows were adapted from (Butler et al., 2018) and shown in Figure 2.13 Daily flows litres per person per day used for determining the dry weather flows. Details of the lift station operating rules and wet well capacity are unknown however it is known that pump has a maximum capacity of 0.9m^3/s and actual daily operating capacity of 0.25m^3/s.

Table 2.6 El Gomork sub-district population density used to derive dry weather flows

Subdistrict Name	Population Density	Areas	Number of Nodes per area
El Gomrok	61,355/km²	1.605 km²	1734
El Manshiyah	38,341/km²	0.6523km^2	1138
Al Labban	42000/ km²	1.042km^2	610

Source: Central Agency for Public Mobilisation and Statistics, Egypt sourced from citypopulation.de.

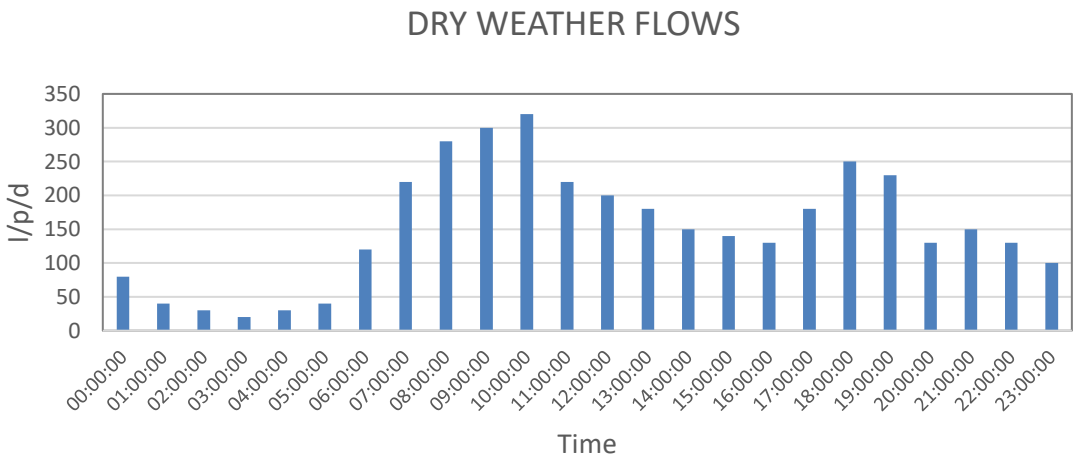


Figure 2.13 Daily flows litres per person per day used for determining the dry weather flows

2.7 A BRIEF INTRODUCTION TO BAYESIAN DECISION THEORY

In Chapter 6, the main objective is to explore how forecast uncertainty can be incorporated into decision frameworks to support anticipatory flood management. A Decision Framework based on Bayesian Decision Theory (Economou et al., 2016; Krzysztofowicz, 1999; Todini, 2017) was developed for the decision and action component of the flood chain.

Decision support tools exist in different forms which can help us to gather, analyse and visualise data, model-based outcomes and likely impacts, stimulate scenarios or communicate uncertainty, all with the goal of providing supporting information to decision makers. In theoretical terms, decisions can be represented in a decision space by alternatives, states of nature, outcomes of the decision and their value, expressed as cost or profit. Alternatives refer to different actions/ strategies that can be implemented. States of nature refer to the possible outcomes of events for example, the future evolution of a weather event and associated flood levels which are beyond the control of the decision maker. The outcome is the consequences of choosing an alternative based on the state of nature.

Bayesian Decision Theory is based on a normative prescriptive decision theory. It provides a systematic probabilistic approach using Bayes' Theorem (Eq. 2.10) to define the posterior probability and a utility function that describes the decision maker's preference for selecting the optimal action over other alternatives. It consists of posterior probability (probabilistic information conditioned on available information and the prior probability is the probability of a particular state of nature before any data or information is available. Bayes' theorem states the posterior probability is proportional to the product of the prior probability and the likelihood of the data given the state of nature. This method is based on choosing the action that maximises (or minimises) the expected utility (DeGroot, 2004; Savage, 1954). The expected value of an action is calculated by weighting the utility of each possible outcome by its probability. It supports informed decision-making under uncertainty by evaluating a finite set of alternative decisions and the full probability distribution, rather than selecting a critical probability threshold to trigger an action.

$$Bayes\ Theorm\colon P(x|y) = \frac{P(y|x)P(y)}{P(x)} = \frac{P(y|x)P(x)}{\sum_{j=1}^{n} P(y|x)P(x)} \qquad (2.10)$$

$$a*(y) = \underset{a\in A}{\operatorname{argmin}} \sum_x L[a,x]\, P(x|y) \qquad (2.11)$$

Mathematically, the decision framework can be represented as Eq 2.11 and visually as Figure 2.14 by the optimal action $a^*(y)$, likelihood $P(y|x)$ is the probability of observing future weather information y given the possible states of nature (x). $P(x|y)$ is the posterior probability of the future state given predictive future weather information, $P(y)$ and $P(x)$

is the prior probability. Selecting a suitable prior P(x) is also an important aspect of Bayesian decision theory. Priors represent initial probability (also known as beliefs), which are computed based on historical data such as occurrences of flood events, the past forecasts model performance or probability of observing rainfall more than a preset value such as 30 mm in a time window such as 3hrs. However, priors can be considered subjective and biased (Banner et al., 2020).

The loss function L(a, x) describes the losses incurred from taking an action a (such as opening a sluice or closing a gate) for state x. The expected loss is obtained by combining the loss function with the probability distribution of possible states. For discrete values, the expected loss is calculated as the sum of the products of the loss and the probability of each state. For continuous values, the expected loss is calculated as the integral of this product over all possible states. In the context of anticipatory actions, P(y) is the probability density function of ensemble forecasts.

In the case of flooding, the Decision Rule (DR) states that the optimal action a*(y) (Eq. 2.11) is the action that minimises the expected value of loss E(L[a,x]) associated with the risk of all states (i.e., the loss multiplied by its probability (Lindley and Smith, 1972). The advantage is in taking an action based on its risk instead of taking the decision based on the most likely state exceeding a probability threshold determined from the ensemble forecast (Harvey et al., 2022).

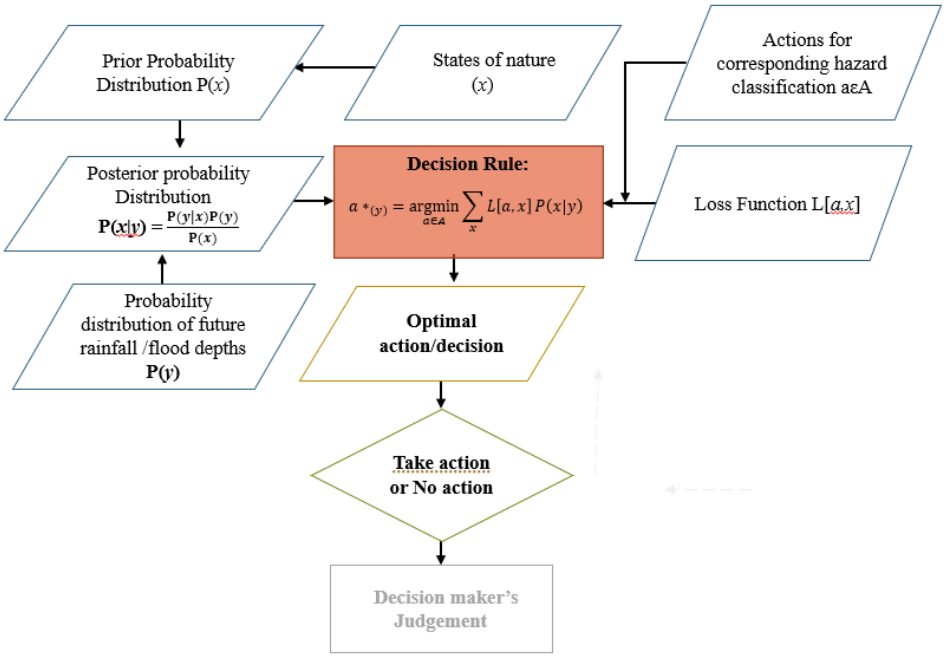


Figure 2.14 Proposed Bayesian Decision Framework used for taking actions utilising inputs of Posterior Probability Distributions, Loss Functions and corresponding hazards.

The framework was calibrated and evaluated using past events. Probabilistic forecast data was derived from the ensemble forecast produced in **Chapter 5**. A corresponding catalogue of actions was developed based on discussions with Alexandria decision-makers and vulnerability data was included to support equitable decision-making. The data descriptions and parameters used in the decision model are presented in **Chapter 6.**

3

A RAINFALL THRESHOLD APPROACH TO EARLY ACTIONS

[1]Rapidly expanding cities in the Middle Eastern and North African (MENA) region are at risk of flooding due to heavy rainfall, insufficient drainage capacity and a lack of preparedness and data to conduct the required studies. A low-regret Early Warning System (EWS) using rainfall thresholds is proposed as a cost-effective short-term solution. This study aims to utilise a probabilistic approach to characterise and predict urban floods by assessing critical rainfall thresholds likely to cause flooding combined with ensemble precipitation forecast in Alexandria, Egypt. Rainfall thresholds were inferred by associating observed rainfall and historical flood information sourced from social media and newspapers. Floods were classified in a colour-coded hazard matrix as no flood (green), minor flood (yellow), significant flood (orange) and severe flood (red). The probability of occurrence of hazard classes was derived by incorporating ensemble rainfall into the hazard matrix to jointly assess likelihood and hazard severity. Results from this study showed that three of four severe events analysed could have been predicted with a high likelihood up to 24 hrs before. The presented approach supports decision-making for issuing warnings and flood control actions with limited data and serves as a model for other data-scarce regions.

---

[1] This Chapter is based on Young, A., Bhattacharya, B. and Zevenbergen, C. (2021) 'A rainfall threshold-based approach to early warnings in urban data-scarce regions: A case study of pluvial flooding in Alexandria, Egypt', *Journal of Flood Risk Management*, 14(2), pp. 1–16. doi: 10.1111/jfr3.12702.

3.1 INTRODUCTION

Real-time simulation forecasts using real-time hydrodynamic and inundation models aim to predict flood extent, locations and impacts. These models demand high-resolution spatial data and copious input parameters into complex models associated with uncertainty and long computational times (Ochoa-Rodríguez et al., 2015). Pre-simulated scenario-based systems, utilise a catalogue of hydrodynamic simulations and the accuracy depends on both model complexity and input data (Henonin et al., 2013). In catchments of short response times and insufficient data, empirical methods offer a simplified approach to issue warnings and predict rainfall depths likely to cause flooding by directly comparing precipitation forecast with critical rainfall thresholds derived from examining rainfall accumulations from previous flood events (Falconer et al., 2009; Georgakakos, 2006; Martina et al., 2006; Parker et al., 2011; S. J. Wu et al., 2015). Comparing critical rainfall thresholds with ensemble rainfall forecasts sourced from NWPs has the additional benefit of providing timely forecasts while allowing decision-makers to incorporate forecast uncertainty. This simplified EWS approach provides an immediate, short-term solution for data-scarce regions that lack the technical capacity and resources to implement complex methods.

Few studies have demonstrated how rainfall thresholds have been used to predict pluvial floods in urban areas (Bouwens et al., 2018; Candela and Aronica, 2016; Jang, 2015) and fewer show how critical thresholds can be compared with probabilistic rainfall forecasts (Hurford, Parker, et al., 2012; Yang et al., 2016). This chapter aims to present a practical approach to incorporating critical rainfall thresholds, historical flood data (crowdsourced) and ensemble precipitation forecasts for forecasting extreme rainfall and flooding and to improve decision-making, especially in data-scarce regions or cities, in the genesis of developing EWS.

Alexandria city in Egypt was selected as the case study for this research. Like many other cities in the MENA region, they suffer from occasional flooding due to runoff accumulation and sewer surcharge, but lack an enabled EWS and are least prepared for the rising threat of floods. In the face of emerging risks, Anticipatory Flood Management (AFMA) has been proposed as a viable solution to increase preparedness and reduce damage (B. Bhattacharya et al., 2018; Zevenbergen et al., 2016).

Ibrahim and Afandi (2014) previously evaluated the use of the Weather Research and Forecasting (WRF) model to predict extreme rainfall in Egypt but did not evaluate the use of ensemble forecasts. The success of such an approach will allow an immediate short-term solution to emerging flood risk in cities in the MENA region. In the interim, cities can acquire more comprehensive data that will contribute to the development of more sophisticated flood forecasting systems, which are otherwise very data and computationally demanding in data-scarce regions.

3.1.1 Threshold-based early warning systems

The complexity of the detection and forecast methods depends on factors related to data availability, severity, frequency and vulnerability to flooding, as well as the lead times required to issue warnings (WMO, 2011). A simple forecasting system can be developed from rainfall forecasts and a historical account of past flood events, with no hydrological/hydraulic modelling (Henonin et al., 2013) and can be implemented at the regional, national, catchment and city scales.

The Flash Flood Guidance System (FFGS), an empirical approach, initially developed by the US National Weather Service (NWS), uses forecast rainfall depths that are likely to result in discharges associated with floods. Georgakakos (2006) studied FFG-based systems and found they have a high probability of detecting flash floods. Such alerts are issued based on different rainfall thresholds, considering the prevalent soil moisture condition. This system has been implemented in many countries, including the new Black Sea and Middle East FFGS and the Central American FFGS (WMO, 2019).

The "FLOODsite" project assessed the advantages of using rainfall thresholds as an alternative to traditional EWS (Borga et al., 2011) and demonstrated success in identifying several flash floods across Europe (Lorenzo Alfieri et al., 2012). Meteo-alarm is another rainfall threshold-based EWS, implemented in approximately 30 European countries and alerts are issued based on comparisons between precipitation and local thresholds. Also, at a regional scale, Alfieri and Thielen (2015) successfully proposed a European Precipitation Index based on simulated Climatology (EPIC), which is calculated using COSMO-LEPS ensemble weather forecasts.

On a national scale, England and Wales initially adopted ERA (Extreme Rainfall Alerts) and Flood Guidance Statement (FGS), which provided warnings of extreme rainfall based on intensities (depth/duration) likely to cause severe surface water flooding (SWF) in urban areas based on a 1 in 30 return period (Hurford, Parker, et al., 2012). The Second-Generation Surface Water Flood Risk Assessment (SWFRA) programme has replaced this system. It empirically estimates surface water flooding risk-based on a weighted score of rainfall probability, spatial extent, soil moisture deficit and proxies for urbanisation (Ochoa-Rodríguez et al., 2015).

Although the application of thresholds is similar, the methodology for applying rainfall threshold methods differs between urban and fluvial areas, particularly in terms of precision and required parameters (Jang, 2015). For example, FFGS thresholds are adjusted based on antecedent moisture conditions (AMC). However, urban areas tend to adopt static thresholds (Figure 3.1) that do not account for AMC; soil moisture becomes less significant as urbanisation increases (Bouwens et al., 2018; Falconer et al., 2009; Hurford, Priest, et al., 2012). The maintenance state of sewers and canals also becomes more critical.

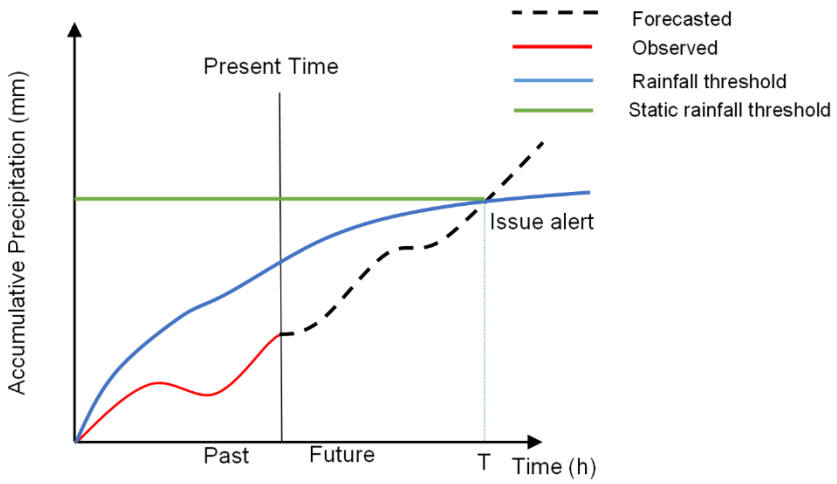


Figure 3.1 Example of rainfall threshold-based approach for the issuance of flood alerts. Modified from Martina et al. (2006) to include static rainfall thresholds

It is acknowledged there are limitations to using this approach as rainfall events of the same return period can produce floods of different extent and water depths; non-linearity between rainfall and floods (Stephens et al., 2015) due to spatial and temporal variability of rainfall, characteristics of the sewer system, topography, storage and runoff characteristics of the catchment (Simões et al., 2015). Current research is aimed at improving methods for determining critical thresholds and trigger definitions and at improving the accuracy and reliability of forecasts by incorporating catchment characteristics. In gauged basins, calibrated rainfall-runoff models and rainfall hyetographs are iteratively used to search for critical rainfall or discharge values (Candela and Aronica, 2016; Jang, 2015; S. J. Wu et al., 2015). Montesarchio (2015) compared methodologies for flood rainfall estimation, including Martina et al. (2006), which adopted a probabilistic approach by analysing the joint probability of rainfall duration and corresponding rainfall depth, combined with a Bayesian utility function. In smaller, semi-gauged basins, empirical methods use of historical data to determine rainfall thresholds and assume rainfall is distributed uniformly in time and space. Bouwens et al. (2018) correlated sub-daily rainfall with citizen incident flood reports and overflow pumping values to determine critical rainfall thresholds. Yang et al. (2016), successfully used a rainfall threshold approach and quantitative precipitation forecasts (QPFs) to evaluate urban inundation risk in Taiwan. This chapter capitalises on the simplicity of this methodology in data-scarce regions and on a local scale.

3.1.2 Early Warning Systems for Decision Support

Lead time refers to the minimum period of warning required for preparatory action to be effective (Carsell et al., 2004; Verkade and Werner, 2011). An important part of an EWS

is allowing sufficient lead-times to issue warnings, or take actions. However, uncertainty exists with predictions at longer lead-times. Probabilistic forecasts are key to quantifying uncertainty in forecasting floods and can be useful in assessing the likelihood of extreme events while providing more consistent successive forecasts (Lorenzo Alfieri et al., 2012; Boelee et al., 2019; Roberto Buizza, 2008; Cloke and Pappenberger, 2009; Dale et al., 2012; M.-H. Ramos et al., 2010; Todini, 2017). However, all authors agree there are several challenges including but not limited to (1) improving forecast with high-resolution data and assimilation; (2) coupling models at differing scales and real-time operational forecasting; (3) how end users can best use this information to improve decision-making for warnings and response. Still, compared to deterministic forecasts, probability forecasts allows an optimal balance between uncertainty and lead-time time (Verkade and Werner, 2011). When issuing warning messages or taking actions, decision-makers require clear informed guidance and decision rules to make confident decisions (Economou et al., 2016; van Andel, 2009).

Traditional yes or no extreme weather warnings are issued once the forecasted rainfall exceeds a critical rainfall threshold. However, such binary forecasts do not indicate the likelihood or severity of an expected hazard. The use of a colour-coded hazard or risk allows a visual expression of a priority matrix (WMO, 2015). When combined with ensemble forecasts, the likelihood of an expected hazard and its potential severity can be used to aid decision-making. The United Kingdom's Met Office and several other countries contributing to the Meteo alarm system (WMO, 2015) are examples of countries using such systems. The UK Met Office's Flood Guidance Statement (FGS) provides daily flood risk forecasts for the UK to assist in strategic and tactical operational planning and decision-making (Flood Forecasting Centre, 2017). Assessments are made and presented in a coloured 4x4 risk matrix based on forecast risk (hazard, vulnerability and exposure: and likelihood (Figure 3.2). When model-based detailed hazard information cannot be produced due to data constraints, a hazard matrix based on hazard severity versus rainfall thresholds derived from historical data may be an alternative. This provides benefits through the selection of suitable threshold severities to provide approximate (often nominal) flood forecasts, which serve as the basis for decision-making regarding flood warning and preparedness. Dale et al. (2012) also proposed using probabilistic flood forecasting and a benefit-cost-inspired decision support framework for flood management actions. However, such an elaborate system is only transferable and compatible with similar flood forecasting systems which have flood impact and cost data which is rarely the case in data-scarce regions.

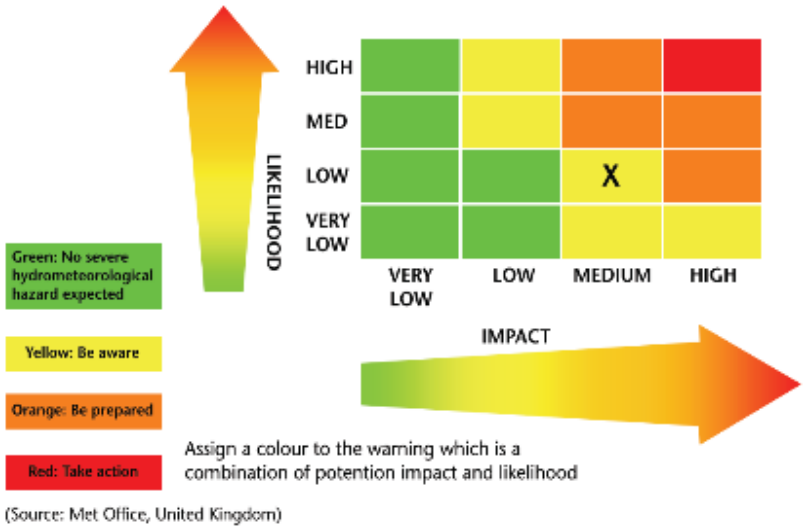


Figure 3.2 Operational application of warning concept, combining impact with the likelihood to create a risk matrix, expressing risk through a simple colour-coded scheme (WMO 2015, UK Met Office)

3.2 METHODOLOGY

A threshold-based framework for supporting early warning decisions based on ensemble forecasts is proposed. Critical rainfall thresholds are defined and used to classify hazards. These are derived by examining rainfall depths associated with historical flood events and the associated hazard severity. The framework used to derive the respective warning classes follows the operational approach of The UK Met office (Figure 3.2). However, exposure and vulnerability data were only sparsely available and not assessed (neither quantitatively nor qualitatively). In lieu, a hazard matrix was developed which allows the likelihood and severity of hazards to be considered in tandem.

3.2.1 Identification of rainfall thresholds

This research adopted an approach using observed rainfall accumulation and rainfall intensity for a training period (using previous floods over a specific duration). Historical flood data were derived from social media mining, archived newspapers, blogs and eyewitness accounts, which have proven useful for assessing evidence of flooding (Chow et al., 2016; Herman Assumpção et al., 2017; J. D. Paul et al., 2017). Sewer design rainfall and knowledge of local drainage conditions were also considered relevant as it is assumed that floods occur once this rainfall depth is exceeded.

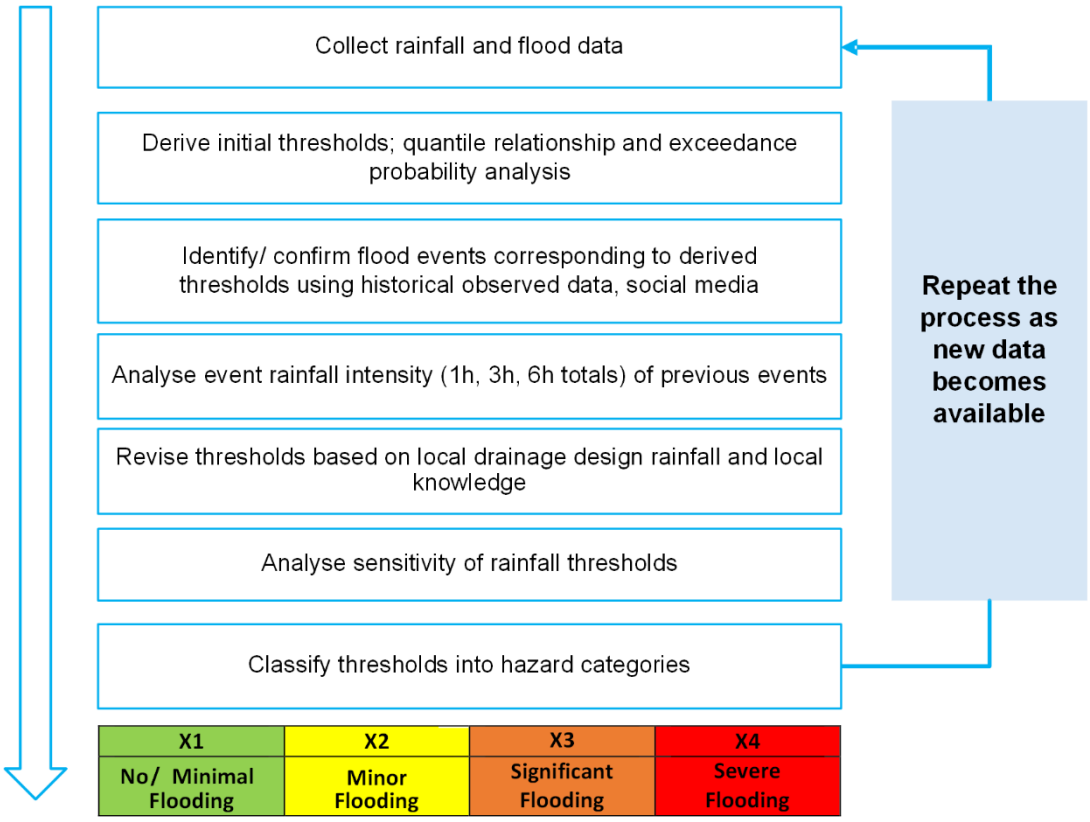


Figure 3.3 Methodology for identifying critical rainfall thresholds using historical data and hazard categories and social media Modified from (D. Wu and Wang, 2009; Yang et al., 2016)

Hazards are classified as "No to minimal flooding", "Minor Flooding", "Significant flooding" and "Severe flooding". Thresholds were determined using data from a suitably identified training period and tested with data from a validation period. The methodology used to identify thresholds is presented in Figure 3.3.

3.2.2 Hazard Matrix

Once the critical thresholds have been determined, an operational system is proposed. This integrates ensemble precipitation forecasts and rainfall thresholds into the hazard matrix to assess the possibility of issuing inundation alerts. The system generates a hazard forecast for various forecast horizons using rainfall forecasts from NWP models, such as the European Centre for Medium-Range Weather Forecasts (ECMWF) or the Global Forecast System (GFS). Rainfall forecast values are then cross-checked against derived rainfall thresholds. To determine the warning classification (hazard class), the likelihood of ensembles is categorised based on probability threshold ranges. The concept of probability thresholds has been used by the UK Met Office and the Rijnland Water Board

(van Andel, 2009). A higher probability threshold will result in more hits but also more false alarms. The method, adapted from Yang et al. (2016), calculates the likelihood of exceeding a particular threshold (Eq. 3.1). The higher the number of ensemble members exceeding the rainfall threshold, the higher the forecasted probability that the threshold will be exceeded.

$$Pr = \frac{1}{N}\sum_{i=1}^{N} f_i \times 100, i = 1,2\ldots N \; and \; N = 50 \qquad (3.1)$$

$$\mathrm{f_i} = \begin{cases} 1, if \; X_j \leq F_i < X_{j+1}, j = 1,2,3,4 \;\; and \\ 0, if \; F_i \; otherwise \end{cases} \qquad (3.2)$$

Where Pr is the likelihood of threshold exceedance, F_i is the forecasted rainfall of the *i*th ensemble member of *N* ensembles and X_j is the rainfall threshold for hazard class *j*. If a threshold is exceeded then f_i is assigned a value of 1; 0 otherwise (*Eq. 3.2*). The *Pr* for each lead-time is evaluated to derive the hazard matrix for the floods occurring. A decision-based rule based on probability thresholds is used to assign the category as high, medium, low, or very low. A high likelihood means a high probability of occurrence and very low likelihood means low to no possibility of occurrence. The Probabilistic Thresholds (PT)s have been adopted from the UK Met Office. However, probability thresholds should be specific to a particular site and updated regularly based on experience. The operational system approach, including the PTs used is presented in Figure 3.4.

3.2.3 Performance Indices

Categorical scoring was used to determine whether the warning class matched the observed event category. Categorical descriptions of a hit, miss and or false alarm are presented in Table 3.1. The methods used for the categorical performance methods are described in Chapter 2, Section 2,1).

Table 3.1 Categorical descriptions of a hit, miss and false alarm

Category	Description
Hit	Forecast warning class (green/yellow/orange/ red) matches the observed class (No flooding/ Minor flooding / Significant flooding / Severe flooding)
Miss	Forecast warning class is anything lower than the observed hazard class (e.g. "Yellow" for significant flooding will be treated as a missed alarm)
False alarm	Observed hazard class is anything lower than the forecast warning class (e.g. "Red" for minor flooding will be treated as a false alarm)

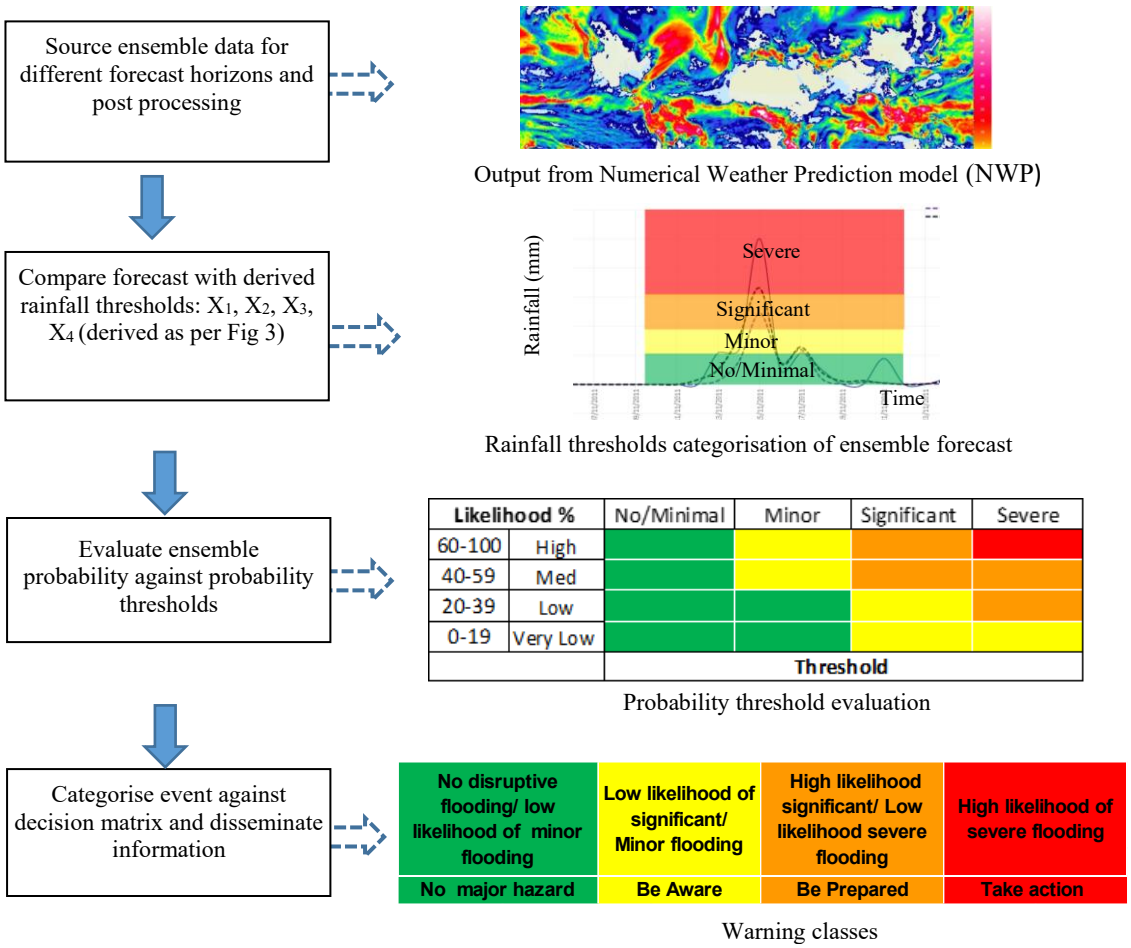


Figure 3.4 Operational flow chart for the proposed urban inundation early warning system.

3.3 Data Preparation

3.3.1 Historically Observed Rainfall and Flood data

Daily rainfall data were sourced freely from the online weather service provider Tutiempo.net for the Nouzha Airport gauge. This data was previously used in several studies including an extensive World Bank study (AASTMT and Egis BCEOM International, 2011). It was further verified by comparing with WMO monthly averages, which are considered a reliable and verifiable source. The selected dataset was then tested for outliers and homogeneity. Only one observed gauge was available, therefore does not represent rainfall spatial variability. A search was conducted to identify days with floods when the observed rainfall exceeded the 90th percentile of daily rainfall data from 2010 to 2012. These rainfall events were used as historical reference events (calibration period). Evidence of flooding was sourced from online newspaper archives, blogs, YouTube videos, social media and other literature, providing valuable flood data. No social media

data or evidence of floods was available before 2010. The rainfall thresholds derived from this period were then tested against events in the 2013-2015 period. To verify intensity, 3-hourly data from the TRMM (Tropical Rainfall Measuring Mission) Multi-satellite Precipitation Analysis 3B42 (version 7) were disaggregated using observed rainfall (Eq. 6). This 3-hourly post-real-time 3B42RT dataset (mm/hr) is estimated monthly and indirectly uses rain gauges by performing monthly bias corrections. A proportional adjustment procedure was applied to disaggregate daily gauge data into three-hourly data using TRMM data (Koutsoyiannis, 2003; Pontien and Bhattacharya, 2011). A summary of data sources used is presented in Table 3.2.

$$X_s = \tilde{X}_j \left(Z \Big/ \sum_{j=1}^{8} \tilde{X}_j \right) \qquad (3.3)$$

X_s - adjusted 3hourly TRMM gauge rainfall, $\tilde{X}_j$ - uncorrected 3hourly TRMM rainfall at time j, Z- daily gauge rainfall, j- sub-period

Table 3.2 Summary of rainfall data used

Historical rainfall				
Data Source	**Station**	**Period**	**No of years**	**Resolution**
tutiempo.net	Nouzha International Airport	1957-2015	53. Years missing 1967-1973	Daily
WMO	Nouzha International Airport	1961-1990	Climate normals	Monthly
TRMM 3B42RT	31.5N, 30 E, 31.5N, 30E	2010-2012	Three years (rainfall intensity)	Three hourly 0.25° x 0.25° grid
Ensemble Forecast data				
Data Source	**Location**	**Period**	**No of ensembles**	**Resolution**
ECMWF	31.5 N, 29.5 E, 31N, 30E	2010-2015	50	0.5° x 0.5° grid

3.3.2 Ensemble rainfall forecast

The Egyptian Meteorological Agency (EMA) forecast rainfall uses a deterministic limited-area Weather Research and Forecasting (WRF) model, derived from the GFS and ECMWF Global models (EMA, 2017). Ensemble rainfall forecast TIGGE (THORPEX International Grand Global Ensemble) from the ECMWF is part of the THORPEX (Observing System Research and Predictability Experiment) and provides medium-range forecasting up to 15 days ahead (ECMWF, 2015). TIGGE precipitation dataset is available as 50 perturbed forecasts and one control forecast. For the period 2010- 2015, 50 ensembles were retrieved from the ECMWF data portal using the Meteorological Archival Retrieval System (MARS) for the corresponding 0.5^0 x 0.5^0 (55km x 55km) grid for the Alexandria study area. Each member represents an equally likely prediction of total precipitation, forecasted at 6, 12, 24, 48, 72 and 96 hrs with a starting time of 0:00 UTC. Ensembles for the test period (2013-2015) were bias-corrected (*Eq.3.5*) using ratios derived from the linear scaling methods *(Eq.3.4*) used by (Crochemore, Ramos, and Pappenberger, 2016; Teutschbein and Seibert, 2013) and then disaggregated to a daily resolution (Arias-Hidalgo et al., 2013) (*Eq.3.6 and3.7*). The monthly mean values from 2010-2012 were used as the training set and applied to the daily 2013-2015 data. The data length used for forecast post-processing was limited and should but should be extended in future studies.

$$K_{i,m} = TP_m / ENS_{tr,i,m} \quad (3.4)$$

$$ENS_{cor,i.m} = K_{i,m}.ENS_{t,i,m} \quad (3.5)$$

$$f_{d,m} = P_{d,m} / TP_{d,m} \quad (3.6)$$

$$ENS_{cor,i,d} = f_{d,m}.ENS_{cor,i,m} \quad (3.7)$$

TP_m - average observed rainfall at month, m; $K_{i,m}$ - ensemble i correction factor for month, m; $ENS_{tr,i,m}$ - uncorrected ensemble i rainfall for the month m for training period; $ENS_{cor,i.m}$ - bias-corrected ensemble i for month, m; $ENS_{t,i,m}$ - uncorrected ensemble i rainfall for the month m for test period; $f_{d,m}$ - disaggregation factor for day d of month m; $P_{d,m}$ - gauge rainfall at day d of month m; $ENS_{cor,i.d}$ - bias-corrected ensemble i for day, d

3.4 Results and Discussion

3.4.1 Rainfall Thresholds

The rainfall threshold methodology used a quantile relationship for rainfall data from 1957 to 2012. Combined with historical flood evidence, 24 hour accumulated rainfall depths were associated with flood incidents (Table 3.3). Rainfall data were fit to a gamma distribution and the percentiles for wet days (i.e., when precipitation was greater than 0mm) were evaluated. The following thresholds were found: 75th percentile: 6.1mm, 90th: 11.94 mm, 95th: 19.04 mm and 99th: 32.1 mm. For the 2010-2012 training period, all events above the 75th percentile are identified in Table 3.3. Of the 19 rainfall events identified, evidence of flooding was found for three events (highlighted blue) in Table 3.3. In the absence of sub-daily data, corrected TRMM 3-hourly data (Figure 3.5) was used to corroborate the high intensities on the days identified in Table 3.3**.** Maximum intensities of 9.8mm/hr and 6.78mm/hr were found for the 2011 and 2010 events, respectively (Figure 3.5). Evidence of the inundation observed for the Dec 2010 event is presented in Figure 3.6**.**

Table 3.3 Summary of rainfall events during 2010- 2012 which exceed the rainfall amount corresponding to the 75th percentile of the daily rainfall data for the period 1957-2012. Some of these days were shown to be clustered events (Events 2, 4, 9, 17and18) which took place over 1 to 3 days

No	Date	mm	No	Date	mm	No	Date	mm	No	Date	mm
1	20/01/10	9.9		13/11/11	10.9		12/01/12	7.8	**17**	05/12/12	6.1
	11/12/10	5.1		14/11/11	11.9	**9**	13/01/12	7.8		06/12/12	21.0
2	12/12/10	6.1	**5**	15/11/11	5.08		14/01/12	6.1	**18**	11/12/12	7.11
	13/12/10	20.1		16/11/11	50.0	**10**	17/01/12	7.8		13/12/12	28.1
3	15/01/11	16		17/11/11	10.9	**11**	25/01/12	7.1		14/12/12	6.1
	16/01/11	7.1	**6**	21/11/11	8.89	**12**	18/02/12	9.9	**19**	21/12/12	7.1
4	19/01/11	7.1	**7**	24/12/11	6.1	**13**	29/02/12	9.9			
	07/02/11	7.8	**8**	02/01/12	10.9	**14**	01/03/12	20.0			
						15	04/03/12	7.1			
						16	25/11/12	7.8			

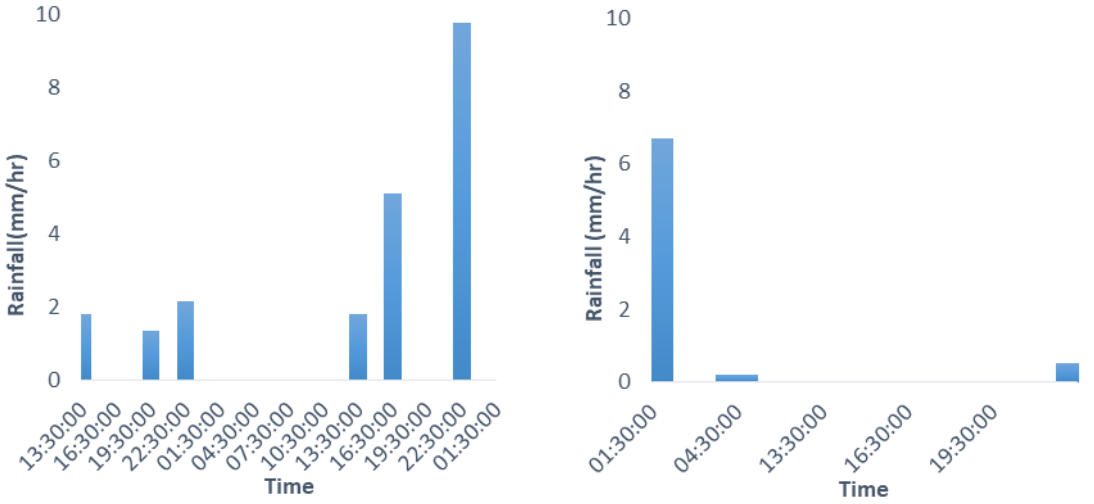


Figure 3.5 (a) Hyetograph for 15th and 16th Nov, 2011 and (b) Hyetograph 12th Dec, 2010 showing rainfall intensity source: TRMM, 2018

Figure 3.6 Flooding on Dec 2010 showing submersion of cars Source: Youtube; SomeOne (2010)

There was significant coverage of the flooding event in 2010 by news agencies and it was also identified as a major event in the World Bank's study (2011), with damage to buildings and disruption of traffic. Casualties were also reported with this event. The observed rain gauge event was associated with a daily rainfall of 20 mm on Dec 12th. The 2011 event was also highlighted as a major flood by the ADSCO. Local drainage capacity (2yr return period) for Alexandria is reported as 26mm/day, considering events lasting no longer than 2hrs. Flooding is assumed to occur once it is exceeded.

On Nov 16th, 2011, the maximum observed rainfall was 50mm, which also caused significant flooding. Rainfall frequency analysis shows that 26mm has a 2-yr return period and 50mm/day has a 10-yr return period (Awadallah et al., 2017). Therefore, we assume a 50mm event would have a higher severity. From this criterion, we infer that rainfall of 20mm leads to significant flooding and severe flooding occurs when rainfall exceeds the 99th percentile, i.e., 32mm. Based on evidence of flooding observed in 2010, 2011 and 2012, the hazard classification was assumed for the following events (Table 3.4). Thresholds were classified as either "No to minimal", "minor", "significant" or "severe".

Table 3.4 Revised rainfall thresholds considering location drainage conditions, non-exceedance probability and return periods.

(mm/day)	0-11.99	12-19.99	20-31.99	> 32
Hazard Class	No to Minimal Flooding	Minor flooding	Significant flooding	Severe flooding

Table 3.5 Summary of major events in 2013-2015 (validation period) to be evaluated including the bias-corrected forecast depths at different lead times

Event	**Date**	**Lead-time**						
		Obs	6	12	24	48	72	96
		(mm)	(mm)	(mm)	(mm)	(mm)	(mm)	(mm)
1	04/11/2015	20.1	8.6	17.9	41.6	35.9	26.6	20.4
	05/11/2015	43.9	43.8	31.4	41.3	24.8	16.7	12.8
	06/11/2015	44.9	5.2	5.8	8.3	5.1	3.5	2.8
2	25/10/2015	32.0	46.9	28.6	36.0	20.	13.8	11.1
3	07/01/2015	20.1	16.7	15.4	12.6	7.1	5.2	6.2
4	06/01/2013	18.0	19.3	16.8	21.4	22.9	20.3	16.4
	08/01/2013	36.1	24.5	22.3	35.7	12.6	9.7	7.7
5	13/12/2013	25.9	8.5	7.1	7.9	4.2	2.9	2.5

3.4.2 Application to 2013- 2015 storms

The assessment was performed for a total of five storm events identified during 2013-2015 (Table 3.5**)**. A visual representation of the threshold categorisation for event #1 and #2 in 2015 is shown in Figure 3.7.

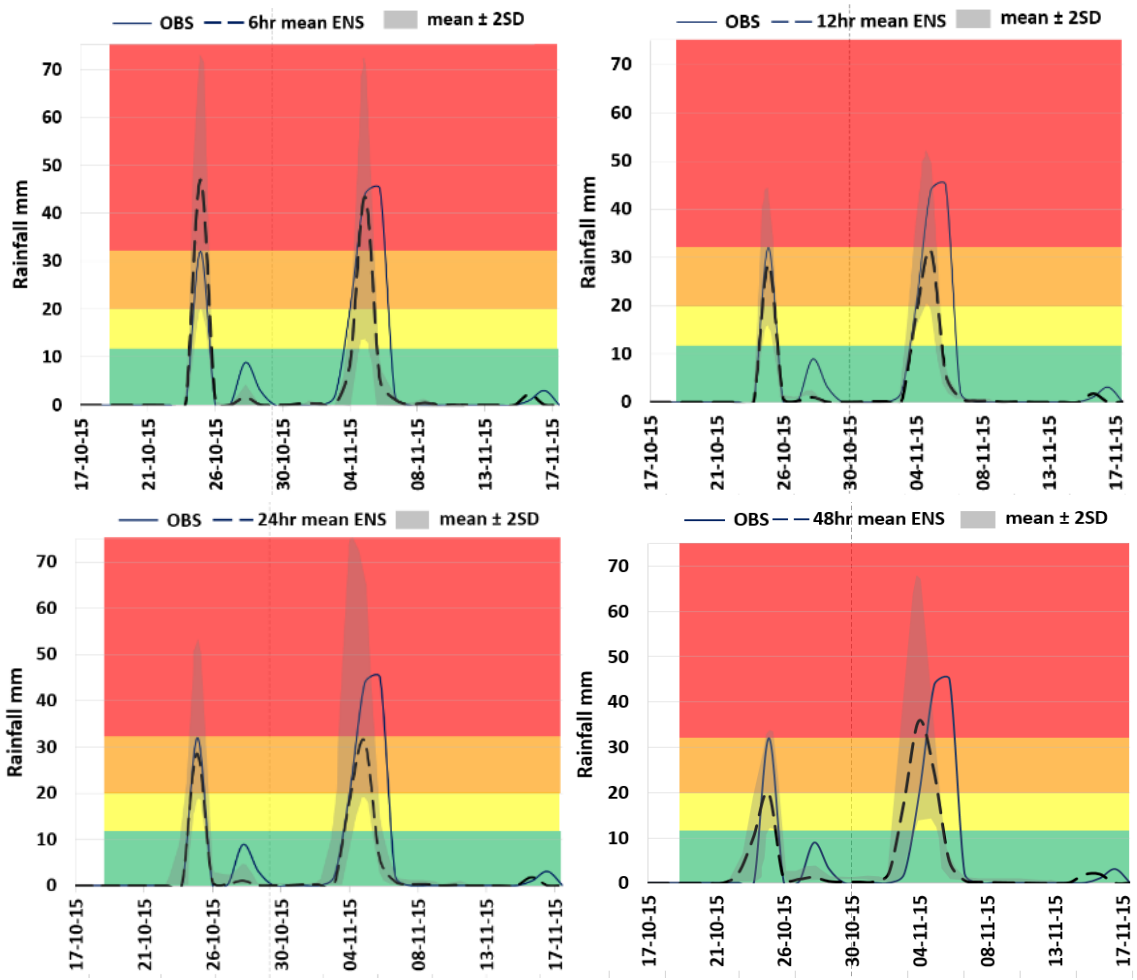


Figure 3.7 Threshold categorisation for Oct 25th and Nov 4th -5th event (mean ensemble and mean +/- 2 standard deviations at different LTs)

3.4.3 Performance Indices

A contingency table was developed to evaluate the ability of this approach in discerning the severity of floods inferred with respect to lead-time. Results of this analysis are presented in Table 3.6 for the validation period 2013-2015. A POD of more than 33% is considered useful (WMO, 2011). The POD and CSI scores for minor events suggest all minor events were correctly forecasted at 12hrs and half at 6hrs. At 24hrs lead-time, severe events received a CSI of 0.75 which suggest at this lead-times ¾ of all severe forecasted events were observed. This is demonstrated for both the Nov 4th and Oct 25th

events. The forecast did indicate some level of persistence as ¾ of all severe events were forecasted as significant events at 48hrs before being upgraded to a severe event at 24hrs. Incidentally, significant events were not well detected at short lead-times but scored a POD of 0.5 at 48 and 72hrs and a FAR of 1 at 24hrs implying that at 24hrs these events were always forecasted at a higher value. All hazard event categories showed poor

Table 3.6 Summary of categorical scores for observed events (2013-2015): Probability of Detection (POD), False Alarm Ratio (FAR) and Critical success ratio (CSI) or Threat score

Lead-Time	***Minor***			***Significant***			***Severe***		
(h)	POD	FAR	CSI	POD	FAR	CSI	POD	FAR	CSI
6	0.5	0	0.5	0	0	0	0.5	0	0.5
12	1	0	1	0	0	0	0	0	0
24	0	1	0	0	1	0	0.75	0	0.75
48	0	1	0	0.5	0	0.5	0	0	0
72	0	1	0	0.5	0	0.5	0	0	0
98	0	0	0	0	0	0	0	0	0

discrimination at the 72 and 98hrs lead-time. Only three years of data were used to calibrate the rainfall thresholds, which is seen as a considerable limitation. The results indicate, even with a relatively coarse spatial resolution, the ECMWF data was still capable of predicting events and performed better for more severe events. These results could be improved using higher spatial resolutions from limited area forecast models and more robust bias correction methods and longer calibration test datasets.

3.4.4 Hazard Matrix

The hazard categorisation, ensemble probability exceedance and the corresponding decision/warnings to be taken are summarised in Figure 3.8. UK Met Office probability thresholds were used in the absence of established probability thresholds (Figure 3.4). For both events, severe events were predicted at both 6h and 24h lead-times with over 60% likelihood, warranting a “take action” decision. At 48hrs the events were predicted with a “high likelihood of significant event” or “low likelihood of a severe event”; “Be prepared” and no major hazard at 96hrs. The agreement on the decision category while definite at some longer lead-times also shows uncertainty at longer lead-times and ambiguity in making a confident decision.

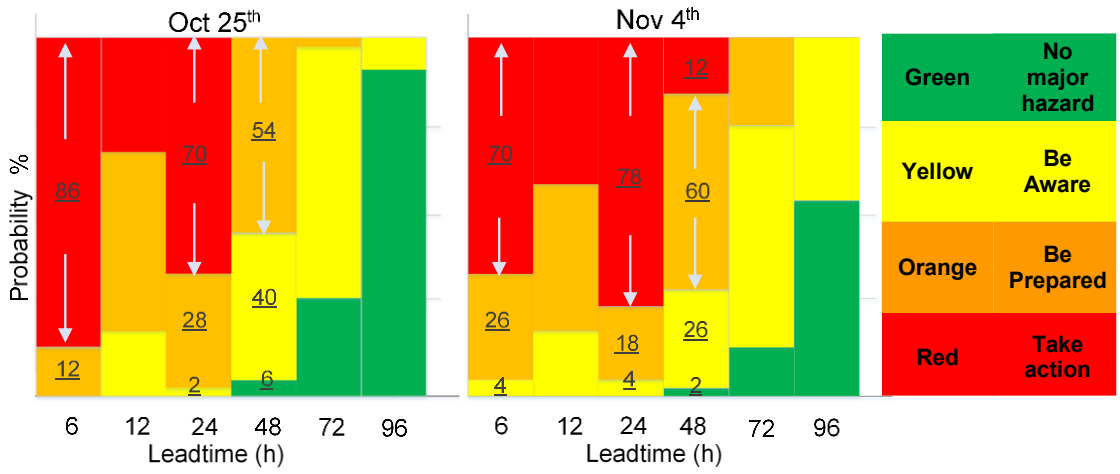


Figure 3.8 Hazard matrix and decision categorisation for Oct 25th and Nov 4th events. Probability of ensemble exceeding a threshold is shown

It is acknowledged that these results are highly dependent on the threshold, as discussed in the next paragraph. Thresholds derived should not be final but need to be updated regularly as more data becomes available. Sourcing evidence of floods was one of the biggest challenges in the research as there are no official existing historical records of floods in Alexandria. In addition to having a limited number of years, there are many circumstantial factors, which contribute to urban flooding such as rainfall intensity, localised flooding in surface depressions, sewer blockages, pump reliability, functionality and changes in permeability over time, of which could not be accessed.

There is uncertainty associated with the derivation of the thresholds. Table 3.7 evaluates how the warning class for these specific events would vary had slight changes to the thresholds been made. In general, there was less variability with the shorter lead-times and at higher severity events. For the Oct and Nov events, the hazard characterisation remained unchanged from a range of 26 to 38mm. For the Jan 6th and 8th event, the classification remained "significant" for most ranges as well. For the purposes of this analysis, the Oct 25th and Nov 4th events were perceived as two independent events but in reality, there is an interdependency as the previous event would have saturated the system, resulting in lower thresholds for the subsequent flood. Thresholds were derived based on daily-accumulated totals however it is also necessary to examine multi-day accumulation thresholds to address clustered events.

Table 3.7 Sensitivity of analysis hazard classification at the 6hr lead time with changing thresholds

No.	Threshold Range (mm)	Hazard classification
1	0-16-25- 36	
2	0-15-24-35	
3	0-14-23-34	
4	0-13-22-33	
5	0-12-20-32	
6	0-11-19-31	
7	0-10-18-30	
8	0-9-17-29	

Jan 2013					Oct 2015				Nov 2015		
5	6	7	8		24	25	26		4	5	6

3.4.5 Experimentation of the system on a real case: the 2018 storm

The above derived rainfall thresholds and hazard classifications are currently being utilised by the Alexandria Sanitation and Drainage Company (ASDCO) to issue warnings and initiate anticipatory actions. In Dec 2018, officials were put on alert for an extreme rainfall event on Dec 5th. The 24 hr rainfall accumulation was forecast to be 33 mm and 27mm at the 48hr and 24hr forecast time horizons respectively. Applying the rainfall threshold categorisations, these events were forecast as "Severe" 48hrs before and "Significant" 24hrs before. The probabilistic forecast at both leadtimes indicated a "high likelihood of significant flooding/ be prepared" would have been issued. ASDCO used the deterministic forecast, combined with their knowledge of the past flood and critical hot spot areas. The city initiated pre-emptive actions including pumping and checking of pumps at specific locations and cleaning of drains. In actuality, the observed rainfall for the event on Dec 5th was observed to be 23mm; a "Significant" flood but no reports of flooding were received. Probabilistic forecasts are not yet being used, but it is anticipated this will the next phase as city officials realise the benefits and how to use probabilistic forecasts. The pre-emptive measures implemented by the city were able to prevent flooding and disruption. Before these threshold categorisations, there was no formal process for pluvial flood warnings or plans for anticipatory actions in the city. Additionally, new pumps have been installed to facilitate anticipatory flood actions. With this approach, the city mitigated the 2018 flood and are now more prepared for future

floods. This is a useful, practical and simple method when proactive actions are not highly objectionable due to false alarms. This is one step towards improved adaptation to the increasing severity and intensity of rainfall events observed in recent years and likely to increase in the coming years. Probabilistic forecasts are not yet used but it is anticipated this will be the next phase. If probabilistic forecasts had been used, "high likelihood significant flooding/ be prepared" would have been issued. In the interim, the city continues to use the threshold classifications.

3.5 Conclusion

In Alexandria and other MENA cities, rainfall variability is high, drainage capacity is insufficient, flood preparedness is low and the risk of pluvial flooding is often unknown. In data-scarce regions the threat is imminent; there needs to be an interim trade-off between complexity and technical feasibility to effectively reduce impacts until the necessary detailed studies can be done. This research demonstrated and applied a practical approach of predicting extreme rainfall events associated with floods using limited data and no flood models by incorporating available data such as historical flood data from social media and satellite precipitation products (SPPs). In addition to assessing uncertainty, ensembles combined with a rainfall threshold approach allowed characterisation of exceedance likelihood associated with flooding. From an operational forecasting perspective, this probabilistic information was used to support decision-making.

In urban areas, sub-daily rainfall intensity is essential to predicting floods. TIGGE ensemble forecast has a forecast grid of ~ 50x50km and may present difficulties in forecasting in small areas as values are averaged over 2500km^2. The non-linearity of rainfall and flood events, along with the limited number of events and historical data, limited the extent of validation of this method, which is one of the main challenges of using ensembles. However, the results were very favourable and the methodology is presented so that it can be used with higher-resolution NWP models or radar, as available.

The simplified and phased warning and response approach is consistent with operational forecasting agencies. Combined with the ensemble forecasting and rainfall threshold approach, this method adds complexity compared to the widely used deterministic rainfall forecasts. However, it can strike a balance between a simplified approach and forecast uncertainty estimation. It does not assess exposure and vulnerability or indicate possible impacts, but when combined with previous knowledge of "hot-spot" areas, it can prove valuable in supporting decision-making, as demonstrated during the 2018 event. This method is not proposed as an absolute substitute for EWS using stormwater models but rather it allows "buy-in time" to increase preparedness until more data becomes available. Additionally, the benefit of site-specific thresholds applied locally rather than nationally

is that local authorities can make independent decisions and increase preparedness in their cities and communities.

Notwithstanding, warnings still need to be communicated to ensure their effectiveness. Consideration should be given to different durations and to significant changes in catchment characteristics over time. Finally, this rainfall threshold approach emphasises the importance of flexible anticipatory mitigation measures that can adapt to uncertainties in forecasts and other urban drivers.

DATA AVAILABILITY STATEMENT:

The data that support the findings of this study are openly available in ECMWF public data TIGGE Data Retrieval at http://apps.ecmwf.int/datasets/data/tigge/. The data that support the findings of this study are openly available in Tutiempo at https://en.tutiempo.net/ climate/ws-623180.html. The data supporting the findings of this study are openly available at GES DISC at https://disc.gsfc.nasa.gov/datasets/TRMM_3B42_7/summary

Other data underlying this Chapter can be found as Version 1.4TU.ResearchData. dataset. https://doi.org/10.4121/bf9f0902-582e-464f-989f-bf0f4a03fde9.v1

4

INTEGRATING HIGH-RESOLUTION RAINFALL FORECAST AT DIFFERENT SCALES

[2] Regional Numerical Weather Prediction (NWP) models increase the resolution and extend forecasts by several days when initialised with global models. These models produce forecasts at higher spatial and temporal resolutions but are computationally demanding and do not necessarily result in more accurate forecasts, especially in data-scarce cities. This research evaluated how rainfall forecasting model scale dependencies satisfy high-resolution hazard forecasting requirements with two different flood forecasting approaches. Rainfall forecasts of different spatial resolutions, cumulus schemes and lead times from a high-resolution Weather Research and Forecasting (WRF) model were first evaluated and then used as input in a rainfall threshold and 1D–MIKE–urban drainage model for flood forecasting in the data-scarce city of Alexandria City, Egypt. Results indicate the flood forecast severity class and flood model simulation results vary with the neighbourhood size, forecast horizon and chosen cumulus configuration but in general the smallest resolutions evaluated did not improve the hazard estimation for both flood forecasting approaches. Therefore, trade-offs must be made regarding model configurations, resolution, lead times and how the forecast output will be used. This study demonstrates the opportunities and limitations for better integrating high-resolution meteorological for the development of a rainfall threshold-based and model-based flood forecasting in cities with similar conditions. It also highlights the need to align the selected model configuration with the goals of the flood forecasting application which is critical for effective early warning systems and anticipatory flood management. also highlights the need to align the selected model configuration with the goals of the flood forecasting application which is critical for effective early warning systems and anticipatory flood management.

---

[2] This Chapter is based on Young, A., Bhattacharya, B., Daniëls, E. and Zevenbergen, C. (2025) 'Integrating WRF forecasts at different scales for pluvial flood forecasting using a rainfall threshold approach and a real-time flood model', *Journal of Hydrology*, 656, p. 132891. doi: 10.1016/j.jhydrol.2025.132891.

4.1 INTRODUCTION

The effectiveness of flood emergency actions depends on the information provided by rainfall forecasts and flood forecasting approaches, which include rainfall thresholds, pre-simulated scenarios and/or real-time simulation approaches. Flood Forecasting and warning approaches in urban pluvial catchments require accurate predictions about peak rainfall rates and the magnitude, timing, location and impacts of flooding and require high spatial and temporal rainfall resolution forecasts suitable for small-scale urban flood modelling, which is often unavailable in data-scarce regions.

Further, delivering urban flood forecasts with sufficient lead times is challenged by uncertainties in forecasting rainfall at high spatial and temporal resolutions and running flood inundation models in real-time. Radar-based nowcasting has been used to improve the spatial and temporal resolution of rainfall for triggering warnings and flood forecasting (Flack et al., 2019; Iqbal, 2017; J.-R. René et al., 2013; Schellart et al., 2011; Thorndahl et al., 2013). However, radar-based nowcasting for flood forecasting applications is provided at short lead times (< 6 hours), which may be insufficient for implementing actions requiring more than 6 hours lead time or unavailable, especially in data-scarce regions where nowcasting is unavailable.

When initialised with global data, the Weather Research and Forecasting Model (WRF) (Skamarock et al., 2008) forecasts rainfall at longer lead times and allows downscaling to smaller high-resolution domains. The smaller domains capture high-scale features and processes which can better simulate precipitation characteristics (Kendon et al., 2017; Lean et al., 2008; Prein et al., 2015). Increasing the simulation domain resolution alone is not enough to significantly improve the model performance. Care must also be given to the treatment of convection (Fowler et al., 2016) and the microphysics and planetary boundary layer (PBL) schemes(Clark et al., 2015; Kain et al., 2008). Many studies have investigated the impacts of different model configurations on rainfall distribution (Jankov *et al.*, 2007; Goswami *et al.*, 2012; Wang *et al.*, 2016. Jankov *et al.,*(2005) found the most significant variability in forecasts from changes in the choice of Cumulus Parametrisation Scheme (CPS) which influences the average rain rate. Umer *et al. (*2021) evaluated different CPS schemes and demonstrated the potential of WRF models as a valuable asset to capture peak rainfall intensity for flash flood modelling in a city where high quality direct and remotely sensed observations of rainfall are limited. Therefore, in high-resolution WRF modelling, the model configuration has a significant impact on rainfall quantities and distributions and subsequently on flood forecasts.

There are interdependencies between the different components of the flood forecasting chain; meteorological forecasting input, hazard estimation methodologies and decision support systems that trigger a warning or action (Figure 4.1). The flood forecasting approach or hazard estimation method uses empirical methods such as rainfall threshold

or real-time flood simulations and data-driven models to identify the critical rainfall or flood depths which will trigger a warning or action. Physically based hydrodynamic flood forecasting models operate in real-time, integrating forecasted rainfall or observed flows to update conditions when available. The suitability of the hazard estimation approach is influenced by several factors: forecasting scale, quality of information and models available, triggering mechanism, requisite forecast variables and end-user needs (Henonin et al., 2013; L. J. Speight et al., 2021). Data-driven models have the advantage of fast computational times but the lack of observed time series in data-scarce regions makes their application limited.

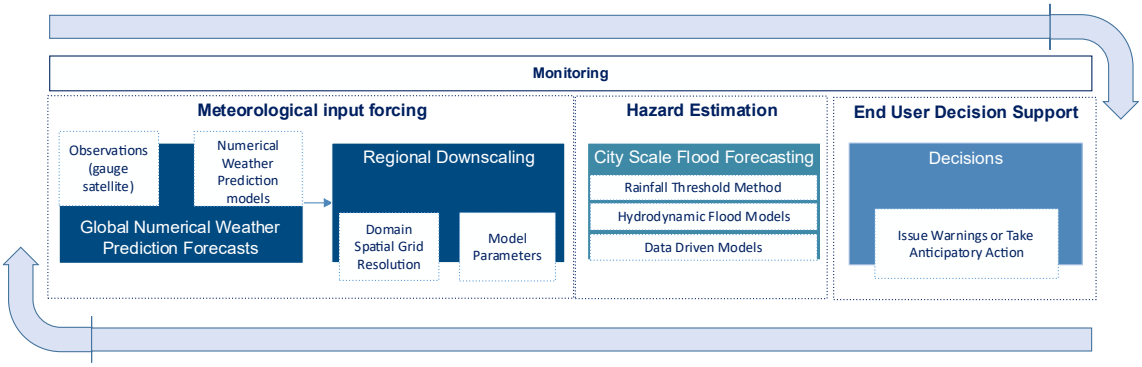


Figure 4.1 Components of a pluvial flood forecasting system for early warnings and anticipatory flood management. Observations and models are used to create Global Numerical Weather Prediction Forecasts. Regional Downscaling is done with the WRF model combining domain variable spatial grid resolutions with different convection parametrisations. Hazard estimation methods using a rainfall threshold approach, real-time flood or data driven models are combined with decision support systems to trigger warnings and actions

Using high horizontal resolution output from Numerical Weather Prediction models such as WRF with rainfall threshold approach or in real-time simulations can be valuable for forecasting urban pluvial floods at various lead times and has been done by other authors (Brendel et al., 2020; Davis et al., 2022; Ming et al., 2020; Thorndahl et al., 2016; H. Wang et al., 2022). The main problem this chapter aims to address is whether increasing the resolution of WRF rainfall forecasts to satisfy urban flood modelling requirements will lead to improved quantification of rainfall magnitude and distribution for flood forecasting in urban pluvial cities.

This will be done by assessing the usability and performance of high-resolution WRF rainfall output to produce flood forecasts for implementing anticipatory flood management actions 12- 72h before an event. Given the influence of the model configurations, WRF rainfall is evaluated at different lead times, domain resolutions, treatments of convection and neighbourhood sizes. Further, the study investigates the influence of increasing the resolution and model configuration with two FF approaches:

a rainfall threshold approach when urban models are unavailable and a real-time hydrodynamic model to forecast flood depths. Whereas previous research has assessed the influence of increased rainfall resolution with radar nowcasting or post-processing with observations, the novelty of this research is in jointly exploring the scale dependencies of both the meteorological forecasting model and different flood forecasting approaches, which has not been explored sufficiently, especially in data-scarce cities.

The research has been carried out for extreme events in the Mediterranean coastal city of Alexandria in Egypt. Every winter, Alexandria experiences extreme weather events that cause significant flooding and disruption to lives throughout the city (Zevenbergen *et al.*, 2016). This research does not aim to determine the optimal model configuration for the WRF model given the challenges in data-scarce cities but instead explores the trade-offs in model accuracy and usefulness of rainfall forecasts to generate flood forecasts in data-scarce regions. Evaluating high-resolution models and their application for rainfall and flood forecasting in data-scarce cities can highlight the interdependencies of the flood forecast chain's meteorological and hydrological aspects, which are often evaluated separately. The findings can inform the design, development and improvement of existing and future FF and early warning systems for increased preparedness by implementing anticipatory actions in cities similar to Alexandria.

4.2 EVENTS

This research evaluates three extreme rainfall events, on October 25th, 2015, Dec 5th, 2018 and Nov 20th 2020 during the winter months which resulted in significant flooding and a state of emergency was declared in the governorate. Several studies using WRF have been carried out in Egypt using the National Centre for Environmental Prediction's (NCEP), Global Forecasting System (GFS) and the European Centre for Medium-range Weather Forecast's ERA5 and ERA interim data sets. Robaa and Wahab (2019) and Ibrahim (2020) evaluated the sensitivity of WRF to convection schemes and both found there was an overestimation of rainfall but results improved with finer resolution. Eltahan and Magood (2018) investigated rainfall sensitivity to microphysics schemes and found all schemes produced higher precipitation than observed. The study further found the Morrison double moment scheme achieved the closest to observed rainfall while the Thompson scheme successfully simulated the cloud pattern. El Afandi and Morsy (2020)found favourable results using a Single-Moment 3-class WSM3 scheme and Kain-Fritsch scheme for the development of an early warning system in the Sinai Peninsula. Ibrahim and Afandi (2014) and Cools *et al. (*2012) previously evaluated the use of a WRF to predict extreme rainfall over the Sinai Peninsula and Egypt. Most recently, Liu *et al. (*2021) used a progressive multi-metric configuration optimisation method and ERA5 reanalysis data to find an optimal configuration of the WRF model for the study area. All previous studies evaluated different resolutions. To the best of our knowledge, no studies

have evaluated the use of WRF with resolutions less than 3km which are the resolutions suggested for flood forecasting in urban scale modelling (Berne *et al.*, 2004).

4.3 Data and Methods

This research analysed WRF rainfall for different lead times, domain resolutions and treatments of convection at different neighbourhood sizes and evaluated with gridded rainfall estimates. The WRF rainfall was then used at different scales with two flood forecasting approaches: a rainfall threshold approach and an urban drainage flood model to evaluate how well they forecasted a hazard. A summary of the research framework is presented in Figure 4.2 The following section briefly describes the WRF model setup and the flood forecasting approaches used.

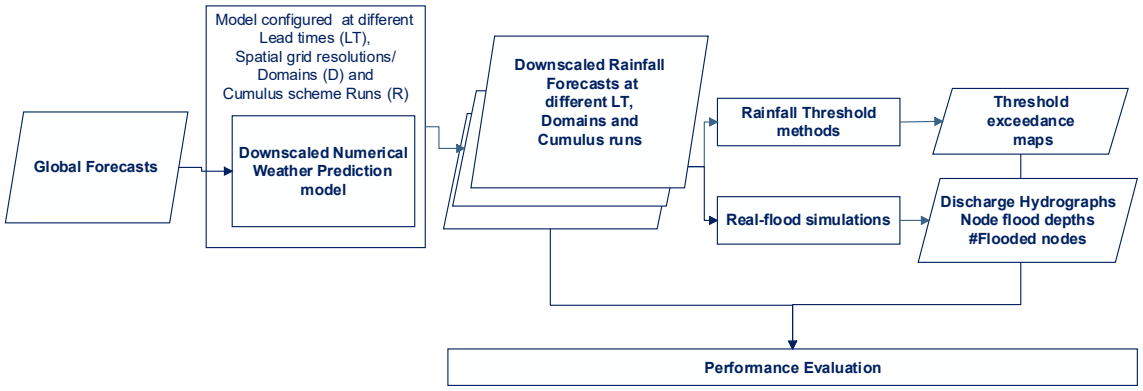


Figure 4.2 Research Framework illustrating the methods for coupling WRF model to produce high-resolution rainfall forecasts and hazard estimation using a rainfall threshold at the city scale and real-time urban flood simulations for part of the city

4.3.1 Data and WRF Setup

The Advanced Research WRF (ARW-WRF), (Skamarock *et al.,* 2008) is a convection-permitting dynamic downscaling model configured using grid parameterisations and physics schemes and nested domains at finer resolutions. This research used 0.25° Global Forecast Grids Historical Archive ds084.1 for lateral boundary conditions, MODIS15 land-use map and two-way nesting. Analyses were first carried out using three domains and then four domains to evaluate if forecast skill is improved by increasing the number of domains and higher spatial resolution. The domains used boundary conditions that were updated every 3 hours, a nesting ratio of 1:3:3:3 and a spatial resolution for domain 1 (D1) 10km, domain 2 (D2) 3.3km, domain 3 (D3) 1.1 km and domain 4 (D4) 0.37km (Figure 4.3). The outermost domain (D1) covered Egypt and the wider Mediterranean. This is consistent with the mother domain used by the Egyptian Meteorological Agency (EMA, 2017). All spatial domains had a vertical resolution of 45 layers and the model used a

timestep of 60s with adaptive timestep. The selection of the physics schemes (Table 1) was guided by the most recent study of Lui et al. (2021) along with the consideration for convection-permitting scales, (1-2km) which can run without convection schemes because the grid spacing is small enough to explicitly resolve convection. However, in the convective grey zone (3-10km) cumulus-cloud processes become partially resolved and traditional closure assumptions break down and scale-aware schemes can be used (Grell and Freitas, 2014; Jeworrek et al., 2019; Zheng et al., 2016). There is still some disagreement on how CPS should be treated in smaller domains. Paul et al. (2018) found for the Mumbai's coastal areas turning the CPS on in a 1km simulated best results, whereas Han and Hong, (2018) found better simulation when CPS in turned off in a 3km domain. Therefore, additional sensitivity experiments were tested for Domain 2 and 3 in which convection was parameterised and cumulus was kept off in D4 (Table 4.1). R1, R2 and R3 correspond to runs with 3 domains and R1_4d, R2_4d and R3_4d correspond to runs with 4 domains. The runs were initialised at the 72h, 48h, 24h and 12h lead times with a spin-up time of 12 hours for three events. The objective of this research was not to determine the optimal model configuration for the WRF model but to investigate if increasing the resolution improves rainfall forecasts and to determine the usability of high-resolution WRF output in flood forecasts and early warning systems. A summary of the WRF setup and methodology used is shown in Figure 4.

Table 4.1: WRF model parameterisation is used and cumulus parametrisation is used for the different runs. Parameterisation schemes are kept the same for all runs except cumulus where the Grell Freitas scheme cu=3 is used or when no cumulus scheme is used cu=0. R1, R2 and R3 correspond to runs with 3 domains (3d runs) and R1_4d, R2_4d and R3_4d correspond to runs with 4 domains (4d runs)

PHYSICS SCHEMES (Based on Lui *et al.,* 2021) same for all runs	**NAME**	**SCHEME #**
Microphysics	Single-moment 6-classWSM6	6
Longwave Radiation	RRTM	1
Shortwave Radiation	RRTM	1
Boundary Layers	Mellor-Yamada-Janjic	2
Surface Layers	Eta Similarity Scheme	2
Land Surface	Unified Noah Land Surface Model	2
Cumulus	Grell Freitas	

Runs (scheme configuration)		**3d runs**			**4d runs**		
	Domain	**R1**	**R2**	**R3**	**R1_4d**	**R2_4d**	**R3_4d**
	D1	3	3	3	3	3	3
	D2	0	3	3	0	3	3
	D3	0	0	3	0	0	3
	D4	-	-	-	0	0	0

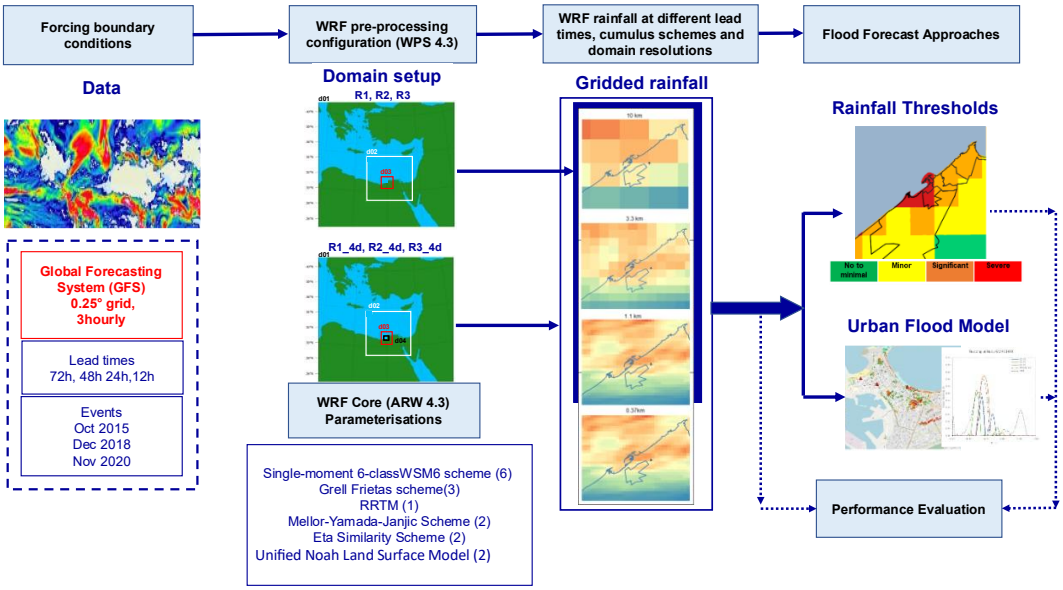


Figure 4.3 Summary of data and methods used for the WRF model setup, rainfall thresholds and real-time simulation

4.3.2 Flood forecasting approaches rainfall threshold approach

Thresholds for the 24h accumulation were selected based on previous analysis of critical rainfall thresholds (Young et al., 2021). These correspond to the severity of flooding: "No to minimal flooding" with rainfall within 0 to 11mm, "Minor flooding" within 12 to 19mm, "Significant flooding" within 20 to 31mm or "Severe" with rainfall ≥ 32mm. Young *et al.* (2021) derived thresholds using observed daily rainfall accumulation and rainfall intensity for a 3-year training period from 2012- 2015 (using previous floods over a specific duration). Local drainage capacity (2-year return period) for Alexandria is reported as 26 mm/day for 2-3 hr events and flooding assumed to occur once exceeded (AASTMT and Egis BCEOM International, 2011). This knowledge of the system drainage capacity was combined with historical flood data derived from social media, archived newspapers, blogs and eyewitness accounts, which have proven useful in assessing evidence of flood. Only one daily rainfall gauge was available for the study thus spatially varying thresholds were not considered. The thresholds were applied to the spatially varying rainfall forecasts to derive hazard classes over the entire governorate.

Table 4.2 Thresholds (in mm) used to trigger warnings

Flood Hazard Classification and Corresponding Rainfall Depth Thresholds			
0-11.99mm	12-19.99mm	20-31.99mm	>=32mm
No to minimal flooding	Minor flooding	Significant flooding	Severe flooding

4.3.3 Urban flood model setup

A flood model based on the 1D MIKE Plus model (from the Danish Hydraulic Institute), capable of real-time FF, was developed to simulate flood inundation. Alexandria city is a very densely urbanised area. A land cover GIS dataset comprising buildings, roads and green infrastructure was used to derive imperviousness for the rainfall-runoff model. The model was developed and calibrated for the Gomork district (~5km^2) (Mahmood, 2021) using available drainage network data.

In the real-time simulation, the hydrodynamic model was run with forecasted spatially varying gridded rainfall at 10km, 3.3 km, 1.1 km and 0.37km resolution to simulate flooding in Gomrok. The combined sewer and stormwater network model was used to simulate flows and water levels in urban storm drainage. It consists of a rainfall-runoff model which transforms rainfall time series to runoff hydrographs to be used as input into the hydraulic network model. The Gomork network comprises 3482 local and collector manholes connected by sewer pipes ranging from 300 to 2000mm in diameter. Pipes flow via gravity to two sanitary lift stations; Outlet #1 (Kaytbay lift station) and Outlet #2 (El Meena). The combined sewage is then pumped from the lift stations and treated at the West Treatment plant before discharging into Lake Maryut.The model setup, location of the West Treatment Plant and Lake Maryut are shown in Figure 4.4. Tides and wave action are not considered to be contributing to flooding in the model setup.

Outlet discharge, total volume and node flooding simulated using the WRF rainfall were evaluated against simulation results generated using MSWEP and IMERG and flood depths were compared with photos taken on the day of the flood. Rainstorms in Alexandria range from a duration of 1 to 9 hours (AASTMT and Egis BCEOM International, 2011; Mahmood, 2021) whereas the MSWEP data has a temporal resolution of 3hs. To simulate extreme cases, the 3h MSWEP rainfall was assumed to occur in a 1h interval. IMERG rainfall, which has a half-hourly resolution was aggregated to hourly temporal resolution.

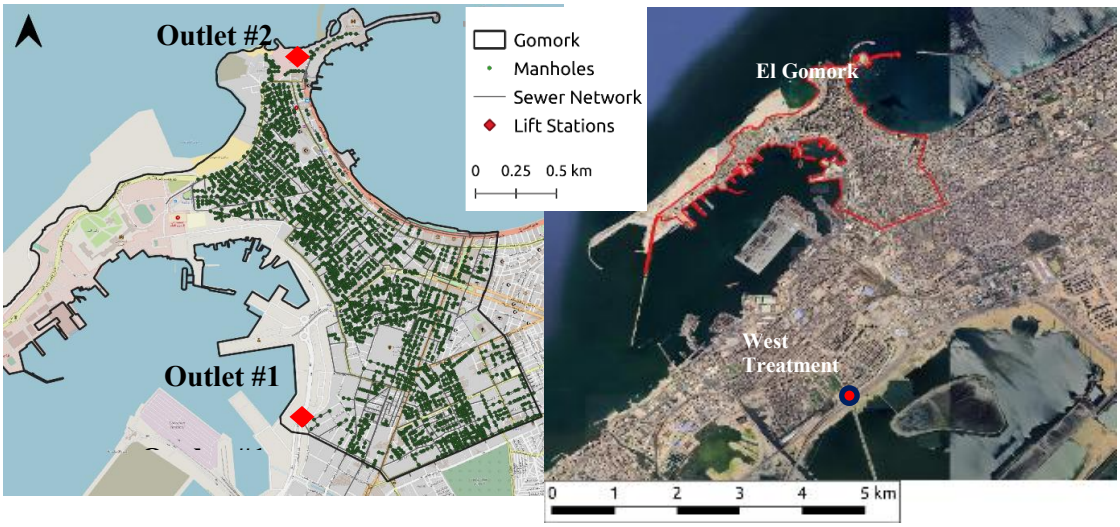


Figure 4.4 Mike Urban plus drainage model set up for El Gomork district showing manholes, sewer pipe layout and location outlet lift stations (left) which pumps combined stormwater and sewage to the West Treatment Plan

4.3.4 WRF Performance Evaluation

This section discusses the methods used to evaluate the performance of the WRF model. Given the large-scale nature of the simulated events, results were evaluated at different neighbourhood sizes to understand the forecasting skill in relation to flood forecasting applications. Neighbourhood locations were determined based on past knowledge of weather systems. In general, rainfall quickly reduces as it moves inland. The larger neighbourhoods 1000 km and 500km size cover the phenomenon scale, whereas the smaller sizes give coverage along the coast (200km), the governorate (100km), the city (50km) and specific districts within the city (20km and 10km) Figure 4.5

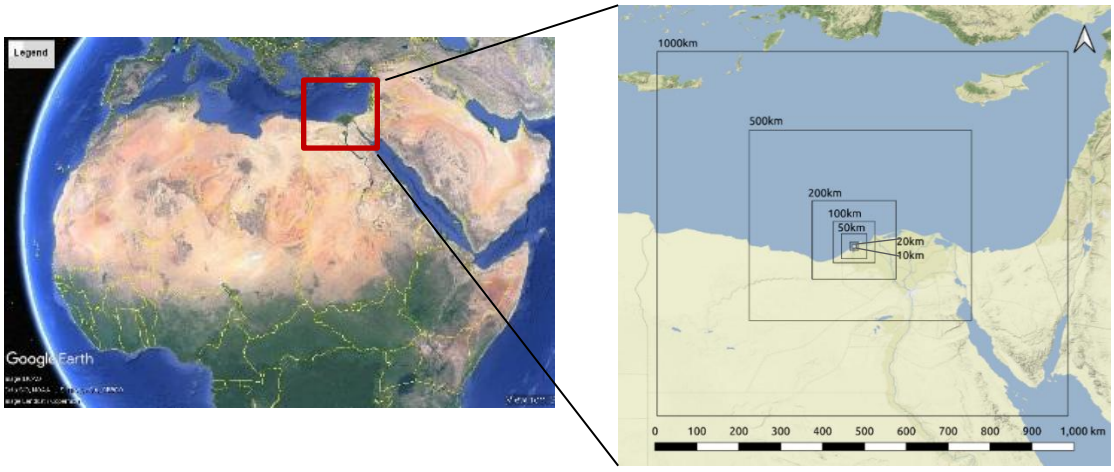


Figure 4.5 Visual representation of the neighbourhood spatial sizes used for FSS calculations

In the absence of reliable sub-daily gauge data, results were compared with the gridded precipitation dataset: Multi-Source Weighted-Ensemble Precipitation MSWEP V2, which merges several gauge, satellite and reanalysis-based data (Beck *et al.*, 2019) and GPM IMERG Final Precipitation Products V7(Huffman et al., 2019) henceforth referred to as IMERG or observed rainfall estimates. There is uncertainty associated with the use of gridded rainfall products, thus both products were compared with the observed rainfall gauge at the Nouzha airport (Table 4.3) and the forecasts were verified against the product which matched the observed rainfall best.

Table 4.3 Comparison of the 72hr accumulated rainfall at Nouzha Airport gauge rainfall with MSWEP and IMERG precipitation products

	Nouzha Airport	MSWEP	GPM
	mm/3days		
Oct 25th, 2015	32.1	38.13	21.97
Dec 5th, 2018	56.13	33.13	59.33
Nov20th, 2020	34.03	15.31	32.13

Both products have a spatial resolution of 0.1 degrees and highest temporal resolution of three hours and half-hour, respectively. Both datasets have been found suitable for detecting rainfall events in North Africa (Mekonnen et al., 2023; Nashwan et al., 2019).

The performance of the WRF model was evaluated on measures of comparison such as scatter plots, Coefficient of Variation and Pearson Correlation Coefficient (CC), Standard Deviation (SD) and presented in a Taylor diagram and the derived Centred Root Mean Square Error (CRSME) (K. E. Taylor, 2001). Enhancing the small-scale detail can decrease the precipitation forecast skill due to the double penalty of erroneous rainfall placement and timing, even though the representation of rainfall accumulation and intensity may look more realistic. To overcome this, the neighbourhood spatial verification Fraction skill score (FSS) (Roberts, 2008) was used. This method directly compares the fractional coverage of forecast and observed rainfall, which exceeds a specified threshold for a neighbourhood size. FSS is calculated for each neighbourhood size (Figure 4.4) using Eq (4.1) where F_i and O_i are the forecast rainfall fraction and observed rainfall fraction exceeding specified thresholds, respectively and N is the number of spatial grids in a neighbourhood size. An FSS of 1 is perfect, an FSS above 0.5 is considered useful and an FSS of 0 is not a skill.

$$\boldsymbol{FSS} = 1 - \frac{\frac{1}{N}\sum_{i=i}^{N}(F_i - O_i)^2}{\frac{1}{N}\left(\sum_{i=1}^{\mathrm{N}}(F_i)^2 + \sum_{i=1}^{N}(O_i)^2\right)} \qquad (4.1)$$

In Eq. (4.1), FSS was computed and evaluated for 24h accumulations for Domain1 (10km), Domain2 (3.3km) and Domain3 (1.11km) at the, 72h, 24h and 12h lead time for both runs with a convective scheme and no convective scheme in respective domains (Table 4.1). All forecast resolutions were aggregated to a 0.1° grid to match observed rainfall estimates.

Rainfall Threshold Urban Flood Model Performance Metrics

For the rainfall thresholds, categorical scoring was used to evaluate if the warning class was consistent with the observed event category. Categorical descriptions of a hit, miss and or false alarm **(Chapter 2).** The Probability of Detection (POD), False Alarm Ratio (FAR) and Critical Success Index (CSI) or threat score takes both false alarms and missed events into consideration and is sensitive to hits. The metrics were evaluated for the 10km, 1,1km and 0,37km spatial resolutions for all runs.

The Normalised Root Mean Square Error (NRMSE) and the Kling-Gupta Efficiency score (KGE) (Eq 4.3) were used to evaluate the performance of the 1D flood model. NRMSE measures the average difference between a statistical model's predicted values and the observed values where a RMSE of 0 is considered a good score. The equation NRMSE is shown as equation 5, where P_i is the model-predicted value, O_i is the observed value, $\bar{O}$ is the mean of the observed value and N is the number of observations. KGE is a single combined expression of correlation, variability and bias into one score and represents a goodness of fit score. A KGE of 1 is considered a perfect score. The metrics were evaluated for the 10km, 1,1km and 0,37km spatial resolutions for all runs for the outlet discharge and node flooding at a known flood location.

$$NRMSE = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{n}(P_i - O_i)^2}}{\bar{O}} \qquad (4.2)$$

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \qquad (4.3)$$

$$KGE = 1 - \sqrt{(r-1)^2 + (\frac{\sigma_{sim}}{\sigma_{obs}} - 1)^2 + (\frac{\mu_{sim}}{\mu_{obs}} - 1)^2} \qquad (4.4)$$

where r is the Pearson correlation between observations and simulations, α is the variability ratio and β is the bias ratio. This is further expressed in Eq. 4.4 where σobs is the standard deviation in observations, σsim the standard deviation in simulations, μsim the simulation mean and μobs the observation mean.

4.4 Results

Results are presented and discussed for WRF rainfall at different spatial resolutions 10 km (D1), 1.1km (D3) and 0.37km (D4) for runs with different treatments of convection. Results are presented for the three events Oct 25th 2015, Dec 5th 2018 and Nov 20th, 2020. The analysis first examined how results vary with lead time and neighbourhood size, followed by the influence of the horizontal grid resolution and cumulus configuration in each domain. The WRF rainfall is then evaluated using rainfall thresholds for forecasting hazard classes and for use in a real-time model simulation for forecasting discharge and node flood depths.

4.4.1 WRF Rainfall Forecast Analysis

Leadtime and Neighbourhood Size Analysis

Figure 4.6 shows how the rainfall event on October 25th 2015 is forecasted at the 1000km and 50km neighbourhood sizes for the 10km grid resolution for Run2 (R2). The 50km neighbourhood gives a closer snapshot of the city scale. While the event's occurrence is captured at all lead times, it shows higher forecasted values at 72h and 24h and lower accumulation and extent at 12h.

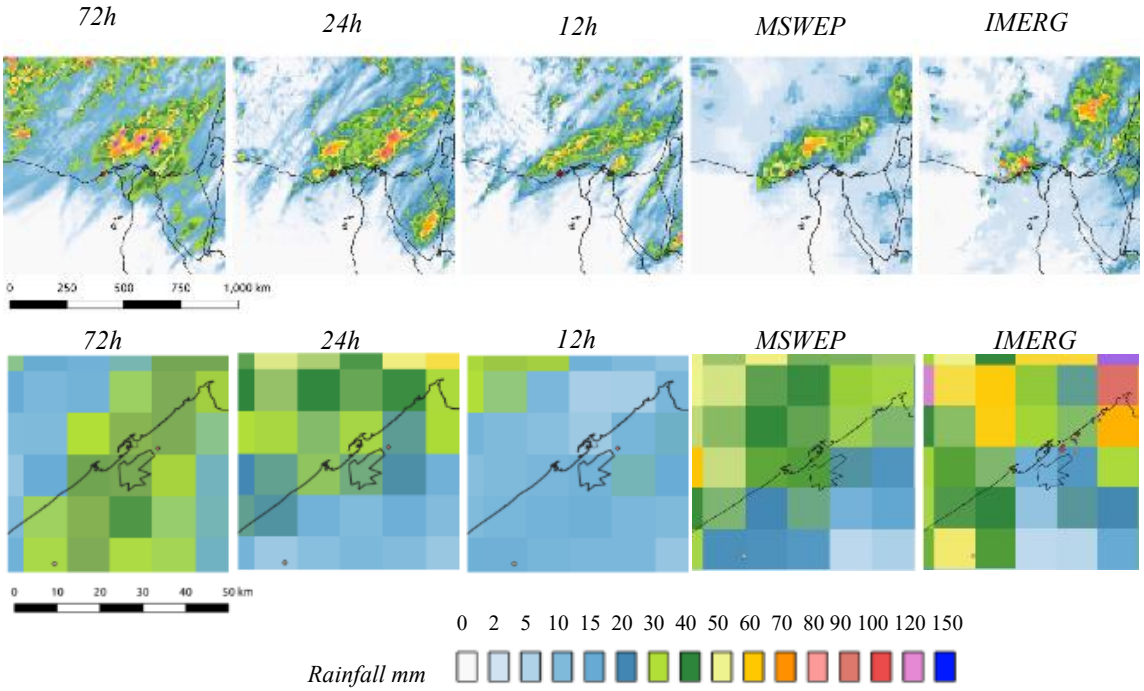


Figure 4.6 Images show how the weather phenomenon is predicted for 24h accumulated rainfall from 21:00 Oct 24th to 21:00 Oct 25th at 72h, 24h and 12h lead time for the 1000km (top) and 50km (bottom) neighbourhood at the 10km resolution and R2. MSWEP and IMERG are shown for the same period.

Generally, the forecasts show similar extent and areas of high accumulation compared to MSWEP whereas the IMERG product showed significantly higher rainfall values offshore but rainfall reduced over the land, less than MSWEP.

Figure 4.7 shows scatterplots of forecasted and MSWEP area average rainfall for the event on Oct 25th, 2015 and Dec 5th 2018 for the 72h, 24h and 12h lead time, 10km domain resolution (D1) and different neighbourhood sizes for R1, R2 and R3. The 12h lead times for all runs performed poorly, heavily under-forecasting rainfall for neighbourhood sizes smaller than 200km. Similar trends are presented in Figure 7. Given the similarity in the trend for lead times in all runs, it can be assumed that the poorer performance was caused by the initial and lateral boundary conditions used to initialise WRF.

Compared to the 12h lead time, forecast results showed high skill at the 72h and best overall performance at the 24h lead time at all sizes, particularly at the smaller neighbourhood sizes. Noticeably, the rainfall was under-forecasted at the 10km scale for the 72h lead time for R1 and R3, but it performed well for R2. R3 showed more similarity for the 24h lead time at the 10km neighbourhood scale. Results for the Dec 2010 and Nov 2020 can be found in the Appendix A.

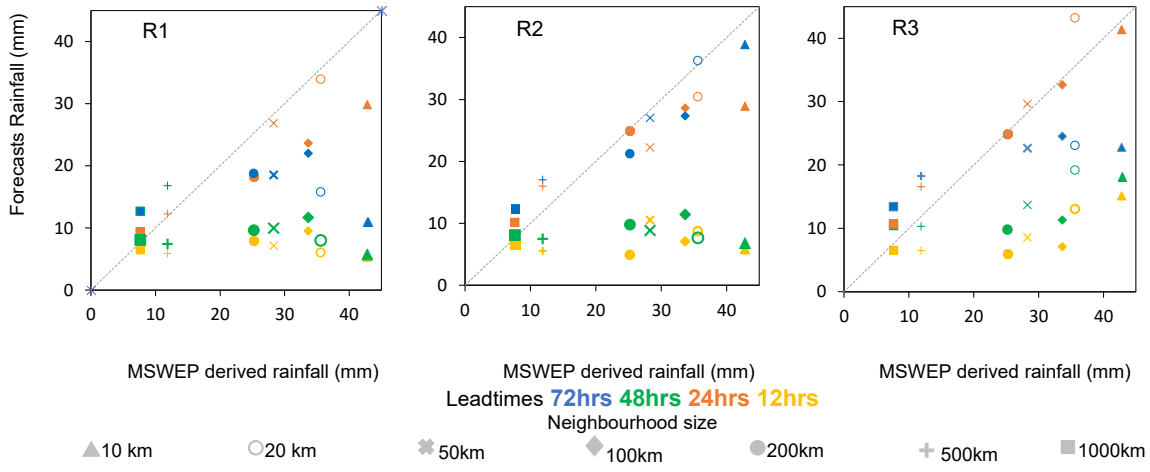


Figure 4.7 Comparison of 24h rainfall accumulation from Oct 24th 21:00 to Oct 25th 21:00 showing the accumulated area averages for the 10km resolution (D1) for different neighbourhoods at lead times of 12, 24,48 and 72 hours for R1, R2 and R3

For the spatial verification metric FSS, there were minor differences between the domain spatial resolutions and runs, but results varied with lead time, neighbourhood size and thresholds. Consistent with the previous analysis, the forecasts performed best at the 24h and 72h lead time for all runs and all thresholds except for the 32mm threshold which shows no skill at the 10km neighbourhood for the 12h lead time. This suggests the WRF model is over-forecasting in some grids of these neighbourhood sizes. Results presented for R3 in Figure 4.8 also show that despite performing poorly for the 20mm and 32mm

thresholds for smaller neighbourhood sizes, the FSS for 12h lead time still showed skill at the 500km and 1000km. The WRF model was able to show a threat was imminent and could confirm the persistency of a threat, which most often may be good enough information for early warning and anticipatory actions for reducing disaster risk. However, the forecasts were not useful (<0.5) at the 10km neighbourhood size at the 72h lead time, which would be more relevant to urban scale modelling. Given the computational requirements of running models at higher resolutions, running the models at a coarser resolution and at longer lead times is preferential. This would allow protocols to be put in place and the model can run at higher resolutions for shorter lead times.

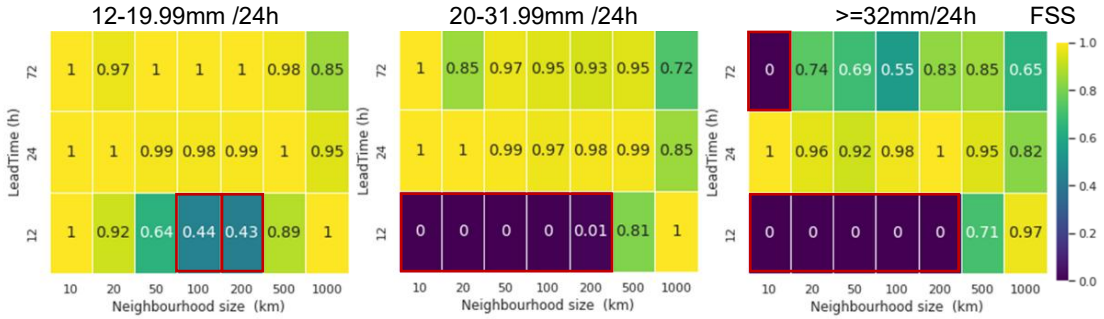


Figure 4.8 Fraction Skill Score (FSS) for D1 (10km) using different thresholds and neighbourhood sizes for run R3. Neighbourhood sizes and lead times without a useful skill are highlighted in red.

Influence of cumulus configuration and horizontal grid resolution

Given the favourable results at the 24h lead time, the results are presented only for this lead time for all model runs, including downscaling to a 4th domain (0.37km) which aims to meet the requirement for higher-resolution rainfall for urban flood forecasting. For R1_D4, R2_D4 and R3_D4 run (see Table 4.1), the runs use the same configuration as the corresponding R1, R2 and R3 but the cumulus scheme was turned off in the 4th domain. The results are presented for the 50km neighbourhood size which showed useful for the 24h lead time will be most pertinent to city administration. Figure 4.9 shows the spatial rainfall variability for each run across the 50km neighbourhood size for the Oct 25th event. In general, the 24h accumulated rainfall across grid resolutions varies by about 5-20mm. Notably, there is a tendency for the finer grids to show areas of more intense rainfall and higher rainfall accumulation than D1 of the same run.

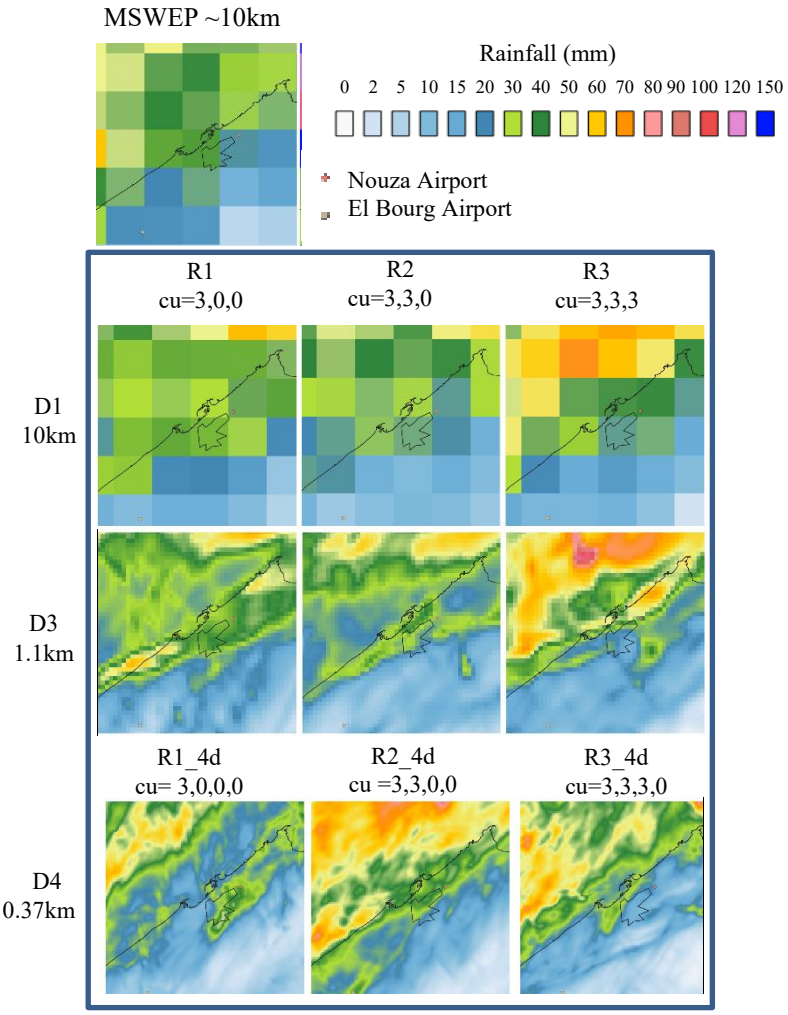


Figure 4.9 Rainfall accumulation from 21:00 UTC on Oct 24th to 21:00 UTC on Oct 25th 2015 for MSWEP, WRF resolutions and different cumulus treatments cu=3 represents the Grell Freitas Cu scheme, 0 indicates no cumulus scheme in the respective domains

It was found that all runs correctly predicted rainfall occurrence, but there was some variability in rainfall locations, magnitudes and intensities across runs and domains. For the 50km neighbourhood size, R1 had the lowest coefficient of variation, 30% and R3 had the closest coefficient of variation 40% compared to MSWEP 43%. This suggests grid points are more spread out across the rainfall mean than observations with smaller variability across grid points. Moreover, in Figure 4.9, rainfall accumulation tends to increase when the cumulus scheme is turned on in the smaller domains for 3 domain runs. Compared to R1 and R2, R3 shows areas of higher rainfall accumulation in the magnitude of 70mm off the coast and parts of the city, particularly on the eastern side of the city. There are also noticeable differences between the horizontal resolutions. For both R1 and R3, D1(10km) averages the rainfall over larger areas detecting lower accumulated rainfall

in the city compared to D3 which shows greater similarity. However, D3 captures greater rainfall variability in detail. In general, there appear to be fewer areas of extreme rainfall over the governorate in R1_D4 and R3_D4, whereas R2_D4 shows areas of increased rainfall compared to R2.

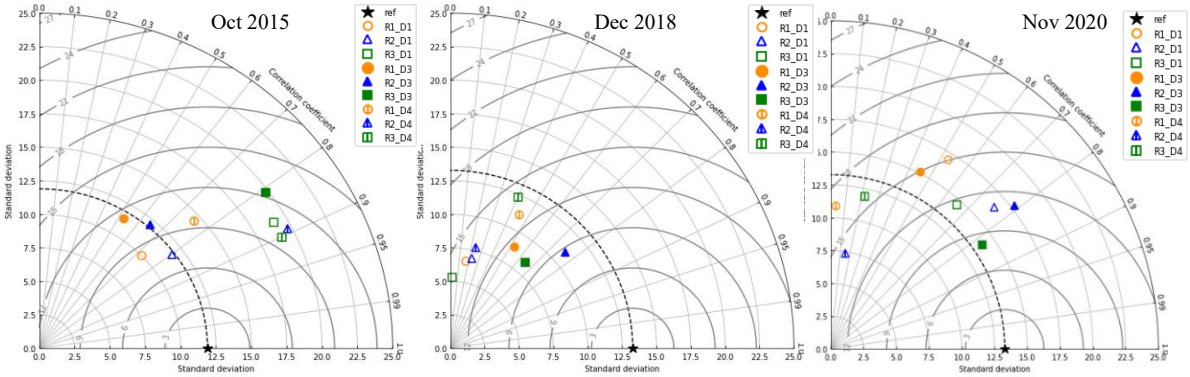


Figure 4.10Taylor diagram for all runs and three domains/resolutions (D1 10km, D3-1.1km and D4 0.37km) at the 50km neighbourhood for all three events. R1(3,0,0), R3 (3,3,3), R1_D4(3,0,0,0), R2_D4(3,3,0,0) and R3_D4(3,3,3,0) for three events.

The differences in the runs are further displayed using a Taylor Diagram (Taylor, 2001) (Figure 4.10) which compares and summarises the standard deviation (SD) and correlation coefficient (CC) of all runs and domains with the SD and CC of MSWEP or IMERG, which is referred to here as the reference data. Forecasts were re-gridded to match the observed resolution. In Figure 4.10, overall the runs vary with events. R2 D3 agree best with observed rainfall values since these runs are closest to the observed ref (dashed) line and correlation coefficients 0.65-0.95 and CRMSE ranging from 8-11mm/24h. the general pattern well with high correlation values (>0.8) but the spread of forecast values is larger than the observed data (above the dashed line). This is primarily because these runs produce areas of extreme rainfall offshore that skew the variability in the data. There is less agreement on the other runs have correlation values between 0 and 0.8. This indicates that while it produces a range of values similar to those observed, the forecast for these runs does not match the precise pattern or distribution of the observed data, leading to a low correlation and suggesting a bias in the forecast or the observation. Domain 4 (D4) (0.37km) for cumulus configurations runs for the three events showed lower correlation coefficients and higher CRMSEs for all runs for all events. The results further highlight that D4 runs performed poorly compared to the D3 and D1 runs for all cumulus configurations.

Rainfall temporal distribution over the Gomork area

The previous analysis focused on how rainfall varies spatially and with different neighbourhood scales. Further analysis examined how the temporal distribution varied within 10 km^2, corresponding to the area used in the MIKE + model and immediate

surrounding areas. This gives insight into the way node flooding might vary with intensity. Results were compared with 3-hourly MSWEP (MSWEP_3h), hourly MSWEP (MSWEP_h) and hourly IMERG (IMERG). Figure 12 presents the rainfall time series for all runs. MSWEP_h was constructed as a worst-case scenario by assuming the 3h accumulated rainfall fell. For the Oct 25th event, the worst-case scenario assumed the 3h accumulation fell in the last hour of a 3h accumulation and the first hour of the subsequent 3h accumulation.

Rainfall is represented as either a single peak or a bi-modal distribution with two peaks. R3 and R2 show primarily single peaks similar to MSWEP_h which assumes a single peak distribution but R3 and R2 both have the highest peak intensities (above 20mm/h) for D3 (1.1km). When downscaled using a 4th domain to 0.37km with no cumulus scheme (Figure 4.11), the rainfall distribution is bimodal with rainfall peaks less than 16mm. Similar results were analysed for the Dec 2018 event, where the 4th domain runs all showed lower rainfall for all domains, whereas for the Nov 2020 event, the D3 and D4 domains both showed high rainfall peak timing. Although the results show similar peak intensities in Dec 2018 and Nov 2020 notably, there is a mismatch in the peak of about 3-4hr when compared with IMERG where the WRF simulated rainfall matching the timing of the MSWEP but not the magnitude (See Appendix A).

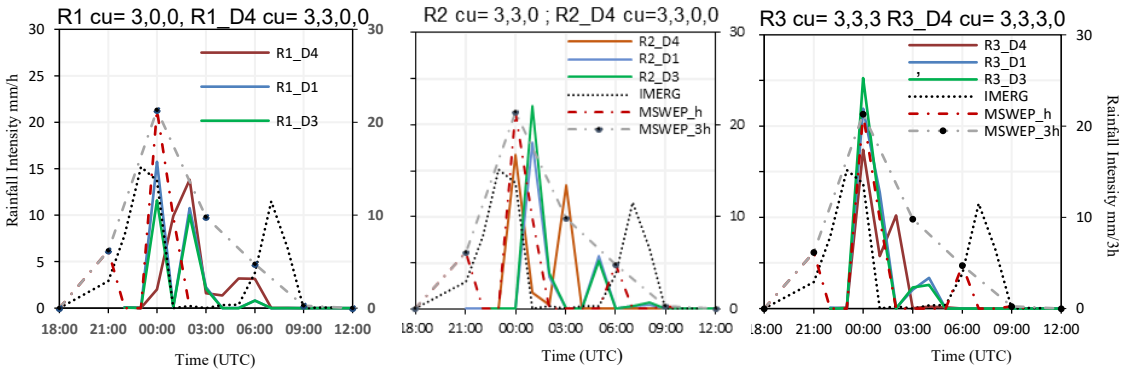


Figure 4.11 Rainfall time series for all runs D1(10km), D3(1km) and D4 (0.37km) from 18:00 Oct 24th to 12:00 Oct 25th 2015 at a known flood location. MSWEP is shown for 3h accumulations and an extreme case 1h accumulations. R1, R2and R3 correspond to runs with 3 domains and R1_D4, R2_D4 and R3_D4 correspond to 4 domains runs. Cu=3 represents the Grell Freitas Cu scheme and 0 indicates no cumulus scheme in the respective domains. The left axis mm/h and the right axis shows mm/3h corresponding to MSWEP_3h

Overall, the three domain runs all matched the accumulated observed rainfall, whereas the D4 runs either underestimated or overestimated in the case of Nov 2020. These results suggest that while the inner domains exhibit higher rainfall accumulations than the parent domain, the rainfall peak does not always increase with the higher resolution, as shown with D4. This could be explained since having no cumulus scheme in the smaller domains tends to decrease rainfall in those domains. Given two-way nesting was used, feedback between domains reduces the rainfall in the outer domains and perhaps influences the distribution in the outer domain. Although the 4th domain (except for R1_4d), better matches the peak intensities of the satellite-derived rainfall forecasts, there is a lower total rainfall forecast compared to MSWEP and IMERG.

4.4.2 Rainfall Threshold Analysis

Hazards are classified as "No to minimal flooding", “Minor Flooding”, “Significant flooding” and “Severe flooding” based on 24h accumulated rainfall (Table 4.2). Analysis of thresholds showed that results varied with the treatment of convection for the different runs. Classifications are shown for different districts in Alexandria for MSWEP and IMERG (Figure 4.12). For the event on 25th October 2015, the forecast hazard classes, derived based on Table 2, varied from a significant to severe flood event along the coastal areas (Figure 4.13). While MSWEP and IMERG indicated a hazard class “Significant flooding” in Montazah, Sharaq and Wasat districts, Gomork and parts of Almeria were predicted to be more extreme. There were minimal differences when comparing the 3.3km, 1.1km and 0.37km grid sizes. Compared to the 10km grid the smaller resolutions detected more spatial variability in the hazard classification and areas of higher rainfall. The 10km grid size tends to average out the rainfall over the grid. Figure 4.13 shows how the hazard warning class varied for the 10km and 1.1km for the 3-domain run and 0.37km for the 4-domain run for the different runs and cumulus configurations for a 24h accumulation for different districts in the city. Most of the runs were able to predict this extreme class except R1_4d and R3_4d, which predicted a less severe but significant class for Gomork and neighbouring districts. R1, R3 and R2_4d over-predicted a more severe event in Montazah and neighbouring areas. Some runs showed variable classes compared to MSWEP and IMERG, but all runs correctly predicted a reduction in the total rainfall inland and different runs were able to distinguish different hazard classes at districts and subdistrict levels, which is useful when location-specific measures need to be taken. This will aid decision-makers in prioritising, coordinating and allocating resources to areas at most risk, combined with prior knowledge of hot spot areas. This could be achieved even without building a stormwater model of the city.

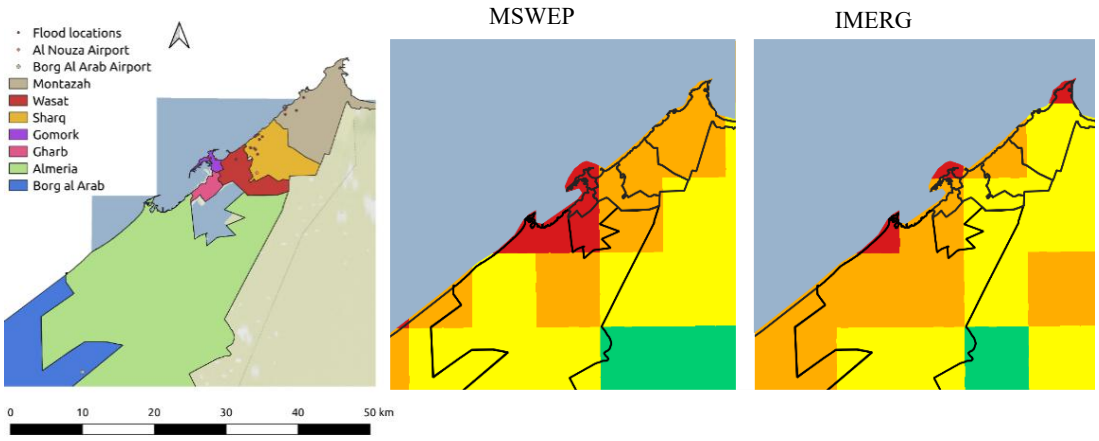


Figure 4.12. Left: Alexandria Districts; the colours have been arbitrarily chosen to display districts. Right: Hazard classification for Alexandria districts for the 25th October 2015 flood event for 24h accumulation for the Oct 25th event based on 3-hourly MSWEP and hourly IMERG data. The colour codes signifying hazard classes are also shown in Table 4.2

Categorical scoring was used to evaluate if the warning classes in Table 4.2 were consistent with the observed event category. A POD and CSI of 1 and a FAR of 0 are considered a perfect forecast whereas a POD and CSI above 0.75 considered useful. The results (Table 4.4) show that the D3 (1.1km) 3 domain runs consistently had higher POD and CSI scores, meaning these runs were better at detecting an event and lower false alarms. Both runs with the 4th domain R1_D4 (10km) and D4_4d (0.37km) achieved lower scores with all the runs with all the thresholds. Within the 3domain runs, results vary with thresholds and cumulus configuration. In general, higher FAR scores were found for rainfall accumulations over 20 and 32 mm. However, either Run2 or R3 performed better when the cumulus is turned off and on in the smallest domain. All runs had good performance above the 50mm threshold rainfall which indicates all forecasts had agreement in not predicting “very extreme” rainfall accumulations.

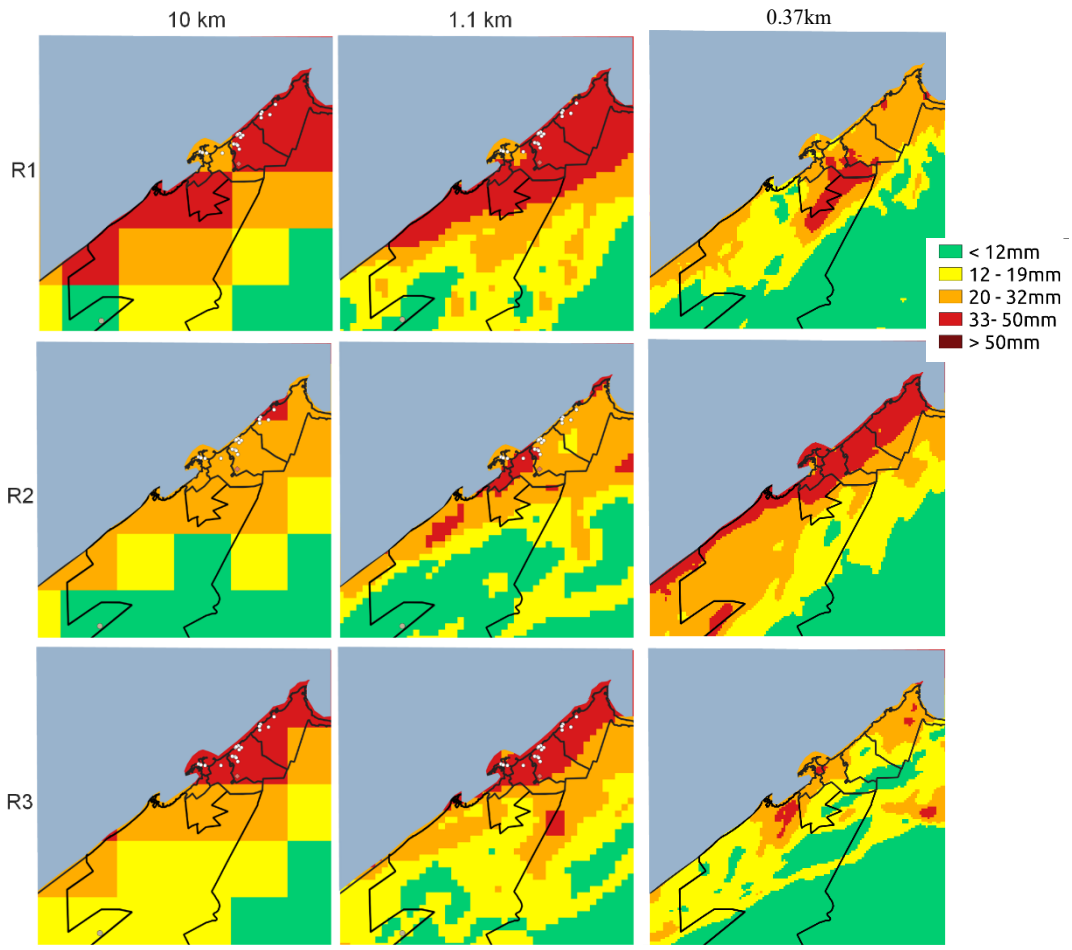


Figure 4.13 Forecast of hazard classes with a lead time of 24 h for the Oct 25th, 2015 flood event for different runs based on the 24h rainfall accumulations. Left:(10 km) and Middle 1.1km grids for run downscaled to 3 domains. Right 0.37km grids downscaled to 4 domains

Table 4.4 Average performance metrics; Probability of detection (POD). False Alarm Ratio (FAR) and Critical Score Index (CSI) for) all runs for all events for the 50km neighbourhood scale. Results are presented for the inner most domains 1.1km and 0.37km respectively.

	0.1			**12**			**20**			**32**			**50**		
	R1	R2	R3	R1	R2	R3	R1	R2	R3	R1	R2	R3	R1	R2	R3
								POD							
D1	1.00	1.00	0.95	0.73	0.68	0.72	0.70	0.57	0.66	0.63	0.64	0.72	0.91	0.93	0.81
D3	1.00	1.00	0.96	0.76	0.71	0.76	0.68	0.55	0.73	0.63	0.67	0.73	0.87	0.93	0.91
D4	0.96	0.96	0.97	0.65	0.69	0.68	0.49	0.51	0.49	0.61	0.56	0.57	0.92	0.88	0.88

	0.1			12			20			32			50		
							FAR								
D1	0.00	0.00	0.05	0.29	0.35	0.26	0.28	0.45	0.30	0.33	0.36	0.31	0.11	0.07	0.08
D3	0.00	0.00	0.04	0.25	0.33	0.27	0.34	0.43	0.31	0.41	0.33	0.24	0.10	0.07	0.20
D4	0.04	0.04	0.03	0.33	0.28	0.31	0.51	0.48	0.49	0.43	0.41	0.45	0.08	0.12	0.16
							CSI								
D1	1.00	1.00	0.91	0.53	0.61	0.56	0.41	0.49	0.38	0.49	0.49	0.53	0.87	0.83	0.87
D3	1.00	1.00	0.93	0.64	0.56	0.62	0.53	0.61	0.55	0.43	0.51	0.61	0.82	0.88	0.68
D4	0.93	0.93	0.95	0.54	0.61	0.57	0.35	0.37	0.36	0.41	0.43	0.39	0.86	0.79	0.73

4.4.3 Coupled WRF and Urban flood forecasting

The MIKE plus model was used to evaluate peak discharge and volume at the outlet of the El Gomork district. The model was run with gridded forecasts rainfall at different resolutions and runs corresponding to different cumulus configurations: D1(10km), D2(3.3km),D3 1.1km and D4(0.37km), with a 24-hour lead time. In the absence of measured flow data, results were verified against model simulation runs with satellite-derived IMERG and MSWEP 1h derived rainfall intervals and historical knowledge of floods. The lift station pump at outlet #1 has a design capacity of $0.9m^3/s$ therefore, it is assumed that flooding will occur once this flow is exceeded. Peak discharge reached 30mins after peak rainfall.

Discharge and volume at outlet

The results for Oct 2015 in Figure 4.14 show that forecast rainfall corresponding to R3 led to the highest simulated discharge for all resolutions and showed higher peaks compared to MSWEP and IMERG. The discharge for each run followed a similar pattern of the rainfall distribution. The discharge at the outlet simulated with the rainfall forecast using R2 and R3 for the 10km (D1) and 1.1km(D3) showed peak discharges similar to that with MSWEP_h and IMERG (~$1.5m^3/s$). The simulated discharge for R1 forecasts showed a different pattern with two lower peaks. Simulations using rainfall forecasts based on R3 and R2 for D3 and D1 gave peak discharges similar to MSWEP. The runs also gave the highest KGE scores (0.4-0.78) and the lowest NRMSE of 0.43 when compared with the other runs and D4 (Table 4.5). The exception being Dec 2018 R3_D3 which gave had a KGE of 0.02 and NRMSE of 0.98 Figure 4.14 also shows discharge simulations using rainfall forecasts based on 10km, 1.1km (D3) and 0.37km (D4) resolutions, the latter which are the innermost domains of WRF runs. All runs underestimated the discharge with the rainfall forecasts using the 0.37km resolution with NRMSE

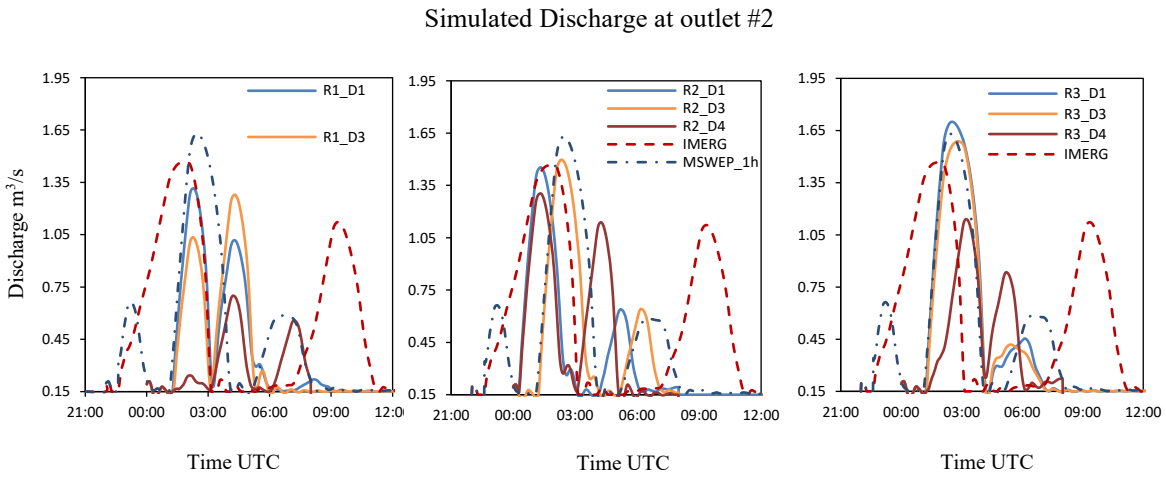


Figure 4.14 Simulated Discharge at outlet #2 on Oct 25th, 2015 for R1, R2 and R3 for D1 (10km), D3 (1.1km) and (D4) 0,37km which corresponds to runs with 4 domain and the cumulus scheme is turned off in the 4th domain

ranging from 0.83 to1.09. Similar results were found for the Oct 2018 and Nov 2020 events which gave lower NRMSE values and higher KGE for 1km domain and higher NRMSE values and lower KGE scores for the 0.37km runs. However, except for Oct 2015, all the runs and domains resulted in high NRMSE values, which means the models still have challenges predicting the timing of the rainfall but the higher KGE indicates some agreement in the correlations and the spread. The NRMSE and KGE values for the three events in shown in Table 4.5.

Table 4.5 NRMSE and KGE values for simulated discharge at outlet #2 for the three events. D1_R3 (10km) performed the best which uses the adaptive cumulus scheme in the smallest domain whereas the highest error was shown the 0.37km domain when the cumulus scheme is turned off. The best score for each run per event is highlighted

NRMSE

	R1_D1	R1_D3	R1_D4	R2_D1	R2_D3	R2_D4	R3_D1	R3_D3	R3_D4
2015	0.80	1.10	1.09	1.12	0.49	1.15	0.42	0.43	0.83
2018	0.97	0.97	1.04	1.01	1.01	0.99	0.98	0.98	1.05
2020	0.83	0.95	1.06	0.96	0.96	0.97	0.93	0.64	1.07

KGE

	R1_D1	R1_D3	R1_D4	R2_D1	R2_D3	R2_D4	R3_D1	R3_D3	R3_D4
2015	0.43	-0.37	-0.35	-0.02	0.68	-0.08	0.77	0.78	0.31
2018	0.18	0.11	0.06	0.23	0.40	-0.09	0.22	0.02	0.09
2020	0.22	0.29	-0.04	0.21	0.24	0.27	0.20	0.66	-0.26

The simulation using rainfall forecasts based on the R3_D3 (1.1km) grid produced the highest total volumes at the outlet. Simulated discharge based on R3 forecasts gave values closest to MSWEP_h, however, simulated discharge with R1_D4 and R3_D4 forecasts gave considerably lower volumes compared to MSWEP_h and IMERG-based simulations (Table 4.6). Although IMERG resulted in a peak discharge similar to MSWEP_h, the total runoff volumes were higher compared to that with MSWEP_h simulated at the outlet because of the higher total accumulated rainfall.

Table 4.6 Total volume of water discharged at the outlet for the Oct 2015 event for the different WRF runs and resolutions

Input rainfall /resolutions	IMERG	MSWEP_h	WRF R1	WRF R2	WRF R3
	Total Volume at Outlet m^3				
10km	21544	14234			
10km			10635	10103	14171
1.1km (D3)			10476	10741	12374
0.37km (D4)			4955	10929	9206

Figure 4.15 Known flooding locations in El Gomork along the Corniche Road in Alexandria. On October 25, 2015, this location experienced flooding depths ~0.15 - 0.3m.

Node flooding

The flood results above focused on the simulation of the entire system for El Gomork and found that the R2and R3 produced good simulations of the peak of flooding for the 10km and 1.1km. This is also observed when analysing node flooding at a known flood location /.in the city. During the October 25, 2015 event, flood depths of 0.15 – 0.3m (Figure 4.15) were reported at this location and in the vicinity. Figure 4.16 shows the simulated node

flooding at a selected node with the rainfall forecasts for the 10km, 1.1km and 0.37km with the hourly MSWEP and IMERG rainfall. The rainfall intensity and spatial distribution of grid rainfall used for the 10km, 1.1km and 0.37km simulations are shown in Figure 18. For both the R2 and R3, the 1.1km grid shows a flood depth of ~0.4m, which is similar to the depth simulated using MSWEP but approximately 0.1m above the flood depth simulated using IMERG derived rainfall and the reported 0.3m flood depth. However, compared to the discharge for the system which showed similar peak values and distributions, for the 10km and 1.1km, higher node flooding was simulated at the manhole. This is because of the 1.1km runs showed higher rainfall values and greater variability in the rainfall which is not detected in the larger 10km grid as shown in Figure 4.17. Although the timing of the flood was well predicted for the Oct 2015 event, there was a delay in the timing of the flood was also observed in the Oct 2018 and Nov 2020 event.

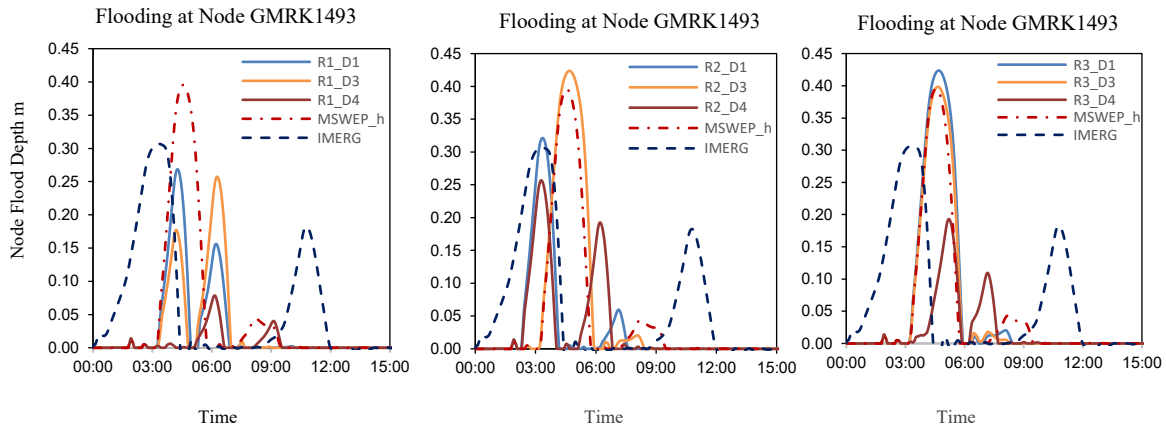


Figure 4.16 Node flooding depths (m) at known flood location in Gomork on Oct 25th, 2015. For all runs at D1(10km), D3(1.1.km). and 0.37km- 4 domains (orange). The horizontal dotted line indicates the observed flood depth at 0.3m

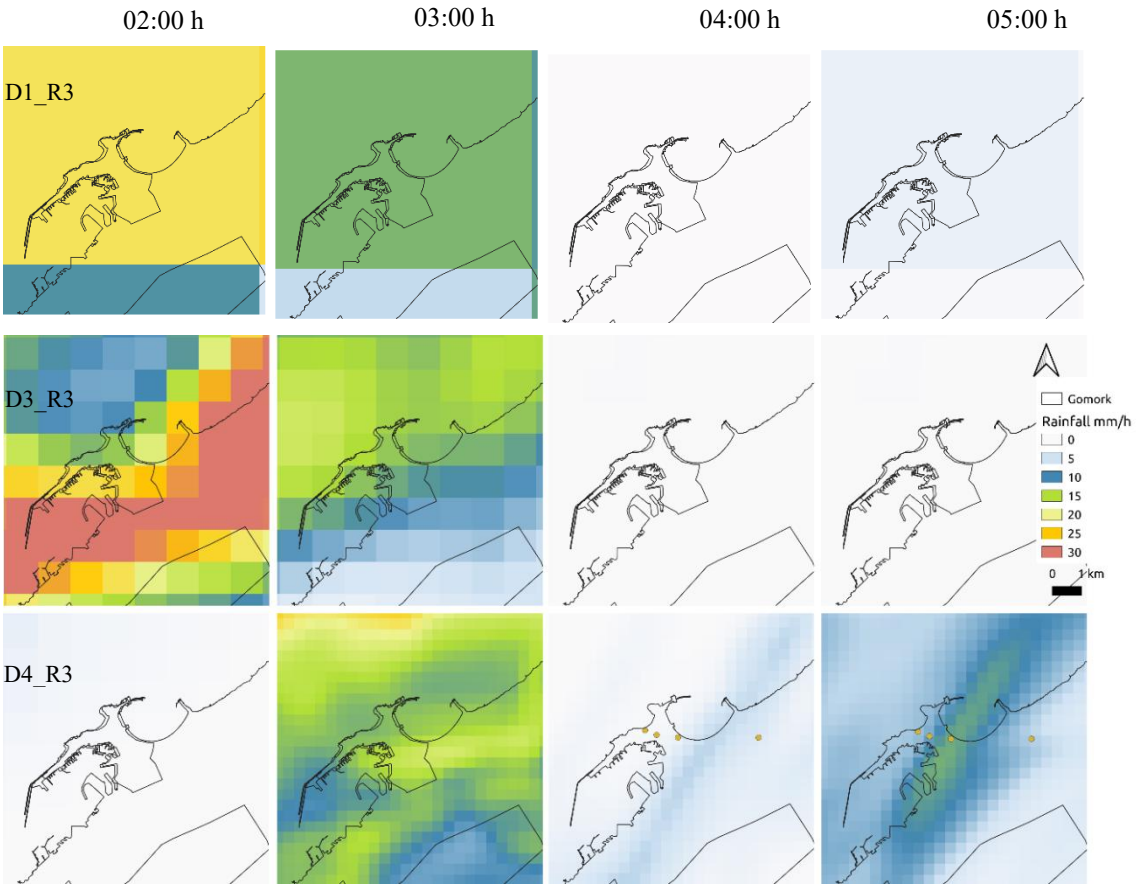


Figure 4.17The rainfall intensity and spatial distribution of grid rainfall used in the MIKE model for 10Km, 1.1 km grid resolution for run R3 and 0.37km R3_D4.

Figure 4.18 shows a comparison of node flooding greater than or equal to 0.3m for R3 and R3_4d runs for the 1.1km grid. R3_D3 shows a higher number of nodes (101 nodes) compared to MSWEP 83 flooded nodes, R3_D1 118 nodes, R3_D4 shows 32 nodes flooded above 0.3m, approximately 1/3 of nodes flooded for R3. Table 4.7 summarises the difference in the number of nodes flooded between the simulations with observed rainfall and the R3 forecast for the Dec 2018 and Nov 2020 events. The findings show that D4 either under forecasted or over forecasted the number of flooded nodes compared to the other grid sizes. Knowledge of which nodes are forecasted to flood can help decision-makers to implement protective measures before an event and coordinate response during an event However, false alarms on the number and location of flooded nodes can lead to the misallocation of resources, flooding and further traffic disruptions.1D-2D models could be run to give a more accurate characterisation of flood inundation however there are still challenges in reducing computational time for real-time applications and high-resolution data requirements.

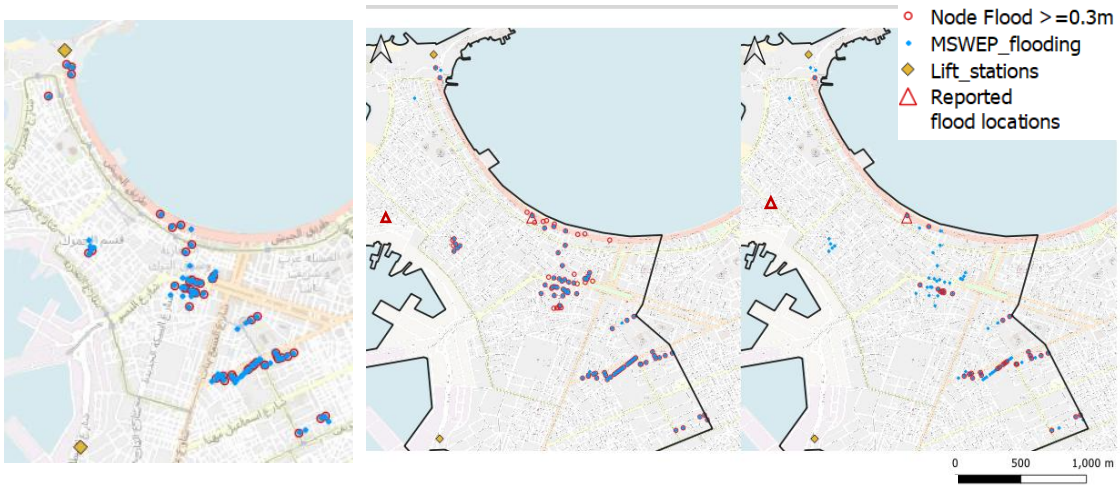


Figure 4.18 For the Oct 2015 event. Maximum node flooding >= 0.3m for R3. The 10km grid (left), 1.1km grid for runs (middle) and R3_4d(right). MSWEP simulated flood locations are shown in blue. More locations were detected for R3_D3 (101 flooded nodes) compared to R3_D1 (118 nodes) and R3_D4 (32 flooded nodes) and MSWEP (83 flooded nodes)

Table 4.7 Difference in the number of flooded noted between simulated 1D runs with MSWEP or IMERG and R3_D1 (10km), R3_D3(1.1km) and R3_D4(0.37km)

	R3_D1(10km)	R3_D3(1.1km)	R3_D4(0.37km)
2015	30	13	-56
2018	1	34	-8
2020	-10	-8	108

4.5 Discussion

4.5.1 Trade-offs in model lead time, domain resolution, cumulus configuration and flood forecasting methodology

Our results reinforce different factors, can influence rainfall structure, magnitude, location and timing in high-resolution WRF simulations, such as spatial resolution, number of domains, lead time and model parametrisation.

This research evaluated the forecast skill at different neighbourhood sizes related to the scale of the phenomena and city level. Different neighbourhoods were evaluated to highlight how forecasts vary with the neighbourhood scales being considered. The 1000km neighbourhood scale detected the presence of a weather system however, this

was not very useful for the coastline. However, the metrics vary considerably at scales below 200km. Overall, the 24 LT performed best and the forecasts showed skill up to the 72hr LT but less at the 48h and the 12hr LT. Although the 72h rainfall forecast is less accurate than the 24h rainfall forecast, the FSS scores showed there is still some skill associated with this forecast. More variability is observed at neighbourhood scales below 200km and highlights the difficulty in evaluating forecasts at smaller neighbourhood scales and longer and shorter lead times which may be more important for flood forecasting and emergency managers.

While many urban flood forecasters may not consider the treatment of convection, it becomes an important factor in rainfall and flood forecasting given the scale dependence of cumulus schemes at high resolutions needed for grid-based hydrodynamic models. This was especially observed in modelling the flood depths. The Grell- Freitas convection scheme was selected given it is a scale-aware scheme suitable for resolutions <10km - 3km. Our results indicate that turning the CPS on and off influenced simulated rainfall spatial and temporal patterns, which is in line with previous studies (Jeworrek et al., 2019). All runs and events forecasted the occurrence of an extreme rainfall event but with variability in magnitude and placement, especially at neighbourhood scales less than 200km. The effect of the parameterisation varied with events, which is expected as effects vary with the size and characteristics of the event and depend on the role of convection in the event formation. One notable difference is the mismatch of the timing of the peak, which results in a delay in peak flood time. When comparing the 10km and 1.1km resolution it was shown that even though the inner domains are exhibiting higher rainfall accumulations than the parent domain, rainfall peaks do increase when the CPS in turned on in the smallest domains (R3). This is similar to the findings of Robaa and Wahab (2019) and Ibrahim (2020) who evaluated the sensitivity of WRF to convection schemes and found there was an overestimation of rainfall when cumulus schemes are turned on in the smaller domains.

When a 4th domain was used to capture a smaller resolution, it appears the runs either under-forecasts or over forecasted. In general, downscaling to 0.37km using a 4th domain did not improve the rainfall forecasts but instead resulted in lower or higher rainfall totals and flood depths when compared to MSWEP and IMERG. While the findings indicate using a 4th domains to achieve finer resolution did not provide any improvement. It was not possible to draw a definitive conclusion about whether R2 or R3 gave better results as the results varied according to the events and their characteristics. However, when the cumulus schemes are turned off in R1, forecasts intended to under forecasts which suggest that CPS scheme should be used for grid sizes less than 10km.

4.5.2 Varying scales with flood forecasting approaches

Generally, in urban flood simulation and forecasting, accuracy is a key factor (H. Wang et al., 2022). The results show challenges in using WRF for flood forecasting, especially when trying to predict flooding at manholes. Although there is a mismatch in the timing, there is better agreement on the magnitudes of the peak. Despite these differences in the results, the models were able to predict an extreme event which resulted in flooding. For both the rainfall threshold method and flood model, the runs were successful in detecting an extreme event but performed better with the 10km and 1.1km (D3) resolutions for the higher thresholds 20 mm and 32 mm while the 0.3km gave the lowest POD and CSI scores and higher scores for the FAR. All runs performed well above the 50mm rainfall, indicating that all forecasts agreed with not predicting "very extreme" rainfall accumulations. For the 1D real-time simulated discharge, the 10km and 1.1km(D3) grid also performed best with NRMSEs of 0.42 and 0.43 and KGE above 0.4 for R2 and R3 but in general, the 0.37km performed worst Our study indicated no gain in increasing resolution below 1 km when evaluating the flood model results irrespective of the cumulus configuration for Alexandria, Egypt Therefore, while it is assumed that higher resolutions will improve the spatial extent and durations of rainfall outputs by capturing more realistic sub-grid model dynamics, the finer spatial resolution can also introduce errors which does not justify the increased computational resources required for running such high-resolution models (Kain et al., 2008; Roberts, 2008; Schwartz et al., 2009).

There are some challenges to overcome and improvements to gain in achieving the 1km resolution. The big advantage of this resolution is that it is suitable for running small-scale flood simulation models and can capture the variability in rainfall, thereby improving the quality of the flood forecasts and the opportunity to increase the effectiveness of flood risk management actions at specific locations. However, when more city scale actions are to be taken, such as issuing warnings, the coarser 10km is suitable. Previous studies (Liu et al., 2012; Woodhams et al., 2018; Goodarzi et al., 2019; Liu et al., 2021) also found good rainfall simulation results at 3–4 km horizontal grid spacing, but it should be noted these studies did not consider a hazard estimation method. Thorndahl *et al. (*2016) highlighted the limitations and trade-offs of using a 3 km grid NWP model vs 0.5 km radar rainfall for urban pluvial flood forecasting to capture the rainfall variability and intensity needed for accurate flood predictions in urban environments. While the study found the 3 km grid NWP rainfall was limited, our study found the 1km grid resolved NWP rainfall had some skill in detecting the correct flood class, flood depths and number of flooded manholes, especially at lead times up to 24 hrs. Therefore, when developing a flood forecasting system, lead time, cumulus configuration and domain resolutions should also be aligned with the flood forecasting approach, the capability of local decision-makers and the requirements for specific actions.

4.5.3 Making imperfect forecasts useful

The rainfall threshold method is quick, simple and convenient, giving forecasted flood classes and does not require complex models. The 1D model, forecasted discharge, flood depths and the number of flooded manholes. The analysis results indicate challenges in using WRF forecasts for flood forecasting at the city scale for both methods. Therefore, even though the results do not always match the flood classes per pixel or flood depths of the observed runs exactly, the forecasts have sufficient skill to raise alert and preparedness levels. Lead time is essential for forecast-based actions considered in the analysis and depends on the specific action. Hazard-reducing actions such as cleaning drains and increasing storage via pumping can range from a few days to a few hours, depending on the event's severity, whereas exposure-reducing measures such as traffic measures and evacuations require hours.

Results found that although the 72h rainfall forecast is less accurate than the 24h rainfall forecast, there is still enough skill to use these forecasts to detect the phenomena and the city's scale. However, below the 50km scale, these other lead times are shown to be less skilful. Therefore, instead of using one method, the rainfall threshold and flood simulation results can be combined and used by decision-makers to make a decision especially when incorporated with predefined emergency preparedness protocols and triggers. Once a threat is detected in these forecasts, the rainfall threshold methods can be used to identify at-risk areas at the governorate scale. Given the skill of the 24h lead time, the WRF can then be run at a higher resolution at this lead time so that rainfall forecasts can be used in the hydraulic model to forecast areas/ districts/ manholes at risk of flooding.

While our research only examined three events, further studies are essential to evaluate more events and the robustness of thresholds that capture the non-linearity of flooding. The real-time flood model requires more data, computational resources and model calibration and uncertainty in model inputs especially when physically-based modelled are used to represent hydrological processes. The advantage of real-time approaches lies in their ability directly model spatial variability and its impacts at city scale which is beneficial to areas that do not have high-resolution observations (Hofmann & Schüttrumpf, 2019; L. J. Speight et al., 2021). Therefore, while the accumulated rainfall. Given the reliance of the peak intensity on and as an indicator, it may be more useful to use peak intensity vs accumulation or timing as an indication for potential flooding (Umer et al., 2021). The practical value of these results depends on how a decision maker values forecast accuracy and which variables are most important to decision-making: peak intensity, peak discharge, flood depth, total number of flooded nodes or the identification of a threat.

This allows decisions to be reviewed and updated in tandem which increases preparedness at longer lead times This can be valuable for complex data-scarce urban areas that lack information to refine models or where long lead times are needed to increase preparedness

(Speight et al., 2021, Corral *et al.*, 2019) compared a rainfall-based forecasting system using hazard level and a rainfall-runoff model to compute the stream flows at pixel scale and found similarities in the use for early warnings for flash flooding.

4.5.4 Limitations and Future Research

The unavailability of sub-daily data is one of the biggest challenges in performance evaluation in data-scarce regions. Gauge data is only available at a daily resolution and the satellite-derived rainfall estimates and reanalysis data are at coarser resolutions than the forecasts and have their own bias. This was highlighted by the differences in MSWEP and IMERG data. Still, our results showed that MSWEP and IMERG data were able to corroborate the occurrence of extreme rainfall and node flooding for respective events. However, there was some variability in the severity. This further supports the use of satellite and reanalysis data in the absence of gauge data (Mekonnen et al., 2023; Nashwan et al., 2019) for the rainfall threshold approach, sensitive to thresholds and in the absence of measured discharge data to evaluate the forecast data.

This research evaluated three events. It did not aim to make a conclusive statement about best model configurations and setup but rather to assess how hazard classes vary when high-resolution precipitation at different scales are used with different flood forecasting approaches in data-scarce regions. Post-processing and data assimilation of forecast precipitation are methods used to improve the accuracy, sharpness and reliability of the forecast by removing systematic errors (Verkade *et al.*, 2013; Crochemore *et al.*, 2016), but it is still reliant on the availability of robust observational data at suitable resolutions which were not available for this study. Data assimilation and bias correction methods such as Kalman Filter or quantile mapping (Bárdossy et al., 2021; Mapiam et al., 2022) should be incorporated into the flood forecasting chain to update forecasts when real-time data is available.

Future research should consider how performance varies with rainfall and storm characteristics. Sources of uncertainty have not been elaborated in this research. In particular, there is uncertainty in initial conditions and urban model uncertainty, which are important components. Probabilistic or ensemble approaches can be used to explore uncertainty in the placement and distribution of rainfall (Böing et al., 2020; Brendel et al., 2020; Ravazzani et al., 2016; Rico-Ramirez et al., 2015; Yang et al., 2016; Young et al., 2021). The ultimate goal will be more accurate, higher resolution spatial and temporal rainfall estimates through weather radar or an increased number of gauges and collection of discharge data. The present study was limited by the availability of photos and impacts to validate the models. In the future, a database of historical flood hazards and impacts will remain invaluable and should be prioritised for verifying and determining expected impacts.

4.6 CONCLUSIONS

This research employed a limited area Weather Research and Forecasting (WRF) model in the context of urban-scale flood forecasting for Alexandria, Egypt, to evaluate if increasing the spatial resolution along with varying treatments of convection improves flood forecasting using a rainfall threshold and real-time urban model. It was found that the WRF model effectively identified extreme weather phenomena and the rainfall threshold method correctly distinguished hazard classes at the district level. However, rainfall magnitudes, distributions and classes depended on the model configuration and cumulus scheme used, particularly at convection-resolving resolutions (>3km). A timing delay of peak flows and flood depths was observed, however, the 1.1km grid resolution was found to provide the best performance, especially when comparing flooding at specific manholes. The rainfall threshold gave POV and CSI values above 0.5 for all events for the 20mm threshold and the outflow discharge KGE values ranged from 0.4-0.8 for the 1.1km. For both the rainfall threshold and urban flood modelling approaches, there was no improvement in using a fourth domain at (0.37km) over the 1.1km or 10km grid resolution regardless of the cumulus configuration. POV and CSI values were below 0.5 for all events for all runs and KGE values below 0.27. Both the 1.1km and the 10km resolution require fewer computing resources than the 0.37km. Therefore, while higher rainfall resolutions are required for urban scale modelling, this study highlighted that smallest resolution did not lead to improved hazard estimation. Additionally, the runs exhibited good skill in detecting extreme events at the 72LT but less skill at the 12hLT lead times and at various useful neighbourhood sizes, with the rainfall threshold method for all runs at coarser resolutions.

Despite challenges in using high-resolution rainfall for flood forecasting, the good performance of both the rainfall threshold and real-time simulation methods suggests the potential for their combined use rather than using one over the other. Cities can use real-time simulation methods as more data becomes available to build models. Combining WRF forecasts with rainfall threshold methods allows early warning and flood forecasting at longer lead times without having to run complex flood models. When flood models are available, they can be coupled with WRF forecasts to provide more detailed information on the specific flood locations and depths at shorter lead times, but these results are also sensitive to the WRF model configurations as smaller resolutions do not necessarily predict more accurate forecasts. Combining a rainfall threshold and real-time forecasting model with a suitable cumulus configuration capitalises on the strengths of each method, which is invaluable for decision-makers in coordinating resources and implementing location-specific measures, particularly in conjunction with prior knowledge of high-risk areas.

Limitations for high-resolution modelling have been focused on the limits of computational resources but this research highlights there are also limits on quality. It is

best to understand how methods can complement each other in data-scarce regions while being mindful of lead time, model configurations, cumulus schemes and resolutions. Forecasters and end-users must determine the useful resolutions and neighbourhood sizes that align with the specific goals of flood forecasting applications. As technology continues to advance and more real-time data becomes available, the use of high-resolution rainfall models hold promise for improving local forecasts and enhancing our ability to mitigate flood-related risks. Overall, this study advances our understanding of urban-scale flood forecasting in data-scarce regions: it stresses the importance of considering lead time, neighbourhood size, resolution, cumulus configuration and integration with drainage models. The findings offer a foundation for the development of effective flood forecasting methodologies for anticipatory flood management and decision support tools despite the challenges in data-scarce cities. Finally, this research reiterates the need to understand the independencies and limitations of the aspects of the flood forecast chain; the meteorological output, the flood forecasting method and the data and methods used for verification with the overall goal for improved decision-making and preparedness.

DATA AVAILABILITY STATEMENT

The GFS forecasts data that support the findings of this study are openly available from the National Centre for Atmospheric Research (NCAR) Research Data Archive (RDA) at https://rda.ucar.edu/datasets/ds084-1/. MSWEP data is openly available from GloH2O at https://www.gloh2o.org/mswep/. GPM IMERG Final data is openly available from the National Space Agency (NASA) at https://gpm.nasa.gov/data/directory

Other data underlying this Chapter can be found as Version 1. 4TU.ResearchData. dataset. https://doi.org/10.4121/a98b55f5-548c-4f0f-ae29-da1552774b5f.v1

5

ENSEMBLE RAINFALL VERIFICATION

5.1 INTRODUCTION

In cities, the success of flood forecasting requires predictions of flood occurrence, water levels, flood depths, timing and locations, with sufficient information and lead time to issue warnings and implement early actions. The term predictive uncertainty refers to the probability of a future event given the information available up to that time (Todini *et al.,* 2018). Predictive information can be obtained from meteorological forecasts for use in rainfall threshold methods or for feeding into hydrological and hydraulic models. Predictive uncertainty can be represented through statistical post-processors, multi-model ensemble forecasts, model updating through data assimilation and expert forecasters' judgment.

While statistical methods apply a post-processor to a deterministic forecast, ensemble approaches generate multiple runs, whereby every run is considered equally likely and plausible. However, they are inherently associated with large uncertainties, reducing their reliability and value, especially in data-scarce regions. The primary sources of uncertainty arise either from input data (i.e., physical measurement errors, differences in spatio-temporal scales between model and measurements and meteorological forecasts) or from the model itself through the mathematical simplification and parameterisation of physical processes (Borga et al., 2008).

Several studies have provided guidance on appropriate uncertainty methodologies (Pappenberger et al., 2006) and frameworks for identifying, estimating and communicating uncertainty (Beven et al., 2018; Doyle et al., 2019; Hall & Solomatine, 2008). With the adoption of ensemble weather predictions from numerical weather prediction models, probabilistic approaches have been used to quantify uncertainty for risk-based approaches. Each ensemble can be translated into a hazard state (Chapter 4) or ensembles of surface runoff, discharge or flood depth, giving a representation of uncertainty and increased confidence in the forecast (Boelee et al., 2019; Cloke & Pappenberger, 2009; Duan et al., 2018; W. Wu et al., 2020).

In section 1.5, we have argued for the use of probabilistic and uncertainty information within a structured, formalised risk-based decision-making process. Building on this, the next chapter (Chapter 6) proposes to incorporate the ensemble forecasting uncertainty into a Bayesian Decision Framework for anticipatory flood management. The main objective of this chapter is to assess the skill of the ensemble rainfall forecasts and to develop probability distributions conditioned on critical rainfall thresholds that will be used in the Bayesian Decision Framework (Chapter 6). Therefore, the focus of this chapter is limited to presenting the ensemble forecast verification to be used in the next chapter. This chapter first outlines the model setup and data used for the ensemble forecasts verification metrics.

5.2 DATA & METHODOLOGY

This research analysed simulated ensemble rainfall over the city of Alexandria, Egypt at the 24h lead time. The Weather Research Forecast (WRF) model was initialised using the Global Ensemble Forecast System (GEFS), which provides 21 ensemble members. The ensemble rainfall was first verified using GPM-IMERG Final Precipitation Products V7 (Huffman et al., 2019) henceforth referred to as IMERG. The WRF setup used for the ensemble rainfall runs is described in the next section.

5.2.1 WRF model and configuration

Ensemble forecasts were sourced from the National Oceanic and Atmospheric Administration's (NOAA) GEFS archive. Records are available from 2017 to the present. Before 2017, historical forecasts were archived at 1-degree resolution; since 2017, they have been available at 0.5 degrees. The WRF model was set up with three domains, with a nesting ratio of 1:3:3 and spatial resolutions of 30 km for domain 1 (D1), 10 km for domain 2 (D2) and 3.3 km for domain 3 (D3) (see Figure 5.1). The inner domain (D3) covers Egypt's north coast and the wider Mediterranean. All spatial domains had a vertical resolution of 45 layers. The ensemble runs used the same parameterisation and physics schemes as the deterministic runs, adopted from Lui et al. (2021) and Young et al.,(2025). The national high-performance computing system Snellius, operated by SURF in the Netherlands, was used to perform ensemble simulations. The system runs on a Linux-based operating environment, which was used to configure the WRF Pre-processing systems (WPS) and WRF models before execution on the SURF nodes. Model runs were distributed across multiple computing nodes, each equipped with dual AMD Epyc Rome processors (64 cores per processor, 128 cores per node) and 256 GB of RAM, enabling large-scale parallel computing for different parameterisations, domain configurations and multiple ensemble forecast runs.

5.2.2 Events

Events were selected from the period Oct 2018 – Jan 2021 based on the availability of historical flood records. Events were selected using a criterion of daily observed gauge rainfall exceeding 10mm/day at the Nouzha airport and the presence of a flooding record. In total, 15 events exceeding 10mm/day were identified for the period Oct 2018 to Jan 2021. Given the large computational requirements for ensemble runs, the events were reduced to 9, which provide a mix of low- and high-intensity events with a record of flooding. For the selected period, three (3) events were chosen for the winter period, Oct 2018 – April 2019, four (4) for Oct 2019- April 2020 and two (2) for Oct 2020-Jan 2021 (Table 5.1). The timing of the events was verified on the weather website Tutiempo, which provides a record of weather conditions as part of daily Meteorological Aerodrome Reporting (METAR). Event 0 was used as the calibration event. Analyses were done for the 12, 24, 48 and 72 h lead times, but only the results for the 24 h lead time are presented.

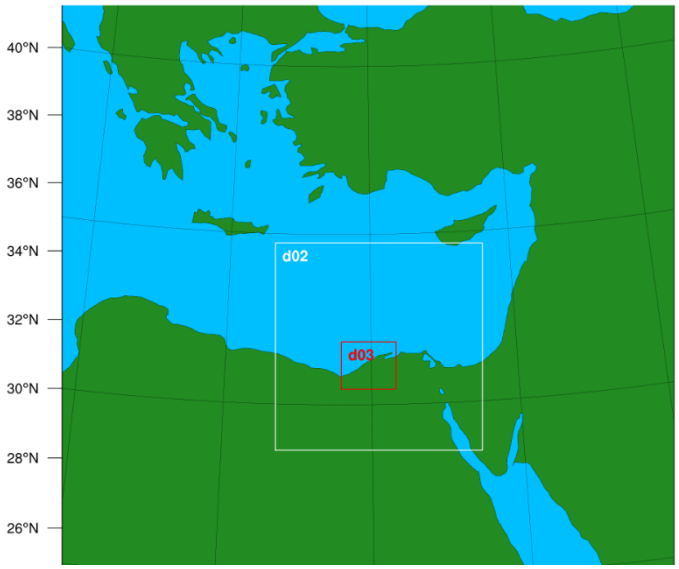


Figure 5.1 Ensemble run domain configuration for D1, D2 and D3 GEFS Ensemble runs

Table 5.1 Summary of rainfall events used in the hindcast analysis from 2018-2020. Event 0 is used as the calibration event. Gauge rainfall source from https://www.tutiempo.net/

Event #	Year	mm/dd	Gauge mm/d	Event #	Year	mm/dd	Gauge mm/d
0	2015	10/25	32.1	5	2019	12/25	17.02
					2019	12/26	21.08
1	2018	11/09	17.02	6	2020	01/03	10.92
					2020	01/04	36.07
2	2018	12/04	29.97				
	2018	12/05	23.11	7	2020	03/12	7.87
					2020	03/13	27.94
3	2019	01/14	10.92	8	2020	11/07	13.97
						11/08	10.92
4	2019	10/24	12.50	9	2020	11/20	13.97
	2019	10/25	28.81		2020	11/21	18.03

5.2.3 Rainfall Verification

Categorical Scores

Ensemble rainfall forecasts were verified using categorical scores: Probability of Detection (POD), False Alarm Ratio (FAR), Critical Success Index (CSI) and Accuracy.

A 2x2 contingency table was used to quantify forecasts into a single value score indicating whether the forecast predicted an event or a non-event. POD measures the system's ability to detect when an event, or, in this case, a specific threshold, is exceeded. POD values are undefined when hits and misses =0, in the case where neither value exceeds the defined threshold. In this case, Accuracy is used to measure both events and non-events. Further details and the equations used to determine the categorical scores are described in Chapter 2, Section 5**.**

Continuous Ranked Probability Score (CRPS) and Brier Score

Given that the study was carried out on a limited number of events, only verification suitable for event-based metrics was used. The Continuous Ranked Probability Score (CRPS) measures the difference between the cumulative distribution function (CDF) of the ensemble forecasts and the CDF of the observations. The CRPS reduces a single value of the mean absolute error for an ensemble forecast, which means it can be used directly to compare the accuracy of an ensemble forecast with that of a deterministic forecast (Casati et al., 2008). A perfect Brier Score is zero and a smaller score is desirable, which occurs when the forecast probability equals the observed frequency. Equations used to calculate CRPS and Brier Score are described in Section 2.5.

The verification analysis was carried out at 10 specific rainfall grids that correspond to known flood locations and grids covering the Gomok area of Alexandria, against GPM IMERG rainfall, which is available at 0.1 degree (~10km) resolution but was resampled to 3.3km (Figure 5.2). Verification was carried out for the maximum 1h rainfall intensity over a 24h period, rather than the 24h accumulated rainfall, to account for variability in the timing of peak rainfall intensity. The 1h rainfall intensity was used to capture the short-duration high-intensity rainstorms that occur in Alexandria. Thresholds for the Brier score and categorical scores were derived from Alexandria's Drainage System IDF return period for 1h rainfall intensity with 1-year, 3-year, 10-year and 28-year return periods, respectively (Table 5.2).

Table 5.2 Maximum 1h Rainfall accumulation thresholds in the forecast verification

Threshold #	1	2	3	4
Rainfall (mm)/1h	>0mm/h	≥6mm/h	≥9mm/h	≥14/h

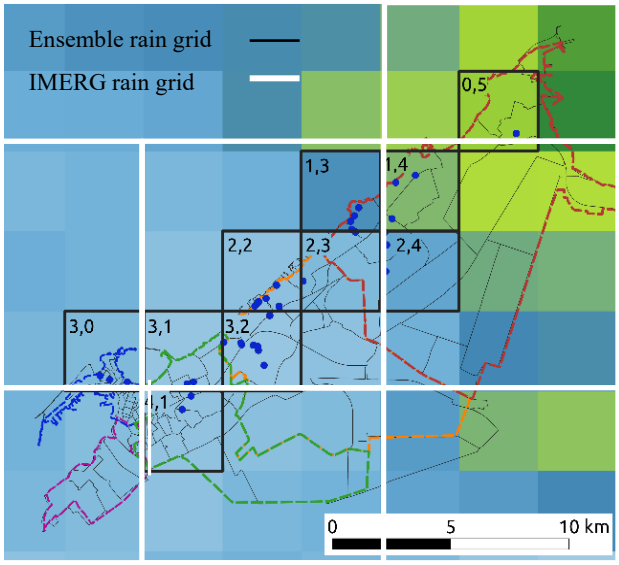


Figure 5.2 Location of Ensemble Rainfall grids 3.3km in Domain 3 and satellite precipitation products IMERG grids 0.1 degree

The ensemble results were then compared with metric scores for the deterministic forecasts. The deterministic forecasts were developed using the model setup described in Chapter 4 for the 3.3km grid. Finally, the Probability Density Functions (PDF) for the rainfall grids are presented. Probability Density Functions are used when representing continuous random variables. For discrete random variables such as the hazard categories used in the study, the probability mass function is applicable. However, for the purposes of generalisation, the term Probability Density Function (PDF) will be used.

5.3 Results and Discussion

5.3.1 Evaluation of the ensemble rainfall forecasts

Verification results are presented and discussed for the ensemble WRF rainfall at a 3.3km spatial resolution. Results are presented for 10 events (Table 5.1). The WRF rainfall is then evaluated using rainfall thresholds for forecasting hazard classes. The scatter plots in Figure 5.3, 5.4 and 5.5 show the 21 ensemble members, the ensemble mean and the deterministic forecasts for the 10 events and 10 grids against observed values. For Figure 5.3, values from the 21 ensembles are plotted against each observed value, reflecting the ensemble variability. Points scattered about the diagonal indicate a perfect forecast, while points above and below the line suggest overforecasting and underforecasting, respectively.

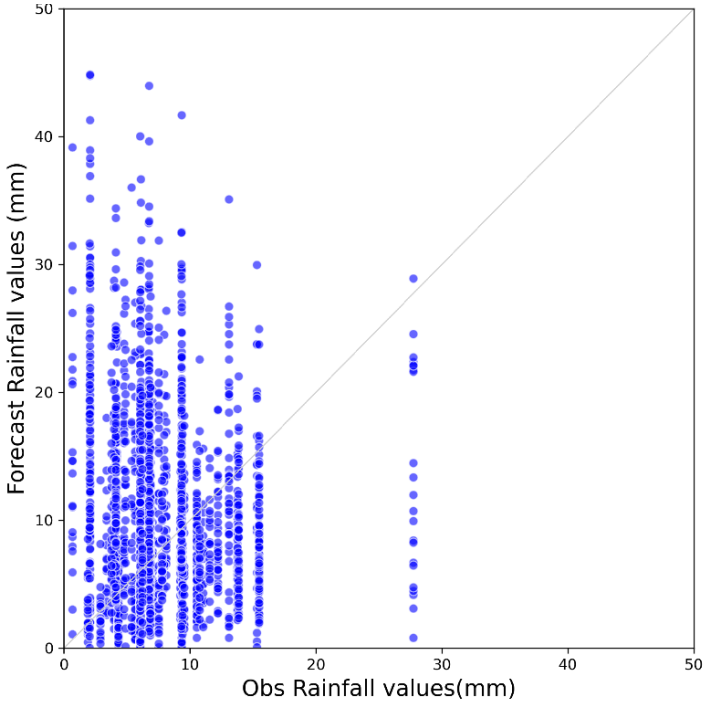


Figure 5.3 Scatterplot of the maximum 1h intensity of ensemble rainfall for all grids against observed rainfall

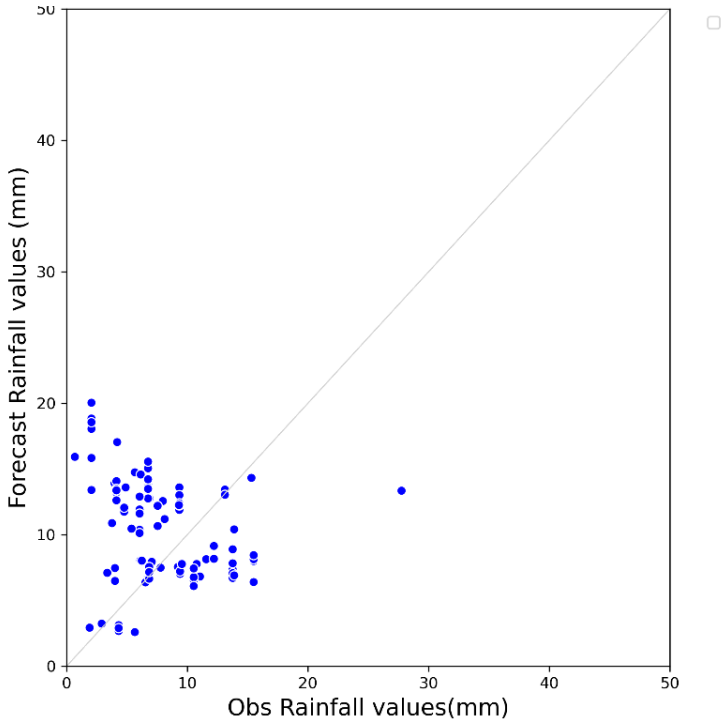


Figure 5.4 Scatterplot of the maximum 1h intensity of the mean ensemble rainfall for all grids against observed rainfall

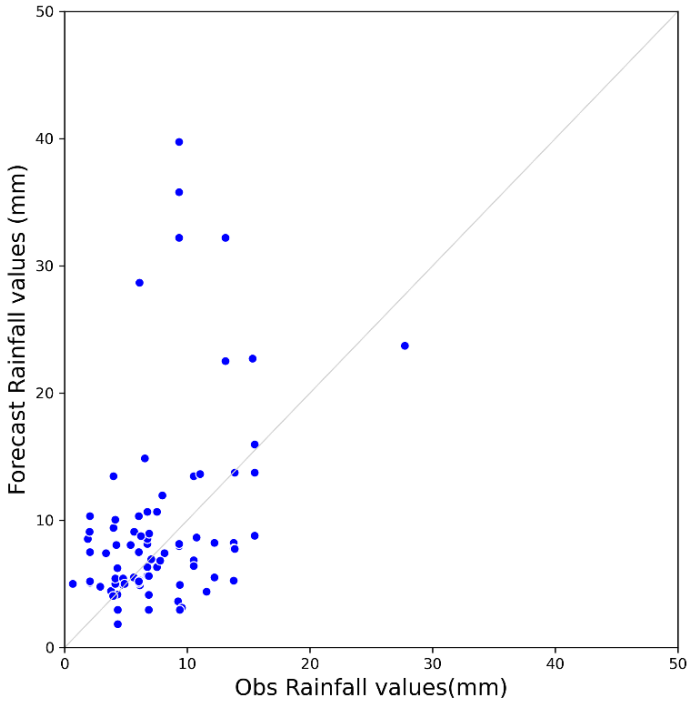


Figure 5.5 Scatterplot of the maximum 1h intensity of deterministic rainfall for all grids against observed rainfall

The deterministic forecasts plot (Figure 5.5) shows several outliers above 20mm and several points above the 1:1 diagonal, suggesting the model tends to overpredict for some grids and events. Compared to the deterministic forecasts, the mean ensemble plot (Figure 5.4) shows greater clustering of points around the 1:1 line and reduced overestimation, indicating better overall agreement between the forecasted and observed events. However, the mean ensemble appears to overly smooth higher values (compared to Figure 5.3), leading to an underestimation of high rainfall intensities. Overall, the ensemble mean reduces the sensitivity to high-intensity events.

Categorical scores

A perfect score for POD, CSI and Accuray is 1 and 0 for FAR (the indices are defined in § 2.5). Values are undefined when hits, misses and false alarms are 0, for example, in cases where neither the forecasted nor observed value exceeds the defined threshold. Figure 5.6 shows the verification results for the categorical scores across all events for different thresholds, as well as a comparison of the mean ensemble scores and the mean deterministic scores across events.

The results show how scores vary across events and evaluated thresholds. At the lowest threshold (>0mm), all events achieve perfect scores, indicating consistent rainfall detection. For thresholds ≥6mm, ensembles achieved moderate scores with no significant difference between the mean ensemble and deterministic scores.

At higher thresholds (≥9 mm), both POD, CSI and Accuracy decrease, while FAR increases, reflecting the increasing challenge of detecting moderate to high rainfall. However, >14mm showed higher accuracy than the 9mm/h threshold, which can be attributed to a higher proportion of correct negatives for higher-intensity events.

In comparing the ensemble mean scores and the deterministic average scores, there was no significant difference between them; however, for 14mm/h, the deterministic forecasts performed better overall, except for FAR, where both performed poorly.

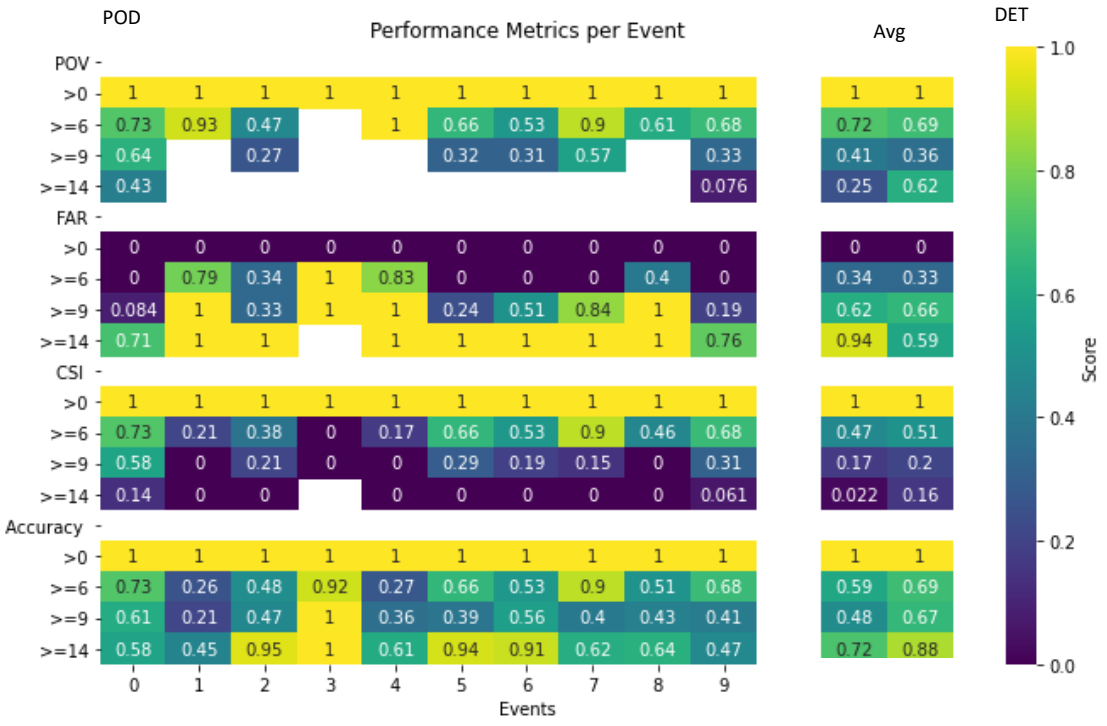


Figure 5.6 Results for POD, FAR, CSI and Accuracy for ensemble rainfall for 10 events. The last two columns show the average and comparison with Deterministic values (DET)

The Brier score measures the mean squared error of probabilities, with a perfect Brier score of 0. Table 5.3 presents the Mean Brier Scores for all 10 events for the different thresholds. Lower thresholds have higher scores, but the 14mm/h threshold performs better than the 9mm/h threshold, except for events 1 and 4, which consistently show lower scores.

Table 5.3 Ensemble Mean Brier Scores where a lower score indicates higher skill.

	Event									
Thresholds	0	1	2	3	4	5	6	7	8	9
1 (>0 mm)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2 (≥6 mm)	0.27	0.74	0.51	0.08	0.73	0.34	0.47	0.1	0.48	0.31
3 (≥9 mm)	0.39	0.79	0.53	0.0	0.64	0.61	0.44	0.6	0.57	0.6
4 (≥14 mm)	0.42	0.55	0.05	0.0	0.39	0.06	0.09	0.38	0.36	0.51

CRPS (§2.5) evaluates the forecast distribution against the observed values and represents the ensemble mean error across the grids. Figure 5.7 presents how CRPS values vary across the 10 grids for each of the 10 events. Each boxplot shows how CRPS values vary spatially across grid locations (see Figure 5.2). The medium CRPS ranges between 2mm and 4mm across most events, indicating consistent forecast skill. Events #1, 5 and 6 showed lower CRPS medians and narrower interquartile ranges, suggesting greater reliability and less variability in ensemble performance. Events #0, 4 and 7 display wider spreads, with event 7 showing the largest variability and the highest outliers above 11mm. However, several other events also showed high outlier values and greater uncertainty and error spread, indicating higher errors.

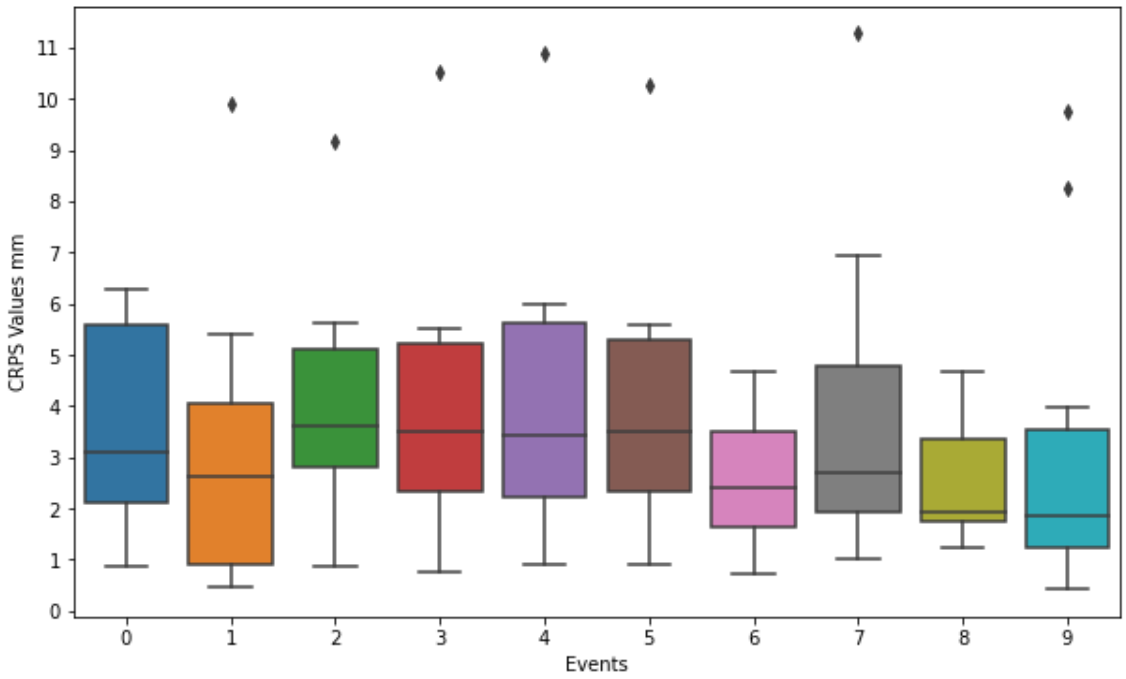


Figure 5.7 Box plot of Ensembles Forecast CRPS values for 10 (#0-9) events analysed over 10 rainfall grids

Comparing the mean CRPS with the deterministic forecasts (Table 5.4), the ensemble forecasts appear to perform slightly better. Ensemble CRPS values ranged from 1.0 to 4.84, with an outlier of 8.6mm, whereas deterministic forecast values ranged from 1.48 to 4.88, with an outlier of 19.2mm, indicating it is difficult to establish which performs better. Further, approximately half of the events had lower ensemble forecast scores, while the other half had lower deterministic forecast scores.

Table 5.4 Comparison of mean CRPS values for ensemble rainfall forecasts and mean Deterministic forecasts across grids lower scores are highlighted in blue.

Event	0	1	2	3	4	5	6	7	8	9
ENS(mm)	3.3	8.6	2.1	1.0	4.8	3.2	1.4	2.7	3.4	4.4
DET(mm)	19.2	3.9	4.7	1.9	1.5	4.9	3.3	1.8	2.3	3.0

5.3.2 **Grid Rainfall Distribution**

Figure 5.8 displays the distribution of the maximum 1h rainfall accumulation for the 10 selected grid points (Figure 5.2) across 21 ensembles for the Oct 25th event. Some grids show the presence of outliers (grid (3,2), (2,2), (2,3) and (2,4)) and wide boxplot interquartile ranges, indicating greater forecast uncertainty and higher variability in the ensemble predictions. Grid (1,3), (1,4) and (0,5) show higher maximum rainfall values above 25mm/h and median values ~10mm/h. Other ensemble members have lower rainfall values but more agreement with smaller values. Grids (4,1), (3,0) and (3,1) have a smaller spread, but grid (3,0) is skewed toward higher values, with a median of ~11mm/h.

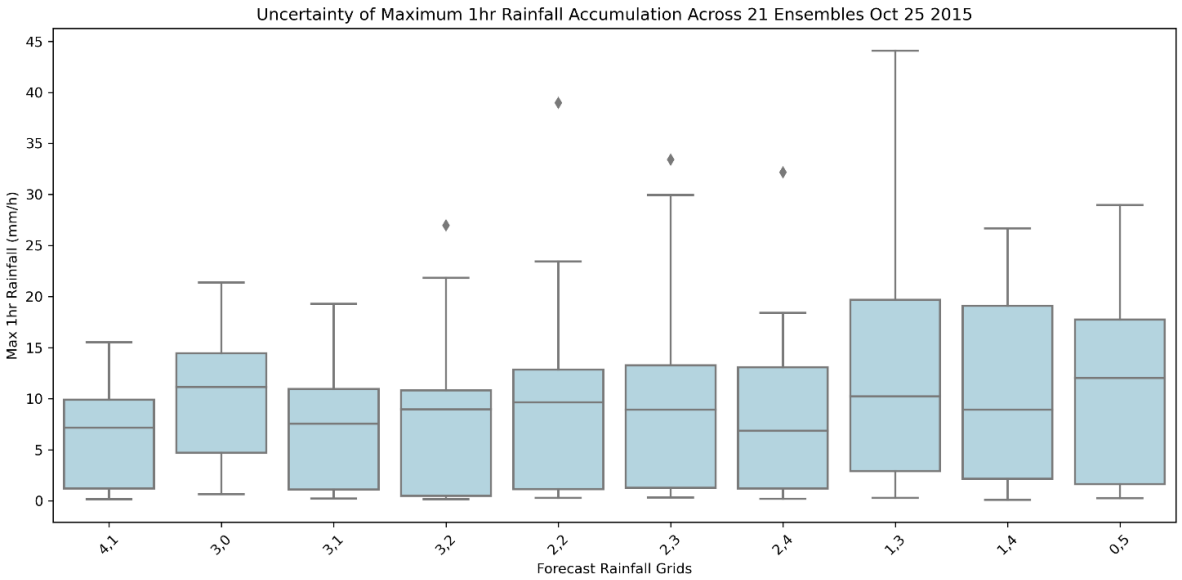


Figure 5.8 Box plot of maximum 1h rainfall accumulation over 24h for the Oct 2015 event, 21 ensembles across 10 selected rainfall grids across Alexandria city. Grid (1.3), (1.4) and (0.5) on the northeastern coast show greater rainfall intensity variability, as indicated by wider bars. The rainfall grids are displayed in Fig. 5.2

5.3.3 Probability Density Functions

The threshold defined in Table 5.5 shows four hazard classes with the following rainfall ranges: 1-5.99, 6-8.99, 9-13.99 and >=14mm/h. These hazard classes are defined as: Green–low, Yellow–moderate, Orange–high, Red–very high.

Table 5.5 Hazard classes and Maximum 1h Rainfall accumulation thresholds used to define hazard state categories

Hazard Class	Low	Moderate	High	Very high
Rainfall (mm)/1h	0-5.99	6-8.99	9-13.99	≥14

The Probability Distribution for the discrete categories is shown below (Figure 5.9). The ensemble rainfall probability distributions and corresponding hazard classifications (Green–low, Yellow–moderate, Orange–high, Red–very high) for 10 rainfall grids (Figure 5.2) coinciding with flood locations on 24 October 2015 at a 24-hour lead time. The ensemble forecasts exhibit substantial variability across grids, with probabilities distributed across multiple hazard categories. Grids (3.0), (1.3) and (0.5) show a maximum probability in the very high/Red class, indicating high rainfall intensity. Grids (4,1), (3,2) and (2,3) exhibit higher probabilities in the Green or Yellow classes, reflecting lower-intensity events. The figure also illustrates the spatial variability of ensemble

forecasts of adjacent rainfall grids, exhibiting the importance of higher resolution forecasts.

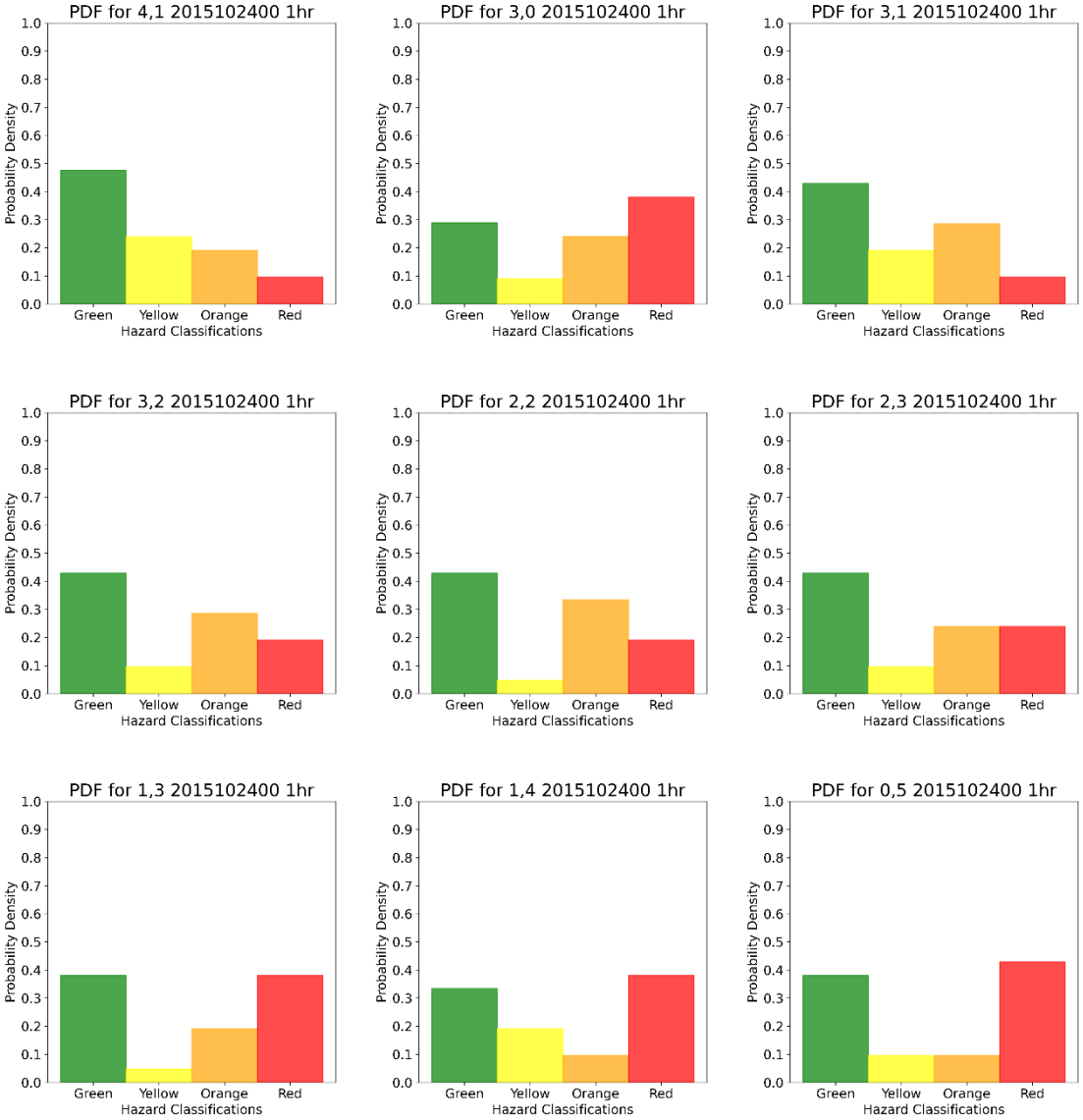


Figure 5.9 Ensemble forecasts distributions and respective hazard categories for 9 of 10 rainfall grids which coincides with flood locations (Figure 5.2) for the 24h lead time on Oct 24th, 2015

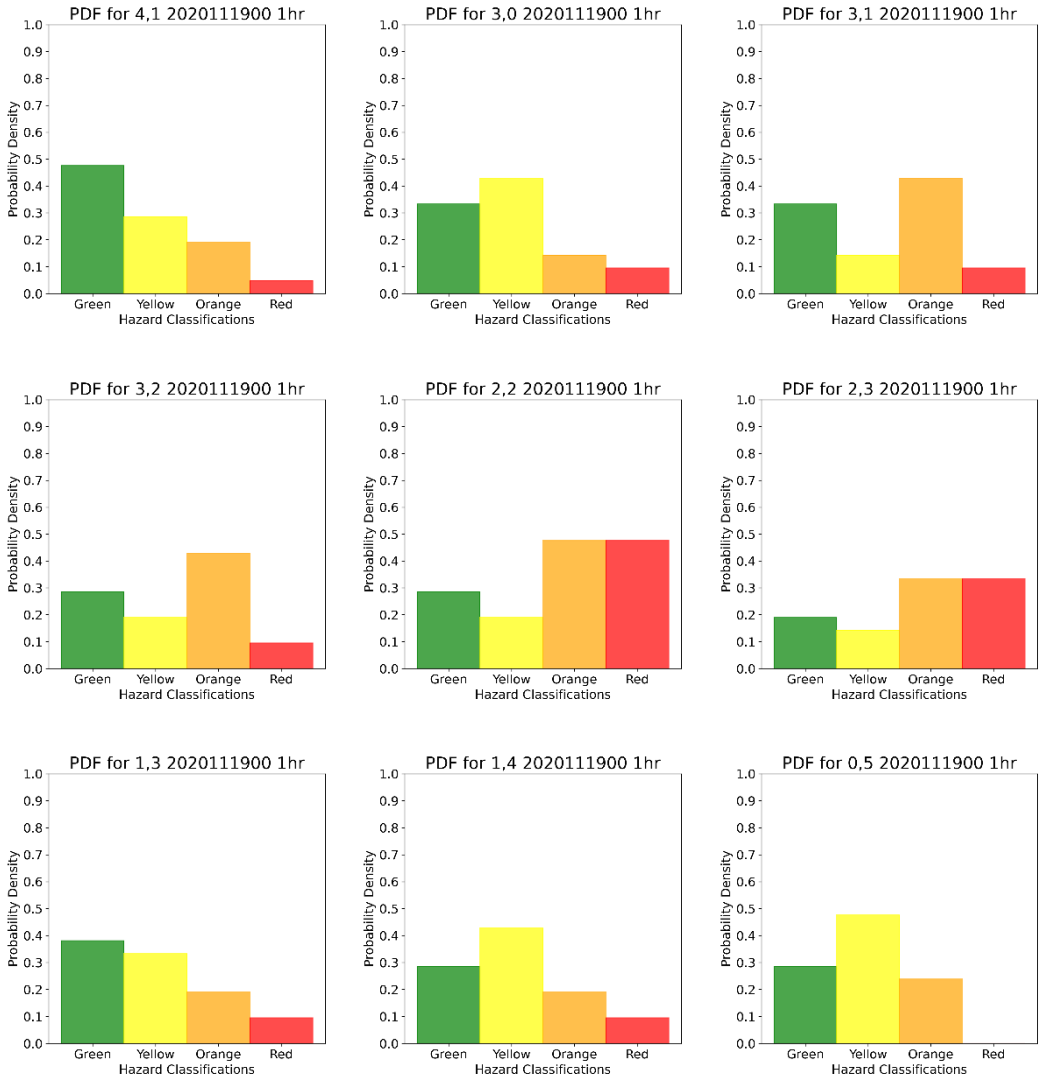


Figure 5.10 Ensemble rainfall distribution 24h lead time on Nov 19th, 2020 and respective hazard categories (Table 5.5) for 9 of 10 rainfall grids which coincide with known flood locations(Figure 5.2)

Compared to the 2015 event (event #0), the ensemble forecasts for the Nov 2020 (event #9) show relatively narrower distributions, with most grids showing dominant probabilities in the yellow and orange categories (Figure 5.10). Grids (2,2) and (2,3) indicated higher probabilities for the red category, suggesting localised higher rainfall intensities in these areas. However, the absence of dominant probabilities across several grids indicates continued forecast uncertainty, especially in distinguishing between significant (Orange) and severe (Red) rainfall intensities, as well as between the green and yellow categories.

The ensemble captures a wide range of possible outcomes. However, the absence of a dominant hazard category in several grids can limit the use of the maximum probability category or the use of categorical or probability thresholds needed to trigger anticipatory flood actions.

5.4 CONCLUSION

This chapter investigated the skill of ensemble grid rainfall forecasts for 10 events across the city at specific known flood locations and compared them with deterministic forecasts. The ensemble rainfall forecasts were simulated with the WRF model forced by GEFS members and verified against GPM-IMERG rainfall products. The evaluation used categorical metrics (POD, FAR, CSI and Accuracy) and probabilistic scores (Brier Score and CRPS) to quantify forecast skill across rainfall thresholds and spatial grids.

The ensemble and deterministic forecasts performed comparably across most thresholds showing skill and limited skill depending on the thresholds and metrics. When comparing the CRPS score, which represents mean error and accuracy, there was no clear winner. However, the deterministic forecasts showed slight improvement over the ensembles when evaluated using the categorical scores. Lower thresholds achieved higher scores, whereas forecast skill decreased with increasing rainfall intensity, consistent with the increased difficulty of forecasting high-intensity events. CRPS and Brier scores showed variability across events and spatially, highlighting the sensitivity of ensemble performance to both event characteristics and location. Therefore, overall, it is difficult to make a conclusive statement as to which product performed best.

Some events were also harder to predict, partially due to variable dynamics, model physics, spatial variability, or boundary conditions. Ensemble hindcasts are retrieved at 1 degree, which is relatively coarse compared to the deterministic forecasts stored at 0.25 degrees and this can also result in reduced performance. For the two events analysed, the probability distributions of hazard classes demonstrated high spatial variability and ensemble spread. In some grids, ensembles showed a dominant hazard class, while in others, probabilities were distributed across multiple hazard classes.

In the future, ensemble rainfall forecasts could benefit from bias correction at smaller spatial grid scales or as more gauge data becomes available. The analysis considered ensemble rainfall uncertainty using only the rainfall threshold method hazard classification. A decision was made not to simulate ensemble flood forecasting using the urban model due to difficulty calibrating the model for both high- and low-intensity

rainfall events and the variability in peak timing. Ensemble flood forecasting is especially challenging in modelling complex sewer networks, which have daily flow variability depending on the time of day, as well as complex pump networks and operating rules.

The results highlight the challenges and opportunities of quantifying predictive uncertainty and linking probabilistic rainfall forecasts to hazard likelihoods. In Chapter 4, we explored the skill of the high-resolution deterministic forecasts and the influence of spatial scales. The outcomes of this chapter indicated that while deterministic forecasts have skill, this skill varies with events and across metrics. These results support moving from deterministic to probabilistic approaches to gain confidence in forecasts and support the adoption of risk-informed decision frameworks. The next chapter builds upon these findings to explore how such probabilistic information can be integrated into anticipatory flood management systems and decision support processes.

Data Availability statement

The NOAA Global Ensemble Forecast System (GEFS) data that support the findings of this study are openly available from the https://registry.opendata.aws/noaa-gefs/

GPM IMERG Final data is openly available from the National Space Agency (NASA) at https://gpm.nasa.gov/data/directory

Other data underlying this Chapter can be found in Version 1.4TU.ResearchData. dataset. https://doi.org/10.4121/0ee90a1c-bb09-4c3a-968c-7fced78c8bea.v1

6

A BAYESIAN DECISION FRAMEWORK FOR ANTICIPATORY FLOOD MANAGEMENT

[3] Alexandria City in Egypt is increasingly affected by pluvial flooding due to extreme rainfall events, rapid urbanisation and inadequate sewer infrastructure. When combined with other flood mitigation measures, flood forecasting and early warning systems can reduce damage by improving preparedness and providing decision support for forecast-based actions before an event. Ensemble prediction systems (EPS) have been proposed to quantify forecast uncertainty, but the extent to which they are useful for supporting decision-making in flood anticipatory actions in an urban, data-scarce city has not been fully explored. This research proposes to go beyond forecasting and demonstrate how a Bayesian decision-theory framework can identify optimal anticipatory actions that minimise expected losses associated with different flood-vulnerability characterisations for selected neighbourhoods (Shyahkas) across Alexandria. Results show that risk-based approaches can support decision-making under uncertainty, compared with relying solely on ensemble means. However, careful consideration is required to define suitable loss functions that reflect emergency managers' risk preferences and the local flood context in Alexandria. This study proposes to close the gap in translating forecast uncertainty into actionable information for local-scale emergency response. It highlights the potential of risk-based decision support systems and provides insight into their future applications for operational use. This work contributes to improving flood resilience against pluvial flooding in Alexandria and other flood-prone cities in the region, despite limited data.

---

[3] This Chapter is based on Young, A., Bhattacharya, B., Daniëls, E. and Zevenbergen, C. 'Incorporating Forecast Uncertainty into Anticipatory Flood Management using a Bayesian Decision Support Framework (under review)

6.1 INTRODUCTION

Making decisions under uncertainty poses a challenge to all decision makers. High-resolution data and forecasts are required to identify local-scale targeted actions. However, city and district emergency managers and officials face challenges about what, where and when to implement actions (Ernst et al., 2018). Given the uncertainty in the forecasts' magnitude, location and timing of extreme rainfall and floods, expected impacts, or the likelihood of those impacts, there may also be a mismatch between the forecast scale and the scale at which decisions must be made. In many cases, city-scale decisions are made using too coarse forecast information, simple heuristic rules combined with emergency manager's experience of past events. However, this can lead to missed events or false alarms, associated losses, and, over time, deterioration in trust due to incorrect actions.

Much of the scientific research addresses the use of probabilistic methods over deterministic forecasts to quantify the uncertainty of input and model parameters on predicted values and increase confidence, but independent of the decision-making process. For example, exceedance of probabilistic or ensemble thresholds and risk matrices have even been used to determine responses or issue warnings (Dale et al., 2012; Ponzano et al., 2025). Alternatively, de Perez *et al.,*(2016) used the False Alarm Ratio (FAR) of probabilistic forecasts to trigger humanitarian actions. However, Morss et al. (2010) and Kox et al. (2015) found that, in the use of warnings by emergency services, no single probability threshold could be identified to trigger mitigation measures and respondents tended to select actions based on low-probability forecasts.

Joslyn and LeClerc (2013) found that the use of numeric uncertainty estimates in weather forecasts increases trust, given a range of possible outcomes and the amount of uncertainty, but the advantage of uncertainty estimates depends on the user's cognitive process and decision context. Therefore, how probabilistic forecasts are best used by and communicated to end users or translated into actions is less agreed upon and it remains a challenge on how to support decision makers (Crochemore, Ramos, Pappenberger, et al., 2016; Demeritt et al., 2016; Fundel et al., 2019; Leskens et al., 2014; Morss et al., 2010; Rabb et al., 2019; M.-H. Ramos et al., 2010; L. J. Speight et al., 2021).

Probabilistic Decision Frameworks

Probabilistic decision frameworks are one such way to provide decision support via a formal and systematic method of incorporating uncertainty through probabilistic forecasts, different alternatives and damage, to evaluate risk and different alternatives (Simpson et al., 2016). Prescriptive decision-making focuses on how decisions should be made, versus descriptive decision-making, which aims to explain how decisions are made. Several researchers have used probabilistic forecasts and prescriptive decision frameworks to issue warnings or take actions by assessing the economic value (Hirschberg et al., 2011; Mylne, 2002; van Andel et al., 2014; Verkade and Werner, 2011). The cost-loss ratio

(Murphy, 1977), in its simplest form, seeks to evaluate the economic value of making decisions by comparing the cost, C, of a measure/warning to the losses, La, that would be avoided. It has since undergone several iterations to accommodate probabilistic warning thresholds and false-alarm intolerance (Roberto Buizza, 2008; Mylne, 2002; Richardson, 2000; Roulin, 2007; Roulston and Smith, 2004; Verkade and Werner, 2011).

Dale et al. (2012) also proposed a risk-based, probabilistic benefit-cost decision framework that similarly accounts for uncertainty and the consequences of adverse outcomes. Each ensemble member is assigned a monetary value of expected flood loss and the average loss is then compared with the costs of taking a flood risk mitigation action.

Decision frameworks have also been applied for humanitarian actions and economic value. Bischiniotis et al. *(*2019) used a two-stage decision-tree action system to evaluate forecast information at varying qualities and lead times to assess the economic value of Early Warning Early Action systems. Lopez et al. (2018) explored a valuation approach based on expected value to selected forecasts (magnitude, probability and lead time) and setting probability thresholds to trigger humanitarian actions. However, these frameworks are considered risk-neutral and the uncertainty around defining the probability thresholds that trigger an action and their suitability for flood forecasting, is debated (Addor et al., 2011; Matte et al., 2017). To consider the risk attitudes of disaster managers and humanitarians (Matte et al., 2017) *and* (Lala et al., 2021)*,* respectively both considered utility functions which incorporate end-user preferences by including risk aversion, which refers to a decision maker willing to pay a certain amount of money to avoid any risk associated.

Bayesian Decision Framework (BDF) using Bayesian Decision Theory (BDT) helps achieve an optimal decision strategy under uncertainty (De Groot, 2004) and allows flexibility in incorporating the strengths of end users' risk-taking preferences (see Section 2). Bayesian decision theory uses prior probability distributions, probability distributions of the future state of nature (hazard), i.e., the forecast and a utility or loss function that incorporates end-user preferences for possible actions. Krzysztofowicz (1993) used Bayesian decision theory in searching for the optimal decision rule using a binary utility function in flood forecasting and warnings and others have followed in advocating for its usage in issuing warnings and taking actions (Biondi and Todini, 2018; Economou et al., 2016; Martina et al., 2006; Todini, 2017).

The decision problem is multidimensional and multidisciplinary, determined by spatial and temporal variability and the quality of information, which varies with lead times, cascading uncertainty in forecasts and varying vulnerabilities. Taking the wrong anticipatory flood actions could result in losses due to acting in vain or missed events. Further, the challenge continues on how to translate national and regional forecasts to local actions in cities. Therefore, decision-making could be improved by using a

structured decision framework which reduces losses and considers uncertainty, decision-making at relevant spatial scales, adaptable to the information available, a range of actions used and end user requirements.

This research aims to demonstrate how probabilistic information can be incorporated into local-scale anticipatory action decisions by combining probabilistic forecasts with risk-based information to trigger actions. Risk-based approaches that consider the likelihood of the event and the consequences, linked not only to physical characteristics but also to socio-economic characteristics. This chapter uses a BDF to incorporate forecast uncertainty for local-scale decision-making and trigger spatially varying anticipatory flood decisions using a rainfall threshold hazard classification method. Further, this research explores how optimal actions vary across different decision parameters, vulnerability considerations and event characteristics.

This chapter, first details the methodology used to develop the variables for the BDF: probabilistic forecasts, decision variables, alternative actions and losses. The BDF is then tested on hindcast data of past events. The research is carried out in the coastal city of Alexandria, Egypt, which occasionally experiences extreme weather events that cause significant flooding and disruption to life throughout the city (Zevenbergen et al., 2016). City officials presently take actions using country-region-scale forecasts and experience from past flood events. However, there is an opportunity to improve the decision-making process, especially given the uncertainty associated with forecasts. Although BDT has been proposed by others as discussed above, its use in the flood and disaster risk management field is limited. This research contributes to the field by demonstrating the application of BDT using uncertainty from local-scale ensemble forecasts and considering decision makers' attitudes and vulnerability for data-scarce cities, which, to our knowledge, has not been done previously.

6.2 A Bayesian Decision Framework for Anticipatory Actions

Bayesian decision theory uses Bayes' Theorem (Eq 6.1) to define the posterior probability and a utility function that describes the decision maker's preference for selecting the optimal action over other alternatives.

$$Bayes\ Theorm\colon P(x|y) = \frac{P(y|x)P(y)}{P(x)} = \frac{P(y|x)P(x)}{\sum_{j=1}^{n} P(y|x)P(x)} \qquad (6.1)$$

$$a*(y) = \underset{a\in A}{\operatorname{argmin}} \sum_x L[a,x]\, P(x|y) \qquad (6.2)$$

P(x|y) is the posterior probability of the future state given predictive weather information; P(y) and P(x) are the prior probability and the likelihood (probability) P(y|x) of the

possible states of nature (x). The objective of the decision model is to minimise damage by selecting the optimal action given the forecast's uncertainty.

The loss function L(a, x) describes the losses incurred from taking an action a (such as opening a sluice or closing a gate) for state x. The expected loss is obtained by combining the loss function with the probability distribution of possible states. For discrete values, the expected loss is calculated as the sum of the products of the loss and the probability of each state. For continuous values, the expected loss is calculated as the integral of this product over all possible states. In the context of anticipatory actions, P(y) is the probability density function of ensemble forecasts.

A brief introduction to Bayesian Decision Theory was presented in Section **2.7**.

In the case of flooding, the Decision Rule (DR) states that the optimal action a*(y) (Eq. 6.2) is the action that minimises the expected value of loss E(L[a,x]) associated with the risk of all states (i.e., the loss multiplied by its probability (Lindley and Smith, 1972). The advantage is taking an action based on its risk rather than making a decision based on the most likely state exceeding a probability threshold determined from the ensemble forecast (Harvey et al., 2022).

6.2.1 Defining the Loss Function

In BDT, the concept of utility is used but for the application of anticipatory action, utility is translated into a loss function. Loss function L[a,x], presented as a loss matrix (Table 6.1) is a concept used widely in many fields for mapping values of loss by assigning a numerical value of loss to each possible outcome associated with an action given the true state of nature (Harvey et al., 2022). In flood risk management, depth-damage curves per land-use class are used to estimate direct flood damage based on a hazard probability (Huizinga *et al.*, 2017). Dutta *et al. (*2003) proposed parameter-based depth-damage curves based on depth, duration and velocity, fitted to historical data for urban, rural, agricultural, infrastructure, or traffic damage. Alternatively, in urban areas, flood depths from pluvial floods are typically lower than fluvial floods (Martínez-Gomariz et al., 2020). Therefore, previously developed traditional depth-damage curves for fluvial floods lead to systematic biases in loss estimates (Porter et al., 2023). When depth-damage curves exist for urban pluvial flooding, they are usually available in high-income, data-rich countries with access to insurance claims or other sources and cannot be applied to data-poor regions. Observed damage claims and impacts are often unavailable, especially in already data-scarce regions, which limits the use of depth-damage curves to accurately reflect flood damage.

IPCC (2022) defines loss and damage as 'the harm from (observed) impact and (projected) risks which can be economic or non-economic'. Therefore, while depth-damage curves and other economic models have been used to quantify damage losses for flood risk

management and adaptation strategies, the reliance on economic and damage data and models to solely motivate actions may not be suitable for triggering and activating early actions.

Table 6.1 Example of a loss matrix where a_i represents the set of all possible actions with (i=1,2,.., m, xi represents all possible hazard states xj (j=1,2,.., n). L(ai., xj) represents the loss of each possible action ai given the actual state of the hazard xj .

Action	x_i	…	x_n
a_1	L(a_1,x_i)	…	L(a_1,x_n)
...	…	…	…
a_m	L(a_m,x_1)	…	L(a_m,x_n)

These models prioritise higher property values while underestimating the risk to poorer, more vulnerable communities, which take longer to recover (Botzen et al., 2019; Kind et al., 2020)**.** Kind *et* al. (2020) argue that cost-benefit analysis disregards the unequal distribution of flood damages that result in greater financial loss and greater impact in lower vulnerability communities. Small increases in stress or hazard intensity can lead to disproportionately greater damage and thus, the prioritisation of socially equitable emergency actions should be considered.

Loss functions have the advantage of incorporating the decision makers' preferences. From the perspective of flood loss, this relates to decision-makers who consistently take actions regardless of forecast severity and to risk-neutral or risk-seeking users who avoid taking actions. Alternatively, risk-averse could refer to a decision maker who prioritises actions in higher-vulnerability areas. (Lockwood et al., 2024) proposes the use of equity weights in benefit-cost analysis of flood mitigation measures and adaptation planning to shift investments towards the poorer population with higher social vulnerability.

Sarlin (2013) used policymakers' preference parameter μ and observation-specific weights to model false or missed alarms during financial crises. Similarly, emergency management decision-makers will have preferences regarding the relative severity of a false alarm versus a missed alarm. Missing a severe flood is typically far more costly than issuing a false alarm. Also, varying vulnerability reflects different preferences, as a missed alarm would entail higher losses; therefore, balancing false alarms and missed alarms is central to decision-making (Sarlin, 2013). In the case of flood actions, higher weights can be applied to areas with higher population density or critical infrastructure to reflect different risk attitudes.

The loss function framework allows these preferences to be quantified by a parameter μ, which represents the relative weight assigned to the cost of missing a flood versus the cost of a false alarm. A higher μ would indicate a greater concern for missing floods. Therefore, in the absence of depth-damage functions, the loss function/matrix will be used

to incorporate vulnerability as a proxy for "loss" and to assess the value of user actions and expected damage estimates, to determine actions that minimise expected losses.

6.3 Study Area and Data

Rainfall and flood hazard, combined with Alexandria's high population density, informal settlements, ageing infrastructure, poor building conditions and inequitable service distribution, contribute to the increased risk of flooding (Elboshy et al., 2019). Abdrabo et al. (2023) previously carried out a spatially varying flood vulnerability assessment of Alexandria at the neighbourhood level (*Shyakhas*), considering physical, social and economic indicators (Figure 6.2). The study utilised both Principal Component Analysis (PCA) and the Analytic Hierarchy Process (AHP) methods. In general, both methods showed similar trends in overall vulnerability, with El-Gomrok region of Alexandria having the highest vulnerability, followed by the Wasat (Middle) District, the El-Gharb (Western) District, the Northern District and El-Montazah. The AHP tended to have higher vulnerability values than PCA, but the minor differences showed a similar trend. Similarly, Elboshy et al. (2019) for pluvial food risk assessment for El Montazah considered institutional coping capacity measures and similarly found higher social and physical vulnerabilities in built-up areas, including informal settlements.

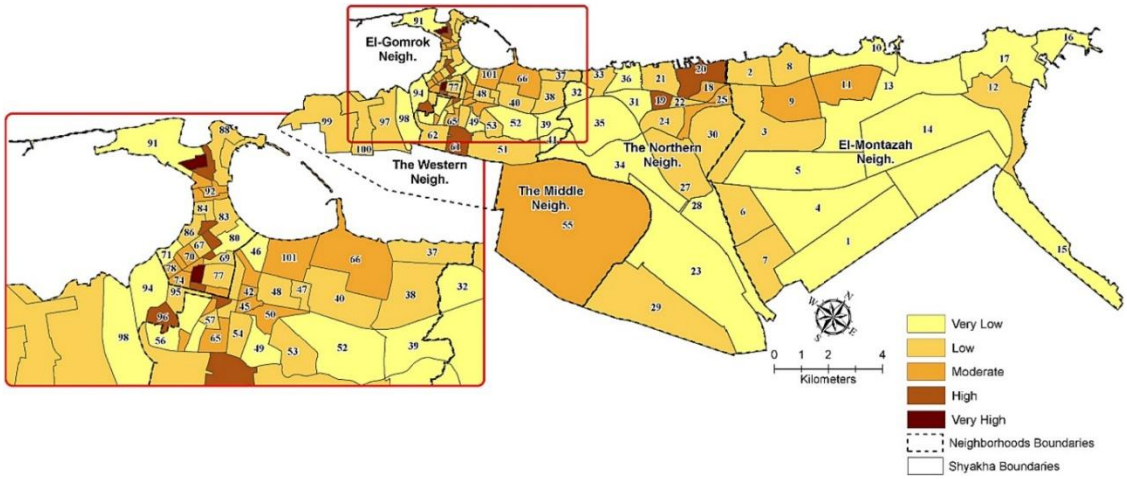


Figure 6.1 Overall flood vulnerability for Alexandria using the PCA method. Source (Abdrabo et al., 2023)

Anticipatory actions in Alexandria

In this research, discussions were held with local decision-makers to better understand the types of decisions and actions taken and how emergency managers currently comprehend and utilise forecast information and local knowledge to issue warnings and take action. The Alexandria Sanitation and Drainage Company (ASDCO) is responsible for coordinating actions along with the Governorate and the Ministry of Irrigation. The

Ministry of Irrigation monitors the surrounding irrigation drainage canals and lakes and manages pumping at El Max point after receiving forecasts from the Egyptian Meteorological Agency (EMA). Further details of anticipatory flood actions are described in Chapter 2, Section 2.2.

In Alexandria, the challenge is how to translate national-scale warnings into local-scale actions that allow tailored actions. Alexandria officials require information at least 24-72 hours ahead of time, including the intensity of rainfall provided by EMA, the location of flooding to determine where and which action to take. The type of actions taken by ASDCO include hazard-reducing measures shown in Table 6.2:

Table 6.2 Examples of emergency management actions taken by emergency managers before a rainfall event

	List of Actions		**List of Actions contd..**
1.	Staff and equipment preparedness / putting staff on stand-by, coordinating and increasing staff before an event	7.	Emptying lift station wells
2.	Issuing warnings to the public	8.	School cancellations
3.	Deployment of staff to respond	9.	Road closures
4.	Dispatching of suction trucks to known hotspot locations before event.	10.	Sourcing additional pump trucks.
5.	Checking and cleaning of known hotspot manholes and lift stations	11.	Issuing a state of emergency
6.	Deployment of temporary structures	12.	Lowering levels in irrigation drainage canals to receive additional stormwater

There is no formalised systematic decision framework in Alexandria. Instead, actions and emergency protocols are taken using an informal process based on the notification that an extreme weather event will occur from the EMA and historical knowledge of hot spots and expected impact. Knowledge of hot spot areas is key, as emergency managers and the Drainage company have a good understanding of areas that flood often, but there is still uncertainty about the magnitude, severity and prioritisation of locations. Following the severe flood events in Oct 2015, city officials have become risk-averse and eager to avoid damage, but there is still the difficulty of taking actions in vain, failing to sufficiently mitigate the threat, or losing trust over time.

6.3.1 Ensemble Rainfall Forecasts and Observed Rainfall Data

This research proposes a Bayesian Decision Framework (BDF) that uses probabilistic forecasts of rainfall, flood depths and vulnerability. The methodology is adaptable to different hazard estimation methods or flood forecasting approaches and adaptable to different Actions. The BDF was applied to a rainfall thresholds method where only the ensemble rainfall forecast is available for the entire city. The analysis was carried out for several events with low to high probability and severe to moderate severity to evaluate the influence of different event characteristics.

This analysis used both historical and forecast rainfall data accessed through open-source online platforms. Historical gauge rainfall records at Nouzha Airport are available from the National Centre for Environmental Information and the Global Historical Climatology Network daily (GHCNd). Gridded rainfall data were sourced from GPM IMERG Final V7 to capture the rainfall spatial variability. The location of the GPM IMERG grids is shown (with white borders) in Figure 6.3. Predictive hazard information y was sourced from the downscaled rainfall forecasts (see Chapter 5). For the period Oct 2018 – Dec 2020, observed rainfall events greater than 10mm/day were selected. This criterion was based on cross-analysis of historical gauge rainfall IMERG data and known events for the same period. Ensemble rainfall forecasts of past events were sourced from the National Centres for Environmental Prediction Global Ensemble Forecasting System (GEFS). The weather model produces 21 forecasts (ensemble members) to represent uncertainties in the input data, such as limited coverage, instrument or observing system biases and the limitations of the model itself (NCEP, 2024). The future values of y vary spatially across the study area and temporally with lead time. GEFS ensembles were downscaled over the study area to a spatial resolution of 3.3km and 1-hour temporal resolution using the Weather Research Forecasting (WRF) Numerical Weather Prediction (NWP) model (Young et al., 2025). The ensemble rainfall grids used in the analysis are shown in Figure 6.3. Forecasts were simulated for 00:00 and 12:00 UTC.

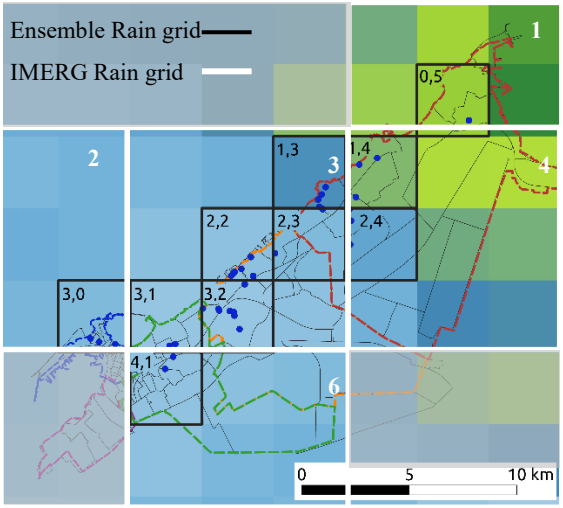


Figure 6.2 Location of Ensemble Rainfall grids 3.3km and satellite precipitation products IMERG grids (in white).

6.3.2 Flood Vulnerability, Locations and Impacts

Rainfall varies spatially and, therefore, varies with vulnerability classifications. The vulnerability classes were sourced from the study by Abdrabo et al. (2023) who previously computed vulnerability classes for Alexandria at the neighbourhood level (Shyakhas), considering physical, social and economic indicators using Principal Component Analysis (PCA). Only four vulnerability classes were selected, as they overlap with historical flood locations. Abdrabo et al. (2023) used three physical components: building connectivity to infrastructure, building structure and residential built-up area density. Four social components included population structure, population social status, population with poor mobility and overcrowding. Finally, the economic components consisted of the population's work activities, service activities, poverty and dependency. The final flood vulnerability class for each administrative unit was calculated by summing the weighted factor scores for the physical, social and economic components. In the absence of loss data for Alexandria, vulnerability is used as a proxy for loss as discussed in Section 6.2.

Shyakha Vul Class 1	Rain grid#	Shyakha Vul Class 2	Rain grid#	Shyakha Vul Class 3	Rain grid#	Shyakha Vul Class 4	Rain grid#
52 Ambrooz and Moharm Bek	4,1	**40** Bab Sharki	3,1	**9** Sidi Bishr Qebil	2,4	**93** Safr	3,0
91 Ras El Tin	3,2	**83** El Magharba	3,0	**11** El Mandara Qebil	1,4	**20** San Stephano	2,3
32 El Reyada	3,2	**21** Fleming	2,2				
31Abou El Nawateer	3,2	**3** El Seyouf Qebli	2,4				
35 Ezbat Saad	3,2	**8** Sidi Bishir Bahri	1,3				
36 Mostafa Kamel & Boloky	2,2	**12** El Mamorah	0,5				
10 El Mandara Bahri	1,4						

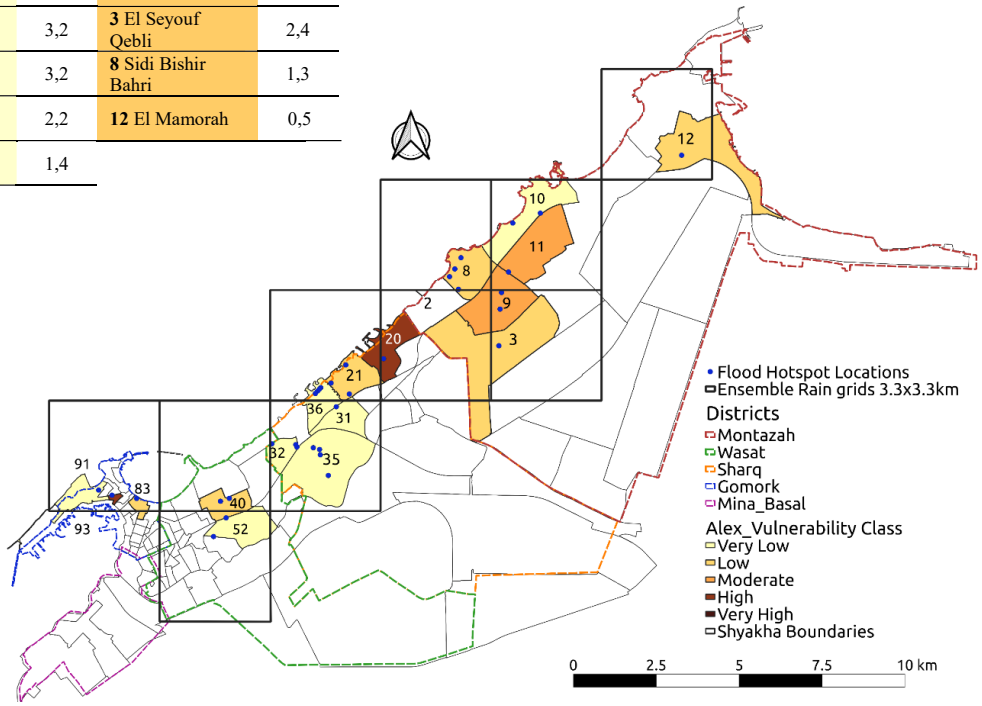


Figure 6.3 Shyakha Neighbourhood vulnerability classifications and location of rainfall grids 3.3km covering known flood locations (Blue dots) (adapted from Abdrabo et al., 2023).

Several flood hotspot locations are known in the city. The Governorate reported over 93 flood hotspot locations in 2020 (Youm7, 2020). For this study, 17 known flood hotspot locations were identified after consultation with Alexandria informants and identification of flood areas from past flood events. The flood locations were overlaid with the flood vulnerability classes and the ensemble rainfall grids in Figure 6.4. The impacts (Table 6.3) for the different hazard levels were determined through discussions with the Alexandria drainage company and the collection of historical impacts on social media and newspapers.

Table 6.3 Flood Impact Table for Alexandria for different hazard levels

Flood depths, Flood hazard Classification and Impacts			
Minimal	**Minor**	**Significant**	**Severe**
- Wet roads - Some pooling of water on the roadways	- Localised flooding the roadway	- Flooding of roadways - Flooding exceeds curbs and can enter buildings, physical	- Prolonged flooding of roadways and road closures

- Day to day activities not affected - Does not hinder traffic	- Slower traffic and short term disruption - Minimal congestions - Impedes pedestrian traffic	damage to building assets and contents - Vehicles at risk of stalling - Road closures and disruption of transportation networks - Short-term flooding in tunnels - Clean up required	- Traffic comes to a standstill and vehicles are at risk of stalling - Flooding exceeds curbs and can enter buildings, Physical damage to building assets and contents - Flooding of ground floors and lower elevation areas below road level - Flooding in tunnels and under bridges - Communities displaced - Significant clean-up required

6.4 BAYESIAN DECISION FRAMEWORK SETUP

6.4.1 Model setup

Risk-based approaches utilise hazard, exposure and vulnerability to categorise a risk. This approach uses probabilistic hazard information, combined with neighbourhood-scale vulnerability, to determine the action that minimises expected loss. The following section further describes the steps taken to develop the BDF:

Step 1: Select prior probabilities of the hazard states.

Step 2: Compute the posterior probability of flooding based on lie of hazard given state.

Step 3: Define the possible actions for the corresponding hazard and vulnerability classification.

Step 4: Define the loss functions and the loss matrix as functions of the actions, state and damage.

Step 5: Iteratively, apply all possible actions to identify the optimal action leading to the minimisation of loss.

The steps (**Figure 6.5**) are presented using a rainfall threshold hazards classification approach (Young et al., 2021) **a**nd further described in detail in this section.

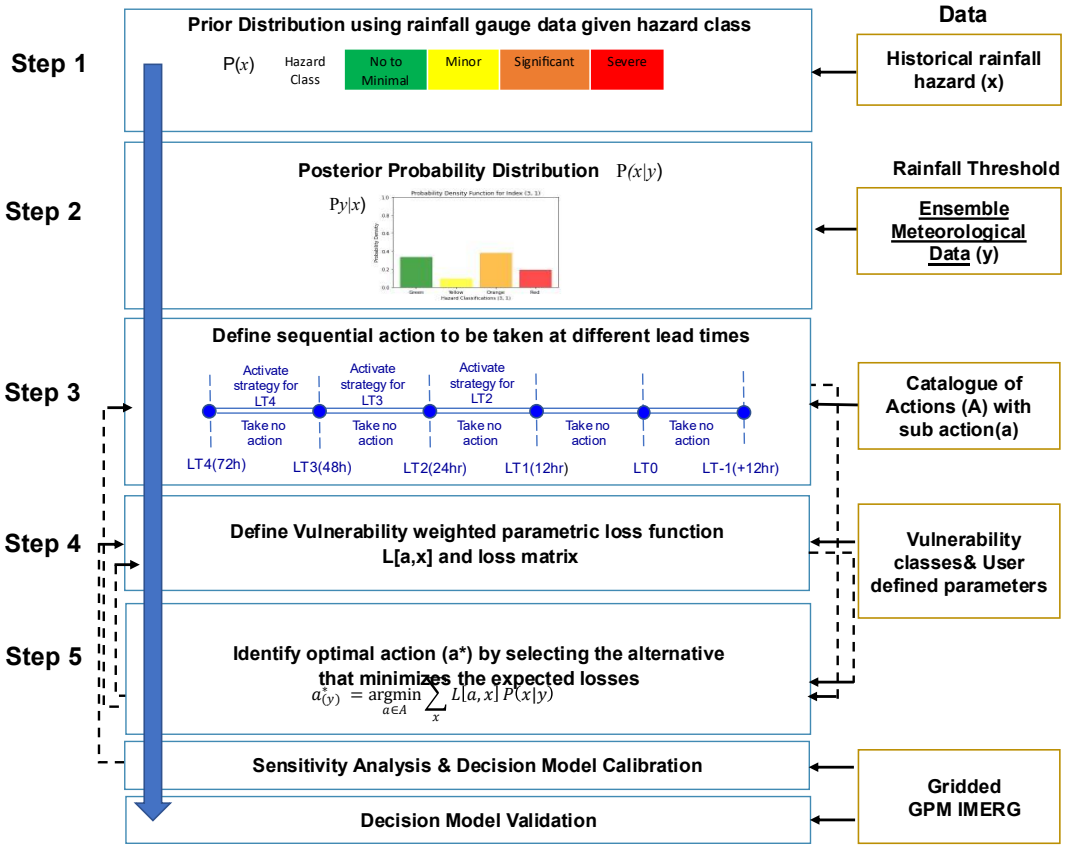


Figure 6.4 Summary of steps used to develop the inputs of the Bayesian decision framework

Step 1: Determine Prior Probability P(*x*)

The marginal probability P(*x*) is the probability of the state occurring, given information collected before new data is received. For example, the probability, computed from historical data, that a specific hazard class will occur. In a case where four hazard states *i*=1,2,3 and 4 (No to Minimal Flooding), yellow (Minor), amber (Significant) and red (Severe)) are equally likely, P(*x*) = 0.25. For this analysis, P(*x*) was determined using Alexandria's intensity-duration curve. Each rainfall event was classified into one of the four hazard classes: 1, 2, 3 and 4, each associated with a corresponding return period for 1hr rainfall intensity: 3-year, 5-year, 10-year and 28-year, respectively (Table 6.4).

Table 6.4 Summary of Prior Probability P(x) for different Hazard classes and Maximum 1h Rainfall accumulation thresholds used to define hazard state categories

	*x*1	*x*2	*x*3	*x*4
Hazard Class	No to Minimal	Minor	Significant	Severe

Rainfall (mm)/1h	1-5.99/h	6-8.99 /h	9-13.99/h	≥14/h
P (*x*)	0.5	0.3	0.15	0.05

Step 2: Compute Posterior Probability State P(*x*|*y*) given forecast probability

The Posterior Probability P(*x*|*y*) is the probability of future probabilistic information, conditioned on available weather information and the prior probability and is determined using Eq. 3. It is the probability that future state severity occurs given predictive information *y.* This translates to the probability of the future hazard severity class given the ensemble forecasts. To calculate the posterior probability P(*x*|*y*), the likelihood P(*y*|*x*) was first determined using the probability distributions of forecasts, where forecast categories *j*=1,2,3,4 match with the hazard classes.

$$P(x_i|y_j) = \frac{p(y_j|x_i)p(x_i)}{p(y_j)} = \frac{p(y_j|x_i)p(x_i)}{\sum_{i=1}^{4} p(y_j|x_i)p(x_i)} \qquad (6.3)$$

$$where\ P(y_j|x_i) = \frac{n_{j,i}}{\sum_{j=1}^{4} n_{j,i}} \qquad (6.4)$$

For the rainfall threshold method, P(*y*|*x*) is the likelihood that the NWP ensemble forecast falls into one of the four hazard classes (Table 6.4) and the relative frequencies derived from the fraction of max 1h accumulations from the 21 ensembles are used as the probability.

Step 3: Define Actions and Decision Rule (DR) based on spatially varying decision-making states

For this analysis, the decisions are based on forecasts at the 24h lead time only because this lead time is critical to decision makers. However, the decision framework can be expanded for sequential anticipatory actions taking place from 0 to 72hrs before an event (Figure 6.6). For sequential decision-making, the strategy is updated using the Decision Rule (DR) at every lead time when new information becomes available.

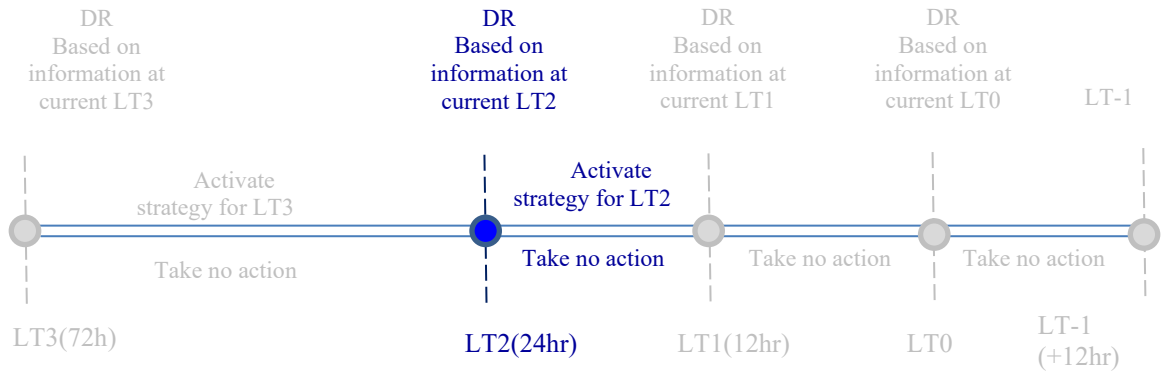


Figure 6.5 Example of sequential action decision chain. The decision rule (DR) is evaluated at each lead time and a decision is made to either activate the Action or take no Action. This is repeated at every lead time as more information becomes available.

Actions consist of low regret, minimal cost such as high-resolution forecasting, preparation of staff and issuing warnings. For higher severity classes, actions consist of dispatching trucks and declaring maximum preparedness. The Actions (*A*) are categorised into *A1, A2, A3, A4, A5, A6* and *A7* which are a combination of a finite number of possible actions (a) corresponding to a hazard and vulnerability class. The Actions (*A*) are predefined as part of the standard operating procedure for each hazard and vulnerability class. For simplicity, each vulnerability class only has the option of 4 Actions which are assigned using a risk matrix. The list of possible actions relating to the lead time at which the action is taken is shown in Figure 6.6. *Green* Actions are taken for all Actions (A1-A7) and Dark Red Actions are taken for A7 only. For the purpose of this study, it is assumed that by taking the correct Action for the corresponding hazard and vulnerability class the expected loss will be reduced to zero.

Step 4: Define the loss matrix for varying vulnerabilities

As discussed in Section 6.2, the loss function represents the user's preference/ utility for each of the possible combinations of the state of nature and is expressed as a loss matrix. The loss function assigns a numerical value to each possible outcome, representing the cost and/or loss associated with that outcome. In the absence of real historical data to define the loss function, we propose to adapt the parametric loss function of L(a,*x*) used by Ecomonou et al., 2016 (Eq. 6.5) to capture the different risk attitudes for different vulnerability classes.

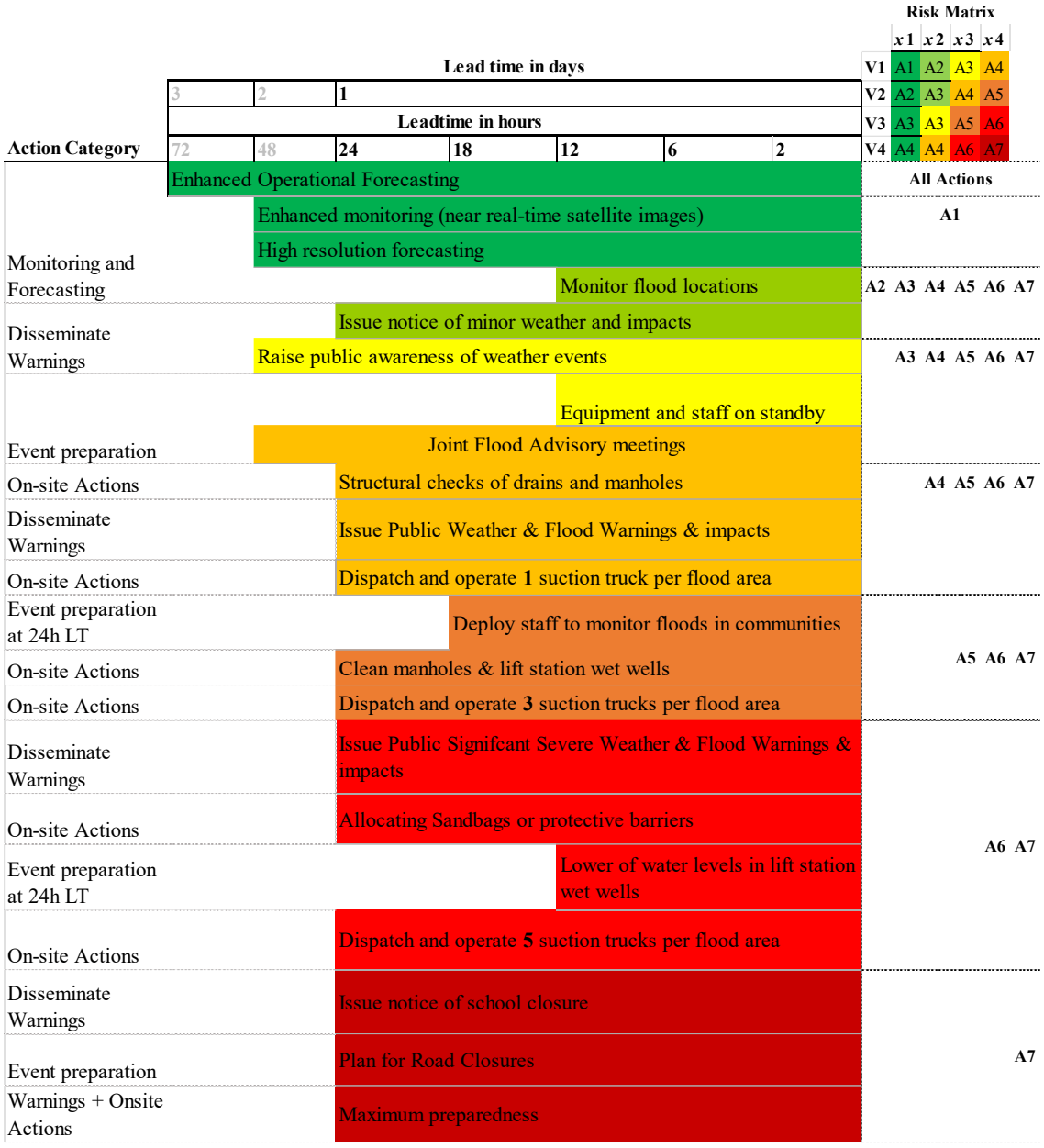


Figure 6.6 Summary of Actions (A) to be implemented as part of Alexandria's Standard Operating Procedure in response to floods

Following Economou et al.,(2016), losses are defined as C(a), the cost of taking a protective action which increases with the severity of the class associated with the action. LR(a)D(x) consist of two terms, where LR(a) is a decreasing function varying between 0 and 1 in which LR(a) outputs a smaller value when a higher action is selected indicating that losses are reduced. D(x) is an increasing function of x where damage losses increase with the severity of the hazard class x indicates the damage when no action is taken. Collectively, this term accounts for losses from damages after an event and accounts for the effectiveness of taking an action.

$$\mathrm{L(a}, x) = \mathrm{Cp(a)} + \mathrm{LR(a)D}(x) \quad (6.5)$$

Discussions with Alexandria decision makers', revealed risk-averse attitudes and relatively low protection cost compared to the potential impacts (Table 6.2). Therefore, instead of protective action costs, a skewed loss function was used to represent reputational costs (C_r) for false alarms when actions are taken in vain and damage losses for a missed alarm which increase in severity and with higher vulnerability classes. This reflects unequal tolerances for false versus missed alarms and the decision makers' preference to avoid large damages.

Similar to Sarlin (2013), a policymaker's preference parameter scaled from 0-1 is used. For this application, the vulnerability weight V(v) which is used to indicate how maximum potential losses scale with vulnerability class. This allows the same hazard to have different impacts in different zones of the city depending on the vulnerability class of a zone. In the absence of real loss values per vulnerability class, hypothetical weights are assumed: *Very low V1:0.3; Low V2:0.5; Moderate V3:0.7; High V4:1.* The maximum loss for a high vulnerability area is assumed to be 100 and a low vulnerability area has 50% of the maximum loss of a high vulnerability area ~ 50. With more data in the future, these numbers can be updated. They are also likely to be sensitive to study areas.

Operational actions are assumed to be based on pre-existing flood preparedness actions shown in Figure 6.6. For simplicity, each vulnerability class has a subset of four actions available which are assigned using the risk matrix (Figure 6.6). For example, in V1 areas, the set of action (A) consists of *A1, A2, A3* and *A4* while for V4, set of available actions includes *A1, A4, A6* and *A7*. It is assumed that when the true state action exactly matches the forecast action, the resulting loss is zero i.e. "Perfect Protection". Translating these criteria into a parametric loss function. The parameters, γl: action reduction (effectiveness of higher actions in reducing damage). γd: damage reduction (how steeply losses increase with hazard severity). l_v: max loss per vulnerability class and c: the max reputational costs. To represent how reputational costs grow with action severity, c_{max} is represented as a ratio of the l_{max} using γc and therefore γc can also be interpreted as a C/L ratio. The components of the loss functions are calculated using Eq. 6.6-6.9 represented as a Loss matrix Table 6.5 showing the subset of actions (Figure 6.7) per vulnerability class.

$$\mathrm{L(a}, x) = C_r(\mathrm{a}, x) + \mathrm{LR(a)D}(x) \quad (6.6)$$

$$C_r(a, x) = c(a - x, 0) \text{ where } \mathrm{c} = \gamma\mathrm{c} * l \quad (6.7)$$

$$LR(a) = l_v(1 - a^{\gamma l}) \; v = vulnerability\ classes\ 1,2,3,4 \quad (6.8)$$

$$\mathrm{D}(x) = x^{\gamma d} \quad (6.9)$$

This parametric loss function does not aim to assign precise monetary values to flood impacts (as in Hudson *et al.,* 2019), but instead to capture preferences and risk attitudes. It reflects how decision makers balance reputational risk, damages and vulnerability

priorities. For instance, a risk-averse decision maker may consistently favour stronger actions, particularly in high-vulnerability areas, whereas a more risk-neutral one may tolerate higher potential damages in exchange for fewer false alarms.

Table 6.5 Loss Matrix and Actions available for the different vulnerability classes

V1 – Low					**V2 – Low**				
	State					State			
	$x1$	$x2$	$x3$	$x4$		$x1$	$x2$	$x3$	$x4$
A1	0	$L(a_1,x_2)$	$L(a_1,x_3)$	$L(a_1,x_4)$	A1	0	$L(a_4,x_2)$	$L(a_1,x_3)$	$L(a_1,x_4)$
A2	$L(a_2,x_1)$	0	$L(a_2,x_3)$	$L(a_2,x_4)$	A2	$L(a_2,x_1)$	0	$L(a_2,x_3)$	$L(a_2,x_4)$
A3	$L(a_3,x_1)$	$L(a_3,x_2)$	0	$L(a_3,x_4)$	A4	$L(a_4,x_1)$	$L(a_4,x_2)$	0	$L(a_4,x_4)$
A4	$L(a_4,x_1)$	$L(a_4,x_2)$	$L(a_4,x_3)$	0	A5	$L(a_5,x_1)$	$L(a_5,x_2)$	$L(a_5,x_3)$	0
V3 -Medium					**V4 -High**				
	State					State			
	$x1$	$x2$	$x3$	$x4$		$x1$	$x2$	$x3$	$x4$
A1	0	$L(a_1,x_2)$	$L(a_1,x_3)$	$L(a_1x_4)$	A1	0	$L(a_1,x_2)$	$L(a_1,x_3)$	$L(a_1,x_4)$
A3	$L(a_3,x_1)$	0	$L(a_3,x_3)$	$L(a_3x_4)$	A4	$L(a_4,x_1)$	0	$L(a_4,x_3)$	$L(a_4,x_4)$
A5	$L(a_5,x_1)$	$L(a_5,x_2)$	0	$L(a_5,x_4)$	A6	$L(a_6,x_1)$	$L(a_6,x_2)$	0	$L(a_6,x_4)$
A6	$L(a_6,x_1)$	$L(a_6,x_2)$	$L(a_6,x_3)$	0	A7	$L(a_7,x_1)$	$L(a_7,x_2)$	$L(a_7,x_3)$	0

Step 5: Identify the optimal action

The final step in the decision model combines the expected value of loss and posterior probability of flooding (or a hazard class) to determine the optimal action a* (Eq. 6.10) which minimises the loss. The optimal action at each flood location is determined based on the vulnerability class defined loss functions and posterior probability derived from the prior and probability of hazard classifications from the ensemble grid rainfall.

$$a*(y) = \underset{a\in A}{\operatorname{argmin}} \sum_x L[a,x]\,P(x|y) \qquad (6.10)$$

6.4.2 Sensitivity Analysis, Model Calibration and Validation

Decision Model Experimental setup

Harvey *et al. (*2022) found that optimal actions are most sensitive to the cost to loss (C/L) ratio. Given the loss function is one of the most important and subjective aspects of the decision model, an initial sensitivity analysis was carried to understand how parameters

influence the decision makers' preferences and to determine the plausible ranges that reflect the decision makers preference for the model calibration. The range of parameters used are shown in Table 6.7. The initial scoping parameters and model calibration was done using the event on October 25, 2015. First, the analysis does not consider vulnerability weights to show when the hazard alone is considered. The analysis assumes all the maximum loss and perceived loss due to false alarms is equal across all grids and only 4 actions with action 4 being the highest. Then vulnerability weights are applied to evaluate how the optimal action changes with the different decision parameters using the Actions per vulnerability class available in Table 6.4.

The model was validated using nine additional events over the 2018-2020 winter season. This includes low-impact, low-severity events to higher-impact, high-severity events. Validating a decision model is a challenge especially for post ante analysis which relies on assessing whether the decision model predicted the correct action and in the case of a spatially varying decision model, the correct decision at the corresponding location. The primary method of validation involved comparing the decision model results with the true-state decision which corresponds to the observed max hourly rainfall accumulation as well as corresponding reports of the actions and impacts reported in newspaper articles and social media post. Observed rainfall is taken from remotely sensed GPM IMERG 1-hrly and thresholds are defined from Alexandria's IDF curve (Table 6.2).

Table 6.6 Range of parameters values used in the experimental setup

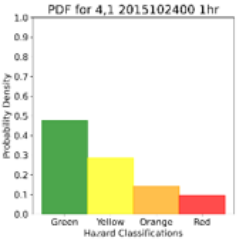
Parameters	Value Range	Step
γ_c	0.1-0.3	0.05
γ_l	0.25-3	0.25
γ_d	0.25-3	0.25

6.5 RESULTS

Results are presented and discussed for probabilistic density function (PDF) for the different ensemble rainfall from the WRF for the 3.3km grid and the results of the experimental setup which was used to determine the parameters for the decision model calibration and validation for the different events as described in the methodology above. The analysis first shows how the PDF varies for each rainfall grid for the Oct 2015 event for the 1hr max thresholds. Using the prior probability P(x) (Step 1 of Section 6.4.1), the ensemble probability distributions (Chapter 5) and the posterior probability P($x|y$) for each grid shows that there is skew towards low intensity events. First, results of the posterior, sensitivity with and without vulnerability are presented, the model calibration based on the suitable parameters from the sensitivity analysis and then results for the 9

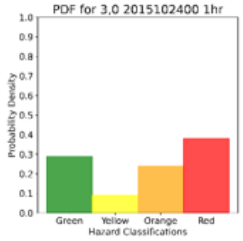
events. Finally, results are presented on how losses due to false and missed alarm vary with parameters.

The ensembles were categorised into different hazard classifications thresholds (Figure 6.8). For example, grid (4,1) shows more ensembles in the green and yellow class compared to (3,0) which has more ensembles in the orange and red. Once categorised into the different hazard classifications, using the rainfall thresholds values which gives the likelihood (P($y|x$), the prior belief P(x) are used to calculate the posterior probability P($x|y$). In this case the prior P(x) is based on the city's drainage IDF curve for the different hazard classes: $x1$, $x2$, $x3$ and $x4$ (Table 6.4) which show a higher belief for lower severity events P (x)=0.5, 0.3, 0.15, 0.05. Therefore, as a result the posterior probability P($x|y$) (Figure 6.8) for each grid tends towards sharper probability distributions which skews to a lower intensity state. For example, Grid (0,3) despite the ensemble having a higher probability for more severe events the priors shift the posterior to a higher belief in lower severity events (from 0.29 to 0.639 and $x4$=0.42 to 0.084).



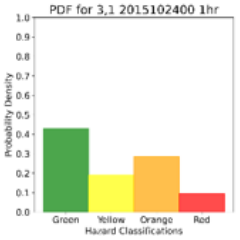
P($x|y$) Grid 4,1

$x1$	$x2$	$x3$	$x4$
0.680	0.245	0.083	0.014



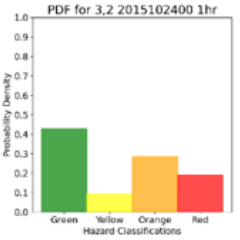
P($x|y$) Grid 3,0

$x1$	$x2$	$x3$	$x4$
0.639	0.119	0.159	0.084



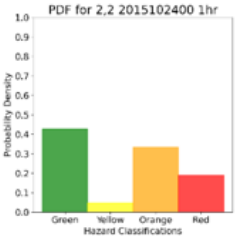
P($x|y$) Grid 3,1

$x1$	$x2$	$x3$	$x4$
0.709	0.17	0.106	0.014



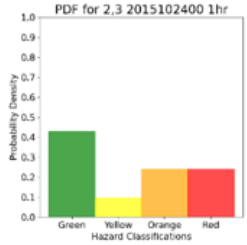
P($x|y$) Grid 3,2

$x1$	$x2$	$x3$	$x4$
0.726	0.097	0.145	0.032



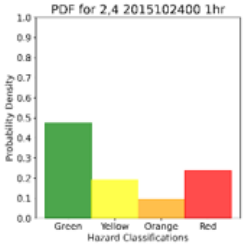
P($x|y$) Grid 2,2

$x1$	$x2$	$x3$	$x4$
0.744	0.05	0.174	0.033



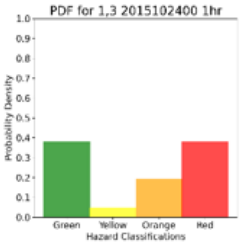
P($x|y$) Grid 2,3

$x1$	$x2$	$x3$	$x4$
0.738	0.098	0.123	0.041



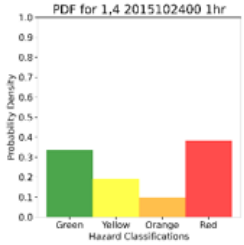
P($x|y$) Grid 2,4

$x1$	$x2$	$x3$	$x4$
0.741	0.178	0.044	0.037



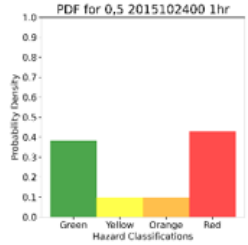
P($x|y$) Grid 1,3

$x1$	$x2$	$x3$	$x4$
0.741	0.056	0.139	0.065



P($x|y$) Grid 1,4

$x1$	$x2$	$x3$	$x4$
0.696	0.209	0.026	0.070



P($x|y$) Grid 0,5

$x1$	$x2$	$x3$	$x4$
0.748	0.112	0.056	0.084

Figure 6.7 For the Oct 25th, 2015 event. Ensemble rainfall distribution given hazard class P(y|x) for each rainfall grid which coincides with flood locations. Posterior Probability P(x|y) for the rainfall grids using the prior probability distribution and evidence from the ensemble rainfall distribution

6.5.1 Sensitivity of Optimal Action with varying loss function parameters (without vulnerability weights)

Using the October 25th, 2015 event, a sensitivity analysis of optimal early actions under varying values of the parametric loss function's parameters: γ_c;C/L (sensitivity to false alarms), γ_l (effectiveness of protective actions) and γ_d (sensitivity of damage to hazard severity) was conducted. This analysis assumed all the maximum loss and perceived loss due to false alarms is equal across all grids. The analysis shows how changes in the parameters **(**Table 6.6) combined with how the different posterior beliefs affect the selection of optimal warning actions across 10 rainfall grids (Figure 6.8).

Changes in γ_c, γ_l and γ_d were associated with transitions between inaction, moderate and higher actions. As γ_c increases, the results shifts towards lower-intensity actions to minimise the cost of false alarms. This trend is especially clear under low-hazard posterior beliefs. Inaction becomes optimal when false alarm costs dominate. Alternatively, as γ_l increases, there are higher penalties for missed alarms associated with more severe events, which generally lead to more protective actions. For increasing values of γl, the influence of this parameter appears to plateau, reflecting diminishing returns.

Similarly, when γ_d decreases, the model assigns greater weight to the consequences of missed alarms, favouring higher-level actions, especially under uncertain or moderate posterior beliefs. Because of the exponential relationship, a lower γ_d indicates greater damage from missed alarms of higher-severity events.

The posterior probability distribution significantly affects the variability of the selected actions. Grid (4,1) shows moderate actions. As γ_c increases, optimal actions shift from action 3 to category 2 and eventually to category 1. Across all grids and parameter combinations, Action 4 was selected when high-severity events are predicted with high certainty, accompanied by low γ_c and high γ_l under the assumed loss structure. When all three parameters (γ_c, γ_l, γ_d) are high, the model typically settles on low-range actions.

Grid (0,5) with a strong belief in low-severity events ($x1 = 0.7481$) tends to select higher level action with lower y_c values, as γ_c increases beyond 0.15 the decision shift to lower-level actions. However, Grid 3,0, although sharing the same belief in $x4$ as 0,5, has a lower $x1$ belief and a stronger probability for $x3$ hazards; therefore, Action 3 is selected across nearly all parameter combinations, indicating a preference for moderate responses compared to 0,5 and 1,3, which exhibit more uncertainty in the ensemble spread.

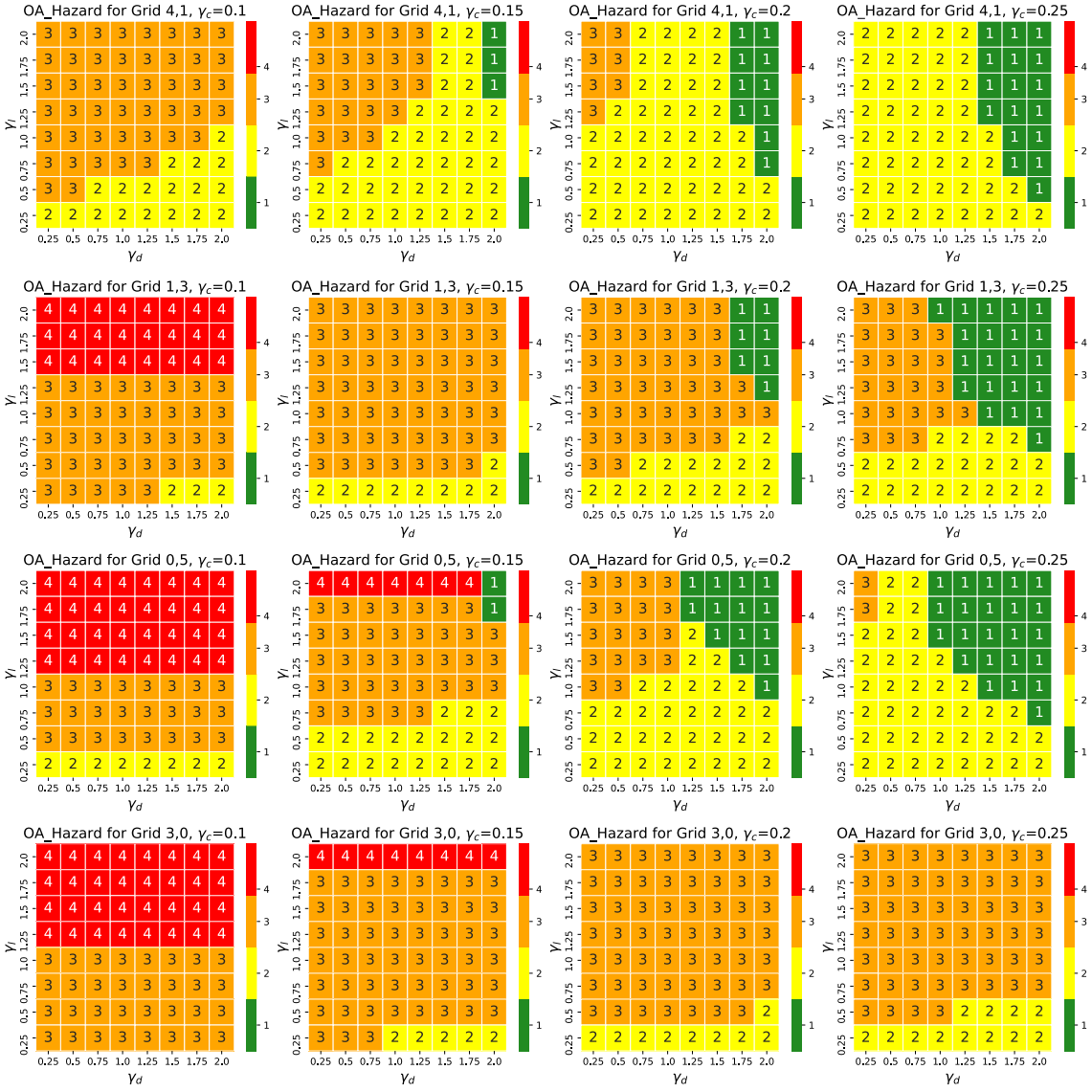


Figure 6.8 Sensitivity of Optimal Action (OA) to loss function parameters γ_c, γ_l, γ_d for 4,1, 1,3, 0,5 and 3,0 rainfall grids at the 24hr lead time using without consideration for vulnerability (Hazard only). Assumes the maximum damage is the same for all neighbourhoods.

A key assumption is that the proposed actions will reduce the loss as defined in the loss matrix. These actions were defined based on existing actions taken by local authorities but further analysis should be done and supported by a flood model where possible.

6.5.2 Influence of vulnerability weights

Figure 6.10 shows how the optimal action varies with the combination of different values of γ_c, γ_l, γ_d for selected rainfall grids when the vulnerability weight is considered. A risk matrix was used to determine Actions to the corresponding hazard class (Figure 6.7).

Each vulnerability class has the option of 4 Actions with different maximum loss and maximum cost false alarm costs as shown in Table 6.5. The trends remain the same as Figure 6.9 but now increasing losses are assigned with increasing vulnerability class resulting in higher Actions category being selected. For example, grid 3,0 neighbourhoods having a higher vulnerability and corresponding higher Actions category. Table 6.8 shows the loss matrix used for γ_c = 0.1, γ_l = 1.25, γ_d=0.7 for the different vulnerability classes. The variability of loss functions parameters for all rainfall grids are shown in Appendix B.

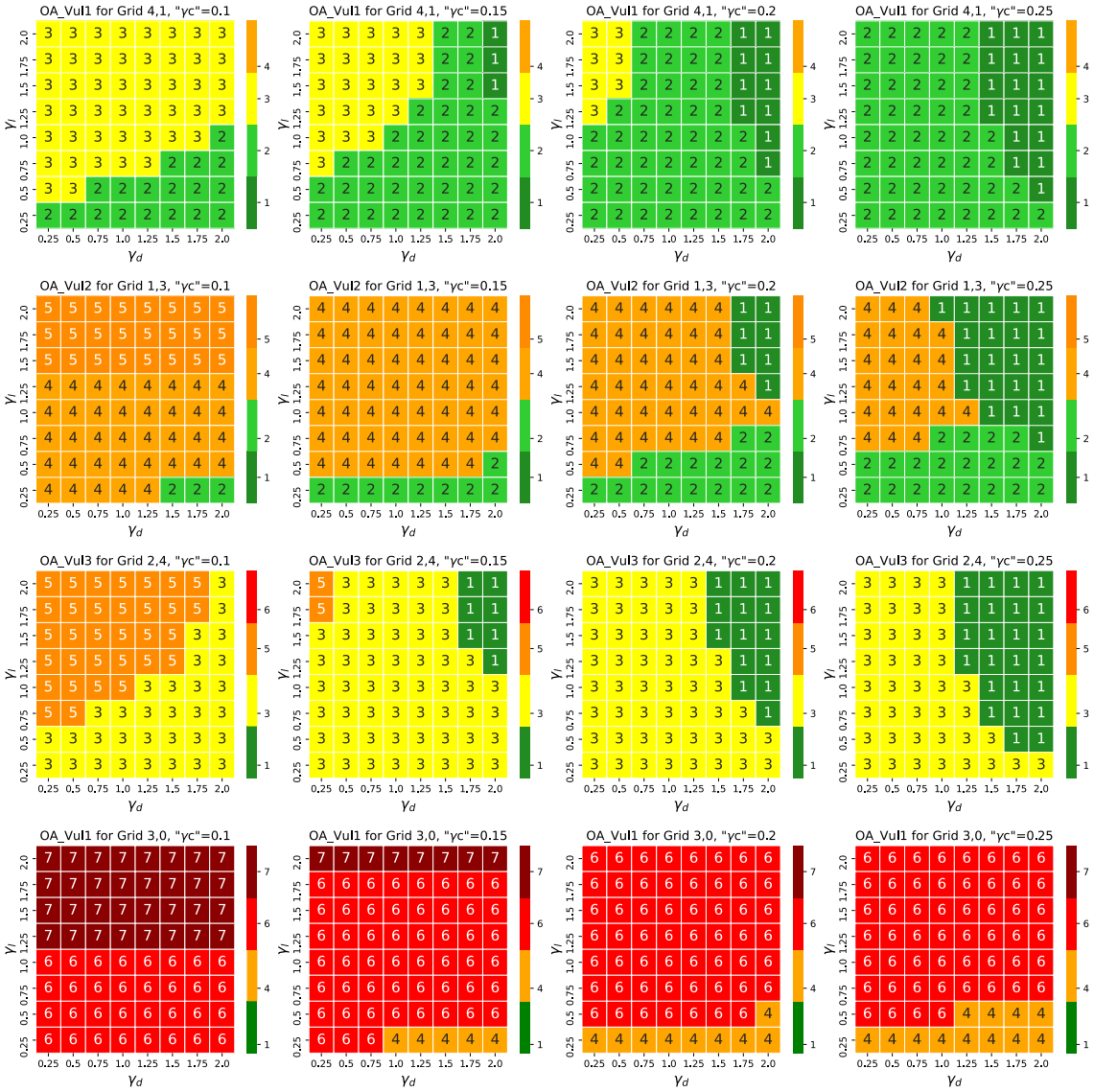


Figure 6.9 Sensitivity of the Optimal Action (OA) to loss function parameters γ_c, γ_l, γ_d for 4,1, 1,3, 2,4 and 3,0 rainfall grids at the 24hr lead time for varying vulnerability classes. Assumes the maximum damage varies with neighbourhood vulnerability class.

6.5.3 Model Calibration using the October 25th, 2015 event

Using the Oct 25th event, the optimal action was determined by varying the different parameters for the different ensemble rainfall grids to minimise overall expected residual loss while balancing missed and false alarms. The "Do-nothing" (Do-not) loss refers to the loss that would occur if no protective actions were taken and the residual loss referred the loss after actions have been taken. The results (Table 6.7) show that the combination of $\gamma_c = 0.1$, $\gamma_l = 1.25$, $\gamma_d=0.75$ achieved the lowest expected residual loss as the optimal actions closely matched the true-state action except for a false alarm for neighbourhood #52.

Table 6.7 Comparison of the True State and Optimal Action (A) Action Categories from 1-7 for all 4 vulnerability groups and corresponding Shyakha neighbourhoods for the Oct 25th, 2015 event. It shows how A* varies with parameters across all grids.*

IMERG rain grid#	6	2			3								4				1	
Forecasts grid	4,1	3,0	3,0	3,0	3,1	3,2			2,2	2,2	2,3	1,3	2,4	2,4	1,4	1,4	0,5	
Vulnerability Class	V1	V1	V2	V4	V2	V1	V1	V1	V1	V2	V4	V2	V3	V2	V3	V1	V2	Expected Loss
Shyakha neighbourhood#	52	91	83	93	40	32	31	35	36	21	20	8	9	3	11	10	12	Do-nothing (Do-not)loss
True State Action IMERG	2	3	5	7	4	3	3	3	3	4	6	4	5	4	5	3	5	644
Parameters γc, γl,γd							Optimal Actions for Forecast											Residual loss
(0.1, 1.25,0.75)	3	3	5	7	4	3	3	3	3	4	6	4	5	4	5	3	5	1
(0.15,1.25,0.75)	3	3	4	6	4	3	3	3	3	4	6	4	3	2	5	3	4	148
(0.2, 1.25,0.75)	2	2	4	6	4	3	3	3	3	4	6	4	3	2	3	2	4	201
(0.4, 1.25,0.75)	2		4	6	2	1	1	1	1	1	1	1	1	1	3	2	1	542

Table 6.8 Example loss functions used for the calibration of the decision model based on the sensitivity of the parameters $\gamma_c = 0.1$, $\gamma_l = 1.25$, $\gamma_d=0.75$

V1					V2					V3					V4				
	*x*1	*x*2	*x*3	*x*4		*x*1	*x*2	*x*3	*x*4		*x*1	*x*2	*x*3	*x*4		*x*1	*x*2	*x*3	*x*4
A1	0	13	22	30	A1	0	22	36.6	50	A1	0	30	51	70	A1	0	44	73	100
A2	1	0	17	23	A2	2	0	27.5	37.5	A3	3	0	38	53	A4	3	0	55	76
A3	2	1	0	12	A4	3	2	0	20	A5	5	3	0	29	A6	7	4	0	41
A4	3	2	1	0	A5	5	3	2	0	A6	7	5	2	0	A7	10	7	3	0

The loss matrix for $\gamma_c = 0.1$, $\gamma_l = 1.25$, $\gamma_d=0.75$ is shown in Table 6.8. However as shown in Figure 6.10 and Table 6.8, as γ_c increases the optimal action tends towards in action as γ_c increases. However, for probability distributions which had more agreement such as

3,0 the optimal action showed minimal changes compared to more uncertain distributions such as Gird (2,4), (1,4), (0,5). Figure 6.11 shows how the Optimal Action also varies spatially in relation to the other neighbourhoods for the Oct 25th, 2015 event. Higher category actions are determined for the eastern and western parts of the city, while lower category actions are determined for the middle. These actions are to be taken at the known hotspot and not the entire neighbourhood.

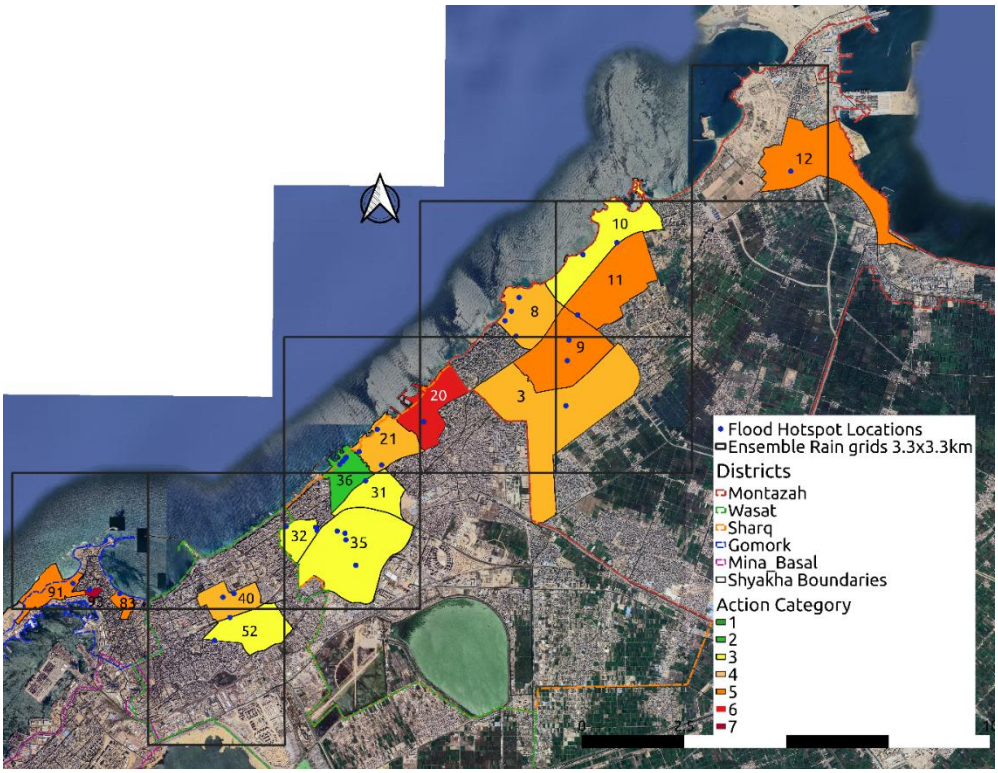


Figure 6.10 Spatial variation of actions to be taken at the flood hot spot location based on the BDF for the Oct 25th, 2015 event. Actions are to be taken at the known hotspot and not the entire neighbourhood.

6.5.4 Bayesian Decision Model Residual Loss

To further evaluate the decision model, the model was run for 9 additional events for the period 2018-2019 and 2019-2020 and the first part of 2020 winter. The Actions based on the GPM-IMERG rainfall data and vulnerability classes is shown in Table 6.9 and optimal action (A) is shown in Table 6.10. The loss values based on the loss function are shown had these actions been taken given the actual event which occurred γ_c = 0.1, γ_l = 1.25, γ_d=0.75parameters. The Table also shows the residual loss changes with parameters (0.2, 1.25,0.75) and (0.4, 1.25,0.75). In general Table 6.10 shows despite using the calibrated parameters there are residual losses which occur except for event #3 which has 0 loss. The Total Do-nothing loss is 3381compared to γ_c=0.1,0.2 and 0.4 which have total residual loss of 368,1103 and 1859 respectively. On comparison of Table 6.9 and 6.10

Action values, it shows that the Optimal actions are higher than the True action state which suggests the BDF model is over forecasting.

Table 6.9 True State Action using GPM-IMERG remotely sensed for the different Shyakha boundaries and vulnerability groups for 9 events. The hypothetical loss that would occur in the Do-nothing scenario (DN) is shown.

IMERG Grd#		6	2			3								4				1	
Forecasts grid		4,1	3,0	3,0	3,0	3,1	3,2	3,2	3,2	2,2	2,2	2,3	1,3	2,4	2,3	1,4	1,4	0,5	
Vul#		V1	V1	V2	V4	V2	V1	V1	V1	V1	V2	V4	V2	V3	V2	V3	V1	V2	
	Shyakha#	52	91	83	93	40	32	31	35	36	21	20	8	9	3	11	10	12	DN
1	20181108	1	1	1	1	1	1	1	1	1	1	1	1	1	2	3	2	1	66
2	20181204	4	2	2	4	4	3	3	3	3	4	6	4	5	1	1	1	1	418
3	20190113	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
4	20191024	1	2	2	4	1	1	1	1	1	1	1	1	1	1	1	1	1	66
5	20191224	2	3	4	6	4	3	3	3	3	4	6	4	5	4	5	3	2	577
6	20200103	2	3	4	6	2	2	2	2	2	2	4	2	3	4	5	3	5	462
7	20200311	2	2	2	4	2	2	2	2	2	2	4	2	3	4	5	3	2	404
8	20201107	1	1	1	1	2	2	2	2	2	2	4	2	3	2	3	2	1	259
9	20201120	4	2	2	4	5	4	4	4	4	5	7	5	6	4	5	3	2	667
																		Total	**3381**

Table 6.10 Optimal Strategy based on the Bayesian Decision Model and the residual loss values based on the loss function , (0.1, 1.25,0.75) had these actions been taken given the actual event which occurred . The Table also shows higher residual loss for parameters (0.2, 1.25,0.75) and (0.4, 1.25,0.75) (see Figure 6.3 for the grid layout)

	6	2			3								4				1	IMERG rain grid#		
	1	2			3	4			5		6	7	8	6	9		10	Forecast Grid #		
	4,1	3,0	3,0	3,0	3,1	3,2	3,2	3,2	2,2	2,2	2,3	1,3	2,4	2,3	1,4	1,4	0,5			
	V1	V1	V2	V4	V2	V1	V1	V1	V1	V2	V4	V2	V3	V2	V3	V1	V2	Residual loss for γc=		
	52	91	83	93	40	32	31	35	36	21	20	8	9	3	11	10	12	0.1	0.2	0.4
1	3	3	4	6	4	3	3	3	3	4	6	4	5	4	5	3	4	52	93	59
2	3	3	4	6	4	3	3	3	3	4	6	4	5	4	5	3	4	21	111	368

	6	2			3								4				1	IMERG rain grid#		
	1	2			3	4			5		6	7	8	6	9		10	Forecast Grid #		
	4,1	3,0	3,0	3,0	3,1	3,2	3,2	3,2	2,2	2,2	2,3	1,3	2,4	2,3	1,4	1,4	0,5			
	V1	V1	V2	V4	V2	V1	V1	V1	V1	V2	V4	V2	V3	V2	V3	V1	V2	Residual loss for γc=		
	52	91	83	93	40	32	31	35	36	21	20	8	9	3	11	10	12	0.1	0.2	0.4
3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
4	1	2	2	4	1	2	2	2	2	2	4	2	3	2	3	2	2	49	105	199
5	3	3	4	6	4	3	3	3	3	4	4	4	5	4	5	3	4	58	153	354
6	3	3	4	6	4	3	3	3	3	4	6	4	5	4	5	3	4	17	92	191
7	4	3	4	6	5	4	4	4	4	5	7	5	5	4	5	3	4	39	48	113
8	2	3	4	6	2	2	2	2	3	4	6	4	5	4	5	3	4	37	205	281
9	3	3	4	6	4	3	3	3	4	5	7	4	5	4	5	3	4	95	296	293
																		368	**1103**	**1858**

6.5.5 Loss per event, Vulnerability and Missed and False Alarms

In addition to the previous sensitivity analysis of loss function parameters, the analysis further examined how losses vary across different vulnerability classes and how the balance between false alarms and missed alarms evolves with changes in false alarm cost (γ_c). Figure 6.12 presents the total loss disaggregated losses by vulnerability class (V1 to V4) for multiple events for two γ_c values: 0.1 and 0.2. When $\gamma_c = 0.1$, the total loss is generally lower across events, with the most vulnerable neighbourhoods (V3, V4) experiencing lower impact. It should be noted that the number of neighbourhoods is not shared equally. There are two neighbourhoods each for V3 and V4 compared to V1 and V2 neighbourhoods which has 6 each. However, as γ_c increases to 0.2, total losses increase significantly, particularly for high-impact events like Oct 25th, 2015 and Nov 19th, 2020. This increase is primarily due to increased losses in V3 and V4 neighbourhoods which indicates that increasing γ_c disproportionately also increases risk in higher vulnerability areas. However total losses still remained less than the Do-nothing losses for most events.

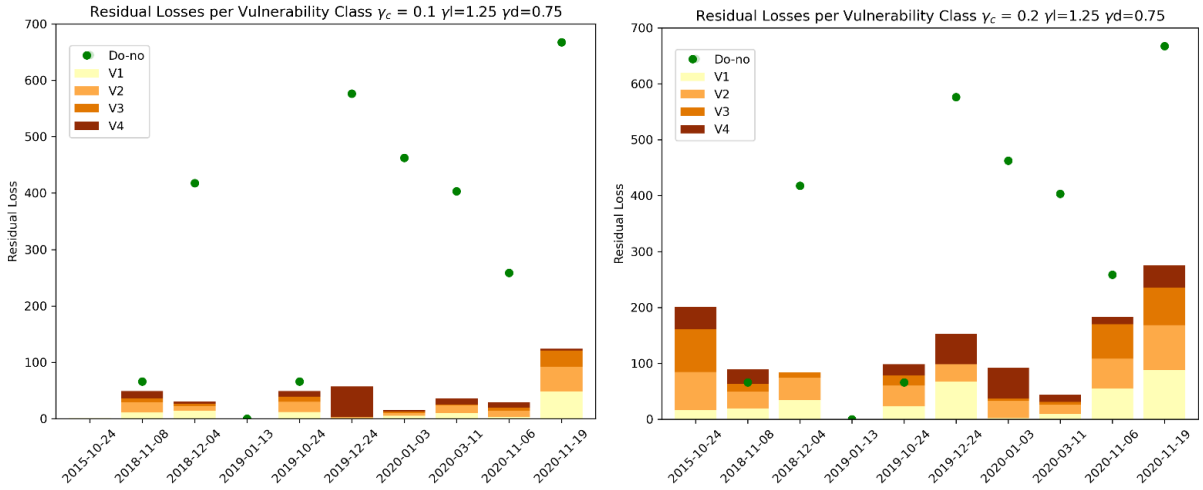


Figure 6.11 Loss for event when γ_c=0.1 (left) and Losses for events when γ_c=0.2 (right). When γ_c increases highest losses for V1andV2

Instead of vulnerability, Figure 6.13 compares the distribution of losses attributed to false alarms and missed alarms across the same set of events and loss function parameters as previously done in Table 6.10. When γ_c value = 0.1 losses are most attributed due to false alarms, as the BDF model mostly correctly matches the true state action and optimal. However, when γ_c = 0.2, the false alarm losses decline sharply and the missed alarm losses increase substantially, particularly during more extreme events. This confirms that raising the C/L ratio leads to more missed alarms, particularly with events that have higher true-state losses and that the increase in the loss γ_c =0.4 shows a very high proportion of missed alarms.

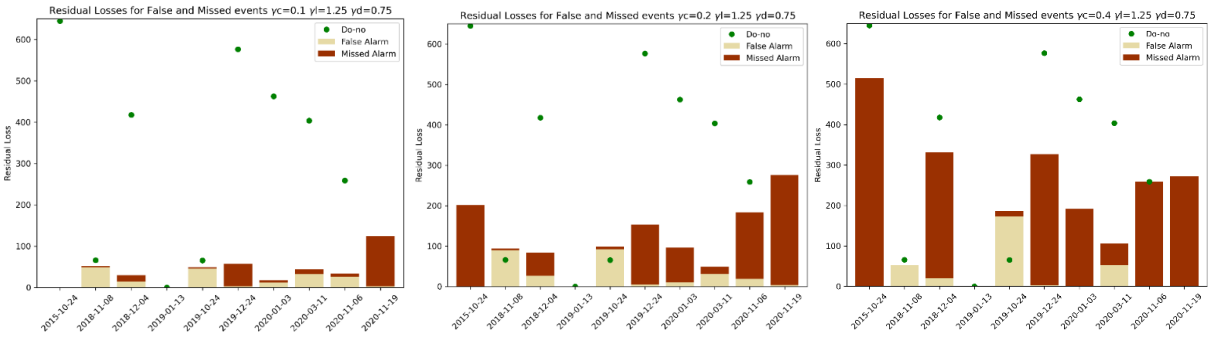


Figure 6.12 Loss disaggregated by False and Missed alarms for γ_c =0.1 and γ_c =0.2

By further disaggregating the false and missed alarm by Vulnerability class, Figure 6.14 shows how the false and missed alarms vary by vulnerability class. For Nov 20th (#9) most losses where in Vul class 1 and 2 while for Oct 25ht, 2019 (#5) missed events in V4 classes. When γ_c =0.2 the missed alarms are not spread across all vulnerability classes and this increases with values of γ_c.

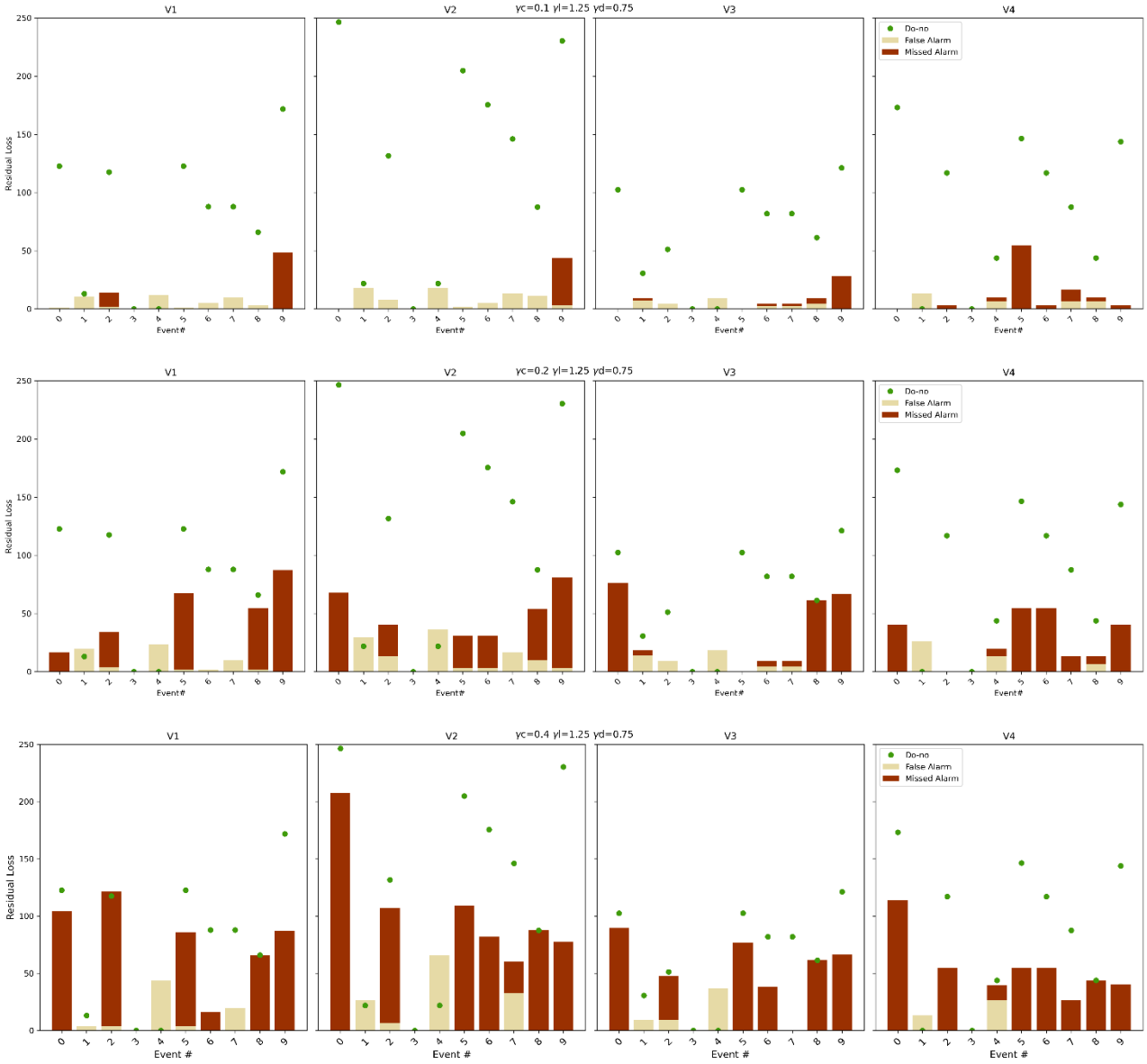


Figure 6.13 (Top) Total false alarm loss and missed alarm losses per vulnerability class for the 10 events for different values of γ_c

6.5.6 Comparison of Bayesian Model with Maximum Probability Class

The cumulative residual loss analysis (Figure 6.14) shows that smaller values of the risk preferene parameter γ_c consistently result in lower total losses when the true optimal station ation is not implemented. Ov r the evaluation period, γ_c=0.1 achieved the lowest umulative loss, followed by γ_c=0.15 and γ_c=0.2, while γ_c=0.4 showed highest losses of the parameters analysed. The max-probability decision rule performed lower losses than

γ_c=0.4 but higher than the γ_c other parameters. The "do nothing" approach resulted in the highest residual loss overall.

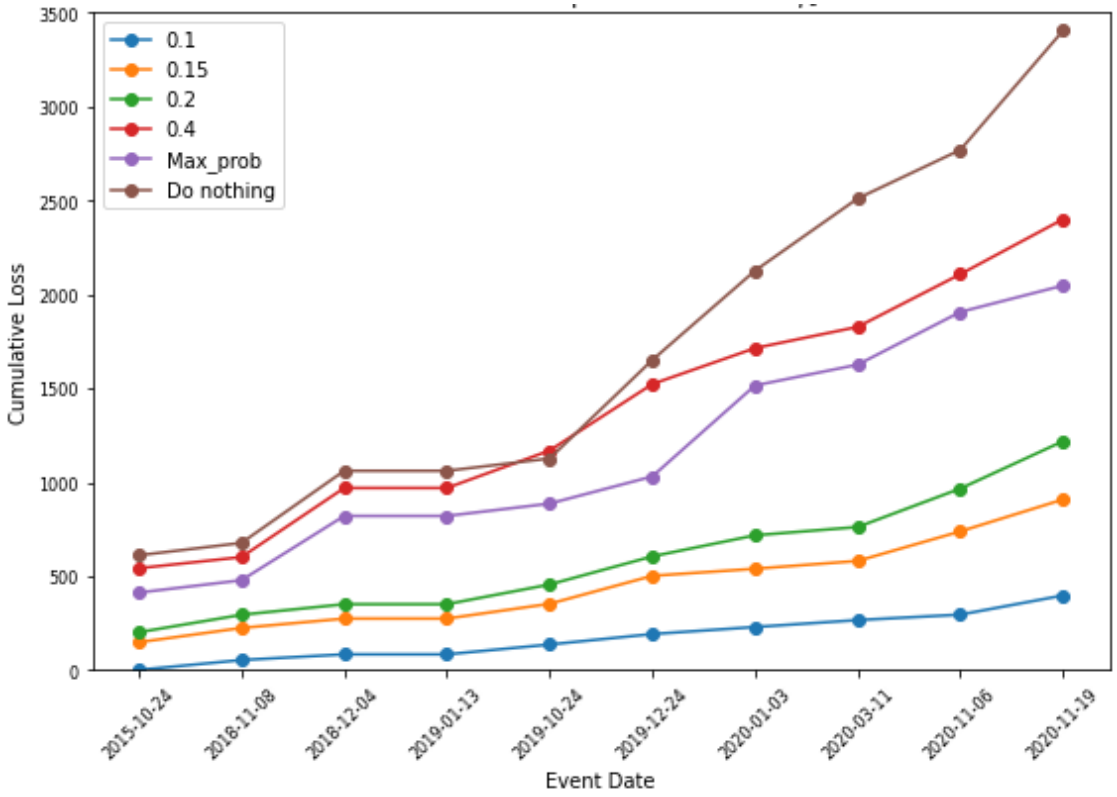


Figure 6.14 Cumulative Losses for varying values of γ_c and with values γ_l, γ_d (1.25,0.75). The Do-nothing approach results in the highest cumulative losses across all γ_c values analysed. The loss using the maximum probability class also shows lower losses than the 'do nothing' approaches, but higher losses than those with γ_c=0.1, 0.15 and 0.2. The Max Prob approach resulted in lower losses than γ_c=0.4, which describes a higher perception of losses for false alarms compared to missed alarms.

6.6 Discussion

6.6.1 Posterior and Parameters

In Bayesian decision theory, the optimal decision/action is the one that minimises the loss. The sensitivity analysis showed that the optimal actions in the Bayesian decision model are highly sensitive to the prior and posterior probability distributions, as well as the loss function parameters. In particular, the optimal action is highly sensitive to the perceived cost-loss balance, especially under differing assumptions about hazard likelihood and severity.

The C/L ratio (γ_c) is particularly influential on the decision as similarly found by (Economou et al., 2016), Busker *et al., (*2023) and Lala *et al., (*2021) who found the C/L

ratio has primary influence over the optimal probabilistic threshold. When the C/L is high, the model becomes increasingly cautious, avoiding higher category actions even when potential damage exists. This is most evident in posterior distributions, where higher confidence is observed in low-severity scenarios and the model rapidly shifts to inaction as C/L increases. This behaviour illustrates how higher cost actions can lead to greater precaution, potentially increasing residual risk, especially when there is uncertainty about the hazard state.

The parameters γ_l and γ_d collectively determine how much the model values protective actions relative to potential damages. Higher γ_l values make protective actions more attractive, especially when damage is also expected to escalate with severity (i.e when γ_d decreases). However, the diminishing returns observed with higher γ_l indicates that there is a threshold beyond which increasing the action no longer justifies higher costs. It is also important to note that the calibration event is considered a high severity event. The validation was conducted on both high and low severity events, which may also explain the BDF's tendency to avoid missed alarms. This suggests that results may be improved if high and low severity events were calibrated separately. Figure 6.14 showed how false and missed alarms vary over the two season how cumulatively the loss is reduced highlighting the need to balancing false and missed alarms when evaluating forecast model skill (Hossain et al., 2025).

Importantly, the prior and posterior distributions themselves significantly influence how sensitive decisions are to parameter changes. In sharper probability distributions, the model leans toward mid-range actions. When the probability distribution reflects greater uncertainty, especially with higher probabilities (e.g., Grid 0,5), decisions often converge to one or two dominant actions.

6.6.2 Incorporating Vulnerability weights

The analyses highlighted the complex interaction between cost-loss ratios, posterior beliefs and vulnerability in selecting optimal early warning actions. The BDF tends to favour moderate actions under uncertain beliefs. The introduction of vulnerability classes supports the adoption of equitable actions by assigning higher protective actions in higher vulnerability communities. Raising the ratio of C/L increases the cost of false alarms, lowers the frequency of protective actions and lowers city wide costs from unnecessary warnings and actions.

However, reducing overall costs comes at the expense of reduced protective coverage, especially for the most vulnerable groups. As seen in Figure 6.12, V3 and V4 experience the most dramatic increases in loss at high γ_c values, suggesting that equity is not considered when warnings are not issued to avoid false alarm cost. While low γ_c reduces the risk of missed events, it increases the frequency of protective action and over-action.

Overall, these findings demonstrate how the model adjusts its behaviour based on both the confidence in the risk estimates (posterior distribution) and the trade-offs and interactions among cost, damage and intervention effectiveness. Once highlighted, the delicate task of balancing missed and false alarms to an end user's priorities.

The decision-making behaviour reflects rational risk management: minimising expected loss by favouring moderate, effective actions under uncertainty, while allowing for inaction only when the belief in safety is strong and action carries high cost.

These results further show the importance of accurately estimating posterior probabilities and calibrating the loss function parameters using additional past events or scenarios and including tangible consequences. Together, these results advocate for decision models that are sensitive to both the distribution of belief (posterior) and the distribution of vulnerability.

6.6.3 Considerations for using a Bayesian Decision Framework

The BDF offers a transparent structured approach to incorporate uncertainty and user preferences, however, important considerations arise in both the methodology and its practical application.

The optimal action

Bayesian decision theory assumes a rational actor who makes decisions to minimise expected loss based on known probabilities and defined utility functions. However, in emergency and humanitarian contexts, decision-making actors may not operate purely rationally, introducing cognitive biases, a concept explored in Prospect Theory (Kahneman and Tversky, 1979) and further examined in emergency management (Z. X. Zhang et al., 2018).

Therefore, each decision maker should ask, "What is an optimal action for them?". This study defines optimality as balancing false and missed alarms to minimise expected loss. It does not fully capture the behavioural aspects of decision-making, including emotional, financial, political and institutional factors that often influence ways of decision-making that diverge from optimality (Frick and Hegg, 2011; Kox et al., 2015; Morss et al., 2010). Thus, the BDF is proposed as an information-driven decision support tool for taking the final action, but it is assumes that expert judgment and experience will determine the final decision. Therefore, it is essential to incorporate decision-makers' concerns into sophisticated models while showing the value of integrating new or more detailed information from models (Leskens et al., 2014).

This research acknowledges the limitations of incorporating vulnerability weights and false alarm attitudes, aims to shift away from purely economic optimality toward greater social equity. It can be further optimised based on what is regarded as an acceptable risk

tolerance, as well as the acceptable residual loss and balance between false and missed alarms, or other attributes important to the decision maker. For example, knowledge of high false alarm ratios or other forecasts biases, recent flood events that may warrant more precautionary actions and knowledge of the time of day of the event.

The Loss Function

This research builds on the parametric loss function proposed by Economou *et al. (*2016) and applies it to Alexandria city, incorporating different assumed vulnerability weights. Using a parametric loss function in a Bayesian decision theory framework offers a structured way to capture decision-makers' risk preferences and to prioritise high-vulnerability areas, even or perhaps especially when historical loss data is unavailable. However, this method also has limitations which should be considered. Parameter choices are subjective with the risk of oversimplifying complex nonlinear impacts. It assumes static preferences, which may change in evolving emergencies and the resulting "loss" may not align with actual social, economic, or political costs. Furthermore, without historical data, validation depends heavily on expert judgement, making it harder to prove the loss function truly represents reality.

Therefore, results must be carefully framed, with sensitivity testing to show the impact of different parameter assumptions, as done in this research. As Economou et al. (2016) has argued, end users have different loss tolerances. Context-specific loss functions should be applied per vulnerability class, aligning with the principle that risk is not experienced equally. In the future, when more data becomes available, the loss functions can be tailored to capture the risk tolerances of specific neighbourhoods. In the future loss function can also explore alternative ways of determining flood losses (Mobini et al., 2021). Co-designed climate services are recognised as best practice from user perspectives and should be context-specific(Fleming et al., 2023; Jacob et al., 2025; Vincent et al., 2021). Parameter values should be co-designed with forecast providers, decision-makers and affected populations to reflect actual preferences, local knowledge of impacts, risk tolerances, trade-offs and actual damage, which aligns with a multi and transdisciplinary approaches. This co-designed process can further explore location-specific thresholds and the impact of consecutive false alarms which is often not captured in loss functions.

The Prior and Thresholds

The prior P(x) is one of the most critical variables in Bayesian decision theory, representing "initial beliefs" based on historical data, past evidence, and/or expert judgment (Wesner and Pomeranz, 2021). However, priors can be considered subjective and may lead to biased information, especially in cases where priors return skewed or smaller probabilities (Banner et al., 2020). For example, uniform or equal priors allow the use of a Bayesian approach however, the posterior distribution is primarily dictated

by the likelihood function, especially when there is limited prior information or expert knowledge (Moala and Ramos, 2024). In the case of Alexandria, the effect of the chosen prior has a significant impact in transforming the probability distribution to the posterior. Further research should explore updating of the subjective priors. For example, using the False Alarm ratio based on hindcast forecast or updating priors with lead time. The peak intensity thresholds used for the hazard classification also has an impact (Young et al.,2021) Therefore further analysis on the sensitivity of the decision to peak intensity thresholds is recommended. Tailoring action thresholds not just by forecast severity but also by community vulnerability can lead to more effective and equitable outcomes.

Future Research

The research focused on a rainfall threshold approach but can be expanded to other hazard classification methods, such as real-time flood models that utilise flood inundation depths. User preference parameters were utilised; however, further consideration is necessary to determine the best methods for calibrating the BDF, particularly when limited validation data is available. Future research should refine parameters by combining actual decisions with the BDF. There is also potential to combine this approach with machine learning techniques to optimise parameters dynamically and to explore the uncertainty of other decision criteria. Additionally, users could also benefit from an interactive decision model that allows further sensitivity testing of loss parameters and other decision variables.

This research only analysed the decision at the 24h lead time which was highlighted as critical for emergency managers, but the evolution of decision should also be considered with increasing and decreasing lead-times. The current framework assumes static decision-maker preferences and does not fully capture dynamic learning or context-dependent decision behaviour over time. In the future, the BDF can be further developed to incorporate dynamic behaviour and sequential decision-making by updating probabilistic information over multiple lead times. Updating over time defines how decisions may evolve with increasing or decreasing accuracy and changes in attitudes (Taskin and Lodree, 2011). A key assumption made is also that the proposed actions will reduce the loss as defined in the loss matrix. These actions were defined based on existing actions taken by local authorities but further analysis should be done and supported by a flood model where possible.

A more user-centric approach would be to integrate it into a GIS-based platform that displays dynamically evolving maps of spatial actions over various lead times. Users can input their own parameters, loss values, risk levels at which actions are triggered, or operational thresholds and dynamically derive the appropriate decision thresholds. At present, the BDF is proposed as a framework; however, training, testing and user feedback from practitioners and decision-makers are recommended for operationalisation.

This method can also be expanded to multiple attributes (Multi-Attribute Utility Theory), for example, minimising traffic disruption or incorporating various stakeholder preferences through sensitivity analyses (Haag et al., 2019; Reichert et al., 2015). Additionally, the loss parameters were specified at a single stakeholder level or a single objective loss function (minimising expected loss). In contrast, actual decisions often involve multi-stakeholder inputs with differing tolerances for risk or sometimes conflicting objectives, such as minimising operational cost while reducing traffic disruptions.

GPM IMERG gridded rainfall data provides an opportunity in data-scarce regions however, its application in urban areas, which require high-resolution data, remains a challenge. Rainfall grids tend to smooth rainfall intensities, potentially underestimating peak rainfall intensities, especially in urban catchments or for convective events. This underestimation may partially explain the high frequency of false alarms as observations underrepresent hazard severity and forecasts tend to overestimate the hazards. In the future, the collection of high-resolution observations or estimates will remain a high priority for developing early warning and flood forecasting models.

6.7 Conclusion

At present, emergency managers in Alexandria use weather forecasts and guidance from the national meteorological agency, along with local knowledge of hotspot areas, to decide on actions. This study developed and showed the value of a Bayesian Decision Framework (BDF) that utilises Bayesian decision theory, prior knowledge of the drainage system design capacity, ensemble rainfall forecasts and a parametric loss function, which assigns weights to different vulnerability classes. The BDF combines well-known concepts of Bayesian decision theory with local hazard and vulnerability knowledge to support knowledge-driven, risk-aware decision-making.

The approach proposes going beyond the likelihood of an event by translating ensembles into predefined actions that have the potential to effectively reduce flood damage. It highlights how forecast probabilities can support proactive measures for a variety of actions, which can be tailored to specific cases and cities. The findings highlight that optimal actions are highly sensitive to both the specification of the loss function parameters (γ_c, γ_l, γ_d) and the shape of the posterior probability distributions. There are limits depending on the certainty with which events are forecasted. In particular, the perceived balance between false alarm costs and potential damage strongly influences whether the model favours precautionary or conservative strategies. The reliance on parametric loss functions without empirical calibration also means that parameter choices remain subjective, potentially oversimplifying complex real-world conditions. Priors, thresholds and corresponding hazard classifications were also shown to influence

posterior probabilities and, consequently, the resulting decisions. This highlights the need for continuous updating of beliefs, calibration with observed data and co-development with decision-makers, communities and stakeholders to ensure that the model reflects local knowledge and acceptable risk tolerances.

The BDF provides a transparent, flexible and theoretically consistent framework but it also has its limitations. The subjectivity of the framework is a strength as the loss function can be tailored to specific contexts, stakeholder priorities and risk perceptions, as shown with the use of vulnerability classes in Alexandria. However, this requires the involvement of stakeholders and increased data collection. The results also showed that BDF does not always result in "better" decisions for all events when compared with the true-state action; however, the cumulative losses over the seasons were reduced when compared with the maximum probability. Additionally, the concepts of assigning numerical values to losses require further attention.

Decision theory for flood forecasting assumes rationality and provides a prescriptive framework for how decisions should be made. However, in practice, expert judgment by emergency managers and forecasters remains essential and often determines the final course of action. Despite the challenges, BDF has shown it can improve the usability of probability forecast information, particularly for decision-makers who might otherwise overlook or be confused by the concepts of uncertainty. The BDF is not being proposed as a replacement of traditional metrics or visualisations used in practice but is proposed as a complimentary tool.

Ultimately, no single decision framework fits all contexts. The BDF's value is maximised when used as part of an integrated decision-making process, where stakeholders can engage with and understand the decision-making process. By acknowledging uncertainty and aligning decisions with both risk profiles and the priorities of emergency managers and communities, when possible, the credibility and usability of flood forecasting systems can be enhanced. This is particularly valuable for anticipatory action planning in data-scarce contexts such as Alexandria and across the Middle East, where similar challenges may arise.

DATA AVAILABILITY STATEMENT

Other data underlying this Chapter can be found in Version 1.4TU.ResearchData. dataset. https://doi.org/10.4121/0ee90a1c-bb09-4c3a-968c-7fced78c8bea.v1

7

CONCLUSION AND REFLECTIONS

7.1 Overview of Main Findings: Research Questions

Pluvial flood forecasting systems rely heavily on complex models that require high-resolution data. However, many cities, especially in data-scarce regions, lack the necessary forecasting capabilities or damage and impact data, which leads to significant challenges in building and maintaining effective flood forecasting and decision support chains. Each component of the Forecast-Decision-Response chain: data, models, knowledge, decisions and emergency response, depends on its own set of data, methods and carries distinct challenges and uncertainties. These uncertainties can have a cascading effect, propagating through the system and impacting not only the accuracy of flood models but also the quality of decision-making.

The overall aim of this research was *to investigate flood forecasting and decision support approaches tailored to data-scarce urban environments, with particular emphasis on assessing the suitability of available data and methods and on evaluating and incorporating uncertainty into anticipatory flood management*. This research was motivated by practitioners' need to take action in advance, despite limited and uncertain data, which is so often the case in urban areas of the Global South.

The research adopted an incremental and sequential approach, in which each research outcome is built upon in the next step and used as input for the next objective. Thereby integrating data, models and knowledge.

First, the research investigated how available data can be used for flood forecasting using a rainfall threshold method, crowd-sourced information and global ensemble forecasts **(Objective 1)**. Next, the research evaluated the suitability and limitations of downscaled, high-resolution, deterministic rainfall forecasts at multiple spatial scales. Two different hazard classification methods were used: The rainfall threshold methods developed in **Objective 1** and an urban flood model (**Objective 2)**. The research then evaluated the skill of high-resolution ensemble rainfall forecasts and explored the application of rainfall uncertainty for probabilistic pluvial flood forecasting (**Objective 3)**. Finally, a Bayesian decision framework was developed and used to incorporate rainfall uncertainty from the ensemble rainfall forecasts. This enables translating probabilistic forecasts into actions (**Objective 4**).

To achieve the research aim, the key research findings and answers to the research questions are presented below:

How can available data be used to improve flood forecasting in data-scarce urban areas? (Objective 1, Chapter 3)

This chapter presents a novel approach of incorporating critical rainfall thresholds, historical flood data (crowdsourced) and ensemble precipitation forecasts for forecasting

extreme rainfall and pluvial flooding to improve decision-making. The research identified extreme rainfall events associated with floods using limited data and no flood models by incorporating available data from social media and satellite precipitation products (SPPs). Combining ensembles with a rainfall threshold approach allowed characterisation of exceedance likelihood associated with flood thresholds and assessment of its uncertainty. This probabilistic information was used to support decision-making from an operational forecasting perspective.

Results were very favourable in correctly forecasting extreme events in Alexandria city. The adopted methodology and phased warning and response approach demonstrated alignment with practices employed by operational forecasting agencies. When combined with the ensemble forecasting and rainfall threshold approach, this method adds some complexity to address the limitations of deterministic rainfall forecasts, which are widely used but suffer from high false alarms and missed events. However, it offers a balance between methodological simplicity and the need to account for forecast uncertainty estimation. It does not assess exposure and vulnerability or indicate possible impacts but when combined with previous knowledge of "hot-spot" areas, can prove valuable in supporting decision-making.

Additionally, the benefit of site-specific thresholds applied on a local scale rather than on a national scale means local authorities can make specific, independent decisions and increase preparedness of their cities and communities. The proposed approach is especially useful in data-scarce areas in the genesis of developing an EWS but do not have the means or data to build detailed hydraulic models. Further, the adoption of this approach can capitalise on high-resolution forecast products such as ECMWF as they become openly available.

How suitable are high-resolution rainfall forecasts at multiple spatial scales for urban flood forecasting? (Objective 2, Chapter 4)

In the previous chapter, global rainfall forecast products were used with a rainfall threshold approach. This research objective aimed to determine whether increasing the resolution of WRF rainfall forecasts to meet urban flood modelling requirements will improve the quantification of rainfall magnitude and distribution for flood forecasting in urban pluvial cities across different flood forecasting approaches (rainfall thresholds and an urban flood model). The results found that rainfall magnitudes, distributions and predicted classes were influenced by lead time and model configuration. In particular, whether scale-aware cumulus schemes or no convection schemes were used to transition between resolved and parameterised convection.

The downscaled model effectively identified extreme weather phenomena to varying degrees at different lead times and neighbourhood sizes. The 24h lead time was found to be the most useful lead time. The runs also exhibited good skill in detecting extreme

events at the 72h lead time, but skill deteriorated at the 12h lead time. The study also highlighted that increasing the resolution does not necessarily lead to higher accuracy. In the study, we observed that the highest resolution (<1km) did not lead to improved hazard estimation.

Despite challenges in using high-resolution rainfall for flood forecasting, the good performance of both the rainfall threshold and real-time simulation methods suggests the potential for their combined use rather than relying on either alone. Combining high-resolution forecasts with rainfall threshold methods enables early warning and flood forecasting at longer lead times without running complex flood models. When available, flood models can be coupled with high-resolution forecasts to provide more detailed information on specific flood locations and depths at shorter lead times. This is valuable for decision-makers in optimising resources and implementing location-specific measures, particularly in conjunction with prior knowledge of high-risk areas.

How well do high-resolution ensemble rainfall forecasts capture intensities compared to deterministic forecasts and what is the potential for their use in probabilistic decision frameworks***? (Objective 3, Chapter 5)***

This chapter investigated the skill of high-resolution ensemble grid rainfall forecasts for 10 events across the city at specific known flood locations and compared them with deterministic forecasts. The ensemble rainfall forecasts were simulated with the WRF model forced by GEFS members and verified against GPM-IMERG rainfall products. The evaluation used categorical metrics (POD, FAR, CSI and Accuracy) and probabilistic scores (Brier Score and CRPS) to quantify forecast skill across rainfall thresholds and spatial grids.

The ensemble and deterministic forecasts performed comparably across most thresholds. When comparing the CRPS score, which represents mean error and accuracy, there was no clear winner. Ensembles showed varying probability of detection for different severity events and rainfall thresholds. In particular, the ensembles showed less skill at moderate thresholds than at higher thresholds. The results highlight the challenges and opportunities of quantifying predictive uncertainty and linking probabilistic rainfall forecasts to hazard likelihoods. The outcomes of this chapter facilitate a move from deterministic to probabilistic and the adoption of risk-informed decision frameworks. The probabilistic distributions were used as input in Chapter 6 to explore how such probabilistic information can be integrated into anticipatory flood management systems and decision support processes.

How can forecast uncertainty be incorporated into a decision framework to support anticipatory flood management? (Objective 4, Chapter 6)

This research component aimed to go beyond the traditional use of probabilistic forecasts for triggering flood early actions. A Bayesian Decision Framework (BDF) was developed

and applied to anticipatory flood management in Alexandria city, illustrating how probabilistic flood forecasts can be combined with local knowledge and user-specific preferences to trigger more equitable early flood actions.

The developed BDF incorporated ensemble flood forecasts with neighbourhood-scale flood observations and vulnerability data, enabling locally tailored anticipatory flood action. It consistently reduced cumulative seasonal losses compared with a maximum-probability decision rule. The results showed that optimal decisions were highly sensitive to the loss function parameters and the posterior probability distributions. Forecast probabilities improved proactive decision-making, but this depended on forecast certainty. Variations in users' risk perceptions highlighted the importance of calibrating loss models and accounting for residual risks' effects on vulnerable groups. The study further concluded the importance of continuously updating priors, thresholds and hazard classifications through co-production with local communities who experience flood impacts firsthand.

The BDF's strength lies in its flexibility and subjectivity which uses loss functions to represent local priorities, stakeholder preferences and perceived risks. In practice, decision theory provides a prescriptive framework that outlines the actions decision-makers should take under uncertainty. However, BDF remains a tool and expert judgment remains key in operational contexts. No single decision framework fits all contexts and the BDF thus serves as a complementary tool that translates probabilistic information into actionable guidance and improves usability, understanding and communication among decision-makers and communities.

7.2 REFLECTIONS

Overall, the methods and research findings developed in this research have the potential to improve flood forecasting and decision support in urban data-scarce areas. First, by overcoming some of the challenges that arise in the genesis of developing flood forecasting and early warning systems and as systems become more sophisticated, adopting higher-resolution forecasts and probabilistic approaches for anticipatory flood actions. Nonetheless, this dissertation and research journey has also uncovered additional insights and reinforced existing challenges.

On Creating value from imperfect data

Creating value from imperfect data and models by ensuring that information is fit for its intended purpose has been a consistent thread throughout this research. In a data-scarce context, data scarcity is a reality and has long been recognised as a limitation. However, it may be more useful to view data scarcity as an operational reality and boundary condition**,** one that requires adaptive and innovative approaches and trade-offs. Rather

than solely aiming for more data and higher forecast accuracy, what is more crucial is understanding how to align available data with model requirements and how to support decision-making across different spatial scales, lead times and user requirements. Moreover, flood forecasting approaches that can explicitly incorporate uncertainty into decision-making, without compromising the usefulness of forecasts to end users, are needed. This has been a central aim of the research. At the same time, we cannot ignore that hydrological and flood forecasting still fundamentally rely on the quality and integration of hydrometeorological data (Pechlivanidis et al., 2025) and this principle extends to decision-making. This highlights the need to not only determine what data is fit for purpose but what level of data quality is fit for purpose. “Purpose” itself also has many dimensions, especially in systems that must balance data requirements, model accuracy, sophistication, computational efficiency and multiple user needs, acceptable level of risk and varying decision context which must be considered.

Developing urban flood models with limited data

Urban flood models play a critical role in flood risk management, scenario-based planning and flood forecasting. Urban environments are dynamic and characterised by heterogeneous built-up areas, sewer systems, localised ponding and complex pump systems. All of which require high-resolution data and regularly updated information to accurately simulate flood behaviour and support timely decision-making. In coastal cities, the interactions between storm surge add an additional level of complexity. Consequently, accurate and reliable urban pluvial flood models remain difficult to develop, particularly in data-scarce settings.

The concepts of “fit for purpose data” and “creating value from imperfect data" have been central themes to this research as discussed in the section above and Chapters 3, 4 and 6. In Chapter 4, we specifically explored how varying rainfall resolutions and variability at different scales influence flood forecasting performance. This provided useful insights into detecting flooding at known hotspots, but the analysis was limited to a small area of the city (< 5 km2).

Our initial objective was to build a 1D–2D flood model capable of capturing surface storage and demonstrate how ensemble forecasts propagate flood uncertainty in flood maps. However, the lack of high-quality data, including a high-resolution DEM and comprehensive drainage network information, knowledge of pump operating rules and observed flows, constrained our ability to develop an urban flood model for ensemble flood forecasting.

While there are many opportunities to improve models (from simplified conceptual models to data assimilation and data-driven approaches) the fundamental challenge remains how to calibrate, validate and operationalise models when the foundational data is poor. Costabile et al. (2020) found that fully dynamic models are essential when the

goal of the urban flood model goes beyond identifying flood locations and involves estimating flood hazard/vulnerability. However, urban flood models for design and extremes differ from operational ensemble forecasting models, which use varying inputs, presenting distinct challenges. This experience has led me to reflect more broadly on the role of modelling in flood forecasting in data-scarce cities. Given the time constraints and limits of research, as modellers, the question becomes not how can we model this system, but rather should we and to what end?

The Optimal Action vs The Effective Decision Maker

As was seen in the floods in Limburg 2021 and Valencia 2024, even when perfect information is available, there is still a weak link in issuing and receiving warnings. The Bayesian Decision Framework (BDF) can identify the optimal action given probabilistic forecasts and defined preferences, but achieving optimal outcomes in practice depends on the decision-maker. The framework supports predictive and prescriptive decision-making, yet its effectiveness depends on emergency managers' ability to interpret, trust and act on its recommendations. In many real-world contexts, decisions are not made in isolation. Decision-makers struggle with interpreting information, competing priorities, distrust in models and inefficient institutional capacities.

Thus, the distinction between the optimal action and the effective decision-maker is crucial: the former is derived analytically, while the latter reflects experience, human judgment, organisational capacity and trust in models and is outside the scope of this research. Strengthening decision-making competence and bridging the gap between science requires training and capacity development in interpreting uncertainty products and institutional support is therefore as important as improving forecast skill or model design.

The BDF should be seen not as a replacement for human judgment or traditional probabilistic information but as a tool to improve decision competence and promote more consistent, transparent and anticipatory flood actions of the effective decision maker. This research did not address last-mile issues in communicating and disseminating warnings to the community, but this remains one of the most important areas.

Revisiting the Forecast-Decision- Anticipatory Response chain

This research adopted inter-disciplinary modelling approaches: meteorological modelling, urban hydrology and decision theory. As a result, it has also allowed me to reflect on each discipline and role in practice as a weather forecaster, flood modeller and emergency manager. Overall, this study advances our understanding of urban-scale flood forecasting in data-scarce regions by emphasising the importance of considering lead time, neighbourhood size and resolution, cumulus parameterisation, integration with drainage models and the value of ensemble forecasting in decision models. The findings reiterate the need to understand the interdependencies and limitations of the flood forecast chain:

the meteorological output, the flood forecasting method and the data and methods used for verification, with the overall goal of improved decision-making and preparedness. Instead of being called a chain, the process should be interpreted more as a loop. To achieve this, an important component of the loop is the "Integrator". The roles of the researchers, practitioner meteorologists, hydrologists, emergency managers and social scientists in the flood forecast and early warning chain are well known, and, in many ways, these roles continue to operate in silos, especially in practice.

The forecasting-decision-response and warning loop is itself a system that can benefit from a systems approach, given the interconnectivity of its components. The role of the system Integrator (an individual vs a tool) is inherently interdisciplinary, responsible for linking the data, models, social and institutional dimensions of the flood forecast chain (now a loop) to ensure coordination among the different components of the flood forecasting loop. The integrator enhances the practical usefulness of forecasts for decision-making. Their primary mandate is to ensure that probabilistic forecasts and vulnerability assessments are translated into timely, equitable and trusted actions, bridging the gap between science, decision-making and response.

7.3 Recommendations and Future Research

This thesis investigated flood forecasting and decision-making approaches suitable for data-scarce urban environments, focusing on data and method suitability and on evaluating and incorporating uncertainty in anticipatory flood management. The previous section reflected on some of the challenges and more practical aspects of this research. Despite these challenges, there are opportunities to further develop uncertainty-aware pluvial flood forecasting and decision support for anticipatory flood management.

Rethinking Anticipatory Flood Management in Urban Data-Scarce Areas

Anticipatory Flood Management (AFMA) represents a shift from reactive to proactive approaches in flood risk reduction. It shares similarities with humanitarian Anticipatory Action (AA), including monitoring, forecasts, triggers and predefined standard operating procedures. However, it differs in its pre-financing operational mechanisms. There is a growing push to incorporate anticipatory actions within broader disaster risk management frameworks to enhance their sustainability (IFRC, 2025). Likewise, AFMA should be integrated into existing long-term urban resilience and climate adaptation plans. Moreover, anticipatory interventions should be considered in the design and optimisation of new multifunctional urban infrastructure, such as pumping systems and drainage canals. The limitations of real-time flood forecasting could be mitigated by using pre-simulated scenarios and/or stress-testing to identify potential impacts and disruptions in the most critical areas. However, these endeavours must be supported by increased knowledge of risk, exposure and vulnerability.

Uncertainty-aware pluvial flood forecasting and decision support in urban areas with data scarcity should be reconceptualised not merely as a limitation, but as a challenge to be addressed within the constraints of imperfect or incomplete data. This reframing enables the development of flexible, adaptable decision support tools that prioritise action over data accuracy. Future research should explore how to systematically integrate multiple data sources and co-produce fit-for-purpose solutions with end users.

A bricolage approach, initially introduced by Levi-Strauss (1966), is described as an adaptive strategy that combines methodologies, data sources and perspectives to navigate constraints using available resources, agility and the capacity to innovate within limitations. This research, in some ways has adopted an informal data bricolage approach by combining diverse data sources, local flood knowledge and vulnerability information. However, there are opportunities to formalise this into a structured methodological bricolage approach that spans the entire forecast–decision–response loop, as well as institutional arrangements, operational constraints and technologies (Bardhan, 2017; Lotteri et al., 2023). Future research could explore systematic methods for integrating data sources of varying quality and for evaluating the suitability of different methodological combinations.

Data Collection

One of the biggest challenges has been the availability of observed data for model calibration and validation at suitable spatial and temporal scales. A priority will be to obtain more accurate, higher-resolution spatial and temporal rainfall estimates at longer lead times. Data scarcity might decrease over time as satellite and weather radar product resolutions increase and become better calibrated. Bridging this gap by blending NWP with nowcasting is an interesting avenue for research currently. Future research should also explore methods for bias correction using varying scales as data become available over longer timescales.

The prioritisation of historical flood hazard and impact databases will remain invaluable for verifying and determining expected impacts and, in the future, can support impact-based forecasting and warnings (Merz et al., 2020; V. Rözer et al., 2021). Citizen science approaches utilising rain sensors and web reporting to expand local data networks are a possible way to collect data, as done in previous projects such as FloodCitiSense (Verbeiren et al., 2019). Further, apps such as WAZE and Google Maps can provide information on traffic disruptions and historical floods.

User-centric Decision Support

For the BDF, the study concluded the importance of continuously updating priors, thresholds and hazard classifications through co-production with local communities who experience flood impacts first-hand. Decision support systems (DSS) visualise and communicate model and GIS outputs but often do not utilise probabilistic information or

translate forecast information into actions or alternatives. Decision analysis frameworks, such as Bayesian decision approaches, can be built into new or existing DSSs to translate probabilistic forecasts into actionable information while accounting for user preferences and other decision variables. Future research should also focus on assigning real or more representative values to damage, impacts and reputational costs. Additionally, users could also benefit from an interactive decision model that enables further sensitivity testing of loss parameters, thresholds and other decision variables on the action, allowing them to evaluate and refine their actions across multiple events iteratively. The BDF is proposed as a framework; however, training, testing and user feedback from practitioners and decision-makers are recommended for operationalisation. Finally, this research did not address household-level protective action decision-making; however, this remains a critical area for ongoing investigation that could be informed and strengthened by the findings of this study, particularly when integrated with appropriate communication and dissemination mechanisms.

REFERENCES

AASTMT, and Egis BCEOM International. (2011). *Phase 1 Risk Assessment for the Present Situation and Horizon 2030-Climate Change Adaptation and Natural Disasters Preparedness in the Coastal Cities of North Africa-* (Issue July). http://www.abhatoo.net.ma/content/download/61166/1315121/version/1/file/Climate+Change+Adaptation+and+Natural+Disasters+Preparedness+in+the+Coastal+Cities+of+North+Africa.+Phase+1%2C+Risk+Assessment+for+the+Present+Situation+and+Horizon+2030+–+Alexandria+

Abdrabo, K. I., Kantoush, S. A., Esmaiel, A., Saber, M., Sumi, T., Almamari, M., Elboshy, B., & Ghoniem, S. (2023). An integrated indicator-based approach for constructing an urban flood vulnerability index as an urban decision-making tool using the PCA and AHP techniques: A case study of Alexandria, Egypt. *Urban Climate*, *48*, 101426. https://doi.org/10.1016/j.uclim.2023.101426

Acosta-Coll, M., Ballester-Merelo, F., Martinez-Peiró, M., & De la Hoz-Franco, E. (2018). Real-time early warning system design for pluvial flash floods—a review. *Sensors (Switzerland)*, *18*(7). https://doi.org/10.3390/s18072255

Addor, N., Jaun, S., Fundel, F., & Zappa, M. (2011). An operational hydrological ensemble prediction system for the city of Zurich (Switzerland): Skill, case studies and scenarios. *Hydrology and Earth System Sciences*, *15*(7), 2327–2347. https://doi.org/10.5194/hess-15-2327-2011

Aldridge, T., Gunawan, O., Moore, R. J., Cole, S. J., Boyce, G., & Cowling, R. (2020). Developing an impact library for forecasting surface water flood risk. *Journal of Flood Risk Management*, *13*(3). https://doi.org/10.1111/jfr3.12641

Alfieri, L., & Thielen, J. (2015). A European precipitation index for extreme rain-storm and flash flood early warning. *Meteorological Applications*, *22*(1), 3–13. https://doi.org/10.1002/met.1328

Alfieri, Lorenzo, Salamon, P., Pappenberger, F., Wetterhall, F., & Thielen, J. (2012). Operational early warning systems for water-related hazards in Europe. *Environmental Science and Policy*, *21*, 35–49. https://doi.org/10.1016/j.envsci.2012.01.008

AON. (2020). *Weather, Climate & Catastrophe Insight: 2019 Annual Report*.

Arias-Hidalgo, M., Bhattacharya, B., Mynett, A. E., & Van Griensven, A. (2013). Experiences in using the TMPA-3B42R satellite data to complement rain gauge measurements in the Ecuadorian coastal foothills. *Hydrology and Earth System Sciences*, *17*(7), 2905–2915. https://doi.org/10.5194/hess-17-2905-2013

Awadallah, A. G., Magdy, M., Helmy, E., & Rashed, E. (2017). Assessment of rainfall intensity equations enlisted in the egyptian code for designing potable water and sewage networks. *Advances in Meteorology*, *2017*(2). https://doi.org/10.1155/2017/9496787

Banner, K. M., Irvine, K. M., & Rodhouse, T. J. (2020). The use of Bayesian priors in Ecology: The good, the bad and the not great. *Methods in Ecology and Evolution*,

11(8), 882–889. https://doi.org/10.1111/2041-210X.13407

Bardhan, R. (2017). Integrating rapid assessment of flood proneness into urban planning under data constraints: a fuzzy logic and bricolage approach. *Area Development and Policy*, *2*(3), 272–293. https://doi.org/10.1080/23792949.2017.1338523

Bárdossy, A., Seidel, J., & El Hachem, A. (2021). The use of personal weather station observations to improve precipitation estimation and interpolation. *Hydrology and Earth System Sciences*, *25*(2), 583–601. https://doi.org/10.5194/hess-25-583-2021

Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., Van Dijk, A. I. J. M., McVicar, T. R., & Adler, R. F. (2019). MSWep v2 Global 3-hourly 0.1° precipitation: Methodology and quantitative assessment. *Bulletin of the American Meteorological Society*, *100*(3), 473–500. https://doi.org/10.1175/BAMS-D-17-0138.1

Berkhahn, S., Fuchs, L., & Neuweiler, I. (2019). An ensemble neural network model for real-time prediction of urban floods. *Journal of Hydrology*. https://doi.org/10.1016/j.jhydrol.2019.05.066

Berkhahn, S., & Neuweiler, I. (2024). Data driven real-time prediction of urban floods with spatial and temporal distribution. *Journal of Hydrology X*, *22*, 100167. https://doi.org/10.1016/j.hydroa.2023.100167

Bermúdez, M., Ntegeka, V., Wolfs, V., & Willems, P. (2018). Development and Comparison of Two Fast Surrogate Models for Urban Pluvial Flood Simulations. *Water Resources Management*, *32*(8), 2801–2815. https://doi.org/10.1007/s11269-018-1959-8

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

Beven, K. J., Aspinall, W. P., Bates, P. D., Borgomeo, E., Goda, K., Hall, J. W., Page, T., Phillips, J. C., Simpson, M., Smith, P. J., Wagener, T., & Watson, M. (2018). Epistemic uncertainties and natural hazard risk assessment - Part 2: What should constitute good practice? *Natural Hazards and Earth System Sciences*, *18*(10), 2769–2783. https://doi.org/10.5194/nhess-18-2769-2018

Bhattacharya, B., Zevenbergen, C., Young, A., & Radhakrishnan, M. (2018). Extreme flooding in Alexandria: Can anticipatory flood management be a solution? *13th International Conference on Hydroinformatics*.

Bhattacharya, Biswa, Zevenbergen, C., Young, A., & Radhakrishnan, M. (2018). Extreme Flooding in Alexandria: Can Anticipatory Flood Management be a Solution? In G. La Loggia, G. Freni, V. Puleo, & M. De Marchis (Eds.), *HIC 2018. 13th International Conference on Hydroinformatics* (Vol. 3, pp. 252–257). EasyChair. https://doi.org/10.29007/wvth

Biondi, D., & Todini, E. (2018). Comparing Hydrological Postprocessors Including Ensemble Predictions Into Full Predictive Probability Distribution of Streamflow. *Water Resources Research*, *54*(12), 9860–9882. https://doi.org/10.1029/2017WR022432

Bischiniotis, K., van den Hurk, B., Coughlan de Perez, E., Veldkamp, T., Nobre, G. G.,

& Aerts, J. (2019). Assessing time, cost and quality trade-offs in forecast-based action for floods. *International Journal of Disaster Risk Reduction*, *40*(May), 101252. https://doi.org/10.1016/j.ijdrr.2019.101252

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

Böing, S. J., Birch, C. E., Rabb, B. L., Kay, |, & Shelton, L. (2020). A percentile-based approach to rainfall scenario construction for surface-water flood forecasts Meteorological Applications Science and Technology for Weather and Climate. *Meteorol Appl*, *27*. https://doi.org/10.1002/met.1963

Borga, M., Gaume, E., Martina, M., & Thielen, J. (2008). *Real time guidance for flash flood risk management*. http://www.floodsite.net/html/grenoble_workshop/documents/T16_08_02_Guidance_FF_Risk_D16_1_V1_6_P01.pdf

Botzen, W. J. W., Deschenes, O., & Sanders, M. (2019). The Economic Impacts of Natural Disasters: A Review of Models and Empirical Studies. *Review of Environmental Economics and Policy*, *13*(2), 167–188. https://doi.org/10.1093/reep/rez004

Bouwens, C., Ten Veldhuis, M.-C., Schleiss, M., Tian, X., & Schepers, J. (2018). *Towards identification of critical rainfall thresholds for urban pluvial flooding prediction based on crowdsourced flood observations*. https://doi.org/10.5194/hess-2017-751

Brendel, C. E., Dymond, R. L., & Aguilar, M. F. (2020). Integration of quantitative precipitation forecasts with real-time hydrology and hydraulics modeling towards probabilistic forecasting of urban flooding. *Environmental Modelling and Software*, *134*, 104864. https://doi.org/10.1016/j.envsoft.2020.104864

Brown, J. D. (2016). *Ensemble Verification Service (EVS) User's Manual Version 5.6*. https://amazon.nws.noaa.gov/ohd/docs/evs/v5.6/EVS_5.6_MANUAL.pdf

Buizza, R., Milleer, M., & Palmer, T. N. (1999). Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, *125*(560), 2887–2908. https://doi.org/10.1002/qj.49712556006

Buizza, Roberto. (2008). The value of probabilistic prediction. *ATMOSPHERIC SCIENCE LETTERS Atmos. Sci. Let*, *9*, 36–42. https://doi.org/10.1002/asl.170

Busker, T., de Moel, H., van den Hurk, B., & Aerts, J. C. J. H. (2023). Impact-based seasonal rainfall forecasting to trigger early action for droughts. *Science of The Total Environment*, *898*, 165506. https://doi.org/10.1016/j.scitotenv.2023.165506

Butler, D., Digman, C., Makropoulos, C., & Davies, J. W. (2018). *Urban Drainage, Fourth Edition*. CRC Press. https://doi.org/10.1201/9781351174305

Campolo, M., Soldati, A., & Andreussi, P. (2003). Artificial neural network approach to flood forecasting in the River Arno. *Hydrological Sciences Journal*, *48*(3), 381–398. https://doi.org/10.1623/hysj.48.3.381.45286

Candela, A., & Aronica, G. T. (2016). Rainfall thresholds derivation for warning pluvial flooding risk in urbanised areas. *E3S Web of Conferences*, *7*, 18016. https://doi.org/10.1051/e3sconf/20160718016

Carsell, K. M., Pingel, N. D., Asce, M., & Ford, D. T. (2004). Quantifying the Benefit of a Flood Warning System. *Natural Hazards Review*, *5*(3). https://doi.org/10.1061/(ASCE)1527-6988(2004)5:3(131)CE

Casati, B., Wilson, L. J., Stephenson, D. B., Nurmi, P., Ghelli, A., Pocernich, M., Damrath, U., Ebert, E. E., Brown, B. G., & Mason, S. (2008). Forecast verification: current status and future directions. *Meteorological Applications*, *15*(1), 3–18. https://doi.org/10.1002/met.52

Chen, H., Lim, S., Chandrasekar, V., & Jang, B.-J. (2017). Urban Hydrological Applications of Dual-Polarization X-Band Radar: Case Study in Korea. *Journal of Hydrologic Engineering*, *22*(5). https://doi.org/10.1061/(ASCE)HE.1943-5584.0001421

Chow, W. T. L., Cheong, B. D., & Ho, B. H. (2016). *A Multimethod Approach towards Assessing Urban Flood Patterns and Its Associated Vulnerabilities in Singapore*. *2016*. https://doi.org/10.1155/2016/7159132

Clark, A. J., Coniglio, M. C., Coffer, B. E., Thompson, G., Xue, M., & Kong, F. (2015). Sensitivity of 24-h Forecast Dryline Position and Structure to Boundary Layer Parameterizations in Convection-Allowing WRF Model Simulations. *Weather and Forecasting*, *30*(3), 613–638. https://doi.org/10.1175/WAF-D-14-00078.1

Cloke, H. L., & Pappenberger, F. (2009). Ensemble flood forecasting: A review. *Journal of Hydrology*, *375*, 613–626. https://doi.org/10.1016/j.jhydrol.2009.06.005

Codo, M., & Rico-Ramirez, M. A. (2018). Ensemble Radar-Based Rainfall Forecasts for Urban Hydrological Applications. *Geosciences*, *8*(8), 297. https://doi.org/10.3390/geosciences8080297

Cools, J., Vanderkimpen, P., El Afandi, G., Abdelkhalek, A., Fockedey, S., El Sammany, M., Abdallah, G., El Bihery, M., Bauwens, W., & Huygens, M. (2012). An early warning system for flash floods in hyper-arid Egypt. *Natural Hazards and Earth System Science*, *12*(2), 443–457. https://doi.org/10.5194/nhess-12-443-2012

Corral, C., Berenguer, M., Sempere-Torres, D., Poletti, L., Silvestro, F., & Rebora, N. (2019). Comparison of two early warning systems for regional flash flood hazard forecasting. *Journal of Hydrology*, *572*(January 2018), 603–619. https://doi.org/10.1016/j.jhydrol.2019.03.026

Costabile, P., Costanzo, C., De Lorenzo, G., & Macchione, F. (2020). Is local flood hazard assessment in urban areas significantly influenced by the physical complexity of the hydrodynamic inundation model? *Journal of Hydrology*, *580*, 124231. https://doi.org/10.1016/j.jhydrol.2019.124231

Cristiano, E., Veldhuis, M. C. T., & Van De Giesen, N. (2017). Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas - A review. *Hydrology and Earth System Sciences*, *21*(7), 3859–3878. https://doi.org/10.5194/hess-21-3859-2017

Crochemore, L., Ramos, M.-H., & Pappenberger, F. (2016). Bias correcting precipitation forecasts to improve the skill of seasonal streamflow forecasts. *Hydrol. Earth Syst. Sci*, *20*, 3601–3618. https://doi.org/10.5194/hess-20-3601-2016

Crochemore, L., Ramos, M.-H., Pappenberger, F., Andel, S. J. van, & Wood, A. W. (2016). An Experiment on Risk-Based Decision-Making in Water Management Using Monthly Probabilistic Forecasts. *Bulletin of the American Meteorological Society*, *97*(4), 541–551. https://doi.org/10.1175/BAMS-D-14-00270.1

Dale, M., Wicks, J., Mylne, K., Pappenberger, F., Laeger, S., Taylor, S., Mylne, K., Pappenberger, F., Laeger, S., & Taylor, Á. S. (2012). Probabilistic flood forecasting and decision-making: an innovative risk-based approach. *Nat Hazards*, *70*, 159–172. https://doi.org/10.1007/s11069-012-0483-z

Davis, S., Pentakota, L., Saptarishy, N., & Mujumdar, P. P. (2022). A Flood Forecasting Framework Coupling a High Resolution WRF Ensemble With an Urban Hydrologic Model. *Frontiers in Earth Science*, *10*. https://doi.org/10.3389/feart.2022.883842

DeGroot, M. H. (2004). *Optimal Statistical Decisions*. Wiley. https://doi.org/10.1002/0471729000

Demeritt, D., Stephens, E. M., Créton-Cazanave, L., Lutoff, C., Ruin, I., & Nobert, S. (2016). Communicating and Using Ensemble Flood Forecasts in Flood Incident Management: Lessons from Social Science. In *Handbook of Hydrometeorological Ensemble Forecasting* (pp. 1–30). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-40457-3_44-1

DHI. (2017). *MIKE 1D Reference Manual*. https://manuals.mikepoweredbydhi.help/latest/Water_Resources/MIKE_1D_reference.pdf

Difrancesco, K. N., & Tullos, D. D. (2014). Flexibility in Water Resources Management: Review of Concepts and Development of Assessment Measures for Flood Management Systems. *Journal of the American Water Resources Association*, *50*(6), 1527–1539. https://doi.org/10.1111/jawr.12214

Doyle, E. E. H., Johnston, D. M., Smith, R., & Paton, D. (2019). Communicating model uncertainty for natural hazards: A qualitative systematic thematic review. In *International Journal of Disaster Risk Reduction*. https://doi.org/10.1016/j.ijdrr.2018.10.023

Duan, Q., Pappenberger, F., Wood, A., Cloke, H. L., & Schaake, J. C. (2018). Handbook of Hydrometeorological Ensemble Forecasting. In *Handbook of Hydrometeorological Ensemble Forecasting*. https://doi.org/10.1007/978-3-642-40457-3

Duncan, A., Chen, A. S., Keedwell, E. ., Djordjević, S., & Savic, D. A. (2011). Urban flood prediction in real-time from weather radar and rainfall data using artificial neural networks. *Weather Radar and Hydrology*. https://ore.exeter.ac.uk/repository/handle/10871/18406

Dutta, D., Herath, S., & Musiake, K. (2003). A mathematical model for flood loss estimation. *Journal of Hydrology*, *277*(1–2), 24–49. https://doi.org/10.1016/S0022-

1694(03)00084-2

ECMWF. (2015). *User guide to ECMWF forecast products*. https://www.ecmwf.int/sites/default/files/elibrary/2015/16559-user-guide-ecmwf-forecast-products.pdf

Economou, T., Stephenson, D. B., Rougier, J. C., Neal, R. A., & Mylne, K. R. (2016). On the use of Bayesian decision theory for issuing natural hazard warnings. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *472*(2194), 20160295. https://doi.org/10.1098/rspa.2016.0295

El Afandi, G., & Morsy, M. (2020). Developing an Early Warning System for Flash Flood in Egypt: Case Study Sinai Peninsula. In A. M. Negm (Ed.), *Flash Floods in Egypt* (pp. 45–60). Springer International Publishing. https://doi.org/10.1007/978-3-030-29635-3_4

Elboshy, B., Kanae, S., Gamaleldin, M., Ayad, H., Osaragi, T., & Elbarki, W. (2019). A framework for pluvial flood risk assessment in Alexandria considering the coping capacity. *Environment Systems and Decisions*, *39*(1), 77–94. https://doi.org/10.1007/s10669-018-9684-7

Eltahan, M., & Magooda, M. (2018). Sensitivity of WRF microphysics schemes: Case study of simulating a severe rainfall over Egypt. *Journal of Physics: Conference Series*, *1039*(1). https://doi.org/10.1088/1742-6596/1039/1/012024

EMA. (2017). *Joint WMO Techincal Progress Report oN Egyptian Meteorological Authority Cairo Numerical Weather prediction centre*. https://www.wmo.int/pages/prog/www/DPFS/documents/Egypt_2017.pdf

Ernst, S., LaDue, D., & Gerard, A. (2018). Understanding emergency manager forecast use in severe weather events. *Journal of Operational Meteorology*, 95–105. https://doi.org/10.15191/nwajom.2018.0609

Falconer, R. H., Cobby, D., Smyth, P., Astle, G., Dent, J., & Golding, B. (2009). Pluvial flooding: New approaches in flood warning, mapping and risk management. *Journal of Flood Risk Management*, *2*(3), 198–208. https://doi.org/10.1111/j.1753-318X.2009.01034.x

Filianoti, P. G. F., Nicotra, A., Labate, A., & Zema, D. A. (2020). A Method to Improve the Flood Maps Forecasted by On-Line Use of 1D Model. *Water*, *12*(6), 1525. https://doi.org/10.3390/w12061525

Flack, D., Skinner, C., Hawkness-Smith, L., O'Donnell, G., Thompson, R., Waller, J., Chen, A., Moloney, J., Largeron, C., Xia, X., Blenkinsop, S., Champion, A., Perks, M., Quinn, N., & Speight, L. (2019). Recommendations for Improving Integration in National End-to-End Flood Forecasting Systems: An Overview of the FFIR (Flooding From Intense Rainfall) Programme. *Water*, *11*(4), 725. https://doi.org/10.3390/w11040725

Fleming, A., Bohensky, E., Dutra, L. X. C., Lin, B. B., Melbourne-Thomas, J., Moore, T., Stone-Jovicich, S., Tozer, C., Clarke, J. M., Donegan, L., Hopkins, M., Merson, S., Remenyi, T., Swirepik, A., & Vertigan, C. (2023). Perceptions of co-design, co-development and co-delivery (Co-3D) as part of the co-production process – Insights

for climate services. *Climate Services*, *30*, 100364. https://doi.org/10.1016/j.cliser.2023.100364

Flood Forecasting Centre. (2017). *Flood Guidance Statement User Guide*. http://www.ffc-environment-agency.metoffice.gov.uk/services/FGS_User_Guide.pdf

Fowler, L. D., Skamarock, W. C., Grell, G. A., Freitas, S. R., & Duda, M. G. (2016). Analyzing the Grell–Freitas Convection Scheme from Hydrostatic to Nonhydrostatic Scales within a Global Model. *Monthly Weather Review*, *144*(6), 2285–2306. https://doi.org/10.1175/MWR-D-15-0311.1

Frick, J., & Hegg, C. (2011). Can end-users' flood management decision making be improved by information about forecast uncertainty? *Atmospheric Research*, *100*(2–3), 296–303. https://doi.org/10.1016/j.atmosres.2010.12.006

Fundel, V. J., Fleischhut, N., Herzog, S. M., Göber, M., & Hagedorn, R. (2019). Promoting the use of probabilistic weather forecasts through a dialogue between scientists, developers and end-users. *Quarterly Journal of the Royal Meteorological Society*, *145*(S1), 210–231. https://doi.org/10.1002/qj.3482

García, L., Barreiro-Gomez, J., Escobar, E., Téllez, D., Quijano, N., & Ocampo-Martinez, C. (2015). Modeling and real-time control of urban drainage systems: A review. *Advances in Water Resources*, *85*(June 2018), 120–132. https://doi.org/10.1016/j.advwatres.2015.08.007

Georgakakos, K. P. (2006). Analytical results for operational flash flood guidance. *Journal of Hydrology*, *317*, 81–103. https://doi.org/10.1016/j.jhydrol.2005.05.009

Ghosh, I., & Hellweger, F. L. (2012). Effects of Spatial Resolution in Urban Hydrologic Simulations. *Journal of Hydrologic Engineering*, *17*(1), 129–137. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000405

Gofa, F. (2010). Guidelines for Verification of Ensemble Forecasts. In *MeteoSwiss Report*. http://www.cosmo-model.org/content/tasks/pastProjects/versus2/docs/epsDocument_task5.pdf

Golding, B., Mittermaier, M., Ross, C., Ebert, B., Panchuk, S., Scolobig, A., & Johnston, D. (2019). A value chain approach to optimizing early warning systems. In *Contributing Paper to Global Assessment Report on Disaster Risk Reduction 2019*. United Nations Office for Disaster Risk Reduction (UNDRR).

Golding, B. W. (1998). Nimrod: A system for generating automated very short range forecasts. *Meteorological Applications*, *5*(1), 1–16. https://doi.org/10.1017/S1350482798000577

Golding, B. W. (2009). Uncertainty propagation in a London flood simulation. *Journal of Flood Risk Management*, *2*(1), 2–15. https://doi.org/10.1111/j.1753-318X.2008.01014.x

Goodarzi, L., Banihabib, M. E., & Roozbahani, A. (2019). A decision-making model for flood warning system based on ensemble forecasts. *Journal of Hydrology*, *573*(February), 207–219. https://doi.org/10.1016/j.jhydrol.2019.03.040

Goswami, P., Shivappa, H., & Goud, S. (2012). Comparative analysis of the role of domain size, horizontal resolution and initial conditions in the simulation of tropical heavy rainfall events. *Meteorological Applications*, *19*(2), 170–178. https://doi.org/10.1002/met.253

Grell, G. A., & Freitas, S. R. (2014). A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. *Atmospheric Chemistry and Physics*, *14*(10), 5233–5250. https://doi.org/10.5194/acp-14-5233-2014

Haag, F., Zürcher, S., & Lienert, J. (2019). Enhancing the elicitation of diverse decision objectives for public planning. *European Journal of Operational Research*, *279*(3), 912–928. https://doi.org/10.1016/j.ejor.2019.06.002

Hall, J., & Solomatine, D. (2008). A framework for uncertainty analysis in flood risk management decisions. *International Journal of River Basin Management*, *6*(2), 85–98. https://doi.org/10.1080/15715124.2008.9635339

Hammond, M. J., Chen, A. S., Djordjević, S., Butler, D., & Mark, O. (2015). Urban flood impact assessment: A state-of-the-art review. *Urban Water Journal*, *12*(1), 14–29. https://doi.org/10.1080/1573062X.2013.857421

Han, J.-Y., & Hong, S.-Y. (2018). Precipitation Forecast Experiments Using the Weather Research and Forecasting (WRF) Model at Gray-Zone Resolutions. *Weather and Forecasting*, *33*(6), 1605–1616. https://doi.org/10.1175/WAF-D-18-0026.1

Hansson, S. O. (2011). Decision Theory: An Overview. In M. Lovric (Ed.), *International Encyclopedia of Statistical Science* (pp. 349–355). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-04898-2_22

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

Harrison, S. E., Potter, S. H., Prasanna, R., Doyle, E. E. H., & Johnston, D. (2022). Identifying the Impact-Related Data Uses and Gaps for Hydrometeorological Impact Forecasts and Warnings. *Weather, Climate, and Society*, *14*(1), 155–176. https://doi.org/10.1175/WCAS-D-21-0093.1

Harvey, N. J., Western, L. M., Dacre, H. F., & Capponi, A. (2022). Can Decision Theory Help End-Users Take the Appropriate Action in an Emergency? *Bulletin of the American Meteorological Society*, *103*(10), 2176–2187. https://doi.org/10.1175/BAMS-D-21-0258.1

Hasanean, H. M. (2004). Precipitation variability over the Mediterranean and its linkage with El Nino Southern Oscillation (ENSO). *Journal of Meteorology*, *29*(289), 151–160.

He, S., Raghavan, S. V., Nguyen, N. S., & Liong, S. Y. (2013). Ensemble rainfall forecasting with numerical weather prediction and radar-based nowcasting models. *Hydrological Processes*, *27*(11), 1560–1571. https://doi.org/10.1002/hyp.9254

Hénonin, J., Hongtao, M., Zheng-Yu, Y., Hartnack, J., Havnø, K., Gourbesville, P., & Mark, O. (2015). Citywide multi-grid urban flood modelling: the July 2012 flood in Beijing. *Urban Water Journal*, *12*(1), 52–66.

https://doi.org/10.1080/1573062X.2013.851710

Henonin, J., Russo, B., Mark, O., & Gourbesville, P. (2013). Real-time urban flood forecasting and modelling – a state of the art. *Journal of Hydroinformatics*, *15*(3), 717. https://doi.org/10.2166/hydro.2013.132

Herman Assumpção, T., Popescu, I., Jonoski, A., & Solomatine, D. P. (2017). Citizen observations contributing to flood modelling: opportunities and challenges. *Hydrology and Earth System Sciences Discussions*, 1–26. https://doi.org/10.5194/hess-2017-456

Hirschberg, P. A., Abrams, E., Bleistein, A., Bua, W., Monache, L. D., Dulong, T. W., Gaynor, J. E., Glahn, B., Hamill, T. M., Hansen, J. A., Hilderbrand, D. C., Hoffman, R. N., Morrow, B. H., Philips, B., Sokich, J., & Stuart, N. (2011). A Weather and Climate Enterprise Strategic Implementation Plan for Generating and Communicating Forecast Uncertainty Information. *Bulletin of the American Meteorological Society*, *92*(12), 1651–1666. https://doi.org/10.1175/BAMS-D-11-00073.1

Hofmann, J., & Schüttrumpf, H. (2019). Risk-Based Early Warning System for Pluvial Flash Floods: Approaches and Foundations. *Geosciences*, *9*(3), 127. https://doi.org/10.3390/geosciences9030127

Hossain, S., Cloke, H. L., Ficchì, A., Gupta, H., Speight, L., Hassan, A., & Stephens, E. M. (2025). A decision-led evaluation approach for flood forecasting system developments: An application to the Global Flood Awareness System in Bangladesh. *Journal of Flood Risk Management*, *18*(1). https://doi.org/10.1111/jfr3.12959

Houston, D., Werrity, A., Bassett, D., Geddes, A., Hoolachan, A., & McMillan, M. (2011). Pluvial (rain-related) flooding in urban areas: the invisible hazard. *Joseph Rowntree Foundation*, 1–92. https://eprints.gla.ac.uk/162145/

Hsiang, H. (2019). *Exploring the usefulness of near-real time remote sensing data in urban flood modelling*. IHE-Delft.

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E. J., Sorooshian, S., Tan, J., & Xie, P. (2019). *Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG) Algorithm Theoretical Basis Document (ATBD) Version 06* (Issue December).

Huizinga, J., De Moel, H., & Szewczyk, W. (2017). *Global flood depth-damage functions: Methodology and the database with guidelines, EUR 28552 EN*. Publications Office of the European Union. https://doi.org/10.2760/16510

Hurford, A. P., Parker, D. J., Priest, S. J., & Lumbroso, D. M. (2012). Validating the return period of rainfall thresholds used for Extreme Rainfall Alerts by linking rainfall intensities with observed surface water flood events. *Journal of Flood Risk Management*, *5*(2), 134–142. https://doi.org/10.1111/j.1753-318X.2012.01133.x

Hurford, A. P., Priest, S. J., Parker, D. J., & Lumbroso, D. M. (2012). The effectiveness of extreme rainfall alerts in predicting surface water flooding in England and Wales. *International Journal of Climatology*, *32*(11), 1768–1774. https://doi.org/10.1002/joc.2391

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., & Collins, W. D. (2008). Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research*, *113*(D13), D13103. https://doi.org/10.1029/2008JD009944

Ibrahim, A. (2020). A Comparative Study of Two Extreme Cases Hit Egypt in January 2008 and 2009 Using WRF Different Convective Schemes. *Biomedical Journal of Scientific & Technical Research*, *26*(2). https://doi.org/10.26717/BJSTR.2020.26.004329

Ibrahim, S., & Afandi, G. El. (2014a). Short-Range Rainfall Prediction over Nigeria Using the Weather Research and Forecasting Model. *Open Journal of Atmospheric and Climate Change*, *1*(2), 38–51. https://doi.org/10.15764/RESD.2014.02006

Ibrahim, S., & Afandi, G. (2014b). Short-range Rainfall Prediction over Egypt using the Weather Research and Forecasting Model. *Open Journal of Renewable Energy and Sustainable Development*, *2014*(2), 56–70. https://doi.org/10.15764/RESD.2014.02006

Ichiba, A., Gires, A., Tchiguirinskaia, I., Schertzer, D., & Bompard, P. (2018). *Sensitivity of urban storm models to rainfall spatio-temporal variability investigated using two different modelling approaches*. *20*, 15203.

IFRC. (2020). World Disasters Report 2020: Come Heat or High Water. In *World Disaster Report 2020*. https://www.ifrc.org/sites/default/files/2021-05/20201116_WorldDisasters_Full.pdf

IFRC. (2025). *Strengthening National Disaster Risk Management Systems through the Integration of Anticipatory Action*. https://www.ifrc.org/document/strengthening-national-disaster-risk-management-systems-through-integration-anticipatory

IPCC. (2022). Summary for Policymakers [H.-O. Pörtner, D.C. Roberts, E.S. Poloczanska, K. Mintenbeck, M. Tignor A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem]. In *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 3–34). Cambridge University Press. https://doi.org/10.1017/9781009325844.001

Iqbal, A. (2017). *Radar rainfall forecasting for sewer flood modelling to support decision-making in sewer network operations*. https://ore.exeter.ac.uk/repository/handle/10871/32565

Ivanov, V. Y., Xu, D., Dwelle, M. C., Sargsyan, K., Wright, D. B., Katopodes, N., Kim, J., Tran, V. N., Warnock, A., Fatichi, S., Burlando, P., Caporali, E., Restrepo, P., Sanders, B. F., Chaney, M. M., Nunes, A. M. B., Nardi, F., Vivoni, E. R., Istanbulluoglu, E., … Bras, R. L. (2021). Breaking Down the Computational Barriers to Real-Time Urban Flood Forecasting. *Geophysical Research Letters*, *48*(20). https://doi.org/10.1029/2021GL093585

Jacob, D., Lera St. Clair, A., Mahon, R., Marsland, S., Ndebele Murisa, M., Buontempo, C., Pulwarty, R. S., Siddiqui, M. R., Grossi, A., Steynor, A., Mugandani, R., Alexander, L. V., Ruane, A. C., Doblas-Reyes, F. J., List, G., Wolff, M., & Noori, S. (2025). Co-production of climate services: challenges and enablers. *Frontiers in*

Climate, *7*. https://doi.org/10.3389/fclim.2025.1507759

Jahanbazi, M., & Egger, U. (2014). Application and comparison of two different dual drainage models to assess urban flooding. *Urban Water Journal*, *11*(7), 584–595. https://doi.org/10.1080/1573062X.2013.871041

Jain, S. K., Mani, P., Jain, S. K., Prakash, P., Singh, V. P., Tullos, D., Kumar, S., Agarwal, S. P., & Dimri, A. P. (2017). A Brief review of flood forecasting techniques and their applications. *International Journal of River Basin Management*, *16*(3), 329–344. https://doi.org/10.1080/15715124.2017.1411920

Jang, J. H. (2015). An advanced method to apply multiple rainfall thresholds for urban flood warnings. *Water (Switzerland)*, *7*(11), 6056–6078. https://doi.org/10.3390/w7116056

Janjić, Z. I. (1994). The Step-Mountain Eta Coordinate Model: Further Developments of the Convection, Viscous Sublayer, and Turbulence Closure Schemes. *Monthly Weather Review*, *122*(5), 927–945. https://doi.org/10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2

Jankov, I., Gallus, W. A., Segal, M., & Koch, S. E. (2007). Influence of initial conditions on the WRF-ARW Model QPF response to physical parameterization changes. *Weather and Forecasting*, *22*(3), 501–519. https://doi.org/10.1175/WAF998.1

Jankov, I., Gallus, W. A., Segal, M., Shaw, B., & Koch, S. E. (2005). The impact of different WRF model physical parameterizations and their interactions on warm season MCS rainfall. *Weather and Forecasting*, *20*(6), 1048–1060. https://doi.org/10.1175/WAF888.1

Jasper-Tönnies, A., Hellmers, S., Einfalt, T., Strehz, A., & Fröhle, P. (2018). Ensembles of radar nowcasts and COSMO-DE-EPS for urban flood management. *Water Science and Technology*, *2017*(1), 27–35. https://doi.org/10.2166/wst.2018.079

Jean, V., Boucher, M. A., Frini, A., & Roussel, D. (2023). Uncertainty in three dimensions: the challenges of communicating probabilistic flood forecast maps. *Hydrology and Earth System Sciences*, *27*(18), 3351–3373. https://doi.org/10.5194/hess-27-3351-2023

Jeworrek, J., West, G., & Stull, R. (2019). Evaluation of cumulus and microphysics parameterizations in WRF across the convective gray zone. *Weather and Forecasting*, *34*(4), 1097–1115. https://doi.org/10.1175/WAF-D-18-0178.1

Jha, A. K., Bloch, R., & Lamond, J. (2012). *Cities and Flooding*. The World Bank. https://doi.org/10.1596/978-0-8213-8866-2

Jiang, Y., Zevenbergen, C., & Ma, Y. (2018). Urban pluvial flooding and stormwater management: A contemporary review of China's challenges and "sponge cities" strategy. *Environmental Science and Policy*, *80*(September 2017), 132–143. https://doi.org/10.1016/j.envsci.2017.11.016

Jiménez, P. A., Dudhia, J., González-Rouco, J. F., Navarro, J., Montávez, J. P., & García-Bustamante, E. (2012). A Revised Scheme for the WRF Surface Layer Formulation. *Monthly Weather Review*, *140*(3), 898–918. https://doi.org/10.1175/MWR-D-11-00056.1

Joslyn, S., & LeClerc, J. (2013). Decisions With Uncertainty: The Glass Half Full. *Current Directions in Psychological Science*, *Vol. 22*, *N*, 308-315 (8 pages). http://www.jstor.org/stable/44318680

Kain, J. S., Weiss, S. J., Bright, D. R., Baldwin, M. E., Levit, J. J., Carbin, G. W., Schwartz, C. S., Weisman, M. L., Droegemeier, K. K., Weber, D. B., & Thomas, K. W. (2008). Some Practical Considerations Regarding Horizontal Resolution in the First Generation of Operational Convection-Allowing NWP. *Weather and Forecasting*, *23*(5), 931–952. https://doi.org/10.1175/WAF2007106.1

Ke, Q., Tian, X., Bricker, J., Tian, Z., Guan, G., Cai, H., Huang, X., Yang, H., & Liu, J. (2020). Urban pluvial flooding prediction by machine learning approaches – a case study of Shenzhen city, China. *Advances in Water Resources*, *145*, 103719. https://doi.org/10.1016/j.advwatres.2020.103719

Kendon, E. J., Ban, N., Roberts, N. M., Fowler, H. J., Roberts, M. J., Chan, S. C., Evans, J. P., Fosser, G., & Wilkinson, J. M. (2017). Do Convection-Permitting Regional Climate Models Improve Projections of Future Precipitation Change? *Bulletin of the American Meteorological Society*, *98*(1), 79–93. https://doi.org/10.1175/BAMS-D-15-0004.1

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., & Kirschbaum, D. B. (2017). So, How Much of the Earth's Surface Is Covered by Rain Gauges? *Bulletin of the American Meteorological Society*, *98*(1), 69–78. https://doi.org/10.1175/BAMS-D-14-00283.1

Kind, J., Botzen, W. J. W., & Aerts, J. C. J. H. (2020). Social vulnerability in cost-benefit analysis for flood risk management. *Environment and Development Economics*, *25*(2), 115–134. https://doi.org/10.1017/S1355770X19000275

Koutsoyiannis, D. (2003). Rainfall disaggragation methods: theory and applications. *Mathematical Methods for Hydrological Analysis*.

Kox, T., Gerhold, L., & Ulbrich, U. (2015). Perception and use of uncertainty in severe weather warnings by emergency services in Germany. *Atmospheric Research*, *158–159*, 292–301. https://doi.org/10.1016/j.atmosres.2014.02.024

Kox, T., Lü Der 1 •, C., & Gerhold, L. (2018). Anticipation and Response: Emergency Services in Severe Weather Situations in Germany. *International Journal of Disaster Risk Science*, *9*. https://doi.org/10.1007/s13753-018-0163-z

Krzysztofowicz, R. (1993). A theory of flood warning systems. *Water Resources Research*, *29*(12), 3981–3994. https://doi.org/10.1029/93WR00961

Krzysztofowicz, R. (1999). Bayesian theory of probabilistic forecasting via deterministic hydrologic model. *Water Resources Research*, *35*(9), 2739–2750. https://doi.org/10.1029/1999WR900099

Krzysztofowicz, R. (2001). The case for probabilistic forecasting in hydrology. *Journal of Hydrology*, *249*(1–4), 2–9. https://doi.org/10.1016/S0022-1694(01)00420-6

Kumar, N., & Venkatesha Prasad, R. (2018). *A Comprehensive Crowdsourcing Approach to Urban Flood Management Ramesh*. *218*(January), 265. https://doi.org/10.1007/978-3-319-73423-1

Lala, J., Bazo, J., Anand, V., & Block, P. (2021). Optimizing forecast-based actions for extreme rainfall events. *Climate Risk Management*, *34*(September), 100374. https://doi.org/10.1016/j.crm.2021.100374

Lean, H. W., Clark, P. A., Dixon, M., Roberts, N. M., Fitch, A., Forbes, R., & Halliwell, C. (2008). Characteristics of High-Resolution Versions of the Met Office Unified Model for Forecasting Convection over the United Kingdom. *Monthly Weather Review*, *136*(9), 3408–3424. https://doi.org/10.1175/2008MWR2332.1

Leandro, J., Djordjević, S., Chen, A. S., Savić, D. A., & Stanić, M. (2011). Calibration of a 1D/1D urban flood model using 1D/2D model results in the absence of field data. *Water Science and Technology*, *64*(5), 1016–1024. https://doi.org/10.2166/wst.2011.467

Leandro, Jorge, Chen, A. S., Djordjević, S., & Savić, D. A. (2009). Comparison of 1D/1D and 1D/2D Coupled (Sewer/Surface) Hydraulic Models for Urban Flood Simulation. *Journal of Hydraulic Engineering*, *135*(6), 495–504. https://doi.org/10.1061/(ASCE)HY.1943-7900.0000037

Leitão, J. P., Simões, N. E., Maksimović, Č., Ferreira, F., Prodanović, D., Matos, J. S., & Sá Marques, A. (2010). Real-time forecasting urban drainage models: Full or simplified networks? *Water Science and Technology*, *62*(9), 2106–2114. https://doi.org/10.2166/wst.2010.382

Leskens, J. G., Brugnach, M., Hoekstra, A. Y., & Schuurmans, W. (2014). Why are decisions in flood disaster management so poorly supported by information from flood models? *Environmental Modelling and Software*, *53*, 53–61. https://doi.org/10.1016/j.envsoft.2013.11.003

Lindley, D. V, & Smith, A. F. M. (1972). Bayes Estimates for the Linear Model. *Journal of the Royal Statistical Society. Series B (Methodological)*, *34*(1), 1–41. http://www.jstor.org/stable/2985048

Liu, J., Bray, M., & Han, D. (2012). Sensitivity of the Weather Research and Forecasting (WRF) model to downscaling ratios and storm types in rainfall simulation. *Hydrological Processes*, *26*(20), 3012–3031. https://doi.org/10.1002/hyp.8247

Liu, Y., Chen, Y., Chen, O., Wang, J., Zhuo, L., Rico-Ramirez, M. A., & Han, D. (2021). To develop a progressive multimetric configuration optimisation method for WRF simulations of extreme rainfall events over Egypt. *Journal of Hydrology*, *598*, 126237. https://doi.org/10.1016/j.jhydrol.2021.126237

Lockwood, J. W., Oppenheimer, M., Lin, N., & Gourevitch, J. (2024). Socioeconomic distributional impacts of evaluating flood mitigation activities using equity-weighted benefit-cost analysis. *Environmental Research Letters*, *19*(7), 074024. https://doi.org/10.1088/1748-9326/ad4ef8

Lopez, A., Coughlan de Perez, E., Bazo, J., Suarez, P., van den Hurk, B., & van Aalst, M. (2018). Bridging forecast verification and humanitarian decisions: A valuation approach for setting up action-oriented early warnings. *Weather and Climate Extremes*. https://doi.org/10.1016/j.wace.2018.03.006

Lotteri, A., Speake, J., Kennedy, V., & Chester, D. (2023). The holistic bricolage research

approach and disaster-risk reduction. *International Journal of Disaster Risk Reduction*, *97*, 104031. https://doi.org/https://doi.org/10.1016/j.ijdrr.2023.104031

Mabrouk, M., Han, H., Fan, C., Abdrabo, K. I., Shen, G., Saber, M., Kantoush, S. A., & Sumi, T. (2023). Assessing the effectiveness of nature-based solutions-strengthened urban planning mechanisms in forming flood-resilient cities. *Journal of Environmental Management*, *344*, 118260. https://doi.org/10.1016/j.jenvman.2023.118260

Madruga De Brito, M., & Evers, M. (2016). Multi-criteria decision-making for flood risk management: a survey of the current state of the art. *Nat. Hazards Earth Syst. Sci*, *16*, 1019–1033. https://doi.org/10.5194/nhess-16-1019-2016

Mahmood, F. (2021). *Rainfall Characterization and Urban Flood Modelling in Data Scarce Regions: A Case Study of Alexandria, Egypt* (Issue August). IHE Delft Institute for Water Education.

Mapiam, P. P., Methaprayun, M., Bogaard, T., Schoups, G., & Ten Veldhuis, M.-C. (2022). Citizen rain gauges improve hourly radar rainfall bias correction using a two-step Kalman filter. *Hydrology and Earth System Sciences*, *26*(3), 775–794. https://doi.org/10.5194/hess-26-775-2022

Martina, M. L. V, Todini, E., & Libralon, A. (2006). *Rainfall Thresholds for Flood Warning Systems: A Bayesian Decision Approach*. *63*, 203–227.

Martínez-Gomariz, E., Forero-Ortiz, E., Guerrero-Hidalga, M., Castán, S., & Gómez, M. (2020). Flood Depth–Damage Curves for Spanish Urban Areas. *Sustainability*, *12*(7), 2666. https://doi.org/10.3390/su12072666

Matte, S., Boucher, M. A., Boucher, V., & Fortier Filion, T. C. (2017). Moving beyond the cost-loss ratio: Economic assessment of streamflow forecasts for a risk-Averse decision maker. *Hydrology and Earth System Sciences*, *21*(6), 2967–2986. https://doi.org/10.5194/hess-21-2967-2017

Mekonnen, K., Velpuri, N. M., Leh, M., Akpoti, K., Owusu, A., Tinonetsana, P., Hamouda, T., Ghansah, B., Paranamana, T. P., & Munzimi, Y. (2023). Accuracy of satellite and reanalysis rainfall estimates over Africa: A multi-scale assessment of eight products for continental applications. *Journal of Hydrology: Regional Studies*, *49*, 101514. https://doi.org/10.1016/j.ejrh.2023.101514

Merz, B., Kuhlicke, C., Kunz, M., Pittore, M., Babeyko, A., Bresch, D. N., Domeisen, D. I. V., Feser, F., Koszalka, I., Kreibich, H., Pantillon, F., Parolai, S., Pinto, J. G., Punge, H. J., Rivalta, E., Schröter, K., Strehlow, K., Weisse, R., & Wurpts, A. (2020). Impact Forecasting to Support Emergency Management of Natural Hazards. *Reviews of Geophysics*, *58*(4), 1–52. https://doi.org/10.1029/2020RG000704

Mile High Flood District (MHFD). (2024). *Urban Storm Drainage Criteria Manual: Volume 1 Management, Hydrology, and Hydraulics*. https://mhfd.specialdistrict.org/files/13092a81c/USDCM_Volume_1_March_2024.pdf

Ming, X., Liang, Q., Xia, X., Li, D., & Fowler, H. J. (2020). Real-Time Flood Forecasting Based on a High-Performance 2-D Hydrodynamic Model and Numerical Weather

Predictions. *Water Resources Research*, *56*(7). https://doi.org/10.1029/2019WR025583

Mirfenderesk, H., Carroll, D., Chong, E., Jafari, A., Hossain, N., van Doorn, R., & Vis, S. (2016). New generation flood forecasting and decision support system for emergency management. *Australian Journal of Emergency Management*, *31*(2), 31–37.

Moala, F. A., & Ramos, P. L. (2024). Selection of objective priors for Gumbel distribution parameters with application to maximum rainfall data. *REVSTAT-Statistical Journal*, *0*(0). https://doi.org/10.57805/revstat.vi.648

Mobini, S., Nilsson, E., Persson, A., Becker, P., & Larsson, R. (2021). Analysis of pluvial flood damage costs in residential buildings – A case study in Malmö. *International Journal of Disaster Risk Reduction*, *62*, 102407. https://doi.org/10.1016/j.ijdrr.2021.102407

Montesarchio, V., Napolitano, @bullet F, Rianna, @bullet M, Ridolfi, @bullet E, Russo, @bullet F, Sebastianelli, @bullet S, Napolitano, F., Rianna, Á. M., Ridolfi, Á. E., Russo, Á. F., Rianna, M., & Sebastianelli, S. (2015). Comparison of methodologies for flood rainfall thresholds estimation. *Nat Hazards*, *75*, 909–934. https://doi.org/10.1007/s11069-014-1357-3

Morss, R. E., Lazo, J. K., & Demuth, J. L. (2010). Examining the use of weather forecasts in decision scenarios: Results from a us survey with implications for uncertainty communication. *Meteorological Applications*, *17*(2), 149–162. https://doi.org/10.1002/met.196

Morss, R. E., Wilhelmi, O. V., Downton, M. W., & Gruntfest, E. (2005). Flood risk, uncertainty, and scientific information for decision making: Lessons from an interdisciplinary project. *Bulletin of the American Meteorological Society*, *86*(11), 1593–1601. https://doi.org/10.1175/BAMS-86-11-1593

Mosavi, A., Ozturk, P., & Chau, K. W. (2018). Flood prediction using machine learning models: Literature review. *Water (Switzerland)*, *10*(11), 1–40. https://doi.org/10.3390/w10111536

Murphy, A. H. (1977). The Value of Climatological, Categorical and Probabilistic Forecasts in the Cost-Loss Ratio Situation. *Monthly Weather Review*, *105*(7), 803–816. https://doi.org/10.1175/1520-0493(1977)105<0803:TVOCCA>2.0.CO;2

Mylne, K. R. (2002). Decision-making from probability forecasts based on forecast value. *Meteorological Applications*, *9*(3), 307–315. https://doi.org/10.1017/S1350482702003043

Nashwan, M. S., Shahid, S., & Wang, X. (2019). Assessment of satellite-based precipitation measurement products over the hot desert climate of Egypt. *Remote Sensing*, *11*(5). https://doi.org/10.3390/rs11050555

Ngo, H., Pathirana, A., Zevenbergen, C., & Ranasinghe, R. (2018). An effective modelling approach to support probabilistic flood forecasting in coastal cities-Case study: Can Tho, Mekong Delta, Vietnam. *Journal of Marine Science and Engineering*, *6*(2). https://doi.org/10.3390/jmse6020055

Nicklin, H., Leicher, A. M., Dieperink, C., & Van Leeuwen, K. (2019). Understanding the Costs of Inaction–An Assessment of Pluvial Flood Damages in Two European Cities. *Water*, *11*(4), 801. https://doi.org/10.3390/w11040801

NOAA. (2001). *National Oceanic and Atmospheric Administration Changes to the NCEP Meso Eta Analysis and Forecast System: Increase in resolution, new cloud microphysics, modified precipitation assimilation, modified 3DVAR analysis*.

Ochoa-Rodríguez, S., Wang, L.-P., Thraves, L., Johnston, A., Onof, C., & Ochoa-Rodríguez, S. (2015). Surface water flood warnings in England: overview, assessment and recommendations based on survey responses and workshops. *Journal of Flood Risk Management*. https://doi.org/10.1111/jfr3.12195

Ochoa-Rodriguez, S., Wang, L. P., Gires, A., Pina, R. D., Reinoso-Rondinel, R., Bruni, G., Ichiba, A., Gaitan, S., Cristiano, E., Van Assel, J., Kroll, S., Murlà-Tuyls, D., Tisserand, B., Schertzer, D., Tchiguirinskaia, I., Onof, C., Willems, P., & Ten Veldhuis, M. C. (2015). Impact of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: A multi-catchment investigation. *Journal of Hydrology*. https://doi.org/10.1016/j.jhydrol.2015.05.035

Parker, D. J., Priest, S. J., & Mccarthy, S. S. (2011). *Surface water flood warnings requirements and potential in England and Wales*. https://doi.org/10.1016/j.apgeog.2011.01.002

Paul, J. D., Buytaert, W., Allen, S., Ballesteros-Cánovas, J. A., Bhusal, J., Cieslik, K., Clark, J., Dugar, S., Hannah, D. M., Stoffel, M., Dewulf, A., Dhital, M. R., Liu, W., Nayaval, J. L., Neupane, B., Schiller, A., Smith, P. J., & Supper, R. (2017). Citizen science for hydrological risk reduction and resilience building. *Wiley Interdisciplinary Reviews: Water*, *5*(February), e1262. https://doi.org/10.1002/wat2.1262

Paul, S., Ghosh, S., Mathew, M., Devanand, A., Karmakar, S., & Niyogi, D. (2018). Increased Spatial Variability and Intensification of Extreme Monsoon Rainfall due to Urbanization. *Scientific Reports*, *8*(1), 1–10. https://doi.org/10.1038/s41598-018-22322-9

Pontien, N., & Bhattacharya, B. (2011). Hydrological modelling in data-scarce catchments - the example of Ruvubu River Basin in Burundi. *34th IAHR World Congress*, *July*, 348–355.

Ponzano, M., Joly, B., Beau, I., Renard, E., & Fifre, G. (2025). Bridging the gap between ensemble forecasting and end-user needs for decision-making on high-impact events. *Advances in Science and Research*, *22*, 39–52. https://doi.org/10.5194/asr-22-39-2025

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., Keller, M., Tölle, M., Gutjahr, O., Feser, F., Brisson, E., Kollet, S., Schmidli, J., van Lipzig, N. P. M., & Leung, R. (2015). A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Reviews of Geophysics*, *53*(2), 323–361. https://doi.org/10.1002/2014RG000475

Rabb, B., Boeing, S., Shelton, K., & Birch, C. (2019). *Enhanced Surface Water Flooding Forecasts*. *21*, 1–36.

Ramos, M.-H., Mathevet, T., Thielen, J., & Pappenberger, F. (2010). Communicating uncertainty in hydro-meteorological forecasts: Mission impossible? *Meteorological Applications*, *17*(2), 223–235. https://doi.org/10.1002/met.202

Ramos, M. H., Van Andel, S. J., & Pappenberger, F. (2013). Do probabilistic forecasts lead to better decisions? *Hydrology and Earth System Sciences*, *17*(6), 2219–2232. https://doi.org/10.5194/hess-17-2219-2013

Ravazzani, G., Amengual, A., Ceppi, A., Homar, V., Romero, R., Lombardi, G., & Mancini, M. (2016). Potentialities of ensemble strategies for flood forecasting over the Milano urban area. *Journal of Hydrology*, *539*, 237–253. https://doi.org/10.1016/j.jhydrol.2016.05.023

Reichert, P., Langhans, S. D., Lienert, J., & Schuwirth, N. (2015). The conceptual foundation of environmental decision support. *Journal of Environmental Management*, *154*, 316–332. https://doi.org/10.1016/j.jenvman.2015.01.053

René, J.-R., Madsen, H., & Mark, O. (2013). A methodology for probabilistic real-time forecasting – an urban case study. *Journal of Hydroinformatics*, *15*(3), 751–762. https://doi.org/10.2166/hydro.2012.031

René, J. R. (2014). *Probabilistic Real-Time Urban Flood Forecasting Based on Data of Varying Degree of Quality and Quantity*. https://ore.exeter.ac.uk/repository/handle/10871/16510

Richardson, D. S. (2000). Skill and relative economic value of the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, *126*(563), 649–667. https://doi.org/10.1002/qj.49712656313

Rico-Ramirez, M. A., Liguori, S., & Schellart, A. N. A. (2015). Quantifying radar-rainfall uncertainties in urban drainage flow modelling. *Journal of Hydrology*. https://doi.org/10.1016/j.jhydrol.2015.05.057

Ritter, J., Berenguer, M., Corral, C., Park, S., & Sempere-Torres, D. (2020). ReAFFIRM: Real-time Assessment of Flash Flood Impacts – a Regional high-resolution Method. *Environment International*. https://doi.org/10.1016/J.ENVINT.2019.105375

Robaa, S. M., & Wahab, M. M. A. (2019). *Sensitivity of WRF Model to convection schemes for rainfall forecast over Egypt*. 12–17.

Roberts, N. (2008). Assessing the spatial and temporal variation in the skill of precipitation forecasts from an NWP model. *Meteorological Applications*, *15*(1), 163–169. https://doi.org/10.1002/met.57

Rosenzweig, B. R., McPhillips, L., Chang, H., Cheng, C., Welty, C., Matsler, M., Iwaniec, D., & Davidson, C. I. (2018). Pluvial flood risk and opportunities for resilience. *Wiley Interdisciplinary Reviews: Water*, *5*(6), e1302. https://doi.org/10.1002/wat2.1302

Roulin, E. (2007). Skill and relative economic value of medium-range hydrological ensemble predictions. *Hydrology and Earth System Sciences*, *11*(2), 725–737. https://doi.org/10.5194/hess-11-725-2007

Roulston, M. S., & Smith, L. A. (2004). The Boy Who Cried Wolf Revisited: The Impact

of False Alarm Intolerance on Cost–Loss Scenarios. *Weather and Forecasting*, *19*(2), 391–397. https://doi.org/10.1175/1520-0434(2004)019<0391:TBWCWR>2.0.CO;2

Rözer, V., Peche, A., Berkhahn, S., Feng, Y., Fuchs, L., Graf, T., Haberlandt, U., Kreibich, H., Sämann, R., Sester, M., Shehu, B., Wahl, J., & Neuweiler, I. (2021). Impact-Based Forecasting for Pluvial Floods. *Earth's Future*, *9*(2). https://doi.org/10.1029/2020EF001851

Rözer, Viktor, Müller, M., Bubeck, P., Kienzler, S., Thieken, A., Pech, I., Schröter, K., Buchholz, O., & Kreibich, H. (2016). Coping with pluvial floods by private households. *Water (Switzerland)*, *8*(7). https://doi.org/10.3390/W8070304

Sarlin, P. (2013). *Policymakers' Loss functions and the evaluation of Early Warning Systems*. *15*. http://ssrn.com/abstract_id=2208385.

Savage, L. J. (1954). The foundations of statistics. In *Naval Research Logistics Quarterly*. John Wiley & Sons, Ltd. https://doi.org/10.1002/nav.3800010316

Schellart, A., Liguori, S., Krämer, S., Saul, A., & Rico-Ramirez, M. A. (2014). Comparing quantitative precipitation forecast methods for prediction of sewer flows in a small urban area. *Hydrological Sciences Journal*, *59*(7), 1418–1436. https://doi.org/10.1080/02626667.2014.920505

Schellart, A., Ochoa, S., Simões, N., Wang, L., Rico-ramirez, M., Duncan, A., Chen, A. S., Keedwell, E., Djordjevi, S., Savi, D. A., Saul, A., & Maksimovi, Č. (2011). Urban pluvial flood modelling with real time rainfall information – UK case studies. *12nd International Conference on Urban Drainage, Porto Alegre/Brazil*, *2*(September), 10–15.

Schmidt, J., Matcham, I., Reese, S., King, A., Bell, R., Henderson, R., Smart, G., Cousins, J., Smith, W., & Heron, D. (2011). Quantitative multi-risk analysis for natural hazards: a framework for multi-risk modelling. *Natural Hazards*, *58*(3), 1169–1192. https://doi.org/10.1007/s11069-011-9721-z

Schwartz, C. S., Kain, J. S., Weiss, S. J., Xue, M., Bright, D. R., Kong, F., Thomas, K. W., Levit, J. J., & Coniglio, M. C. (2009). Next-Day Convection-Allowing WRF Model Guidance: A Second Look at 2-km versus 4-km Grid Spacing. *Monthly Weather Review*, *137*(10), 3351–3372. https://doi.org/10.1175/2009MWR2924.1

See, L. (2019). A Review of Citizen Science and Crowdsourcing in Applications of Pluvial Flooding. *Frontiers in Earth Science*, *7*(March), 1–7. https://doi.org/10.3389/feart.2019.00044

Silvestro, F., Rossi, L., Campo, L., Parodi, A., Fiori, E., Rudari, R., & Ferraris, L. (2019). Impact-based flash-flood forecasting system: Sensitivity to high resolution numerical weather prediction systems and soil moisture. *Journal of Hydrology*, *572*, 388–402. https://doi.org/10.1016/J.JHYDROL.2019.02.055

Simões, N. E., Ochoa-Rodríguez, S., Wang, L.-P., Pina, R. D., Marques, A. S., Onof, C., & Leitão, J. P. (2015). Stochastic Urban Pluvial Flood Hazard Maps Based upon a Spatial-Temporal Rainfall Generator. *Water*, *7*, 3396–3406. https://doi.org/10.3390/w7073396

Simpson, M., James, R., Hall, J. W., Borgomeo, E., Ives, M. C., Almeida, S., Kingsborough, A., Economou, T., Stephenson, D., & Wagener, T. (2016). Decision Analysis for Management of Natural Hazards. *Annual Review of Environment and Resources*, *41*(1), 489–516. https://doi.org/10.1146/annurev-environ-110615-090011

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., & Powers, J. G. (2008). *A Description of the Advanced Research WRF Version 3*. https://doi.org/10.5065/D68S4MVH

Speight, L., Cole, S. J., Moore, R. J., Pierce, C., Wright, B., Golding, B., Cranston, M., Tavendale, A., Dhondia, J., & Ghimire, S. (2018). Developing surface water flood forecasting capabilities in Scotland: an operational pilot for the 2014 Commonwealth Games in Glasgow. *Journal of Flood Risk Management*, *11*(S2), S884–S901. https://doi.org/10.1111/jfr3.12281

Speight, L. J., Cranston, M. D., White, C. J., & Kelly, L. (2021). Operational and emerging capabilities for surface water flood forecasting. *WIREs Water*, *April 2020*, 1–24. https://doi.org/10.1002/wat2.1517

Stephens, E., Day, J. J., Pappenberger, F., & Cloke, H. (2015). Precipitation and floodiness. *Geophysical Research Letters*, *42*(23), 10,316-10,323. https://doi.org/10.1002/2015GL066779

Taskin, S., & Lodree, E. J. (2011). A Bayesian decision model with hurricane forecast updates for emergency supplies inventory management. *Journal of the Operational Research Society*, *62*(6), 1098–1108. https://doi.org/10.1057/jors.2010.14

Taylor, A. L., Kox, T., & Johnston, D. (2018). Communicating high impact weather: Improving warnings and decision making processes. *International Journal of Disaster Risk Reduction*, *30*, 1–4. https://doi.org/10.1016/j.ijdrr.2018.04.002

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, *106*(D7), 7183–7192. https://doi.org/10.1029/2000JD900719

ten Veldhuis, J. A.E. (2011). How the choice of flood damage metrics influences urban flood risk assessment. *Journal of Flood Risk Management*, *4*(4), 281–287. https://doi.org/10.1111/j.1753-318X.2011.01112.x

ten Veldhuis, Johanna A.E., Clemens, F. H. L. R., & van Gelder, P. H. A. J. M. (2011). Quantitative fault tree analysis for urban water infrastructure flooding. *Structure and Infrastructure Engineering*, *7*(11), 809–821. https://doi.org/10.1080/15732470902985876

Teutschbein, C., & Seibert, J. (2013). Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions? *Hydrol. Earth Syst. Sci*, *17*, 5061–5077. https://doi.org/10.5194/hess-17-5061-2013

Tewari, M., Chen, F., Wang, W., Dudhia, J., LeMone, M. A., Mitchell, K., Ek, M., Gayno, G., Wegiel, J., & Cuenca, R. H. (2004). Implementation and verification of the unified noah land surface model in the WRF model. *Bulletin of the American Meteorological Society*, 2165–2170.

Thorndahl, S., Nielsen, J. E., & Jensen, D. G. (2016). Urban pluvial flood prediction: A case study evaluating radar rainfall nowcasts and numerical weather prediction models as model inputs. *Water Science and Technology*, *74*(11), 2599–2610. https://doi.org/10.2166/wst.2016.474

Thorndahl, S., Poulsen, T. S., Bøvith, T., Borup, M., Ahm, M., Nielsen, J. E., Grum, M., Rasmussen, M. R., Gill, R., & Mikkelsen, P. S. (2013). Comparison of short-term rainfall forecasts for modelbased flow prediction in urban drainage systems. *Water Science and Technology*, *68*(2), 472–478. https://doi.org/10.2166/wst.2013.274

Todini, E. (2009). *Uncertainties in Environmental Modelling and Consequences for Policy Making*. *January 2009*. https://doi.org/10.1007/978-90-481-2636-1

Todini, E. (2017). Flood Forecasting and Decision Making in the new Millennium. Where are We? *Water Resources Management*, *31*(10), 3111–3129. https://doi.org/10.1007/s11269-017-1693-7

Umer, Y., Ettema, J., Jetten, V., Steeneveld, G.-J., & Ronda, R. (2021). Evaluation of the WRF Model to Simulate a High-Intensity Rainfall Event over Kampala, Uganda. *Water*, *13*(6), 873. https://doi.org/10.3390/w13060873

United Nations Department of Economic and Social Affairs (UNDESA), P. D. (2024). *Thirty years after Cairo: Major trends, progress made and challenges ahead*. UN DESA/POP/2023/TR/NO. 8.

US Army Corps, Engineers, & Hydrologic. (n.d.). *HEC-RAS Hydraulic Reference Manual*. 2020. https://www.hec.usace.army.mil/confluence/rasdocs/ras1dtechref/6.1/modeling-culverts/culvert-data-and-coefficients/manning-s-roughness-coefficient

van Andel, S. J. (2009). Anticipatory water management: Using ensemble weather forecasts for critical events [Unesco-IHE and TU Delft]. In *PhD Thesis*. http://www.narcis.nl/publication/RecordID/oai:tudelft.nl:uuid:eab0dcb9-79ee-404c-bb62-650b945db1ad

van Andel, S. J., Price, R., Lobbrecht, A., van Kruiningen, F., Mureau, R., & Cordero, W. B. (2014). Framework for Anticipatory Water Management: Testing for Flood Control in the Rijnland Storage Basin. *Journal of Water Resources Planning and Management*, *140*(4), 533–542. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000254

Van Steenbergen, N., & Willems, P. (2014). Rainfall uncertainty in flood forecasting: Belgian case study of rivierbeek. *Journal of Hydrologic Engineering*, *19*(10), 1–6. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001004

Verbeiren, B., Seyoum, S. D., Lubbad, I., Xin, T., ten Veldhuis, M.-C., Onof, C., Wang, L.-P., Ochoa-Rodriguez, S., Veeckman, C., Boonen, M., See, L., Nalpas, D., O'Brien, B., Johnston, A., & Willems, P. (2019). FloodCitiSense: Early Warning Service for Urban Pluvial Floods for and by Citizens and City Authorities. In *New Trends in Urban Drainage Modelling* (Issue September 2018, pp. 660–664). https://doi.org/10.1007/978-3-319-99867-1_114

Verkade, J. S., Brown, J. D., Reggiani, P., & Weerts, A. H. (2013). Post-processing

ECMWF precipitation and temperature ensemble reforecasts for operational hydrologic forecasting at various spatial scales. *Journal of Hydrology*, *501*, 73–91. https://doi.org/10.1016/j.jhydrol.2013.07.039

Verkade, J. S., & Werner, M. G. F. (2011). Estimating the benefits of single value and probability forecasting for flood warning. *Hydrol. Earth Syst. Sci*, *15*, 3751–3765. https://doi.org/10.5194/hess-15-3751-2011

Vincent, K., Steynor, A., McClure, A., Visman, E., Waagsaether, K. L., Carter, S., & Mittal, N. (2021). Co-production: Learning from Contexts. In *Climate Risk in Africa* (pp. 37–56). Springer International Publishing. https://doi.org/10.1007/978-3-030-61160-6_3

Wang, H., Hu, Y., Guo, Y., Wu, Z., & Yan, D. (2022). Urban flood forecasting based on the coupling of numerical weather model and stormwater model: A case study of Zhengzhou city. *Journal of Hydrology: Regional Studies*, *39*(December 2021), 100985. https://doi.org/10.1016/j.ejrh.2021.100985

Wang, X., Steinle, P., Seed, A., & Xiao, Y. (2016). The Sensitivity of Heavy Precipitation to Horizontal Resolution, Domain Size, and Rain Rate Assimilation: Case Studies with a Convection-Permitting Model. *Advances in Meteorology*, *2016*. https://doi.org/10.1155/2016/7943845

Weerts, A. H., Winsemius, H. C., & Verkade, J. S. (2011). Estimation of predictive hydrological uncertainty using quantile regression: Examples from the National Flood Forecasting System (England and Wales). *Hydrology and Earth System Sciences*, *15*(1), 255–265. https://doi.org/10.5194/hess-15-255-2011

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

Wesner, J. S., & Pomeranz, J. P. F. (2021). Choosing priors in Bayesian ecological models by simulating from the prior predictive distribution. *Ecosphere*, *12*(9). https://doi.org/10.1002/ecs2.3739

Wetterhall, F., & Smith, P. (2019). Hydrological Challenges in Meteorological Post-processing. In Q. Duan, F. Pappenberger, A. Wood, H. L. Cloke, & J. C. Schaake (Eds.), *Handbook of Hydrometeorological Ensemble Forecasting* (pp. 239–253). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-39925-1_15

WMO. (2011). *Manual on flood forecasting and warning* (Issue 1072). http://www.wmo.int/pages/prog/hwrp/publications/flood_forecasting_warning/WMO 1072_en.pdf

WMO. (2015). *THE WMO GUIDELINES ON MULTI-HAZARD IMPACT-BASED FORECAST AND WARNING SERVICES*. https://www.wmo.int/pages/prog/www/DPFS/Meetings/ET-OWFPS_Montreal2016/documents/WMOGuidelinesonMulti-hazardImpact-basedForecastandWarningServices.pdf

WMO. (2018). Multi-hazard Early Warning Systems: A Checklist Foreword. In *Outcome of the first Multi-hazard Early Warning Conference* (Issue May).

https://library.wmo.int/doc_num.php?explnum_id=4463

WMO. (2019). *Flash Flood Guidance System (FFGS) with Global Coverage*. http://www.wmo.int/pages/prog/hwrp/flood/ffgs/index_en.php

WMO. (2021). *WMO Guidelines on Multi-hazard Impact-based Forecast and Warning Services* (Issue 1150).

Woodhams, B. J., Birch, C. E., Marsham, J. H., Bain, C. L., Roberts, N. M., & Boyd, D. F. A. (2018). What is the added value of a convection-permitting model for forecasting extreme rainfall over tropical East Africa? *Monthly Weather Review*, *146*(9), 2757–2780. https://doi.org/10.1175/MWR-D-17-0396.1

Wu, D., & Wang, Y. (2009). A research on using critical precipitation value for taiwan inundation warning system. *Proceedings the 2009 Conference for Disaster Management in Taiwan, Tai Pei 19th December 2009*.

Wu, S. J., Hsu, C. T., Lien, H. C., & Chang, C. H. (2015). Modeling the effect of uncertainties in rainfall characteristics on flash flood warning based on rainfall thresholds. *Natural Hazards*, *75*(2), 1677–1711. https://doi.org/10.1007/s11069-014-1390-2

Wu, W., Emerton, R., Duan, Q., Wood, A. W., Wetterhall, F., & Robertson, D. E. (2020). Ensemble flood forecasting: Current status and future opportunities. *WIREs Water*, *7*(3), 1–32. https://doi.org/10.1002/wat2.1432

Xing, Y., Liang, Q., Wang, G., Ming, X., & Xia, X. (2019). City-scale hydrodynamic modelling of urban flash floods: the issues of scale and resolution. *Natural Hazards*, *96*(1), 473–496. https://doi.org/10.1007/s11069-018-3553-z

Yang, T. H., Hwang, G. Do, Tsai, C. C., & Ho, J. Y. (2016). Using rainfall thresholds and ensemble precipitation forecasts to issue and improve urban inundation alerts. *Hydrology and Earth System Sciences*, *20*(12), 4731–4745. https://doi.org/10.5194/hess-20-4731-2016

Yoon, S.-S. (2019). Adaptive Blending Method of Radar-Based and Numerical Weather Prediction QPFs for Urban Flood Forecasting. *Remote Sensing*, *11*(6), 642. https://doi.org/10.3390/rs11060642

Youm7. (2020). *Alexandria Governor: We have identified 93 hotspots of rainwater accumulation this*. https://www.youm7.com/story/2020/11/29/محافظ-الإسكندرية-رصدنا-93-5090802/بؤرة-ساخنة-لتجمعات-مياه-الأمطار-خلال

Young, A., Bhattacharya, B., Daniëls, E., & Zevenbergen, C. (2025). Integrating WRF forecasts at different scales for pluvial flood forecasting using a rainfall threshold approach and a real-time flood model. *Journal of Hydrology*, *656*, 132891. https://doi.org/10.1016/j.jhydrol.2025.132891

Young, A., Bhattacharya, B., & Zevenbergen, C. (2021). A rainfall threshold-based approach to early warnings in urban data-scarce regions: A case study of pluvial flooding in Alexandria, Egypt. *Journal of Flood Risk Management*, *14*(2), 1–16. https://doi.org/10.1111/jfr3.12702

Yu, D., Yin, J., & Liu, M. (2016). Validating city-scale surface water flood modelling

using crowd-sourced data. *Environmental Research Letters*, *11*(12), 1–21. https://doi.org/10.1088/1748-9326/11/12/124011

Zanchetta, A. D. L., & Coulibaly, P. (2020). Recent advances in real-time pluvial flash flood forecasting. *Water (Switzerland)*, *12*(2). https://doi.org/10.3390/w12020570

Zevenbergen, C., Bhattacharya, B., Wahaab, R. A., Elbarki, W. A. I., Busker, T., & Salinas Rodriguez, C. N. A. (2016). In the aftermath of the October 2015 Alexandria Flood Challenges of an Arab city to deal with extreme rainfall storms. *Natural Hazards*, *October 2015*. https://doi.org/10.1007/s11069-016-2724-z

Zevenbergen, C., Veerbeek, W., Gersonius, B., & Van Herk, S. (2008). Challenges in urban flood management: travelling across spatial and temporal scales. *Journal of Flood Risk Management*, *1*(2), 81–88. https://doi.org/10.1111/j.1753-318X.2008.00010.x

Zhang, Q., Li, L., Ebert, B., Golding, B., Johnston, D., Mills, B., Panchuk, S., Potter, S., Riemer, M., Sun, J., Taylor, A., Jones, S., Ruti, P., & Keller, J. (2019). Increasing the value of weather-related warnings. *Science Bulletin*, *64*(10), 647–649. https://doi.org/10.1016/j.scib.2019.04.003

Zhang, Z. X., Wang, L., & Wang, Y. M. (2018). An Emergency Decision Making Method Based on Prospect Theory for Different Emergency Situations. *International Journal of Disaster Risk Science*, *9*(3), 407–420. https://doi.org/10.1007/s13753-018-0173-x

Zheng, Y., Alapaty, K., Herwehe, J. A., Del Genio, A. D., & Niyogi, D. (2016). Improving High-Resolution Weather Forecasts Using the Weather Research and Forecasting (WRF) Model with an Updated Kain–Fritsch Scheme. *Monthly Weather Review*, *144*(3), 833–860. https://doi.org/10.1175/MWR-D-15-0005.1

Zhou, Q., Mikkelsen, P. S., Halsnæs, K., & Arnbjerg-Nielsen, K. (2012). Framework for economic pluvial flood risk assessment considering climate change effects and adaptation benefits. *Journal of Hydrology*, *414–415*, 539–549. https://doi.org/10.1016/j.jhydrol.2011.11.031

LIST OF ACRONYMS

1D	One Dimensional
2D	Two Dimensional
AFMA	Anticipatory Flood Management
ASDCO	Alexandria Sanitary and Drainage Company
BDF	Bayesian Decision Framework
BS	Brier Score
CRSME	Centred Root Mean Square Error
CRPS	Continuous Ranked Probability Score
CSI	Critical Success Index
DEM	Digital Elevation Model
DSS	Decision Support Systems
ECMWF	European Centre for Medium-Range Weather Forecasts
EPS	Ensemble Prediction Systems
EWS	Early Warning Systems
FAR	False alarm rate
FSS	Fraction Skill Score
GEFS	Global Ensemble Forecast System
GFS	Global Forecast System
GPM	Global Precipitation Measurement Mission
IMERG	Integrated Multi-satellitE Retrievals for GPM

IFRC	International Federation Red Cross and Red Crescent Societies
JAXA	Japan Aerospace Exploration Agency
KGE	Kling-Gupta Efficiency
MSWEP	Multi-Source Weighted-Ensemble Precipitation MSWEP
NCEP	National Center for Environmental Prediction's
NRMSE	Normalised Root Mean Square Error
NWP	Numerical Weather Prediction models
PDF	Probability Density Function
POD	Probability of Detection
QPF	Quantitative Precipitation Forecast
SPP	Satellite Precipitation Products
TRMM	Tropical Rainfall Measuring Mission
TIGGE	THORPEX International Grand Global Ensemble
WMO	World Meteorological Organisation
WPS	WRF Preprocessing System
WRF	Weather Research and Forecasting

LIST OF TABLES

LIST OF FIGURES

APPENDICES

Appendix A

Leadtime and Neighbourhood Size Analysis

Area Average Scatter Plot

Scatterplots of the 10km domain resolution (D1) forecasted and MSWEP or IMERG area average rainfall for the event on Dec 5th, 2018 and Nov 20th 2020, at the 72h, 48h, 24h and 12h lead time and different neighbourhood sizes (1000km, 200km, 100km, 50km, 20km and 10km) for R1, R2 and R3. Similarly, with the Oct 25th event, the results indicate the agreement between forecasts and observed rainfall varies with the cumulus configuration, neighbourhood scales and lead time. In general, the forecasts perform best at the 24h LT, especially for neighbourhood scales up to 50km but at smaller scales, this lead time shows poorer skill. For the other lead times, the forecasts either under forecasts or over forecasts. In particular, the results show that the forecasts did not improve at the 12h LT, but this varies with the event.

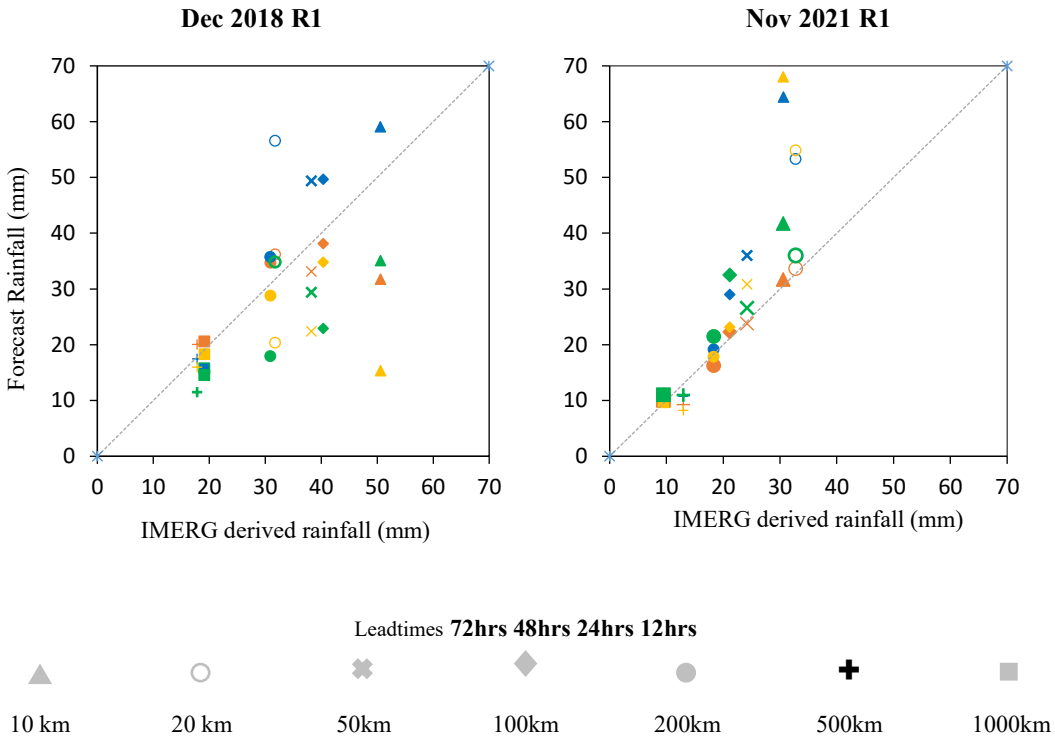


Figure A1 Comparison of 24h rainfall accumulation from Dec5th 9:00 to Dec 6th 9:00 (left) and Nov 19th 21:00 to 20st Nov 21:00 (right) showing the accumulated area averaged rainfall for the 10km resolution (D1) for different neighbourhood sizes (10, 20,50, 100, 200, 500, 1000km) at lead times of 12, 24,48 and 72 hours for domain R1

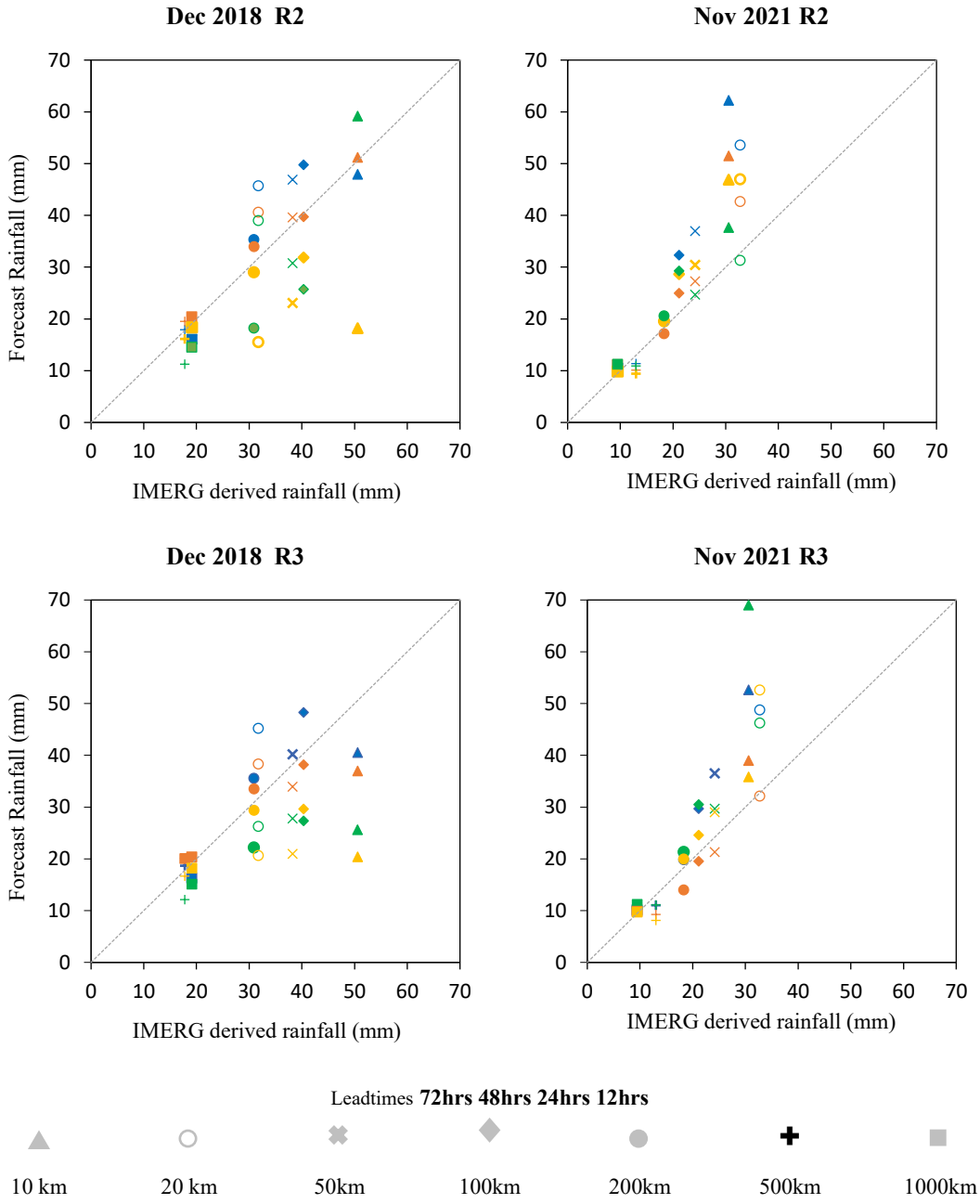
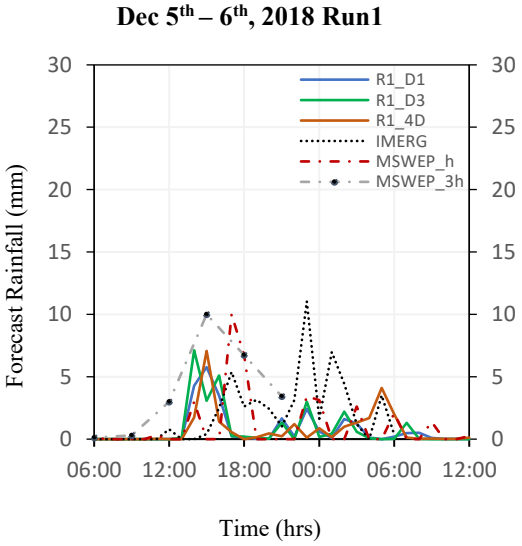
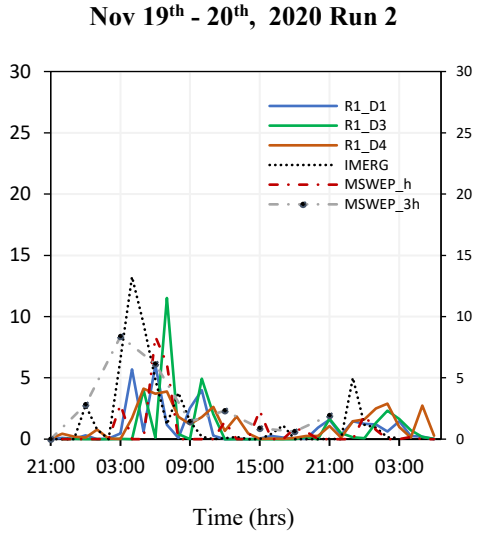


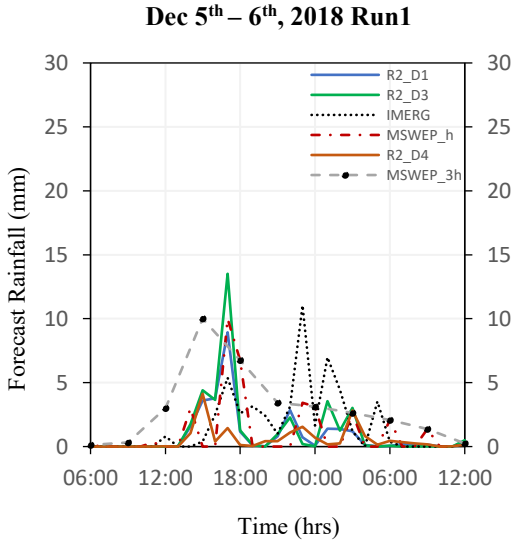
Figure A2 Comparison of 24h rainfall accumulation from Dec5th 9:00 to Dec 6th 9:00 (left) and Nov 19th 21:00 to 20st Nov 21:00 (right) showing the accumulated area averaged rainfall for the 10km resolution (D1) for different neighbourhoods at lead times of 12, 24,48 and 72 hours for R2 and R3

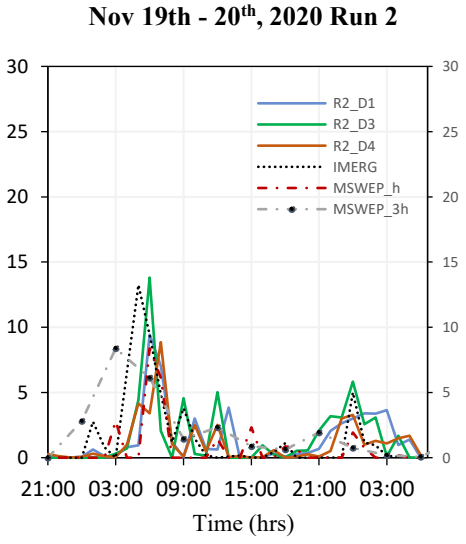
Rainfall Temporal Distribution, Discharge and Node flooding

The temporal distribution varied within 10km2, corresponding to the area used in the MIKE + model and immediate surrounding areas. This gives insight into the way node flooding might vary with intensity. For the Dec 5th, 2018 event the 0.37km (D4) runs produced the lowest rainfall intensities when compared to the IMERG rainfall and MSWEP, the worst-case scenario which assumes the 3h accumulation fell in the last hour of a 3h accumulation and the first hour of the subsequent 3h accumulation. Similarly for the Nov 20th event, the 0.37km(D4) run gave the lowest intensity except for R3, which showed a very high peak.









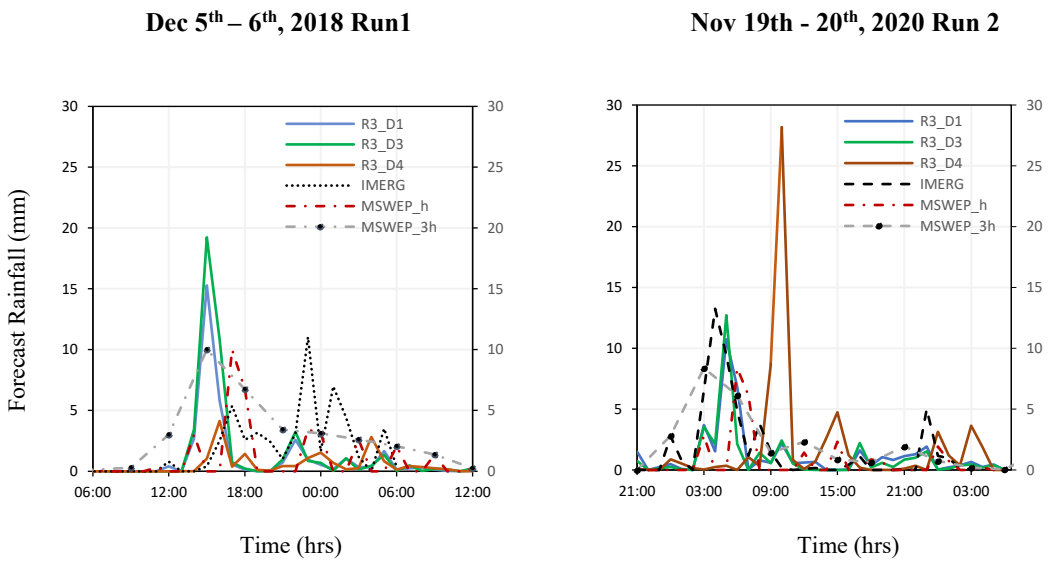


Figure A2 Rainfall time series for all runs D1(10km),D3(1km) and D4 (0.37km) from 6:00 Dec 5thth to 12:00 Dec 6th (above) and 19th Nov 21:00 to Nov21st 6:00 2020 at a known flood location. IMERG and MSWEP are shown for 3h accumulations and an extreme case 1h accumulations. R1, and R3 correspond to runs with 3 domains and R1_4D, R2_D4 and R3_D4 correspond to 4 domains runs. Cu=3 represents the Grell Freitas Cu scheme and 0 indicates no cumulus scheme in the respective domains. The left axis mm/h and the right axis show mm/3h corresponding to MSWEP_3h

Rainfall Threshold Analysis

Hazards were classified as “No to minimal flooding”, “Minor Flooding”, “Significant flooding” and “Severe flooding” based on 24h accumulated rainfall (Table 2 in the main document). Analysis of thresholds showed that results varied with the treatment of convection for the different runs. Classifications are shown for different districts in Alexandria for IMERG for the Dec 2018 and Nov 2020 event. All runs were able to detect areas of extreme rainfall in agreement with the observations but the smaller domains detected more extreme areas as shown in the Figure below.

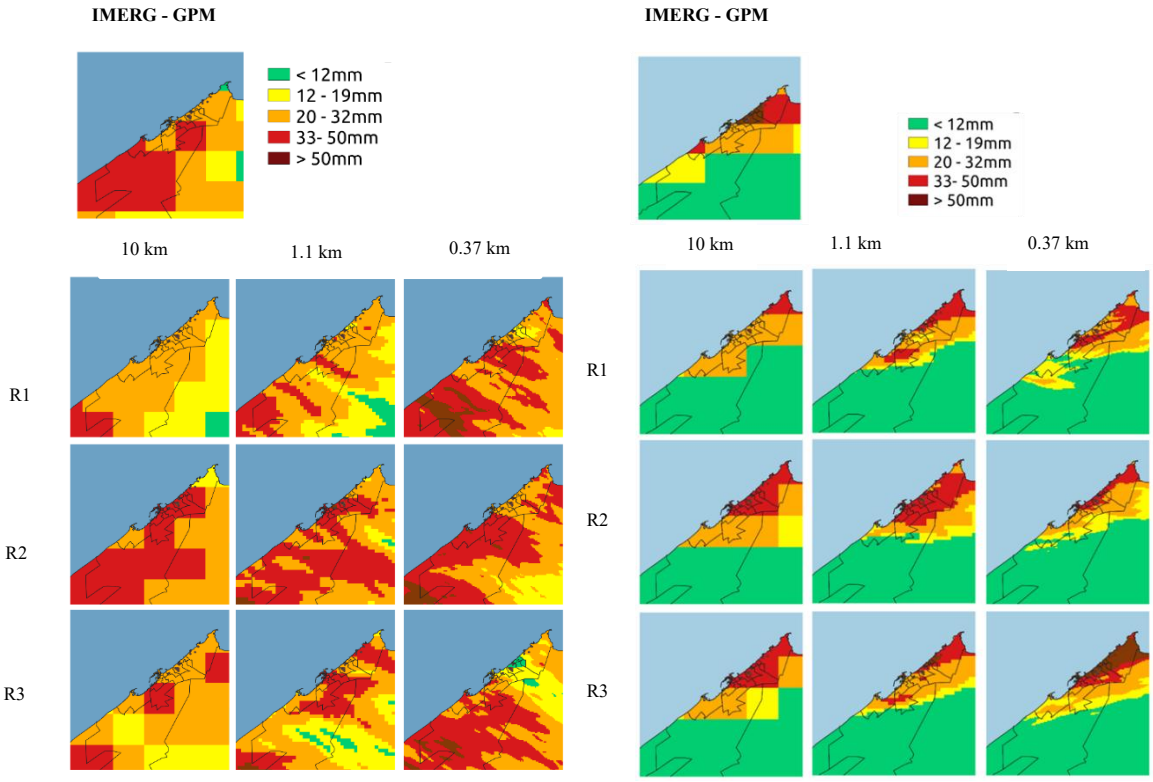
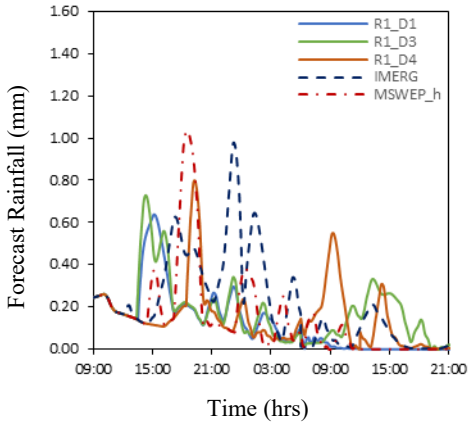


Figure A3 Forecast of hazard classes with a lead time of 24 h for the Dec 5th, 2018 (left) and Nov 20th,2020 (right) flood event for different runs based on the 24h rainfall accumulations. Left:(10 km) and Middle 1.1km grids for run downscaled to 3 domains. Right: 0.37km grids downscaled to 4 domains

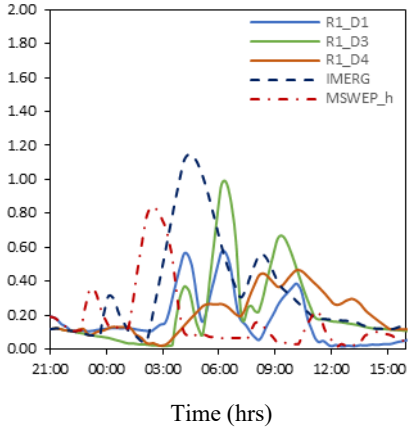
Discharge and Node flooding outlet

Discharge was simulated at the outlet for all runs for the Dec 5th 2018 and Nov 20th 2020 and compared with the simulated results of both the MSWEP and IMERG for the 10km (D1), 1.1km(D3) and 0.37km (D4). Peak discharges for observed MSWEP and IMERG has a similar range ~1 m^3/s. In general, R2 showed a better match of the peak discharge values for both events with a peak of about 1.2m^3/s however the peak is prematurely predicted or delayed for the Dec 2018 and Nov 2020. In the case of R3 Dec 2018, the 1.1km resolutions over estimated the discharge with 1.5m^3/s). This was also reflected in the poor NRMSE and KGE values. Similarly, for the Nov 20th event, the R3_D4 over forecasted the peak discharge. The capacity is 0.9m^3/s so it is assumed flooding occurs once this discharge is exceeded. The simulated discharge for R1 forecasts showed a different pattern with two lower peaks.

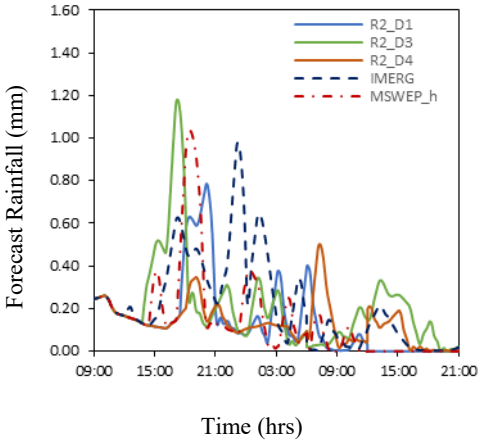
Dec 5th – 6th, 2018 Run1



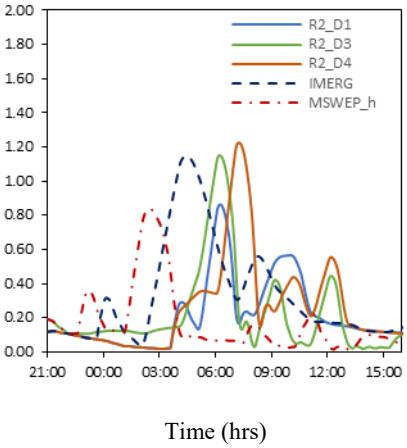
Nov 19th - 20th, 2020 Run 1



Dec 5th – 6th, 2018 Run 2



Nov 19th - 20th, 2020 Run 2



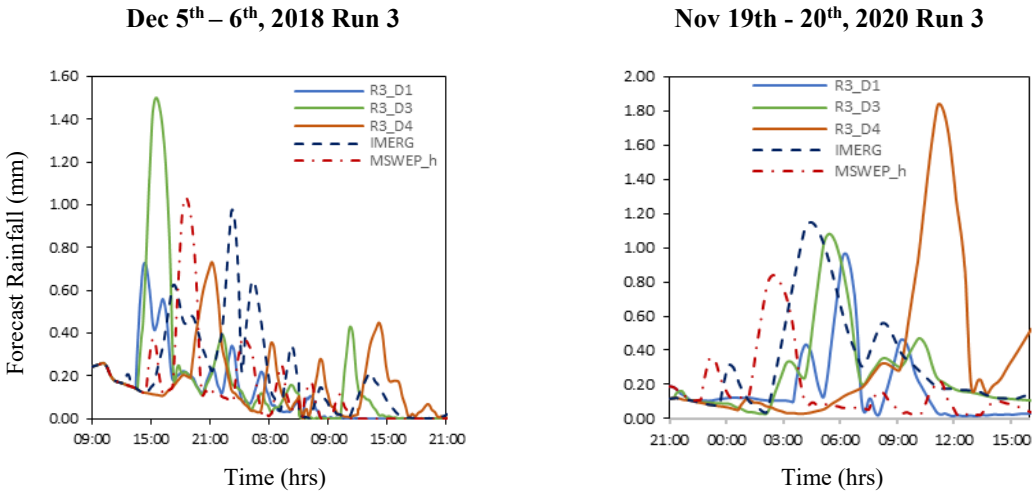


Figure A4 Simulated Discharge at outlet #1 on Dec 5th- Dec 6th 2018 and Nov 19th -21st, 2020 for R1, R2 and R3 for D1 (10km), D3 (1.1km) and (D4) 0,37km which corresponds to runs with 4 domain and the cumulus scheme is turned off in the 4th domain

Node Flooding

The discharge was simulated for the entire system for El Gomork. The following results present the node following at a known flood location in the city

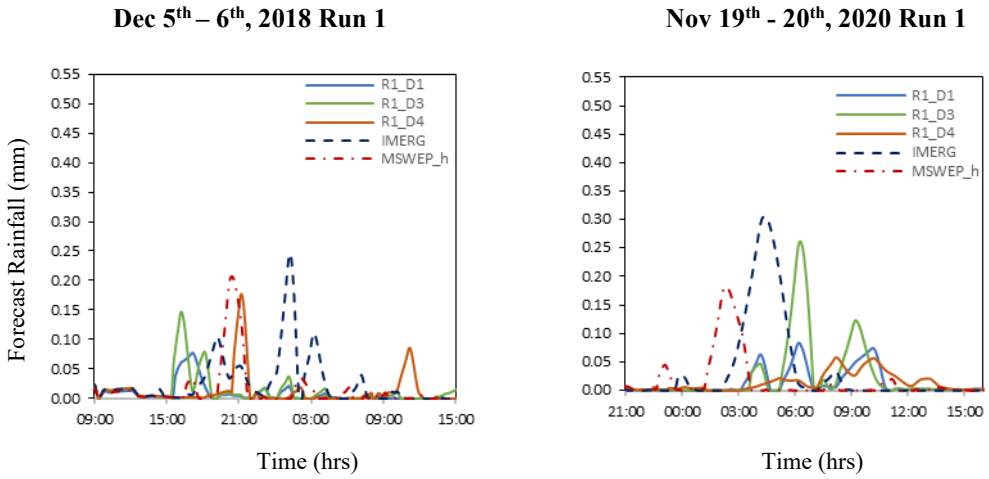


Figure A5 Node flooding depths (m) at known flood location in El Gomork on from Dec 5th 9:00 – Dec 6th 15:00, 2018 (above) and Nov 19th 21:00 to Nov 20th 15:00 2020 (below). For run 1 at D1(10km), D3(1.1.km). and 0.37km- 4 domains (orange). These events were evaluated again IMERG simulated runs (blue dashed line)

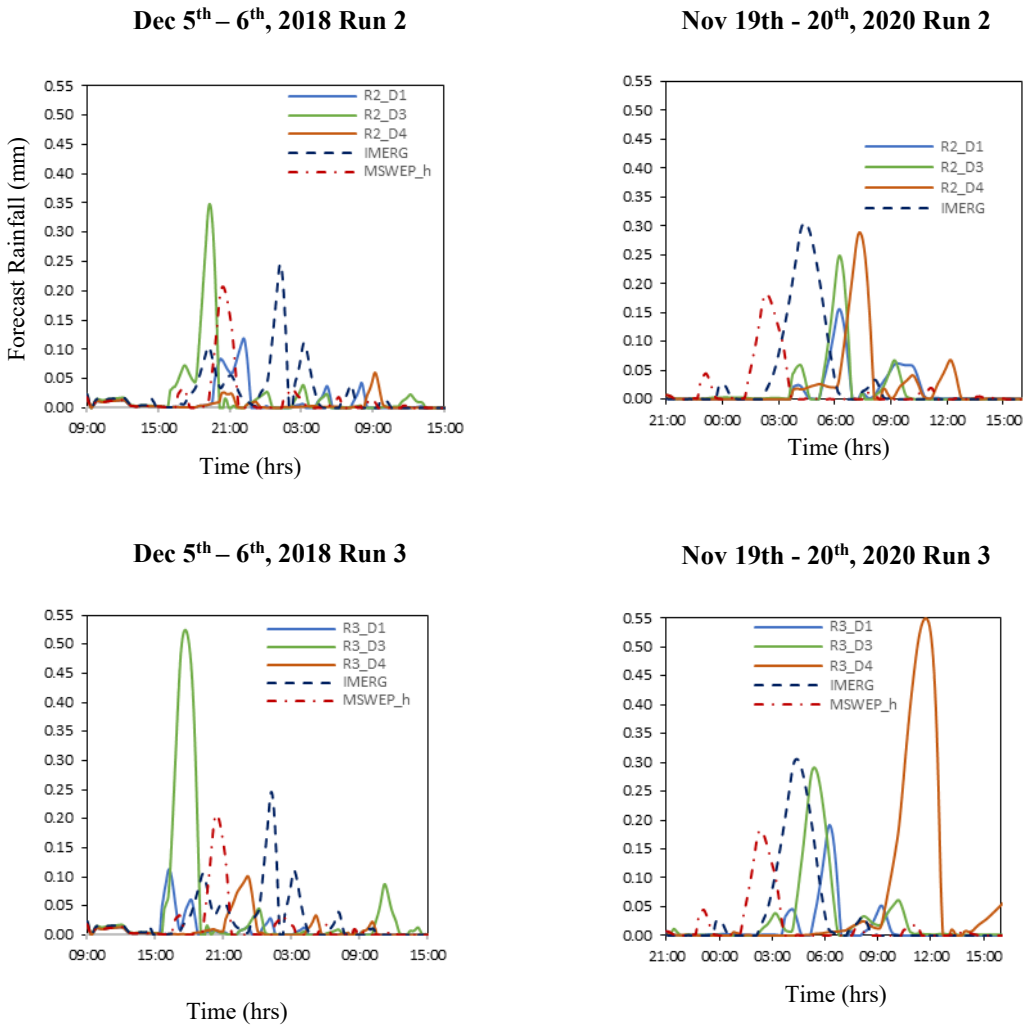
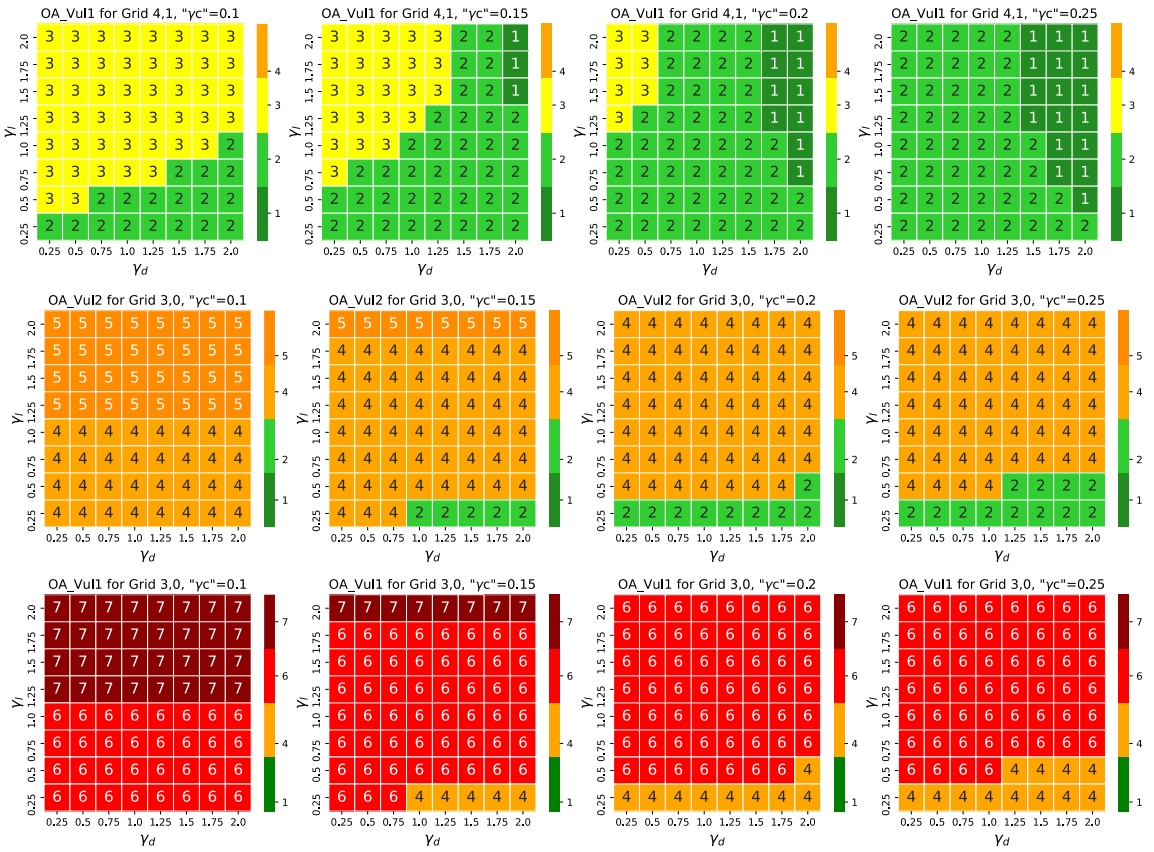


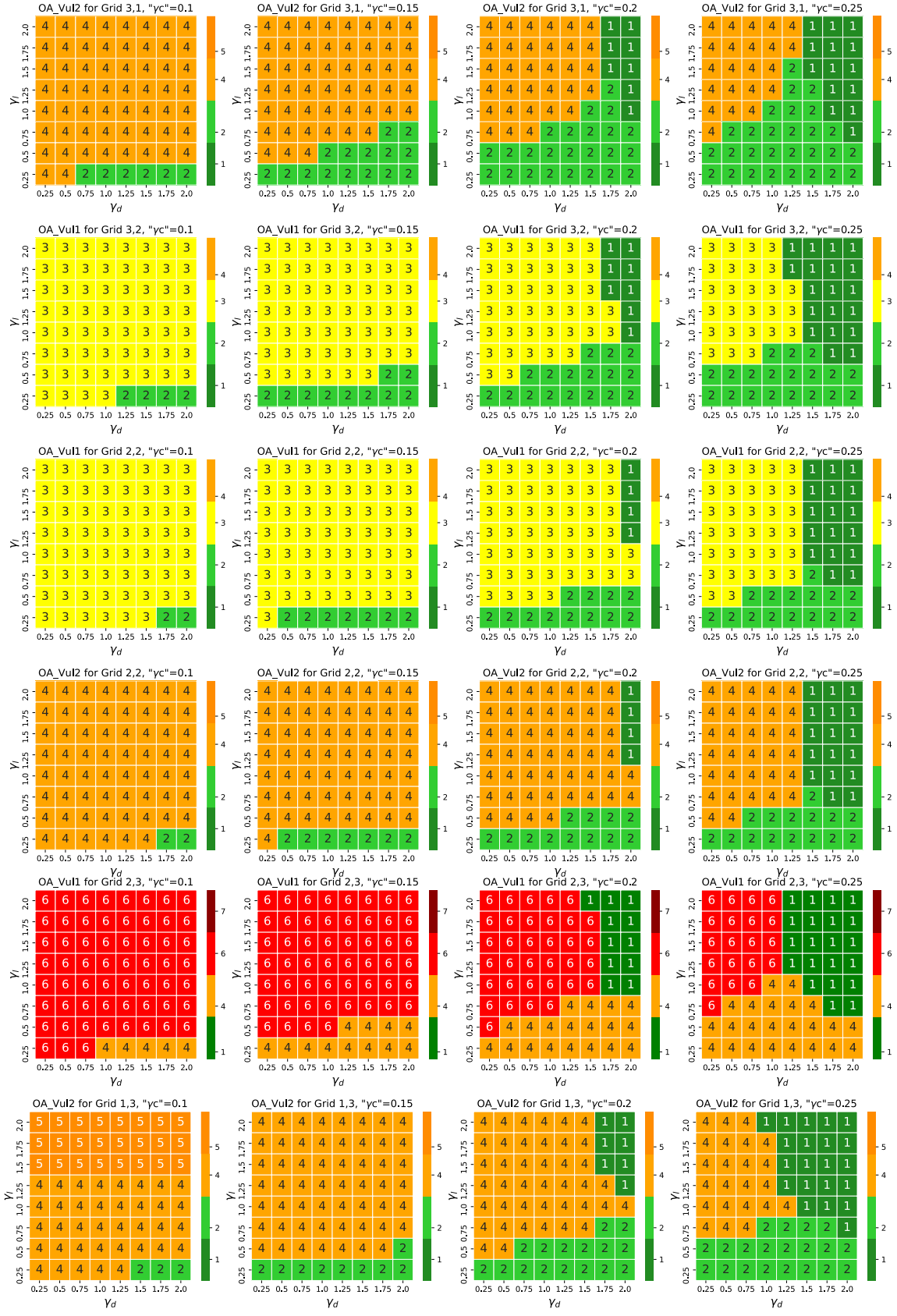
Figure A6 Node flooding depths (m) at known flood location in El Gomork on from Dec 5th 9:00 – Dec 6th 15:00, 2018 (above) and Nov 19th 21:00 to Nov 20th 15:00 2020 (below). For run 2 and 3 at D1(10km), D3(1.1.km). and 0.37km- 4 domains (orange). These events were evaluated again IMERG simulated runs (blue dashed line)

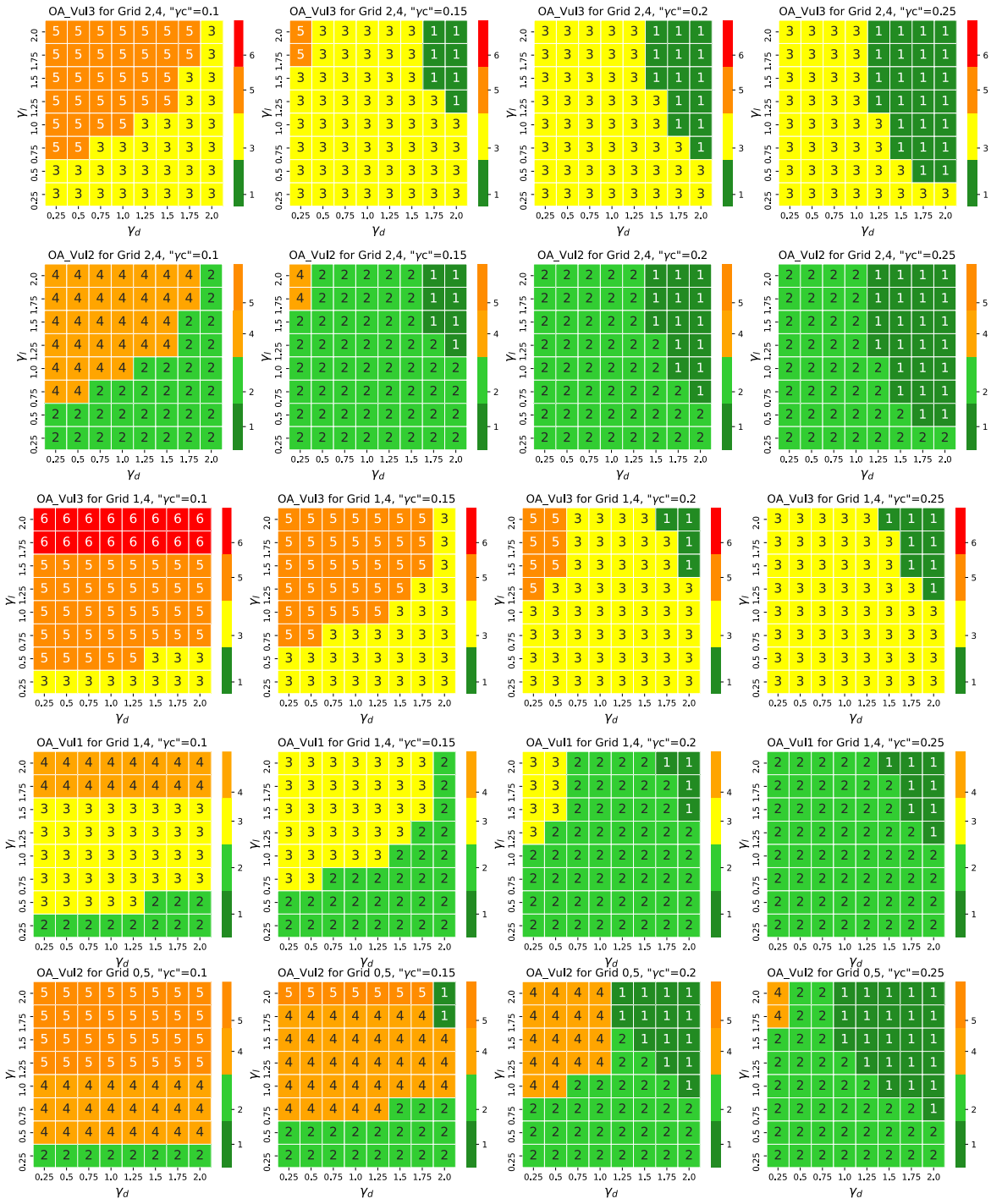
Appendix B

Sensitivity analysis of loss parameters, including the influence of vulnerability weights

Using the Oct 25th, 2015 event, a sensitivity analysis of Optimal Actions (OA) under varying values of the parametric loss function's parameters: γ_c; C/L (sensitivity to false alarms), γ_l (effectiveness of protective actions) and γ_d (sensitivity of damage to hazard severity). The analysis shows how changes in the parameters **(**Table 6.6) combined with how the different posterior beliefs affect the selection of optimal actions across 10 rainfall grids at the 24hr lead time and respective vulnerability classes for the Oct 25th event (See Table 6.7 for the rainfall grid and corresponding vulnerability class). Assumes the maximum damage varies with neighbourhood vulnerability class. The loss function parameters are used to calibrate the decision model using the Oct 25th, 2015 event. Changes in γ_c, γ_l and γ_d were associated with transitions between inaction, moderate and higher actions. As γ_c increases, the results shift towards lower-intensity actions to minimise the cost of false alarms. This trend is especially clear under low-hazard posterior beliefs. Inaction becomes optimal when false alarm costs dominate. Alternatively, as γl increases, there are higher penalties for missed alarms associated with more severe events, which generally lead to more protective actions. For increasing values of γ_l, the influence of this parameter appears to plateau, reflecting diminishing returns.




OA_Vul2 for Grid 3,1, "γc"=0.1
OA_Vul2 for Grid 3,1, "γc"=0.15
OA_Vul2 for Grid 3,1, "γc"=0.2
OA_Vul2 for Grid 3,1, "γc"=0.25
OA_Vul1 for Grid 3,2, "γc"=0.1
OA_Vul1 for Grid 3,2, "γc"=0.15
OA_Vul1 for Grid 3,2, "γc"=0.2
OA_Vul1 for Grid 3,2, "γc"=0.25
OA_Vul1 for Grid 2,2, "γc"=0.1
OA_Vul1 for Grid 2,2, "γc"=0.15
OA_Vul1 for Grid 2,2, "γc"=0.2
OA_Vul1 for Grid 2,2, "γc"=0.25
OA_Vul2 for Grid 2,2, "γc"=0.1
OA_Vul2 for Grid 2,2, "γc"=0.15
OA_Vul2 for Grid 2,2, "γc"=0.2
OA_Vul2 for Grid 2,2, "γc"=0.25
OA_Vul1 for Grid 2,3, "γc"=0.1
OA_Vul1 for Grid 2,3, "γc"=0.15
OA_Vul1 for Grid 2,3, "γc"=0.2
OA_Vul1 for Grid 2,3, "γc"=0.25
OA_Vul2 for Grid 1,3, "γc"=0.1
OA_Vul2 for Grid 1,3, "γc"=0.15
OA_Vul2 for Grid 1,3, "γc"=0.2
OA_Vul2 for Grid 1,3, "γc"=0.25
γl
γd


OA_Vul3 for Grid 2,4, "γc"=0.1
OA_Vul3 for Grid 2,4, "γc"=0.15
OA_Vul3 for Grid 2,4, "γc"=0.2
OA_Vul3 for Grid 2,4, "γc"=0.25
OA_Vul2 for Grid 2,4, "γc"=0.1
OA_Vul2 for Grid 2,4, "γc"=0.15
OA_Vul2 for Grid 2,4, "γc"=0.2
OA_Vul2 for Grid 2,4, "γc"=0.25
OA_Vul3 for Grid 1,4, "γc"=0.1
OA_Vul3 for Grid 1,4, "γc"=0.15
OA_Vul3 for Grid 1,4, "γc"=0.2
OA_Vul3 for Grid 1,4, "γc"=0.25
OA_Vul1 for Grid 1,4, "γc"=0.1
OA_Vul1 for Grid 1,4, "γc"=0.15
OA_Vul1 for Grid 1,4, "γc"=0.2
OA_Vul1 for Grid 1,4, "γc"=0.25
OA_Vul2 for Grid 0,5, "γc"=0.1
OA_Vul2 for Grid 0,5, "γc"=0.15
OA_Vul2 for Grid 0,5, "γc"=0.2
OA_Vul2 for Grid 0,5, "γc"=0.25
γd
γl

ABOUT THE AUTHOR

Adele is from Tobago, the smaller of the two islands that make up the nation of Trinidad and Tobago. She obtained a BSc in Civil Engineering and an MSc in Civil and Environmental Engineering from the University of the West Indies, St. Augustine, in Trinidad. Adele practised for eight years as a civil engineer on several flood hydrological studies and drainage infrastructure design and construction management projects in Trinidad and Tobago before being awarded a Small Island Developing States (SIDS) Fellowship to study in the Netherlands.

In 2018, Adele graduated with an MSc in Water Science and Engineering, specialising in Hydraulic Engineering and River Basin Development, from IHE Delft. Adele decided to take a leap of faith to pursue a PhD (after saying she would never do a PhD). Her research focused on Pluvial Flood Forecasting and Decision Support for Anticipatory Flood Management in Alexandria, Egypt, with the Coastal and Urban Risk Resilience Group. At IHE Delft, she was also a member of the Small Island Research Group and served as a member and Chair of the PhD Association Board (PAB) from 2020 to 2022.

During her MSc., Adele became interested in youth-driven early warning systems and Disaster Risk Reduction initiatives which continued into her PhD. As an active member of the Water Youth Network's Disaster Risk Reduction (DRR) thematic group and a co-coordinator of the EWS Young Professionals group, she has coordinated and hosted several capacity-building and young professional networking sessions promoting the involvement and perspectives of youth and young professionals at the Multi-Hazard Early Warnings System International Conference(MHEWS-IC), European Geo-Science Union (EGU), Understanding Risk and COP 26, UN-Water 2023 Conference, among others.

In the latter part of her PhD, Adele worked part-time as a consultant with the Red Cross Red Crescent Climate Centre on the Resilience and Preparedness to Tropical Cyclones in Southern Africa (REPRESA) project and contributed to a flagship report on Future Fit Anticipatory Action.

Having worked in both research and practice, Adele is looking forward to using her new expertise, along with her past experience, to bridge the science policy practice gap.

Journal publications related to this thesis

Young, A., Bhattacharya, B., Daniëls, E. and Zevenbergen, C. 'Incorporating Forecast Uncertainty into Anticipatory Flood Management using a Bayesian Decision Support Framework (under review).

Young, A., Bhattacharya, B., Daniëls, E and Zevenbergen, C. (2025) 'Integrating WRF forecasts at different scales for pluvial flood forecasting using a rainfall threshold approach and a real-time flood model'. Journal of Hydrology. https://doi.org/10.1016/j.jhydrol.2025.132891

Young, A., Bhattacharya, B. and Zevenbergen, C. (2021) 'A rainfall threshold-based approach to early warnings in urban data-scarce regions: A case study of pluvial flooding in Alexandria, Egypt', Journal of Flood Risk Management, (February), pp. 1–16. doi: 10.1111/jfr3.12702.

Other publications

Jaime, C., Scholz, C., Young, A., Coughlan de Perez, E., & van Aalst, M. (2025). Weathering Conflict: Impacts and Solutions for Protecting Hydrometeorological Infrastructure during Armed Conflict. Weather, Climate and Society (published online ahead of print 2025). https://doi.org/10.1175/WCAS-D-25-0073.1

Jawad, M., Bhattacharya, B., Young, A. and van Andel, S. J. (2024) 'Evaluation of Near Real-Time Global Precipitation Measurement (GPM) Precipitation Products for Hydrological Modelling and Flood Inundation Mapping of Sparsely Gauged Large Transboundary Basins- A Case Study of the Brahmaputra Basin', *Remote Sensing*, 16(10), p. 1756. doi: 10.3390/rs16101756.

Conference presentations

Young, A., Bhattacharya, B., Daniels, E and Zevenbergen, C.(2025). Improving Local Scale Pluvial Flood Anticipatory Actions in Alexandria City, Egypt. Cairo Waterweek, Egypt, 12th -16th Oct, 2025

Young, A., Bhattacharya, B., Daniels, E and Zevenbergen, C.(2025). Probabilistic Decision Support for Anticipatory Flood Actions in Alexandria City, Egypt, 41st IAHR World Congress, Singapore, 22–27 Apr 2025

Young, A., Bhattacharya, B., Mahood, F., Daniels, E. and Zevenbergen, C.: High-Resolution Ensemble Precipitation for Pluvial Flood Forecasting in the Urban Data-scarce city of Alexandria, Egypt, EGU General Assembly 2022, Vienna, Austria, 23–27 May 2022, EGU22-2596, https://doi.org/10.5194/egusphere-egu22-2596, 2022

Huthoff, F., Young, A., Cortes Arevalo, J., Hagedooren, H. and Rudolph, M.: Community-centred Disaster Risk Reduction: Experiences from “Our Flood Mural” in Beira, Mozambique, EGU General Assembly 2022, Vienna, Austria, 23–27 May 2022, EGU22-12476, https://doi.org/10.5194/egusphere-egu22-12476, 2022.

Young, A., Mahood, F., Bhattacharya, B., Daniels, E. and Zevenbergen, C. (2021),High-resolution Precipitation for Pluvial Flood Forecasting in the urban data-scarce city of Alexandria, Egypt-H45Z , presented online at 2021 Fall Meeting, AGU, New Orleans ,LA 13-17 Dec https://agu.confex.com/agu/fm21/meetingapp.cgi/Paper/872858

Young, A., Bhattacharya, B. and Zevenbergen, C. (2020) ‘Pluvial flood forecasting in urban data-scarce regions: Influence of rainfall spatio-temporal data (in)accuracy on decision-making’, in EGU General Assembly Conference Abstracts. (EGU General Assembly Conference Abstracts), p. 9492. https://doi.org/10.5194/egusphere-egu2020-9492

Netherlands Research School for the
Socio-Economic and Natural Sciences of the Environment

D I P L O M A

for specialised PhD training

The Netherlands research school for the
Socio-Economic and Natural Sciences of the Environment
(SENSE) declares that

Adele Therese Young

born on 21 December 1985 in Scarborough, Trinidad and Tobago,

has successfully fulfilled all requirements of the
educational PhD programme of SENSE.

Delft, 23-04-2026

SENSE coordinator PhD education

Dr Ir Peter Vermeulen

The SENSE Director

Dr Jampel Dell'Angelo

The SENSE Research School declares that **Adele Therese Young** has successfully fulfilled all requirements of the educational PhD programme of SENSE with a work load of 54.2 EC, including the following activities:

SENSE PhD Courses

- Environmental research in context (2019)
- Research in context activity: 'Part of the PhD ITeam, organised Nerd Nite and organised a workshop on writing successful research and project proposals (2019-2022)

Selection of Other PhD and Advanced MSc Courses

- The Art of Presenting Science, TU Delft (2019)
- Analytic Storytelling (2019) and Popular Writing (2021), TU Delft
- Foundations of teaching, learning and assessment, TU Delft (2020)
- Time Management - first things first, TU Delft (2019)
- Data visualisation using R, TU Delft (2020)
- Decision support systems in the water domain, IHE Delft (2019)
- Python Workshop, IHE Delft (2020)
- Exploring a research career outside academia, TU Delft (2023)

External training at a foreign research institute

- Forecast Verification, University of Reading and Imperial College London (2021)
- Urban Design Climate Change, Gustave Eiffel University (2022)
- Synergizing Disaster Risk Reduction and Climate Change Adaptation, United Nations System Staff College, Online (2023)

Management and Didactic Skills Training

- Chair of PhD board (2021-2022), member from 2020-2022
- Co-developed and organised short course an "Interdisciplinary approach to Forecasting Early Warnings systems: a roleplaying game" at EGU on '29th April 2021'
- Supervising three MSc students with thesis (2019-2022)
- Teaching in the MSc course 'River Flood Analysis and Modelling', MSc course 'Flood Risk Management'

Selection of Oral Presentations

- *Flood forecasting in urban data-scarce regions: Influence of rainfall spatiotemporal (in) accuracy on decision-making,* IHE Delft PhD Symposium, Oct 2019, Netherlands
- *Crowdsourcing for improving pluvial flood forecast and decision-making in the Caribbean*. Caribbean Water and Waste Water Association Annual Conference, Oct 2020, Virtual /Trinidad and Tobago
- *High-Resolution Precipitation for Pluvial Flood Forecasting in the urban data scarce city of Alexandria, Egypt*. AGU General Assembly, April 2021, Virtual /US
- *A Bayesian decision framework to support flood anticipatory actions in the urban data-scarce city of Alexandria, Egypt*. EGU General Assembly, 2023, Vienna, Austria

IHE DELFT
Institute for Water Education
under the auspices of UNESCO

TU Delft Delft University of Technology

Pluvial flooding is increasing in cities due to the combined effects of climate change, rapid urbanisation, and inadequate drainage infrastructure. Anticipatory Flood Management is proposed as a “low-regret” measure to reduce residual risk and improve preparedness through flood forecasting and early action using existing capacities rather than heavy capital investments. Yet many cities, particularly in data-scarce regions, lack the capacity to produce reliable rainfall forecasts and well-calibrated flood predictions, resulting in cascading uncertainties that hinder decision-making and subsequent action. While previous research has focused on improving model accuracy, less attention has been given to the complexity of the forecast–decision chain. This thesis addresses that gap by utilising diverse data sources, high-resolution deterministic and ensemble rainfall forecasts, different flood forecasting approaches and a Bayesian Decision Framework. Focusing on Alexandria, Egypt, the findings demonstrate that there are limits to forecasting and decision-making due to data availability, urban rainfall forecast scale requirements, and model approaches. Despite limitations, the results also suggest that fit-for-purpose methods and probabilistic decision frameworks can produce actionable information and serve as valuable decision-support tools when end-user requirements and decision-makers’ risk attitudes are considered. Rather than seeking perfect accuracy, the research emphasises uncertainty-aware decision-making to support effective and equitable anticipatory flood action.

This book is printed on paper from sustainably managed forests and controlled sources

9 789073 445796