Optimizing SAR-based Flood Extent Assimilation for Improved Hydraulic Flood Inundation Forecasts

A thesis submitted in partial fulfilment of
the requirements for the degree of

Doctor of Philosophy

by

Antara Dasgupta

Under the Supervision of

Dr. RAAJ Ramsankaran

Dr. Jeffrey Phillip Walker

The course of study for this award was jointly developed by
Monash University, Australia and the Indian Institute of Technology Bombay, India,
and administrated by The IITB-Monash Research Academy

(Year 2020)

Approval Sheet

The thesis entitled “Optimizing SAR-based Flood Extent Assimilation for Improved Hydraulic Flood Inundation Forecasts” by Antara Dasgupta is approved for the degree of **Doctor of Philosophy.**

Prof. D. Nagesh Kumar
(IISc Bangalore)
External Examiner

Prof. Arpita Mondal
Internal Examiner

Prof. RAAJ Ramsankaran
IITB Supervisor

Prof. Jeffrey P. Walker
Monash Supervisor

Prof. G. N. Jadhav
Chairman

Date: 30th April 2020
Place: New Delhi, India

Declaration

I hereby certify that the work embodied in this thesis is the result of original research and has not been submitted for a higher degree to any other University or Institution, and to the best of my knowledge this thesis does not contain any material previously published or written by another person, except where due reference is made in the text. I also confirm that all principles of academic honesty and integrity have been adhered to, and no findings/ideas/data/facts/ have been misrepresented or fabricated or falsified. I understand that any violation of the above can lead to disciplinary action by the Institute and evoke penal action from the sources which have not been properly cited or from whom proper permission has not been pre-emptively acquired.

(Signature)

Antara Dasgupta

154044001

Date: 20/4/2020

This page has intentionally been left blank.

To my parents,

Amita and Shankar Dasgupta,

for their unwavering support

and many sacrifices.

This page has intentionally been left blank.

Acknowledgements

Every time I think about my PhD the following quote from Nelson Mandela's autobiography comes to mind,

"…after climbing a great hill, one only finds that there are many more hills to climb."

Through this long and arduous journey I faltered many times but each time that I did, the support of my family, friends, and colleagues, reminded me how far I had come.

Here I want to thank each person who made this PhD a once in a lifetime experience. The list included here of individuals and their contributions to this PhD is by no means exhaustive, and I want to apologise in advance for missing any names or experiences that I should have mentioned.

First, I must gratefully acknowledge the IITB-Monash Research Academy which funded my research for four years and gave me an opportunity to work with the best schools in India and Australia. Everyone at the Academy from the CEO to the staff, worked tirelessly to make my PhD experience a little less daunting. I would especially like to mention Nancy, Laya, Jayasree, Kiran, Rahul, Pinki, and Bharat, who were always there with a smile on their faces to help me and hug me on hard days.

I was extremely lucky to have found close friends and excellent mentors in both my supervisors, Dr. RAAJ Ramsankaran and Dr. Jeffrey Walker. Over the years I learnt significant lessons about life and a successful academic career from each of them. RAAJ sir is an excellent mentor who cares deeply about his students. What I admire most about him is his curiosity to learn and his ability to adapt and implement the best practices from all over the world. He gave us a truly global experience at IITB and I feel extremely honoured to join the ranks of the alumni of the Hydro-Remote Sensing Applications (H-RSA) Group. The colleagues at the HRSA lab have been less like friends and more like family for these four years, and their support enabled me to overcome many hurdles. Smarika, Swathy, Sangita, Anita, Shruti, Avani, Rajmita, Pratiksha, Prerna, Amol, Ashish, and Sathya have all provided intellectual inputs and emotional support at different times during this journey, for which I am thankful.

Jeff, has also been a wonderfully supportive and nurturing supervisor, mentor, and friend. His teaching philosophy was mostly "nudging me in the right direction", and I strongly believe this led to many of the things I was able to achieve during my PhD. His timely feedback at various stages in my research, helped not only to improve the quality of my research but also contributed immensely to my personal growth. I hope to always imbibe his scientific rigour, dedication, and passion for research and sincerely hope that he continues to support me throughout my career. I also received the support of my colleagues at Monash. Specifically, I am thankful to Simone, Jaya, Adrian, Sabah, Ling, Ying, Mayer, Mayu, Sarah, Gonzalo, Rukshan, and Radhika, for their support and fun coffee break conversations.

I am deeply indebted to, Dr. Stefania Grimaldi and Dr. Valentijn Pauwels, a part of my research team at Monash and my research progress committee, for their prompt, critical, and insightful comments. Stef was responsible for most of the input data collection and synthesis for this project. Thanks Stef for your detailed comments on my work, I have learnt a lot from your dedication. I am also extremely grateful to my research progress committee at IITB, Dr. E. P. Rao, Dr. Y. S. Rao, and Dr. Arpita Mondal, who always encouraged me and gave me constructive feedback during our interactions.

I would also like to sincerely thank our collaborator, Dr. Renaud Hostache, from the Luxembourg Institute of Science and Technology (LIST). Thanks Renaud for the many long discussions about assimilation and for sharing your extensive expertise with me. I am deeply indebted to him for sharing his flood extent assimilation code, which was the first step for me to start developing my own algorithm. It was his timely advice and willingness to share which contributed significantly to my progress, and for that I am eternally grateful.

This section cannot be complete without mentioning Dr. Guy Schumann from the University of Bristol and Remote Sensing Solutions Inc. Guy was my knight in shining armour and *Sensei*, who was always there for me through all my major and minor research successes and failures since the day we first met. Thanks Guy for spending hours discussing the hundreds of research ideas floating around in my mind at all times and for getting as excited about them as me. I can never thank you enough for taking me under your wing and making this gigantic uphill task of earning a PhD feel more achievable. Working with Renaud and Guy was a dream, and in fact it was their papers that inspired me to take on this particular research field. I am truly grateful for having the opportunity to work with them during my PhD.

I would also like to thank my friends like family at IITB, Gowtham, Sohan, and Prithvi, and elsewhere, Supriti, Kashi, Richi, Katie, Karina, Kim, Shiwang, and Ana. It is through the knowledge that they stand behind me as constant pillars of support, that I've been able to reach for many things I previously believed impossible.

I am grateful to my extended family for their constant love, support, and encouragement. Thank you Mama, Mami, Dadabhai, Piya di, Boudi, Auroni, Kuttuni, and Shona Mashi, your pride and excitement for my accomplishments is truly more rewarding than my own.

I feel immensely blessed for my partner and best friend, Harsh, who constantly supported me while nonchalantly dealing with the best and worst in me. Thank you for being my rock and picking me up every time I faltered. Loving and being loved by you, is one of the greatest privileges of my life.

Finally, I would like to thank my parents who stood by my side through all of it. Thanks mom for listening to the details of every model run, every code, and every graph I have prepared over the last four years, whether it bored you or not. Thank you dad for being my rock, for always encouraging me to continue bravely on the path I had chosen, and for teaching me to be patient with myself and my work. I would never have made it anywhere without your love and support.

One day during my first Master's degree in Dehradun it started raining. The rain continued for three days, flooded the entire valley, and killed thousands of people. Later during my second master's, I learnt that the first responders were supported through remote sensing based flood maps generated by the scientists at the institute. It was that day that I first dreamt of joining the ranks of the people who were directly contributing to the fight against floods. This thesis brings us one step closer to achieving the larger dream that I share with many of my colleagues from the field, to increase global flood resilience and reduce its fatal consequences. I would like to conclude with a quote from John Lennon's iconic song *Imagine*,

"You may say that I'm a dreamer, but I'm not the only one."

This page has intentionally been left blank.

Abstract

Climate change, urbanization, and increasing population density in global floodplains have made accurate forecasts of flood inundation vital to designing effective rescue, response, and resource allocation strategies during emergencies. Precipitation forecasts from numerical weather prediction models are typically propagated through hydrological models to generate streamflow at catchment outlets, which are then input to hydraulic models to determine inundation depth and flow. The uncertainties in precipitation forecasts are usually amplified through this modelling chain, and significantly reduce the reliability of the resulting inundation forecasts. Consequently, point hydrometric gauge observations of discharge and water level have traditionally been used to quantify flood risk and increase the reliability of flood inundation model predictions. However, the global decline in gauge networks and lack of spatial representativeness of point measurements have necessitated the exploration of alternative data sources to reduce flood inundation forecast uncertainties.

In this context, remote sensing has emerged as a valuable tool, providing capabilities for synoptic and systematic coverage of large inundated areas. Synthetic Aperture Radar (SAR) sensors are ubiquitous in flood monitoring applications, due to their all-weather/all-day imaging capabilities. SAR-derived flood extents can be converted to floodplain water levels through integration with digital elevation models (DEMs), and subsequently used to inform flood inundation forecasts in real-time using data assimilation. However, this water level derivation process is neither straightforward nor automatic and requires several simplifying assumptions. Moreover, due to SAR data uncertainties in urban and vegetated areas, reliable water level estimates can only be obtained at select shoreline locations. Recent studies have therefore recommended the direct assimilation of flood extents without converting to water levels but have struggled with (a) reducing and quantifying the uncertainties in SAR-based flood maps and (b) defining a cost function that appropriately compares the modelled and observed inundation. Therefore, this thesis first improved the estimation of probabilistic flood extents from SAR data using optimized texture-based methods, and then developed a novel cost function to sequentially assimilate them into a hydraulic model. Independently acquired crowdsourced water levels were used to quantitatively calibrate the hydraulic model, and the

sensitivity of the proposed flood extent assimilation strategy to observation spatiotemporal characteristics was also assessed.

In order to overcome the problem of uncertainties in SAR-based flood extents, a novel texture optimization method was introduced. Image texture or the spatial arrangement of SAR backscatter values, provides additional information about the observed features which can complement intensity information. However, the selection of an appropriate window size and orientation for texture calculations, as well as the subjectivity in choosing optimal application specific textures has impeded widespread application. The proposed method first estimated the appropriate window size through variogram analysis, then second-order uncorrelated statistical textures were estimated and optimized using a dimensionality reduction technique. Finally, an Adaptive Neuro-Fuzzy Information System (ANFIS) was used to classify the texture-enhanced SAR images, which provides the advantage of using neural networks to optimize the backscatter distribution parameters, while a fuzzy expert-based system can be used to define the form of the distributions. The proposed approach was validated using independent aerial photographs, achieving up to 54% improvements in classification accuracy in some areas.

A particle filter based assimilation framework was then implemented with a novel mutual information based cost function to assimilate the fuzzy flood extents, which measures the amount of information contained in one random variable about another. The performance of the proposed assimilation algorithm was evaluated through synthetic experiments and real data. Results revealed the suitability of the cost function to effectively assimilate flood extents, resulting in improved quality of inundation forecasts for lead times of up to a week. The sensitivity of the proposed flood extent assimilation framework to observation spatiotemporal characteristics was also synthetically evaluated. Specifically, questions pertaining to the optimal location, timing, and frequency of satellite images to obtain maximum improvements from model data integration were answered in this section. Observation timing with respect to the flood wave arrival time, emerged as a crucial factor determining the success of the assimilation. In fact, a single image acquired shortly after the flood peak assimilated at an optimal location in the catchment, resulted in forecast improvements comparable to the assimilation of multiple images (±4%). Moreover, this thesis also revealed that the possibility of forecast degradation if early rising limb observation were assimilated or if the correlation between subsequent observations was overestimated.

This research demonstrated that the assimilation of SAR-based flood extents can be optimized to maximize flood inundation forecast improvements and can successfully mitigate streamflow forecast uncertainties. The increased reliability of inundation estimates could pave the way for their inclusion in global flood forecasting systems, potentially leading to predictions relevant at catchment scales which can significantly augment flood resilience.

This page has intentionally been left blank.

Table of Contents

PART IV SUMMARY AND PERSPECTIVES

List of Acronyms and Abbreviations

AHD	Australian Height Datum
BSS	Brier Skill Score
CSK	Cosmo Skymed
CS	Crowdsourced
CSI	Critical Success Index
DA	Data Assimilation
DEM	Digital Elevation Model
DLR	Deutsches Zentrum für Luft- und Raumfahrt (German Space Agency)
EnKF	Ensemble Kalman Filter
EnSRF	Ensemble Square Root Filter
EO	Earth Observation
ETKF	Ensemble Transform Kalman Filter
ERS	European Remote Sensing
ESA	European Space Agency
FP	Floodplain
GFDS	Global Flood Detection System
GPS	Global Positioning System

HAND	Height Above Nearest Drainage
HD	Hydrodynamic
HEC-RAS	Hydrological Engineering Corps – River Analysis System
HH	Transmitted and received waves are horizontally polarised
HV	Transmitted wave is horizontally polarised and received wave is vertically polarised
IRS	Indian Remote Sensing
KF	Kalman Filter
KGE	Kling Gupta Efficiency
LETKF	Localised Ensemble Transform Kalman Filter
MODIS	Moderate Resolution Imaging Spectrometer
MSS	Multi Spectral Scanner
NASA	National Aeronautical and Space Administration
NDVI	Normalised Difference Vegetation Index
NDWI	Normalised Difference Water Index
MNDWI	Modified Normalised Difference Water Index
PF	Particle Filter
RADAR	Radio Detection and Ranging
RMSE	Root Mean Squared Error
RS	Remote Sensing

SAR	Synthetic Aperture Radar
SIS	Sequential Importance Sampling
SIR	Sequential Importance Resampling
SPOT	Systeme Pour l'Observation de la Terre
SRS	Satellite Remote Sensing
SWOT	Surface Water and Ocean Topography
TM	Thematic Mapper
VH	Transmitted wave is vertically polarised and received wave is horizontally polarised
VV	Transmitted and received waves are vertically polarised
WDD	Water Depth Difference
WL	Water Levels
WOfS	Water Observations from Space
WSE	Water Surface Elevation

List of Symbols

A	m^2	Area
Ag_i		Model-observation agreement for each particle
$c.v.$		Coefficient of variation
D		Kullback-Leibler distance
E_{P_X}		Expectation of the probability distribution function
$F_{SAR_{flood}}$	-	Fuzzy membership value to the flood class in a given pixel in the SAR-based flood maps
$F_{SAR_{non-flood}}$	-	Fuzzy membership value to the non-flood class in a given pixel in the SAR-based flood maps
F_s	-	Fuzzy membership value predicted in the class based on SAR analysis
$F_{VAL_{flood}}$	-	Fuzzy membership value to the flood class in a given pixel in the validation flood maps
$F_{VAL_{non-flood}}$	-	Fuzzy membership value to the non-flood class in a given pixel in the validation flood map
F_v	-	Observed proportion of wet cells in the validation flood maps
g	m/sec^2	Acceleration due to gravity
h	m	Root mean squared deviation
$\bar{h}$	m	Mean height
H		Entropy of a random variable
I		Mutual information between two random variables
K	-	Kalman gain matrix
K_{fuzzy}	-	Fuzzy implementation of Cohen's Kappa
n_i	-	Pixel count in each probability bin used for the reliability diagram
n	$m^{-1/3}/sec$	Manning's roughness coefficient

P		Probability distribution function of a random variable
P_o		Observed agreement
P_e		Expected agreement
$p(F\|\sigma^0)$		Conditional probability of flooding given a particular backscatter value
$p(\sigma^0\|F)$		Conditional probability of observing a specific backscatter value given that the pixel is flooded
$p(\sigma^0)$		Marginal probability distribution of backscatter values at a given pixel
Q	m^3/sec	Discharge
$\mathbf{Q}$	-	Model error variance-covariance matrix
r		Pearson's correlation coefficient
$\mathbf{R}$	-	Observation uncertainty
S	-	Spatial similarity
S_{obs}	-	Observed spatial similarity
S_{exp}	-	Expected spatial similarity
S_0	-	Bed slope
S_f	-	Friction slope
t	sec	Time
$\mathbf{w}$	-	Model white process noise
w_i		Global weight for each particle
x	m	Along channel distance
$\mathbf{X_0}$	-	Initial model state vector
$\mathbf{X}$	-	System state vector
$\mathbf{X}_k^b$		Background state vector
$\widehat{\mathbf{X}}_k^a$		Analysis state vector
y	m	Depth of flow

$\mathbf{Z}_k$		Vector of state observations
$\hat{\mathbf{Z}}_k$		Vector of model predicted state values
z	m	Bed elevation
α	-	Stability factor in max. adaptive time step calculation
α_F		Mixing proportion of the flood class in the Gaussian mixture model
α_{NF}		Mixing proportion of the non-flood class in the Gaussian mixture model
α_r		Ratio of the standard deviations of the simulated and observed values
α_t		Temporal decorrelation factor
β		Ratio of the means of the simulated and observed values
Δt_{max}	sec	Maximum stable time step in Lisflood-FP
Δt	sec	Time interval
Δx	m	Model cell resolution
$\varepsilon_{reliability}^{WRMS}$	-	Root mean squared deviation from the 1:1 line of the reliability diagram, weighted based on bin sizes
ρ		Variance growth scaling factor
σ^0		Backscatter value at a given pixel
σ_k		Time varying model error variance
σ		Variance of model errors
$\mathbf{\Sigma}$	-	Model uncertainty
θ_{loc_i}	radians	Local incidence angle
λ	m	Wavelength
τ	days	Temporal correlation length
ϑ		Tempering factor to ensure weight variability

List of Figures

List of Tables

PART I
INTRODUCTION AND BACKGROUND

CHAPTER ONE

"Experiment is the sole source of truth. It alone can teach us something new; it alone can give us certainty."

- Henri Poincare, *The Foundations of Science: Science and Hypothesis, the Value of Science, Science and Method,* translated by Mélanie Frappier, Andrea Smith, and David J. Stump.

1. Introduction

This thesis presents a novel probabilistic flood extent mapping algorithm and assimilation framework, to improve the integration of Synthetic Aperture Radar (SAR) derived flood extents with hydraulic flood inundation models in near real-time. The mapping technique, as well as the model-data integration methods, were developed from the perspective of transferability to current and future remote sensing-based observations of flooding. The efficacy of the mapping and assimilation approaches were demonstrated through a series of numerical experiments in the Clarence Catchment, Australia, using both synthetic and real-world case studies. An investigation of the spatiotemporal observation characteristics that best facilitate model-data integration impact on flood forecasts concludes the thesis.

1.1 Background and Problem Statement

In the absence of accurate flood inundation model forecasts, floods can have extremely expensive and often fatal consequences (CRED and UNISDR 2016). The lack of sufficiently precise global elevation data and inflow uncertainties propagated from precipitation forecasts leads to inherently erroneous flood inundation model outputs, frequently impeding their operational application (Camacho et al. 2015). The worrisome escalating trends in flood risk can primarily be attributed to climate change impacts, such as growing number of extreme weather events and changing precipitation patterns (Quinn et al. 2019). However, these flood impacts are further compounded by anthropogenic factors such as detrimental land-use change (urbanization, deforestation, and agricultural intensification) and an increasing population of people living on river floodplains (Rakovec 2014). As a consequence, flood events disproportionately impact developing nations, due to low climatic resilience of their populations and poor disaster management mechanisms (Uhe et al. 2019).

Accurate forecasts of flood inundation are necessary to augment flood control measures and to ensure effective mitigation of flood impacts. Traditionally, flood predictions have been obtained from stage-discharge relationships obtained from hydrometric river gauges or more recently, from numerical models based on the laws of conservation of mass and momentum (Di Baldassarre and Montanari 2009). Flood inundation models are primarily classified based on the approximations of the shallow water equations which they solve. In principle, 1D models assume that water velocity, height, and discharge vary along the channel flow direction exclusively. In contrast, 2D models account for channel as well as floodplain flows, following the grid defined by the digital elevation model or DEM (Haile and Rientjes, 2007). The hydraulic heads computed at each computational node are inter-compared to determine direction of flow from one grid to the next. These models therefore require inputs such as floodplain topography and channel bathymetry as well as hydrometric data for model initialization and constraint. Long term field observations of discharge and water levels are also required to define the upstream and downstream boundary conditions (Coulthard et al. 2013). The global decline in gauge networks have unfortunately reduced the likelihood of obtaining such data and the developing regions of the world which exhibit the lowest flood resilience are also the ones with the least number of hydrometric gauges (Revilla-Romero et al. 2015). Consequently, alternative data sources which can supplement this lack of hydrometric data for the initialization and constraint of flood inundation models need to be explored.

Surface (e.g., flood inundation and lakes) and sub-surface (e.g., soil moisture and groundwater) water has been monitored using Earth Observation (EO) satellites for decades. The synoptic views provided by Earth Observation satellites is especially suitable for the spatial characterization of floods, which are known to span large areas which are almost impossible to monitor in the field. Satellite, aerial, and drone based remote sensing platforms provide a unique opportunity to obtain distributed flood observations at a variety of spatial scales. The topographic information provided as an essential input to hydraulic models are primarily sourced from satellite based DEMs. Moreover, additional information to support hydraulic flood inundation modelling such as surface roughness (Mason et al. 2003), channel top width (Pekel et al. 2016), and even flow directions (Yamazaki et al. 2019), have all been derived from satellite data in the recent past (Schumann et al. 2012; Andreadis et al. 2013). Satellite images can also be used to extract flood inundation extent and depth (by integrating with DEMs) (Musa et al. 2015). Remote sensing derived flood extents and depths have widely been utilized for the calibration and validation of hydrodynamic models in the absence of gauge information

representative of flood dynamics or to complement hydrometric information (Schumann et al. 2008a; Patro et al. 2009; Di Baldassarre et al. 2009; Wood et al. 2016). Studies have also implemented data assimilation frameworks to update flood inundation models using such observations in real time (Neal et al. 2009; Andreadis and Schumann 2014; Hostache et al. 2018a)

Multiple studies have highlighted the simplicity of flood extent delineation from optical sensors (Schumann et al. 2009b; Grimaldi et al. 2016). In fact, high resolution aerial optical imagery has served as the benchmark validation data for a majority of flood mapping studies (e.g., Giustarini et al., 2016, 2013; Mason et al., 2014, 2010). The acquisition of such data, however, requires meticulous planning, dedicated flights, and is associated with high costs, thus impeding access for countries with limited economic resources. Earth observation satellites offer a cheaper alternative, with higher spatial coverage and varying resolutions. Coarse resolution images are typically available for free from the relevant space agencies, while higher to medium resolution datasets often need to be purchased at high prices. However, during major flood disasters, countries can activate the International Charter on Space and Major Disasters. The charter mandates any satellites positioned around the affected region, to capture the event and provide the data free of cost (Martinis et al. 2015c). The potential of optical remote sensing images, acquired in the visible region of the spectrum, has been demonstrated for flood mapping by a number of studies which were most recently reviewed by Huang et al. (2018).

Freely available satellite data such as from the LANDSAT and MODIS sensors, have also allowed the development of historical databases of observed inundation. The Dartmouth Flood Observatory (http://floodobservatory. colorado.edu/), for example, uses MODIS data at 250 m, in conjunction with appropriate ancillary data, such as the Shuttle Radar Topographic Mission (SRTM) Water Body Data Set. Maps of flood inundation are processed and provided in near real time, as well as added to an open archive of large floods at global scale (Shen et al. 2019). In a similar effort to demarcate permanent water bodies, Geoscience Australia developed the Water Observation from Space (WOfS) database (http://www.ga.gov.au/scientific-topics/hazards/flood/wofs). WOfS uses statistical analysis of historical surface water observations derived from LANDSAT-5 and LANDSAT-7 satellites for all of Australia from 1987 to 2015. Pixels which are inundated in >80% of the images, are classified as permanent water and provided freely as a surface water body dataset (Mueller et al. 2016). The study by Pekel et al. (2016) extended a similar approach to map surface water on a global scale, while

widths derived through drainage area to discharge regression were further translated into bankfull depths using hydraulic geometry theory in Andreadis et al. (2013).

While the use of optical data offers certain advantages due to its high spatiotemporal coverage and ease of interpretation, it can be severely affected by clouds in the ascending limb of the flood hydrograph (Schumann et al. 2009a). In fact, for small to medium catchments, the flood often recedes before the clouds have dissipated (Schumann and Moller 2015). Moreover, optical data are severely limited by dense canopies, as they completely obscure the flooding underneath from the sensor's view (Bates et al. 2014a). When available, cloud-free optical data are still a vital resource for flood management, as the processing is straightforward and does not require extensive expertise (Grimaldi et al. 2016). However, for systematic monitoring of floods, microwave data are preferred as they can penetrate clouds and are independent of solar illumination (Landuyt et al. 2018), making optical data best suited to complement and evaluate microwave images when available (O'Grady et al. 2014).

Passive microwave radiometers measure the thermal radiation naturally emitted by the Earth's surface in the form of brightness temperatures (Lievens et al. 2016). The differences in thermal inertia properties of land and water, allow the detection of inundated regions, as these areas typically exhibit lower brightness temperatures than land (Ahamed et al. 2017). In contrast to optical sensors, radiometers can image flood inundation independent of the prevailing solar illumination, weather conditions, or for longer wavelengths even vegetation cover. However, the detection of inundated dense vegetation is unique to passive microwave imaging, due to the exclusive use of surface emissivity properties. Since, only the brightness temperatures are observed and subsequently used for extracting inundation, the detection of floods is unaffected by obstructions such as tree canopies, which hamper accurate identification of inundation in densely vegetated regions for all other forms of remote sensing. This particular feature is unique to microwave radiometry and has been extensively used to monitor the flood dynamics of ecologically sensitive regions like the Niger Delta and the Amazon (Slinski et al. 2019).

The remotely sensed surface water extent information provided by the Global Flood Detection System (GFDS), hosted by the Global Disaster Alert and Coordination System (GDACS, http://www.gdacs.org/flooddetection/), uses passive microwave sensing to generate global products. First, the brightness temperature for a wet (measurement - M) pixel centred on the channel and the floodplain is extracted, identified based on historical sensitivity to changes in river flow, manifesting as an increase in water extent as a function of discharge. Second, dry

(calibration - C) pixels observed over land are extracted, close enough to the measurement pixel such that, all physical conditions other than the inundation can be assumed to be sufficiently similar. Measured brightness temperature (Tb) values at each pixel usually also include effects of physical temperature, permittivity, surface roughness, moisture, and tranmissivity. While estimating the relative contribution of these factors is challenging, these effects can be normalized by considering a ratio of the measurement and calibration Tb values. River flooding is then defined through a series of thresholds on the observed M/C ratio anomalies with respect to its cumulative frequency in the time series. Major floods are then classified as the 95th percentile and floods as the 80th percentile of its cumulative histogram; anything below that is classified as normal flow. The GDACS portal subsequently issues disaster warnings based on the detected flood levels, currently operating as an experimental project. Similar techniques are employed by the Dartmouth Flood Observatory River and Reservoir Watch project to monitor river discharge (http://floodobservatory.colorado.edu/DischargeAccess.html).

As passive microwave radiometers only record the emitted radiation from the Earth's surface, large pixel sizes often have to be considered such that sufficient amounts of energy can be recorded at the sensor, after the dissipation caused by atmospheric effects (Smith 1997). Consequently, large angular beams with spatial resolutions ranging from 20-100 km are typically used for such systems to ensure signal strength at the antenna. This implies that the use of passive microwave sensors can only be feasible for large catchments with channels at least several kilometres wide (Kim and Sharma 2019). Therefore, despite the potential of passive microwave sensors for flood monitoring, at local scales - especially for small to medium catchments - the use of much higher resolution imagery is often deemed necessary (Bates et al. 2006). The ability of active microwave sensors to provide high resolution images, independent of weather and time of day, makes them uniquely suitable for flood monitoring applications (Schumann et al. 2009a). This led to spate of development in Synthetic Aperture Radar, as they are often considered the only source of reliable information for rivers with sub-kilometre widths (Schumann and Moller 2015).

SAR backscatter intensity is primarily driven by surface roughness, with secondary impacts of electrical permittivity characteristics. As water is smoother relative to the wavelength of the transmitted radar beam, the incident microwaves are reflected away from the sensor in a specular mirror-like fashion, resulting in low recorded backscatter at the receiver (Smith 1997). Land, in contrast, presents a complex cohort of potential scatterers to the radar

beam, which results in high backscatter returned to the sensor. This land-water backscatter contrast is exploited by most mapping algorithms to identify flood inundation (Landuyt et al. 2018). However, as microwaves are highly sensitive to surface roughness changes, wind, emergent vegetation, or urban features which alter the backscatter characteristics of the flooded area, could potentially lead to under- or over-detection (Martinis et al. 2015b). SAR-backscatter over regions of emergent vegetation and urban areas can be quite ambiguous due to the presence of multiple potential scatterers per ground resolution cell. Studies have suggested complementing the SAR backscatter intensity information with interferometric coherence data to counter this problem, demonstrating significant improvements in classification accuracy (Pulvirenti et al. 2015). Fully polarimetric SAR images or those acquired from multiple viewing angles can also facilitate the detection of flood extents in case of ambiguous radar signatures under vegetation (Plank et al. 2017). The use of ancillary datasets such as topographic indicators, land cover information, and reference water masks can further reduce false alarm errors (Moulatlet et al. 2015; D'Addabbo et al. 2016; Grimaldi et al. 2020). Accurate flood extents with objective and reliable estimates of uncertainty, are vital for their integration with flood inundation models through data assimilation, which has the potential to improve flood forecasts and therefore, enhance flood resilience under a changing climate (Hostache et al. 2018a).

SAR-derived flood extents are often intersected with digital elevation models (DEM) to derive water stages, which are subsequently used for model performance evaluation (Hostache et al. 2009) or data assimilation (Hostache et al. 2010). SAR-based water level estimation algorithms assume horizontal water levels across a transverse cross-section of the extent, due to the instantaneous nature of the observations (Schumann et al., 2007). However, this assumption may not hold in complex catchment or channel geometries, land-uses, or when the flow is highly hydrodynamic. Moreover, the retrieval of water level from satellite remote sensing images requires a number of non-trivial processing steps and has often been identified as a source of uncertainty in hydraulic model forecasts (Bates et al., 2014; Schumann et al., 2009a). Assimilating SAR-based flood extents can eliminate the need for water level derivation, as the observed inundated area can be directly used to inform the model forecasts as opposed to the water levels. Although flood extent is only a good proxy for model performance when flows are rapidly varying with respect to change in simulated water depth, especially in low slope conditions where hydrodynamics dominate (Gobeyn et al. 2017). For example, when flow is contained within the banks while the channel is filling up, water levels will vary significantly

with each time step while the extent will remain almost constant within the river network. In order to be useful then, careful considerations of the spatial location, timing, and frequency of image acquisition are also required (García-Pintado et al. 2013).

Studies have investigated the utility of crowd-sourced water level data, to supplement remote sensing, for hydraulic model calibration and data assimilation, as the frequency of remote sensing data can still be a limiting factor (Mazzoleni et al. 2017). Preliminary analysis conducted using synthetic datasets, imply that such data could prove useful in improving flood predictions (Mazzoleni 2016). As crowd-sourced data can be considered independent from the model and remote sensing, adding it to the assimilation framework or even just at the model calibration stage can reduce forecast errors (Assumpção et al. 2018). Crowd-sourced water levels hold potential for those regions and time steps where extents are not sensitive to subtle changes in simulated water depth. The utility of crowd-sourcing needs to be explored as satellite remote sensing acquisitions still face problems of data latency and temporal coverage, which implies that observations of fast-moving floods are largely opportunistic, especially for small to medium sized catchments. Further, crowd-sourcing has the potential to provide accurate local information and is already informing first responders and planners alike (Chini et al. 2019). Since SAR-data are known to be unreliable in urban areas, crowd-sourced data could potentially serve as a valuable complementary dataset for flood inundation modelling. Therefore, in this study preliminary investigations are also undertaken to assess the utility of crowd-sourced data for the quantitative evaluation of hydraulic model predictions.

The increasing number of SAR satellites have enhanced the probability of monitoring flood dynamics from space, but the sensor and surface dependent uncertainties require significant post-processing using ancillary data for reliable inundation estimates. Yet, there are no sensor independent probabilistic mapping algorithms, which attempt to reduce the uncertainty in flood extent delineation using a single SAR image. Accordingly, this thesis develops a novel optimized texture-based approach to improve flood detection from SAR images. Since, the use of SAR-derived flood maps in conjunction with DEMs for the indirect retrieval of floodplain water levels has been identified as a potential source of forecast uncertainty, recent studies have recommended the direct assimilation of flood extent. However, the development of an appropriate likelihood function to estimate the probability of the model given the observation at each assimilation time step remains a scientific challenge. Therefore, this thesis develops a new method for assimilating SAR-based flood extents into hydraulic

model forecasts and investigates the impacts of observation spatiotemporal characteristics on the corresponding assimilation efficiency.

1.2 Objectives, Assumptions, and Scope

The principal objective of this thesis was to develop methods for improved near real-time estimation and forecasting of flood inundation, which was identified as a major gap in literature based on the review (See Section 2.3). Accordingly, a novel approach for 1) satellite SAR-based flood extent delineation and 2) combination with hydraulic model forecasts was developed. Specific objectives included:

1. Improving operational single image SAR-based flood mapping.
2. Testing the utility of crowd-sourced water levels for quantitative hydraulic model calibration.
3. Developing a new cost function for a particle filter-based flood extent assimilation framework.
4. Investigating the sensitivity of assimilation performance to observation spatiotemporal characteristics.

Synthetic as well as actual remote sensing observations of surface inundation extent, were assimilated into the hydraulic flood model, to evaluate the impacts of the proposed flood extent assimilation algorithm on the various model states and fluxes. For instance, the synthetic study allowed a distributed evaluation of the assimilation impacts on simulated floodplain flow velocities as well as water depths. Moreover, impacts on maximum hazard estimates, which are a function of maximum inundation depth, duration, and velocity, were also assessed. In order to ensure reproducibility for real cases and to emulate as realistic a setting as possible, the synthetic satellite images were generated based on uncertainties estimated from the SAR images available to this study. Similarly, the case study using actual satellite observations allowed comparisons with previous studies in literature, while also facilitating an appraisal of the pitfalls faced in a real-world scenario.

The approach selected for experimental designs of the open loop simulations also reflect a similar strategy. The experimental set up considered in this thesis, emulated the case of most developed countries – where reasonably accurate topography and bathymetry are available,

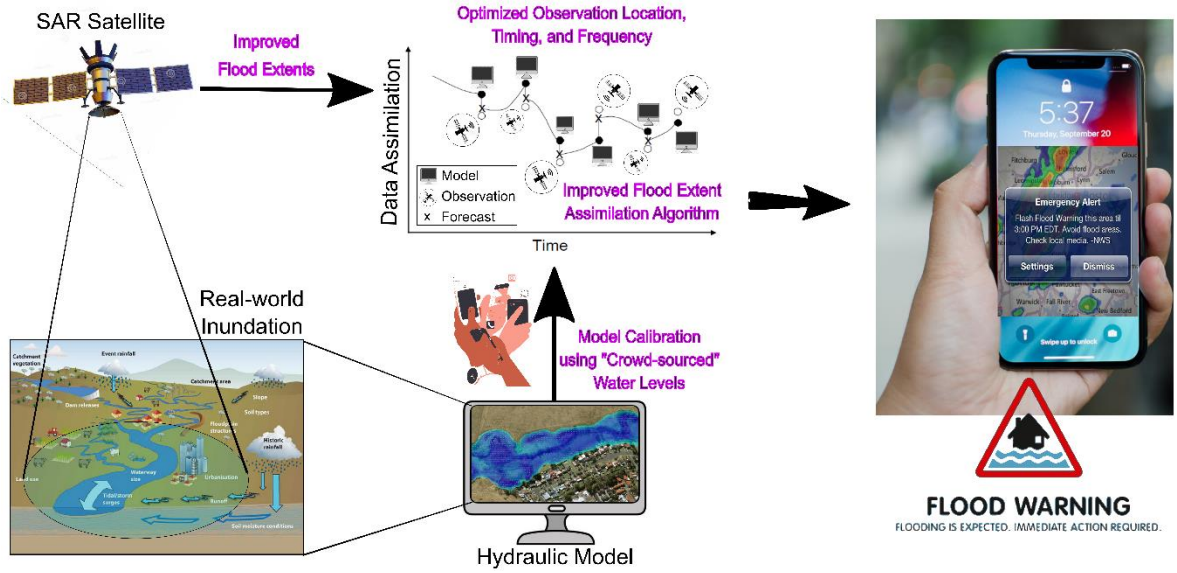

Fig. 1.1 Schematic of overall methodology followed in this thesis. Note that the objectives and contributions have been highlighted in magenta.

implying that forecast streamflow uncertainties dominate. Hydrological model forecasts, input to the hydraulic models mostly dominate forecast errors. This allowed an objective evaluation of the assimilation results independent of the complex and often strongly non-linear relationships between the different sources of uncertainty, including the channel geometry, friction, and topography, for example.

The overall methodology followed in this thesis has been summarized in Fig. 1.1. In order to enhance universal application of the approaches developed in this thesis, all assumptions were based on an operational implementation scenario. For instance, in the SAR-based flood extent mapping, the use of single SAR images is emphasized with the objective being to reduce the dependence on ancillary datasets whose global availability is unknown. Although the approach developed herewith is not completely automatic, methods to automate the process are discussed at length and could be implemented easily in future. Similarly, in the case of the flood extent assimilation, the objective was to develop a framework that could effectively quantify and reduce uncertainty to facilitate local relevance of large-scale forecasts. Although the methods proposed are tested at the catchment scale, an attempt is made to remain cognizant of the uncertainties that might affect global models and represent them sufficiently in the experimental setup. For example, to ensure that the findings from spatiotemporal observation impact on the assimilation performance is not just catchment specific, the assessment is founded on reach hydraulic behaviour. This implied that the findings could easily be extended to

hydraulically similar channels globally and were not confined to the study area (Clarence Catchment, NSW, Australia).

Another key assumption of this thesis is around fuzziness and uncertainty. Throughout the approaches outlined in this thesis, attempts have been made to account for uncertainty in validation methods. In fact, for each experimental output presented herewith, extensive validation is undertaken to ensure scientific rigour. Moreover, a detailed validation exercise for the flood extent assimilation, allowed an examination of the impacts on the delicate and highly non-linear covariability of intermediate states and fluxes within the hydraulic flood inundation model. Finally, the experiments on the location, timing, and frequency were conducted with the outlook that the number of observations available from remote sensing during any individual flood event is likely to increase. Consequently, the thesis questioned, how many observations are necessary and when does assimilation performance become independent of observation frequency. This ensures that the results will remain relevant in future despite consistent evolutionary advances in satellite SAR technology.

1.3 Outline

The research encompassed in this thesis is divided into eleven chapters that has been organized into four parts – Introduction and Background, Data and Models, Flood Extent Assimilation Framework, and Summary and Perspectives. Chapter 2 presents a detailed review of the state-of the art literature in SAR-based flood extent mapping and on the use of data assimilation of Earth observation images to improve hydraulic flood inundation model forecasts. Extensive discussions on literature surrounding the hydraulic modelling of floods and model calibration/validation as well as assimilation of field observations, was deliberately limited in this chapter to maintain focus on the major contributions of this thesis. Accordingly, the theoretical background for each technique employed in this thesis is included in the relevant chapters where the methods are implemented and not in the literature review.

A detailed description of the Clarence Catchment and the data available to this thesis is presented in Chapter 3. The study area presented in this chapter is common to all the experimental chapters in this thesis. Moreover, the data available in the form of remote sensing and field observations, was critical to the development and testing of the methods outlined in Chapters 4 to 9 of this thesis.

The procedure to extract accurate probabilistic maps of flood inundation from single SAR-images is developed and tested in Chapter 4. Theoretical background pertaining to the mapping of floods from SAR images is presented first, followed by a description of the methods and a discussion of the results obtained through a real-world application. The method used a novel texture optimization technique to maximize the information content of the SAR image itself, prior to flood extent delineation using a probabilistic neuro-fuzzy classifier. Finally, a fuzzy validation scheme is introduced to flood mapping literature which facilitates the representation of uncertainties in the optical aerial imagery, popularly used as the benchmark for evaluating SAR-based flood extents.

The hydraulic modelling component of this thesis is described in Chapter 5. First, the principles of hydrodynamic modelling are presented, then the sources and characteristics of the inherent uncertainties in flood models are discussed. The argument for choosing Lisflood-FP is developed subsequently, followed by a description of the full 2D-inertial acceleration solver used in this thesis. The model implementation is presented thereafter, including descriptions of input data and pre-processing. A simple framework to calibrate the channel friction parameter in Lisflood-FP using crowd-sourced observations of maximum flood depth is developed, using the 2013 flood event in the Clarence Catchment. Flood extents simulated using the calibrated model, are finally evaluated against an optical remote sensing-based flood observation.

Chapter 6 develops the flood extent assimilation framework used in Chapters 7 to 9. The principles of data assimilation are first outlined, along with an introduction to the terminology. This is followed by a discussion on the nature of hydraulic model uncertainties and the possible role of data assimilation in addressing them. The classification of data assimilation schemes is then described from a hydrological systems perspective. The chapter concludes with the particle filter-based flood extent assimilation scheme, developed in this thesis using a novel mutual information-based cost function.

In Chapter 7, the performance of the flood extent assimilation algorithm is tested through synthetic experiments. The 2011 flood event in the Clarence Catchment is reproduced with a representation of temporally correlated forecast inflow uncertainties for the open loop, while the calibrated full-2D Lisflood-FP model implementation is assumed to be the “truth”. For the synthetic experiment, binary flood extents are derived from the truth model at the times corresponding to the actual SAR-observations, by sampling from the flood and non-flood backscatter distributions of the real images. These were converted into maps of flood

probability using the approach outlined by Giustarini et al. (2016), for the flood extent assimilation performed subsequently using methods developed in Chapter 6.

The impacts of observation location, timing, and frequency on the assimilation performance are investigated in Chapter 8. Through synthetic experiments, algorithm sensitivity to spatial location, first visit, and revisit frequency, and input uncertainties with respect to reach hydraulic characteristics is analysed. The best-case scenario is also identified and the maximum possible reductions in forecast error using the methods outlined in this thesis are quantified. Chapter 9 demonstrates a real-world application using the 2011 flood event in the Clarence, where the methods developed in this thesis are used to derive SAR-based flood maps and assimilate them into the hydraulic model Lisflood-FP. The experimental design is identical to Chapter 7, except that assimilation impacts on the flood extent and channel depth are evaluated using actual observations.

Chapter 10 summarizes the major contributions of this thesis and the possible implications for the field, along with the limitations of the research. Chapter 11 concludes with a discussion on the future perspectives for each major research component of this thesis. The publications resulting from the research work embodied in this thesis have been listed after the perspectives.

CHAPTER TWO

"The Scientist must set in order. Science is built up with facts, as a house is with stones. But a collection of facts is no more a science than a heap of stones is a house."

- Henri Poincare, *The Foundations of Science: Science and Hypothesis, the Value of Science, Science and Method*, translated by Mélanie Frappier, Andrea Smith, and David J. Stump

2. Literature Review

This chapter reviews the relevant literature pertaining to the two major components of this thesis, namely SAR-based flood extent mapping and the subsequent dynamic integration of these with hydrodynamic models to improve flood inundation predictions in near real-time. The information presented in state-of-the-art literature is synthesized, with respect to the specific objectives of this thesis. First, a detailed review of existing SAR-based techniques for surface water detection is provided, and their limitations with respect to flood mapping in particular are discussed. The second section presents a critical evaluation of the current efforts to update hydraulic flood inundation models using remotely sensed flood information. The need to consider observation characteristics and operators, as well as the uncertainty of both the data and the model used for experimental design is highlighted. Following this, an outline of gaps in current research, leading into the research questions addressed in this thesis along with the hypotheses is presented. Finally, the chapter concludes with a short summary. The two major components of this literature review have been published as book chapters in edited monographs published by AGU-Wiley[1] and Elsevier[2].

[1]**Dasgupta, A.,** Grimaldi, S., Ramsankaran, R., Pauwels, V. R. N., Walker, J. P., Chini, M., Hostache, R. and Matgen, P. (2018). Flood Mapping Using Synthetic Aperture Radar Sensors from Local to Global Scales. In *Global Flood Hazard (eds G. J. Schumann, P. D. Bates, H. Apel and G. T. Aronica).* AGU Books. John Wiley Publications. doi:10.1002/9781119217886.ch4

[2]**Dasgupta, A.,** Grimaldi, S., Ramsankaran, R., Hostache, R., Matgen, P., Chini, M., Pauwels, V. R. N., and Walker, J. P. (2018). Earth Observation for Improved Hydraulic Flood Forecasts. In *Earth Observation for Flood Applications: Progress and Perspectives (ed. G. J.-P. Schumann).* Elsevier.

2.1 Synthetic Aperture Radar-based Flood Extents

Flooding involves large inundated regions which are often inaccessible or ungauged. Remote sensing (RS) data provides an elegant and practicable solution to assess spatiotemporal flood evolution. RS based flood mapping has witnessed significant research breakthroughs over the last decade. In addition to providing stakeholders with timely and spatially distributed information for crisis response (Schumann et al. 2016), RS-based flood maps are now utilized for hydrodynamic model calibration and evaluation, and to improve forecasts through assimilation (Schumann and Domeneghetti 2016).

The cost of high resolution (>4 and ≤10 m) imagery and sparse temporal coverage previously acted as a deterrent to unlocking the full potential of RS for flood management. In 1999, the International Charter "Space and Major Disasters" was initiated to provide a unified system of rapid satellite data acquisition and delivery in the face of major disasters (Martinis et al. 2015c). Floods are so frequently occurring and globally pervasive, that majority of all satellite data requests through the charter were flood related in the past decade, as illustrated in Fig. 2.1. Sensors operating in the visible region of the

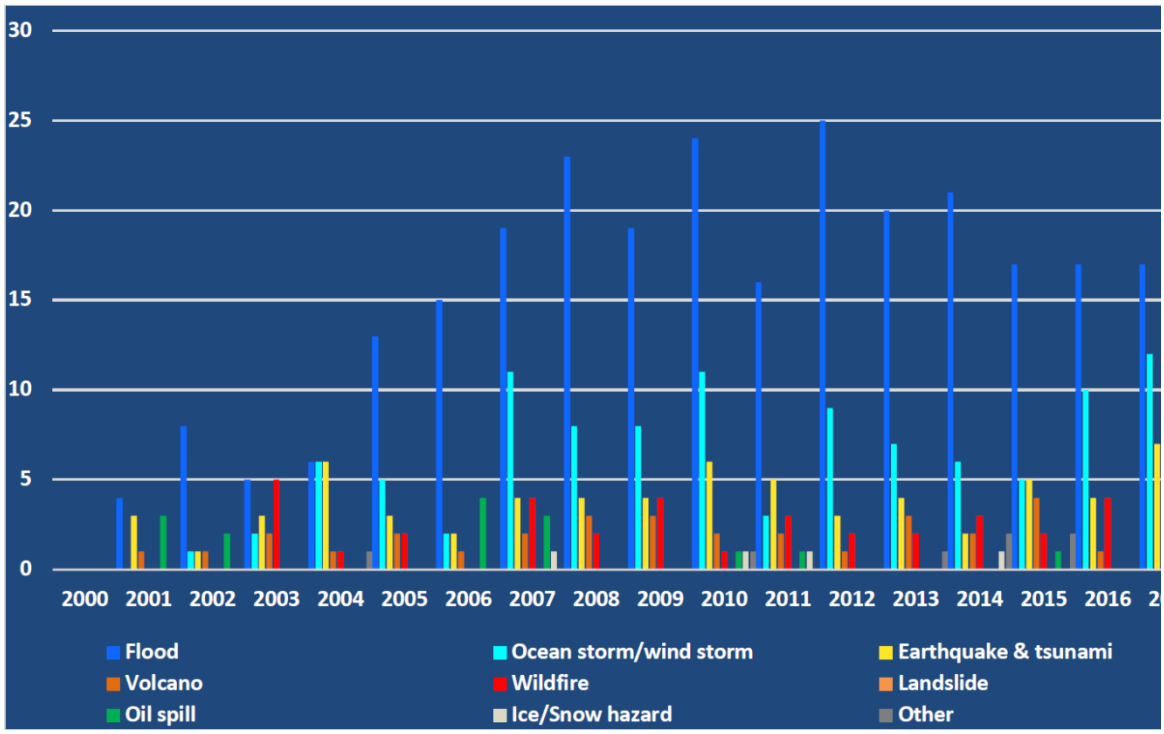

Fig. 2.1 Distribution of hazard types for Charter activations between 2000 and 2017. Source : International Charter "Space & Major Disasters", 2017 Annual Report.

electromagnetic (EM) spectrum offer the most straightforward solution for flood monitoring. Optical RS data is easy to interpret and multiple studies have demonstrated the utility of optical RS for flood mapping (Blasco et al., 1992; Jain et al., 2005; Li et al., 2015; Ogilvie et al., 2015; Ordoyne and Friedl, 2008). The increasing number of optical sensors with comparatively shorter revisit times has improved the spatiotemporal coverage substantially. However, as flood events are frequently characterized by persistent cloud cover, systematic monitoring using optical sensors is challenging.

Microwave remote sensing in the longer wavelength regions is able to penetrate clouds, which obstruct the view of optical sensors. Synthetic Aperture Radar (SAR) sensors use active imaging techniques and therefore can function independently of solar illumination or weather conditions. This property is indispensable for small to medium sized catchments, where flood waters often retreat before the cloud cover dissipates sufficiently (Schumann et al., 2007). Consequently SAR data are currently the only reliable source of information for monitoring riverine floods for small catchments with sub-kilometre channel widths (Schumann and Moller 2015).

Like the response of a mirror, a smooth surface or what is often referred to as a specular reflector in microwave remote sensing, reflects the radar beam back at an angle equal and opposite to the angle of incidence. This causes smooth or level surfaces to appear black in SAR images as the radar return is not recorded by the antenna. Conversely, heterogenous land surfaces appear rough to the sensor and return high backscatter. It is this high contrast in backscatter values recorded for land and water that facilitates surface water detection through SAR imagery. The launch of several high-resolution SAR missions has also contributed to improvements in the spatial and temporal resolutions and global coverage, making their use in flood mapping more practical. A summary of currently operational, historical, and planned SAR missions is presented in Fig. 2.2, with Table 2.1 providing details of the sensor characteristics.

Flood extent information can not only facilitate effective regional prioritization, but also efficient resource allocation, both during and after events. For ungauged catchments and inaccessible regions, SAR based flood maps are frequently used in combination with high resolution topographic data to derive spatially distributed water levels (Frappart et al. 2006; Schumann et al. 2007a; Hostache et al. 2009; Barreto et al. 2016; Matgen et al. 2016). SAR derived water levels (WL) are often used as hydrodynamic model calibration and validation targets (Gupta et al. 1998; Horritt 2000; Horritt and Bates 2002; Hostache et al.

2006; Pappenberger et al. 2007b; Schumann et al. 2008a; Di Baldassarre et al. 2009; Gobeyn et al. 2015, 2017; Wood et al. 2016), or assimilated into the model trajectory for an improved forecasting skill (Lai and Monnier 2009; Hostache et al. 2010; Matgen et al. 2010; Mason et al. 2012b; García-Pintado et al. 2013; García-pintado et al. 2014).

The last few years have seen a massive increase in utilization of spaceborne SAR systems for flood extent mapping, as new high-resolution platforms like TerraSAR-X/TanDEM-X and the COSMO-Skymed constellation became operational (Pulvirenti et al. 2011a, 2012, 2013, 2014b; Pierdicca et al. 2013, 2014; Voormansik et al. 2014; Pradhan et al. 2014, 2016; Martinis and Rieke 2015). Launch of the Sentinel-1A/B twin satellites, which provide global SAR coverage at 20 m spatial resolution with a revisit time of 3 days (and a repeat cycle of 6), marks a shift towards open data sharing in the satellite industry. Moreover, accessibility to fine resolution SAR imagery has already triggered a plethora of automated flood extraction algorithms (Boni et al. 2016; Sala et al. 2016; Twele et al. 2016).

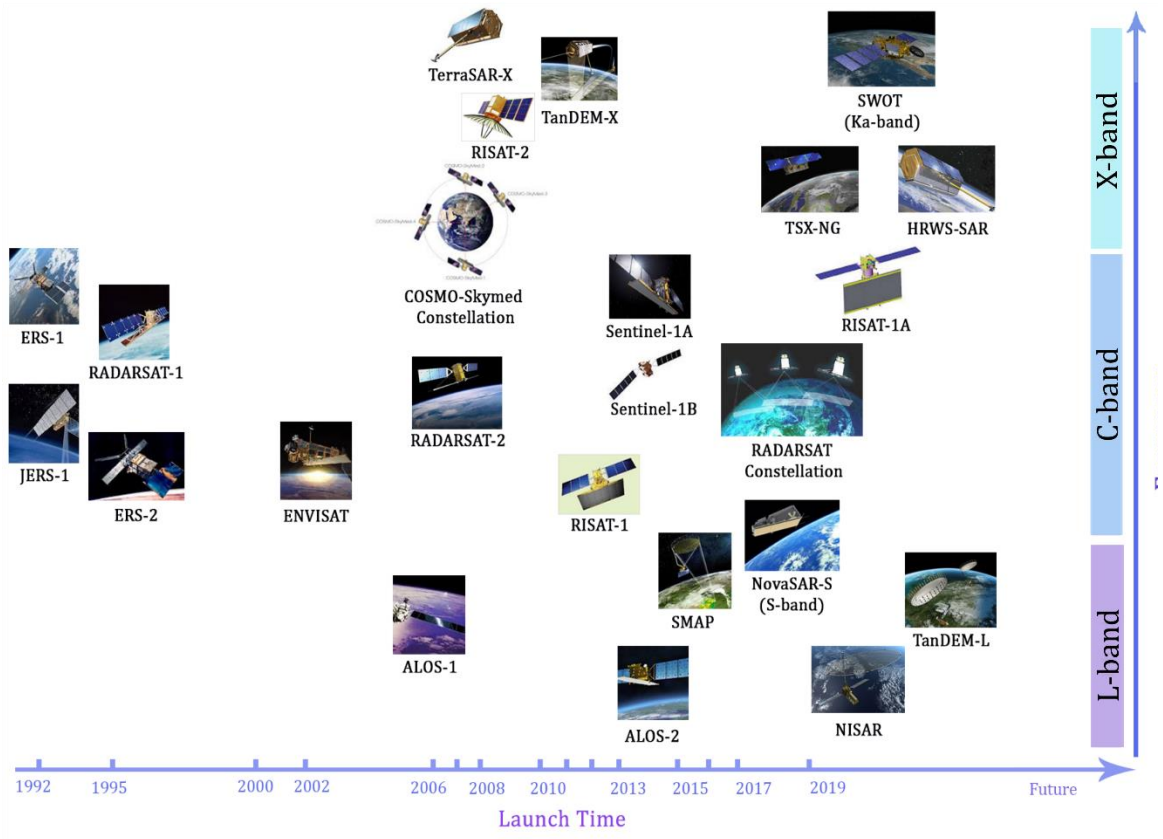

Fig. 2.2 Summary of satellite-based SAR missions which are applicable for flood studies, with corresponding wavelength bands and frequencies illustrated.

It can therefore be envisioned that the next decade will witness an inclusive data sharing environment, conducive to operational SAR-based flood mapping.

As operational flood mapping needs to facilitate rapid image interpretation, low-resolution datasets can be used to assess flood hazards at global scales (Westerhoff et al. 2013; Giustarini et al. 2015a). This diagnostic analysis can be utilized to identify areas of high hydraulic complexity, where finer scale imagery needs to be used. Such analysis can potentially inform variable resolution global flood models, which can be used to optimize the utilization of computational resources (Mason et al. 2015). As high resolution image processing as well as modelling are associated with significant computational costs, it's vital to utilize data at scales appropriate for flood event under investigation. This section provides a discussion on the issues related to operational SAR-based flood mapping at multiple scales and current progress in finding practicable solutions. An overview of the state of the art operational SAR-based flood delineation techniques is presented first, followed by a discussion of challenges in image interpretation and proposed solutions in literature. The section concludes with a discussion on methods to represent SAR-based flood mapping uncertainty.

2.1.1 Common SAR-based Flood Mapping Methods

A large variety of methods have been introduced in the recent past to map water bodies using SAR imagery. Table 2.1 lists some of the commonly used approaches, along with their strengths and limitations. When favourable conditions prevail, a single SAR image acquired during a flood (hereafter referred to as 'flood image') can be sufficient to reliably detect terrestrial water bodies. The single image technique works best when there is no wind roughening the water surface and when the detection is limited to floodwaters on bare soils and scarcely vegetated terrains. However, it is widely recognized that it is preferable to consider at least a pair of images consisting of the flood image and an adequate reference image (i.e. an image acquired in non-flooded conditions). Change detection not only tends to improve the classification accuracy, but also helps distinguishing permanent and transient water bodies (Chini et al. 2017). Eventually, only an approach taking advantage of a dense time series of backscatter recordings derived from tens or hundreds of SAR images acquired over a given area provides all the information that is needed to fully understand a floodplain's backscatter response to changing water levels and to accurately delineate the floodwaters on any given day (Schlaffer et al. 2016).

Table 2.1: Summary of spaceborne SAR missions and sensor characteristics. The italicized entries indicate satellite constellations with identical configurations. (Source: Modified based on Lillesand, T.M. et al., Remote Sensing and Image Interpretation, 5th edn., John Wiley & Sons, New York, 2004; eoPortal, https://directory.eoportal.org/, 2014.; OSCAR, https://www.wmo-sat.info/, 2017)

SAR Platform	Band	Polarization	Look Angle (°)	Swath (km)	Resolution (m)	Revisit Time (days)	Mission Status
ALMAZ-1	S	Single	20–70	350	10–30		Completed
ALOS PALSAR-1	L	Dual	10–51	40–350	6.25–100	46	Completed
ALOS PALSAR-2	L	Quad	8–70	25–350	1–100	14	Active
COSMO-SkyMED	*X*	*Quad*	*20–59.5*	*10–200*	*1–100*	*<1*	*Active*
CSK-2	*X*	*Quad*	-	*40-200*	*0.8-20*	-	*Planned*
Envisat ASAR	C	Quad	14–45	58–405	30–1000	35	Completed
ERS-1/2 AMI	C	Single	23	100	30	35	Completed
ICEYE-X1	*X*	*Single*	*15-25*	*35*	*10*	-	*Proof-of-concept*
ICEYE-X2	*X*	*Single*	*15-35*	*10-120*	*1-20*	*0.125*	*Active*
JERS-1	C	Single	35	75	18	44	Completed
KOMPSAT-5	X	Quad	20–55	5–100	1–20	28	Active
NISAR	L, S	Quad	-	-	-	-	Planned
NovaSAR-S	S	Dual	-	15-150	6-30	-	Planned
PAZ	X	Dual	15–60	5–30	1–6	11	Active
RADARSAT-1	C	Single	10–60	45–500	8–100	24	Completed
RADARSAT-2	C	Quad	10–60	10–500	3–100	24	Active
RCM	*C*	Quad	-	*20 - 350*	*5 - 50*	*1-4*	*Planned*
RISAT-1	C	Quad	12–55	10–225	1–50	4	Active
RISAT-2	X	Quad	20–45	10–50	1–8	4	Active
RISAT-1A	C	Quad	12–55	10–225	1–50	4	Planned
SAOCOM-1A/B	L	Quad	21–50	20–350	10–100	8-16	Active
SEASAT-1	L	Single	20–26	100	25	17	Completed
Sentinel-1A/B	C	Quad	20–45	20–400	5–100	6	Active
SIR-A	L	Single	47–53	40	40	–	Completed
SIR-B	L	Single	15–60	10–60	15–45	–	Completed
SIR-C	X, C, L	Quad	15–60	15–90	15–45	–	Completed
TerraSAR/TanDEM-X	X	Quad	15–60	5–200	0.24–40	11	Active
TanDEM-L	L	Quad	-	~350	1–20	16	Planned
TSX-NG	X	Quad	20-55	5-400	1-30	-	Planned
TerraSAR-X (HRWS-SAR)	X	Quad	-	10-800	0.25-25	-	Planned

Table 2.2 Strengths and limitations of commonly used SAR-based flood extent mapping methods, modified based on Di Baldassarre et al. (2011) and Martinis et al. (2015b).

Flood/Surface Water Detection Method	Subtype	Strengths	Limitations	References
Visual interpretation	-	Straightforward application and high accuracy; single-image approach	Extremely subjective; requires expert knowledge of image processing and flood processes; difficult to implement over large number of images; rarely useful for operational cases; Challenging for images with complex flow paths	Sanyal and Lu, 2004
Thresholding	Manual trial-and-error	Very fast, high potential of automation, moderate complexity, basis for other methods (e.g., change detection, integration of contextual and auxiliary information); single-image approach	Fails when land-water backscatter contrast is low; low flexibility; optimized threshold might not be most appropriate	Henry et al., 2006; Lang and Kasischke, 2008; P. Matgen et al., 2007; Townsend, 2001
	Automatic global			Martinis et al., 2009; Matgen et al., 2011; Pulvirenti et al., 2012; G. J.-P. Schumann and Di Baldassarre, 2010
	Automatic local			Chini et al., 2017; Martinis et al., 2015a, 2011
Change detection	Post classification comparison	Reduction of water look-alikes; improved detection of flooded vegetation areas; separation between flooded and permanent water areas	Availability of reference data of non-flood conditions; selection of appropriate reference image; possibly high complexity; multi-image approach	Clement et al., 2017; Giustarini et al., 2013; Hostache et al., 2012; Li et al., 2012; Mason et al., 2014
	Repeat pass interferometry	Facilitates flood detection in complex land-uses like urban and vegetated regions		Canisius et al., 2019; Chaabani et al., 2018; Chini et al., 2019; Li et al., 2019; Nico et al., 2000; Ohki et al., 2019; Pierdicca et al., 2014; Pulvirenti et al., 2015; Refice et al., 2014

Contextual classification	Texture-based	Accounts for spatial autocorrelation of SAR backscatter; statistical and robust; mimics human interpretation accounting for tonal differences; automation possible; moderate complexity; single-image approach	Window size and area appropriate texture measure selection; thresholding still required to obtain flooded area	Adam et al., 1998; Pradhan et al., 2014; Senthilnath et al., 2013
	Object-based	Avoids dispersed misclassified pixels removing the need for classification post-processing; reduces speckle effects; accounts for backscatter spatial structure; automation possible; single-image approach	Misclassified segments pose a higher risk of over- and under-detection, as compared to pixel-based approaches; post classification region merging might still be necessary; computationally intensive, might slow down flood map delivery during emergency	Evans et al., 2014; Ouled Sghaier et al., 2018
	Active contour models	Statistical and robust; high accuracy; easy automation; mimics inundation processes when integrated with topography; single-image approach	Post-processing necessary; computationally intensive; time consuming so unsuitable for real-time applications	Horritt, 1999; Horritt et al., 2003, 2001; Mason et al., 2010; Mason and Davenport, 1996; Tong et al., 2018
	Markov random fields	Allow integration of spatial, hierarchical, and temporal information; high automation possibilities; single-image approach	Must be applied to all image pixels; enormous computational complexity; might be unsuitable for real-time applications	Martinis et al., 2011; Martinis and Twele, 2010
	Ancillary data fusion	Increases accuracy; adds information content to single channel SAR response;	Might be unavailable for all regions globally; selection of suitable ancillary dataset is subjective, e.g. using HAND-based post processing might erroneously remove areas of pluvial flooding; multi-image approach	D'Addabbo et al., 2016; Pierdicca et al., 2008

2.1.1.1 Single Image Analysis

Histogram thresholding is the simplest and most straightforward single-image classification method. It consists of assigning to the semantic class “flooded” all pixels with a backscattering value lower than a given threshold. To overcome the subjectivity of this method many automated techniques have been proposed in the literature. They are based on either parametric or non-parametric approaches. When the former are applied, water and all other classes are discriminated by approximating the class distributions with predefined statistical models and, as a result, the optimal threshold value can be derived from their parameters. By contrast, the non-parametric approaches do not make any assumption about the classes’ statistical distributions. In 2007, Bazi et al., introduced a representative variant of a parametric thresholding approach that consists in automatically estimating the statistical parameters of the “target” and “background” classes by the expectation–maximization algorithm. The approach is based on the assumption that both the two classes follow a generalized Gaussian distribution. One of the most widely used non-parametric image thresholding techniques is known as Otsu's method. It searches for the threshold that minimises within-class variability while at the same time maximising between-class variability (Otsu 1979). The main advantage of this approach is that it is computationally inexpensive and therefore particularly suitable for rapid mapping applications.

For calm open water surfaces, the results of thresholding approaches are usually reliable and the largest part of an inundation area is detected. However, the effectiveness of pure thresholding methods is reduced when the “target” and “background” classes are unbalanced and/or overlap significantly. This is often the case as flooded areas typically cover a relatively small fraction of a SAR scene and different factors contribute to having relatively high backscatter from inundated terrain. A procedure adopted to render thresholding approaches more robust consists in splitting the entire SAR scene into different tiles that were either manually or automatically selected for their bimodality characteristics (Martinis et al. 2009, 2015a; Chini et al. 2017). To limit the over- and under-detection of flooded areas, backscatter thresholding is sometimes complemented with contextual information. One possible alternative is region growing, an image segmentation method that starts with the selection of seeds which are usually identified via thresholding. The backscatter values of neighbouring pixels are then examined to determine whether they should be added to the “flood” class or not (Giustarini et al. 2013). Similarly, active contour models (ACMs) allow converting incomplete or noisy edge

maps into smooth continuous vector boundaries (Mason and Davenport 1996), and were first described by Horritt (1999) for fluvial flood extent delineation. ACMs identify areas of homogeneous speckle statistics, and supplementary information such as topography and vegetation height maps can also be added as inputs (Mason et al. 2007). A hybrid region growing and active contour modelling approach was adopted by Mason et al. (2010), which allows the flood edge elevations to vary smoothly along the river reach.

2.1.1.2 Multi-temporal Image analysis

A second category of flood mapping algorithms is based on the analysis of two or more SAR scenes acquired over time. In its simplest form, a change image is produced by subtracting the grey values of a flood scene from those in a so-called reference image acquired before the event. Changes in backscatter are assumed to be mostly due to the appearance of floodwater and the delineation of the latter requires the application of a classification algorithm to differentiate the "target" (i.e. changed pixels) and "background" (i.e. unchanged pixels) classes derived from the difference image (Chini et al. 2008, 2013). A variety of algorithms are available to distinguish the changed and unchanged areas from a pair of SAR scenes. The previously described thresholding algorithms represent an option. A more advanced split-based approach for unsupervised change detection was proposed by (Bovolo and Bruzzone 2007), for the identification of tsunami-induced changes obtained from multi-temporal SAR imagery. A fully automated hierarchical split based approach was introduced by (Chini et al. 2017) that searches for tiles that are characterized by histograms of backscatter and change values both depicting an observable bimodality. The process facilitates the parameterization of the two distribution functions, thereby rendering the mapping of the floodwater more accurate and reliable.

The application of change detection requires the availability of SAR scenes acquired under non-flooded conditions. To select an adequate reference image several requirements should be fulfilled: the reference image should have the same viewing geometry and the same polarisation configuration as the selected flood image. Furthermore, it should have been acquired during the same season as the flood image, especially for applications in regions characterized by a pronounced seasonality in moisture and vegetation growth. The difficulties of finding an adequate image in the archive and of correctly interpreting all detected changes in backscatter represent the main limitations of flooding-related change detection. The advent

of missions such as Sentinel-1 providing regular observations at high temporal resolution together with the development of algorithms rendering the retrieval of reference images automatic and objective (Hostache et al. 2012) facilitates the application of these techniques.

Methodologies based on interferometric SAR are starting to be more widely used for flood mapping. These methods also fall in the category of change detection techniques. A well-known problem of SAR-based flood mapping is that the detection of floodwater in built up areas remains problematic. This is because the increase of the double bounce effect resulting from the presence of floodwater between buildings is hardly detectable through the observation of changes in SAR intensity. To address this important issue, Pulvirenti et al. (2015) and Refice et al. (2014) developed algorithms that integrate SAR intensity data with other features extracted from SAR data, such as the coherence. Interferometric coherence can be defined as the degree of correlation between two complex (phase and amplitude) SAR images. It is particularly related to the change in the spatial arrangement of the scatterers within a SAR image pixel (Chini et al. 2015), and thus to geometric changes in the scene. Flooded areas exhibit low coherence, which helps distinguishing them from non-flooded regions where coherence tends to be high. A coherence-based change detection approach thus effectively complements one that is solely based on intensity change detection.

The launch of several constellations of SAR satellites has paved the way to improve flood mapping by making use of multi-temporal as well as multi-angular information. An approach that is based on a time series of backscatter derived from tens or hundreds of images acquired over an area allows characterising floodplains in an unprecedented way. Notwithstanding this recent progress, multi-temporal image analysis of SAR data is still the exception rather than the rule and these techniques are mostly applied to optical data.

In their pioneering study Westerhoff et al. (2013), used multi-temporal ASAR imagery to estimate for each pixel specific probability distributions of water and non-water backscatter. Using these histograms, the probability of a "new" measurement belonging to either one or the other population is derived. In O'Grady et al. (2014) collections of SAR data were used to find a relationship between local incidence angle and backscatter coefficient, that is used to separate water and non-water pixels. Thereby addressing commonly encountered problems with single image techniques and simpler forms of change detection such as under-detection due to waves on water and over-detection due to low backscatter from dry surfaces. A harmonic model was

fit to the backscatter time series on a per-pixel basis and used to generate flood maps from newly acquired SAR scenes in (Schlaffer et al. 2015).

2.1.2 Image Interpretation: Challenges and Solutions

The contrast between water and land backscatter values, is caused by specular reflection which decreases the backscatter returned to the sensor. This allows for the use of thresholding procedures, when the distributions of the two main classes in the image histogram do not exhibit significant mixing (Martinis et al. 2009, 2015a; Boni et al. 2016). These techniques work fairly well for surface water extraction in relatively homogeneous regions such as bare soils. However, as discussed in the previous section, natural environments are rarely so ideal. Flood surfaces are mostly broken by emergent vegetation, roughened by wind effects or by protruding urban structures, each of which contribute to complex scattering responses. As backscatter is affected by all the factors discussed above, inferences based solely on the signal return are often ambiguous. Motivated scientific research in this direction has resulted in some remarkable improvements, however, many open research questions remain. In this section, the challenges in SAR-based flood delineation and state of the art solutions proposed in literature are discussed.

2.1.2.1 Smooth Water Detection

When the water surface is substantially smoother than the adjacent land pixels at the boundary, a clear distinction between land and water backscatter values exists, the conditions for surface water detection become ideal. Some of the most common methods which exploit this property include visual interpretation, histogram thresholding (Matgen et al. 2004), automatic classification algorithms like active contour (snake) (Horritt et al. 2001), and contextual classification (Martinis et al., 2011), which have been reviewed most recently by (Brown and Brownett 2016; Grimaldi et al. 2016). Some hybrid automated techniques featuring backscatter modelling, radiometric thresholding, region growing, and change detection have been proposed to simplify the processing of large number of datasets in a near-real time manner (Matgen et al. 2011; Giustarini et al. 2015b).

Most of the methods listed above work very well for smooth water detection in near ideal cases. Roughness and smoothness are a function of wavelength and angle of the incident radar beam. With present understanding, it intuitively follows that lower wavelength or higher

frequency systems display a higher land-water backscatter contrast ratio leading to better detection of smooth open water. The sensitivity to surface roughness is drastically reduced with increasing wavelengths, as many of the potential scatterers on land also appear smooth to the low frequency sensor, which in turn leads to a reduced contrast between flood and non-flood classes.

The contrast between these two land cover classes increases with the incidence angle, as the scattering from a smooth surface decays more rapidly than that of a rough surface, as a function of the incidence angle (Weydahl 1996; Wdowinski et al. 2008). The drawback of working with higher incidence angles is an increase in the shadow effect or regions where an object point is not reached by any portion of the radar beam (Kropatsch and Strobl 1990). This effect is caused primarily by the side-looking imaging technique of SAR system and primarily occurs in areas with steep reliefs or in the presence of obstacles such as buildings. As the radar views objects at an angle, it cannot "see" a certain region hidden by vertical structures, these appear as dark shadows in the SAR image. This can give rise to some false alarms since its backscattering values is similar to that of water. Moreover, if the incidence angle is reduced, the shadow and layover effects which occur may mask the view of hydraulically important features leading to loss of important domain knowledge and underestimation. For example, a long line of tall buildings orthogonal to the line of sight of the sensor, may obscure the flooding of an important highway or a dyke which may be close to overtopping, which might be critical from a flood management perspective. As the spatial resolution is continuously improved with newer SAR sensors joining the fleet of EO satellites every year, the size of objects which can cause such distortions is getting smaller. For example, in high resolution imagery, riparian vegetation can cast a shadow over the flooded area boundary being detected and add uncertainty to various regions.

Change detection approaches which use the difference between a pre and post flood image acquired with the same geometric characteristics, are often used to deal with this problem as shadowed areas show the same backscattering values. In some cases, the same considerations can apply for very smooth urban surfaces such as tarmac or asphalted areas. However, the shadowed area remains a region where no information is available and thus, hydraulic models or other ancillary data must be used to fill the gap (Pierdicca et al. 2008). The shadow effect can also be pre-calculated if a high-resolution DEM is available to calculate the local incidence angle at each pixel (i.e., incidence angle with respect to the local normal to the surface). Then,

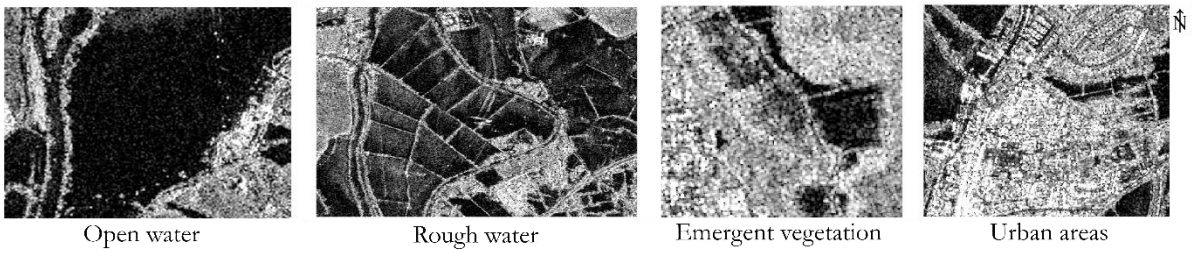

Fig. 2.3 The image shows example subsets of problem areas in SAR based flood mapping taken from a TerraSAR-X (HH, 3m Stripmap) scene acquired on the 25th of July, 2007, covering the Severn River flood event. The urban area shown here lies to the west of Tewkesbury, UK. © 2007 DLR, adapted from (Mason et al. 2012a).

a map of areas in shadow, which occurs when the local incidence angle is larger than $\pi/2$ rad, can be pre-computed to avoid false alarms. One of the more recent approaches suggest using linear regression to derive the slope coefficient between the local incidence angle and backscatter. The study found that the class separability was much higher when using the derived coefficient in place of backscatter (O'Grady et al. 2013). The only limitation of this study was the data intensive approach which requires enough samples of contemporary SAR images to assess the variation in backscatter behaviour with the local incidence angle, which may not be available for most study areas.

In terms of polarization, several studies have assessed the best combination of transmitted and received polarizations for smooth open water detection. As one might expect, the horizontal component of the beam undergoes a perfect specular reflection resulting in low signal returns. Conversely, vertical polarizations are extremely sensitive to surface roughness, especially at higher frequency ranges (Martinis et al. 2015b). This means that a short wavelength radar beam would experience severe scattering over "rough" land surfaces. Although all polarizations can be used for flood mapping of bare terrains, the VV one is more sensitive to the small waves generated by the wind over a water surface, so that the wind effect for flood detection is mitigated in HH polarization. Wind induced surface roughening, is one of the major causes of under detection due to the reduced contrast between backscatter from flooded and non-flooded soils.

The contrast between flood and non-flood regions is usually larger for moist soils, because the volumetric soil moisture content increases the complex soil permittivity. This results in an increase of the contrast between the electromagnetic impedances of air and terrain. In fact for a constant roughness, backscatter increases with an increase in soil moisture

(Pierdicca et al. 2008; Panegrossi et al. 2011). Moreover, because soil permittivity and roughness play the same role for both water and land, the difference in backscattering between floodwater and bare soils is not largely affected by frequency. Although in case of a very calm water surface the contrast may increase with frequency, as the soil would appear equivalently rougher to the sensor and therefore appear brighter in the image. At this point it would be worth to recall the concept of roughness which needs to be considered relative to the radar wavelength. A surface considered smooth at lower frequency (e.g., L-band) may behave as a rough surface at higher one (e.g., X-band).

2.1.2.2 Rough Water Detection

Flooded SAR images are often captured while the associated rainfall event is ongoing which implies that open water surfaces are roughened due to wind and rain effects. Larger water surfaces like oceans are more susceptible to the effects of wind, which generates small amplitude waves over a wide range of wavelengths resulting in Bragg scattering. Smaller inland water bodies such as lakes, reservoirs or flood inundation patches, exhibit a more irregular pattern of roughness leading to some diffuse surface scattering. These factors which increase the surface roughness result in higher signal returns, which in turn reduce land-water class separability and could potentially lead to an underestimation of the flooded area.

Wind-induced gravity waves are still an open issue in case of mapping inland water from SAR because many unknown factors (different depths or obstacles that screen the wind flow) contribute to the problem. This makes modelling the radar signal rather challenging, even in case accurate meteorological information is available. The linear regression approach proposed by O'Grady et al. (2013) to enhance separability between water and lookalike surfaces which exhibit specular scattering behaviours, was extended to solve this problem. Additional parameters like backscatter normalized to 30 degrees (using the fitted model) and the ratio between standard deviations of the time series of backscatter and intercept (SDR) at each pixel, was additionally used to derive a thresholding based flood map (O'Grady et al. 2014). Optimally thresholding the SDR image showed the best class separability, establishing the merit of generating such a database globally. Currently operational and planned high resolution SAR missions, facilitate the development of such a database in the future.

In addition to wind roughening, intense precipitation events (Pulvirenti et al. 2012) and wet snow (Pulvirenti et al. 2014a) also contribute to flood detection problems. The signal

attenuation caused by heavy rain can be observed at higher frequencies, since the amount of absorption and scattering of the signal is due to water drops is higher. Similar to the case of rain, wet snow is also very absorbent and produces very low backscatter which can be easily misinterpreted as floodwater (Pulvirenti et al. 2014a). Ancillary data such as local incidence angle maps, land cover maps, and optical imagery, can be utilized to supplement SAR information and reduce false alarm errors arising from this phenomenon (Pulvirenti et al. 2014a).

2.1.2.3 Partially Submerged Vegetation

Identifying flooding in pixels having a high vegetative fraction is relatively harder due to the complex scattering patterns which result from a combination of volume and double-bounce scattering. Double-bounce scattering represents the key process used to detect flooded vegetation on a SAR image. Radar beams achieve measurable penetration depths directly proportional to the incident wavelength. When the ground is covered by a smooth and very reflective water surface, the intensity of the double-bounce effect increases notably depending on canopy penetration. The returned signal is a combination of dihedral reflection - caused by steep emergent stems for the parts of the radar beam that surpass the canopy, and volume scattering - resulting from the radar beam travelling through air, interleaved with a multilayer canopy, all having different electrical permittivity properties (Richards et al. 1987). However, the difference between the signal returned by a flooded and a non-flooded forest, is strongly dependent on vegetation and sensor characteristics.

The abovementioned factors result in volume scattering which negates specular reflection effects and reduces detectability of flood patches. SAR sensors operating in the longer wavelength regions can partially penetrate vegetation canopies, as penetration is directly related to the signal wavelength. Additionally, the double-bounce can be exploited to detect the flooding, as non-flooded forested areas would return a lower backscatter due to the underlying rough forest floor which does not support corner reflection mechanisms (Pulvirenti et al. 2013). In this case an increase in soil moisture content decreases the contrast with flooded vegetation because the double-bounce effect gets amplified.

Flooded vegetation often results in enhanced backscatter due to a combination of dihedral and volume scattering, as opposed to dry vegetation where the radar beam only undergoes volume scattering. A review of early studies investigating flooding under vegetation

concluded that L-band radar images exhibited higher sensitivity to detect flooding under vegetation (Hess et al. 1990), and that steep incidence angles are better for flood detection as the possibility for the radar pulse to reach the forest floor is increased. Increasing the incidence angle also raises the bistatic surface scattering, however, smaller angles favour larger penetration depths. It is critical then to determine the optimum angle of incidence for the specific domain application being considered, such that the double-bounce enhancement effect is most pronounced.

As longer wavelengths favour greater penetration depths, they are often preferred for monitoring floods in vegetated areas. However, several studies have demonstrated that C-band data could allow flood mapping under lighter canopies (De Grandi et al. 2000). Another factor affecting the penetration depth for forested regions outside the tropics is seasonality. As the shedding of leaves reduces the potential scatterers in each pixel, lower wavelength sensors can also achieve penetration during leaf-off seasons (Townsend 2001). This has been examined by using X-band data to investigate surface water extent under low density canopies (Antonova et al. 2016). An increase in backscatter was noted for all the different forest types in that area under flooded conditions indicating that such approaches could be useful for boreal and temperate environments (Voormansik et al. 2014; Cohen et al. 2016). For cases of increasing complexity where flooded water needs to be distinguished from marshes or mangroves, using multifreqency fully polarimetric data and the use of P-band is suggested for high density canopies (Martinis and Rieke 2015).

Although radar polarimetry can facilitate a reliable detection of the complex scattering mechanisms observable in flooded vegetation, fully polarimetric SAR images are often unavailable during a flood event. Polarimetry is able to isolate the contribution of the double bounce from volume and surface scattering mechanisms, which is beneficial for detecting the changes caused specifically by flood water. This is done by measuring the magnitude and phase difference between co-polarized (i.e., HH, VV) and cross-polarized (e.g., HV backscatter) backscatter (Plank et al. 2017).

It was also demonstrated that using a combination of HH and VV polarized images favoured the separation of flooded and non-flooded forests over the use of HH data alone (Zalite et al. 2014). This can largely be attributed to the increased phase difference observable in the former, as a result of interactions with inundated vegetation. In case of dense vegetation, the combined occurrence of absorption and scattering, can produce an overall attenuation that is

relevant also for the HH polarization. In such cases, the SAR signal is reduced as the biomass increases, penalizing especially higher incidence angles and creating false alarms as well, since the backscattering is drastically reduced.

More recent studies proposed the use of interferometric SAR (InSAR) which measures the level of similarity between two SAR images based on the speckle distribution patterns. Recall that speckle results from multiple scatterers within the same ground resolution cell, meaning that this pattern should be consistent for a given geographic area imaged using identical sensor characteristics. If the difference in acquisition times of the SAR images can be limited such that the only change in terrain properties which occurs within that duration is the flooding, coherence data can supplement intensity information for flood detection under canopies (Refice et al. 2014).

Vegetation normally decorrelates the microwave signal intensities due to volume scattering, which determines low Differential SAR Interferometry (DInSAR) coherence (even for short temporal baselines). This is primarily a result of changes in plant phenology and movement of stalks and/or leaves caused by wind. However, flooded vegetation is characterized by even lower values of coherence due to the added presence of floodwaters (Chini et al. 2012). Flooded waters contribute further to decorrelation due to specular reflection properties, which in the presence of emergent vegetation, manifests as double-bounce scattering (Zebker and Villasenor 1992). When combined with intensity information, flooded vegetation can be identified as areas displaying an increase in backscatter and decrease in coherence (Nico et al. 2000). These approaches primarily detect the change in image texture and dielectric properties, which can be explained as a function of the underlying water level (Pulvirenti et al. 2011a). As the use of these change detection approaches are very sensitive to the temporal baseline of the SAR acquisitions, the use of newer satellites offering relatively lower repeat cycle, like the COSMO-SkyMed (CSK) Constellation, should be explored.

The problem with high frequency SAR systems like CSK is the heightened sensitivity to any potential scatterers. However, for narrow leaved plants these data could still provide useful information (Pierdicca et al. 2017). Furthermore, the capability to provide multitemporal observations can support the monitoring of flood evolution in some areas. If the land-cover distribution of the vegetated areas in the scene is known a priori, vegetation scattering models can be used to predict backscatter behaviour (Pulvirenti et al. 2011a). This information assists in tuning fuzzy models which can then be used to interpret flooded vegetation while

simultaneously acknowledging the uncertainty (Pulvirenti et al. 2013). The study by Grimaldi et al. (2020), attempted to further reduce the dependence on site specific knowledge required for the implementation of electromagnetic models by statistically analysing wet and dry backscatter distributions through probability binning. Implementing the thresholds obtained within a fuzzy classification framework, with the addition of commonly available ancillary data layers, the authors demonstrated an overall accuracy >80% for flood maps derived from L-(ALOS PALSAR) and C-band (Cosmo-Skymed) SAR images.

2.1.2.4 Urban Flooding

Flood detection around urban features is challenging due to a variety of factors. Double bounce scattering from buildings and pavements, along with metals with high dielectric constants having high reflectivity, contribute to the enhanced backscatter observed. The effect of material properties dominates over the soil roughness characteristics, since the smooth urban surfaces like tarmac and asphalt are already smooth with respect to still water (Mason et al. 2010b). Distortion effects caused by the side-looking sensor geometry additionally pose problems for high to very high resolution SAR imagery as high rise buildings tend to cause shadows and layover effects. Smooth urban surfaces cause specular reflection and return a signal, very similar to inundated areas, frequently leading to false alarm situations. However, it must be noted that specular reflection effects are increased for smooth urban surfaces oriented along the sensor's line of sight. Conversely, corner reflection is more likely to occur if the urban features are oriented orthogonally to the incoming radar beam (Pulvirenti et al. 2015).

Change detection (CD) based approaches were found to reduce the misclassifications of other dark image regions by identifying areas of shadow, layover, tarmac, and permanent water a priori (Giustarini et al. 2013). Once these areas are diagnosed and masked out, overestimation can be substantially lowered. Earlier studies investigated the use of active contours in conjunction with region growing. However, each step of the seed selection process required manual intervention. Further studies automated this seed selection process for region growing using LiDAR-based topographic information (Schumann et al. 2011).

While backscatter based approaches like change detection can work in some case studies (Aggarwal et al. 2014), there are some limitations for urban flood detection problems. For instance, the backscatter variation attributable to a change in the surface dielectric constant which is caused by the flooding, is nearly unobservable as the returned signal is too low due to

specular reflection. If the smooth tarmac surfaces like roads and pavements are masked out of the analysis, flooding on and around these features cannot be detected (Pulvirenti et al. 2015). Conversely, in the case of dihedral reflection this intensity change is usually detectable, as the increased surface reflectivity results in a stronger return. However, if the water level is not negligible with respect to the building height, the signal return from corner reflectors might be reduced due to a decrease in the beam cross-sectional area (Thiele et al. 2007). Therefore, the reliability of SAR intensity based approaches appears to be limited for operational urban flood mapping.

The interferometric coherence approach has recently been applied to deal with many of these limitations (Chaabani et al. 2018). Urban areas are generally temporally coherent as the decorrelation is a function of relative motion of scatterers in the pixel over the temporal baseline (separation between the two image acquisitions), which can be considered negligible for high density built-up areas. On average, double-bounce in urban regions correspond to highly coherent targets in the image as a function of their high backscatter in combination with temporal invariance. As spatial decorrelation is inversely proportional to the wavelength of the radar beam, at lower wavelength bands this effect is more pronounced (Zebker and Villasenor 1992).

Standing water is expected to decorrelate the signal more than is expected, a property which may be exploited for urban flood mapping (Matgen et al. 2011; Pulvirenti et al. 2015). In high resolution SAR imagery, some false alarms in the coherence information may be generated from dynamic traffic flows and parking lots. However, assessing three pairs of SAR images acquired within the critical spatiotemporal baselines, one for the dry or pre-event situation, one with images acquired before and after the event and finally a post-event pair, can help in correctly identifying the decorrelation caused by flooding. While this seems to be a data intensive approach at the outset, when the improved temporal resolutions of current and upcoming SAR sensors is considered, it appears more likely that such data may soon be available for analysis globally. Gaps still remain in urban areas, such as shadowed areas structures where SAR backscattering does not provide any information about the presence of water, and they could be filled-up by the assimilation of the flood maps into hydraulic models. In that case, the flood maps should be provided in terms of probability that a certain pixel is flooded in order to characterize the uncertainties that associated with the flood map (Schumann 2019).

2.1.3 Representation of Uncertainties

Regardless of the number of images considered, most SAR image processing methods output flood extent estimates in the form of binary maps and thus do not provide any indication on the uncertainty associated with the pixel classification. A more informative and arguably more advanced method involves estimating the probability of each pixel belonging to the flood class. However, probabilistic flood mapping is still not very common and only a few noteworthy exceptions exist (Westerhoff et al. 2013; Giustarini et al. 2016; Schlaffer et al. 2017). Previous studies on the characterisation of uncertainties in flood extent maps have often been limited to random realisations of potential sources of uncertainty (Hostache et al. 2006, 2009; Schumann et al. 2008d; Di Baldassarre et al. 2009; Refice et al. 2014; Giustarini et al. 2015b).

The uncertainties contributed by ambiguities in backscattering properties cannot be completely removed (Stephens et al. 2012). Approaches which account for this uncertainty in the flood detection process provide fuzzy outputs in which pixel values convey the confidence with which the pixel is correctly classified as flooded. As SAR-based flood extents serve as calibration and evaluation targets for hydrodynamic models, quantifying the uncertainty of flood observations is essential (Schumann et al. 2008a). Studies have successfully illustrated that factors like the acquisition time of the SAR image (Gobeyn et al. 2017), sensor characteristics (Giustarini et al. 2015b) as well as the flood mapping procedure chosen (Di Baldassarre et al. 2009), can impact the calibrated model parameters. These studies demonstrated the need for a shift towards probabilistic mapping, which is the current consensus amongst the flood mapping community (Di Baldassarre et al. 2010, 2011; Schumann and Di Baldassarre 2010b; Alfonso et al. 2016).

The discussion in the previous section, of factors which may contribute to errors in SAR-based flood maps, is important to understand and identify areas in which the model outputs will always be more reliable. However, in practical applications when large amounts of data need to be processed, especially at the regional or global scale, it makes more sense to convey the observational uncertainty to the model for which several techniques have been proposed in literature. For example, Schumann et al. (2008c) investigated the uncertainty in SAR-derived water stages, from a single SAR image and a single flood mapping procedure, and identified two main sources of uncertainty: (i) the parameter value applied to classify a pixel as flooded (i.e. flooded/non-flooded classification threshold) and (ii) geocoding of the image itself. The study conducted by (Refice et al. 2014) implicitly introduced a semi-

automated approach that allows integrating ancillary information to derive a posteriori probabilistic maps of flood inundation, accounting for different scattering responses to the presence of water.

In order to eliminate the subjectivity in the selection of a particular algorithm for flood mapping, several studies investigated the use of a possibility of inundation (POI) map (Schumann et al. 2009b; Di Baldassarre et al. 2009; Schumann and Di Baldassarre 2010b). An ensemble of maps generated using a variety of flood detection algorithms which perform equally well for a given study region are combined to arrive at the POI map. Each pixel in the POI map is assigned a value between 0 and 1 based on how many of the ensemble members classified it as flooded. It can be interpreted as an event-specific flood probability map which can facilitate probabilistic calibration procedures.

Fuzzy approaches which combined SAR based information with available ancillary data sets have since been suggested and exploited for operational flood mapping at various levels (Pappenberger et al. 2007b; Pierdicca et al. 2008; Pulvirenti et al. 2011b, 2013). The fuzzy membership functions, are usually defined based on EM backscattering models for each land-use class, which the classifier may encounter in the SAR scene. These models can be difficult to parameterize if domain knowledge is limited, thus histogram based methods still need to be tested for more generalizability.

Probabilistic mapping approaches which utilize Bayesian posterior probabilities of flooding for each pixel were tested recently, demonstrating a good agreement with the validation data assessed through reliability diagrams (Giustarini et al. 2016; Schlaffer et al. 2017). One of the major limitations of current validation strategies is that the validation data is assumed to be completely true. This is counterintuitive as probabilistic mapping operates on the premise that observational uncertainty needs to be acknowledged. The current challenge is to develop an effective statistical measure to assess fuzzy flood maps with fuzzy validation data.

2.2 Remote Sensing Data Assimilation for Improved Flood Inundation Modelling

Current global flood forecasting systems, primarily consist of a numerical weather prediction and hydrological model cascade, which provide predictions of streamflow (Emerton

et al. 2016). For example, the Global Flood Awareness System (GloFAS, Alfieri et al., 2013) provides forecast probabilities of the streamflow exceeding pre-defined flood severity thresholds, in the global channel network with a lead time of up to a month (Hirpa et al. 2018a). GloFAS is the first globally concerted effort towards operational flood forecasting, which could potentially offset the disproportionate impacts of flooding on developing nations where the necessary expertise and infrastructure might be lacking (Uhe et al. 2019). However, the lack of corresponding inundation information severely prohibits the direct translation of GloFAS forecasts into actionable insights for flood mitigation. As flood extent estimates are not provided, pre-emptively assessing risk and damage based on GloFAS predictions is nearly impossible.

Although there exists a widespread scientific consensus regarding the inclusion of hydraulic flood inundation models into global forecasting chains, there are mainly two reasons why this has not materialized yet. The first, pertains to the availability of sufficient computational resources for operational real-time ensemble forecasts of flood inundation. To put this in perspective, hydrological model calibration studies use thousands (Pappenberger et al. 2005) to a million model runs (Moradkhani et al. 2005a) within a research scenario, while GloFAS (real world scenario) can only use 51. This is primarily because a trade-off between computational time and prediction accuracy, becomes absolutely critical in an operational scenario where time is of essence (Sanders and Schubert 2019). Moreover, in an ensemble forecasting model cascade these 51 GloFAS outputs are inputs to a hydraulic model. A task which requires even more time and even larger computational capabilities. Recent advances in scientific computing prowess suggest that integrating flood inundation models into forecasting chains might now be feasible (Ward et al. 2015; Bates et al. 2017). However, extensive investigations into the computational demand versus accuracy trade-off are essential to realize this operationally.

The second reason is related to the data required for the implementation of flood inundation models (Bates 2012; Bates et al. 2014a). Typically, this necessarily includes information on upstream (rainfall or inflow) and downstream (outflow or water levels) boundary conditions, channel geometry, and floodplain topography. Among these, the floodplain topography is arguably the most critical factor which determines floodplain connectivity and flow pathways, exerting a dominating influence on the resulting inundation patterns (Schumann et al., 2014). The uncertainty contributed by forecast inflows compounds

topographic errors, often leading to highly erroneous predictions of inundation extent (Hostache et al. 2018b). Current global digital elevation models (DEMs), such as those from the Shuttle Radar Topography Mission (SRTM) or from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) mission exhibit large vertical errors to the order of several meters (Chen et al. 2018). These errors are exacerbated over complex topography and cannot resolve microtopographic variations low gradient regions, limiting their utility towards generating flood predictions with acceptable accuracy (Chu and Lindenschmidt 2017; Schumann and Bates 2018; Grimaldi et al. 2019).

Studies have explored a variety of bias correction (Sanders 2007; Pramanik et al. 2010; Kumar et al. 2019), vegetation and artefact removal (Yamazaki et al. 2012, 2017; Hirt 2018), artificial enforcement of drainage networks (Gallant et al. 2011; Yamazaki et al. 2019), and DEM merging techniques (Robinson et al. 2014; O'Loughlin et al. 2016; Yue et al. 2017; Pham et al. 2018; Wang et al. 2018) to improve the quality of global DEMs for flood modelling. For example, the global Multi Error Removed Improved Terrain (MERIT) DEM (Yamazaki et al. 2017) removed absolute bias, stripe noise, speckle noise, and tree height bias from a merged SRTM3 and AW3D elevation product. The resulting MERIT DEM showed significant improvements, especially in the representation of microtopographic variations and channel networks in flat terrains. In spite of this, Schumann & Bates (2018) observe that these improved topographic datasets still exhibit vertical errors much larger than those acceptable for flood inundation forecasting.

Recent literature has explored the possibility of a high resolution global DEM based on TanDEM-X data, reporting sub-meter vertical accuracy in simulated water surface elevations, indicating suitability for flood applications (Archer et al. 2018). Moreover, the TanDEM-X 90 (~1 arc second) product was also found to exhibit lower vertical errors than the MERIT DEM, for all land-use classes except tree-covered regions (Hawker et al. 2019). However, the development and provision of a global product is still a long-term goal which would need extensive testing, as even small vertical errors can strongly impact flood forecasting accuracy especially at local scales (Schumann and Bates 2018). As Hawker et al. (2019) demonstrate, the choice of an appropriate global DEM for floodplain applications should ideally be determined through an assessment of the predominant land cover in the region. While this recommendation is feasible on a case to case basis for research purposes, it also implies that when using any one of the global DEMs for operational forecasting, large vertical errors may

still exist in one or more of the underlying land cover regions. Consequently, in order to prepare for the imminent integration of flood inundation models into global forecasting systems, methods to reduce the uncertainty contributed by topography need to be developed urgently (Fleischmann et al. 2019).

Earth Observation (EO) data has the potential to provide independent observations of surface water flood dynamics (Bates et al. 2014a), which can be used to improve flood inundation predictions in near real-time as shown in Fig. 2.4. The primary objective of data assimilation is to nudge the model trajectory towards observed data, improving the agreement between them whenever new observations become available (Schumann et al. 2009a). Data assimilation (DA) can be interpreted as a dynamic calibration technique in that sense, where the observations are used to update model forecasts in an online manner (Smith et al. 2011). The implementation of DA techniques also allow extending the temporal coverage of instantaneous satellite-based flood observations, through integration with a continuous dynamic process-based hydraulic flood inundation models (Lahoz et al. 2010). In fact studies have shown that forecast errors can be improved by up to ~50% at the assimilation time step, while

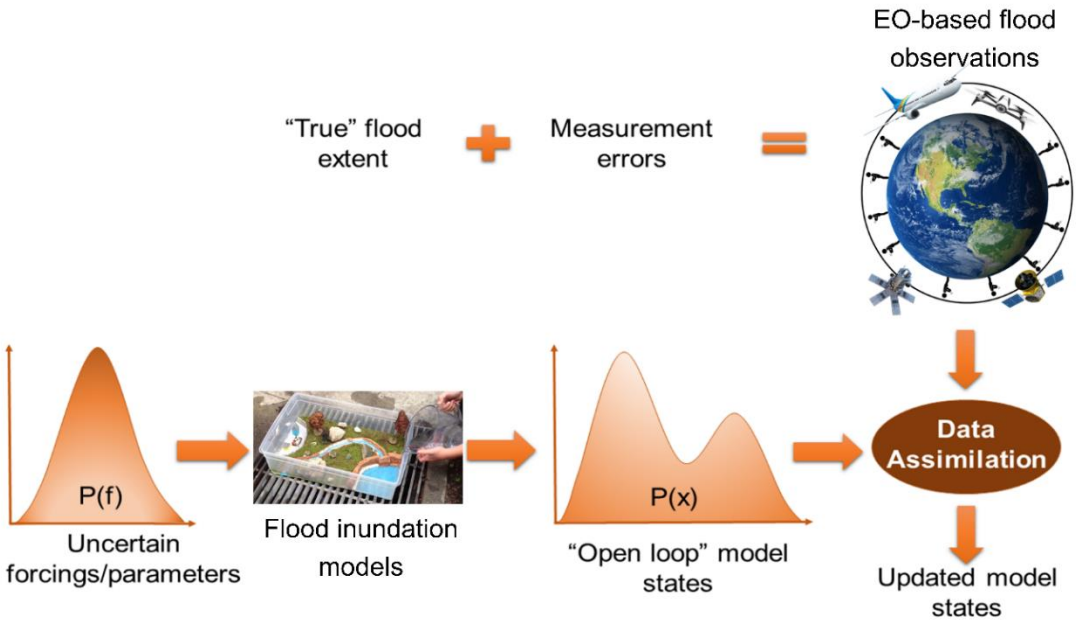

Fig. 2.4 Schematic of the Earth Observation data assimilation problem in hydraulic modeling, adapted after Moradkhani (2008). Here, Earth Observations are interpreted as "truth" plus errors, as satellite-derived flood extents are expected to encompass the "true" flood extent even though a major component of measurement errors are also expected to be present in the observation.

improvements can persist for up to 48 hours (Hostache et al. 2018b). EO-datasets which can be assimilated into flood models to improve forecast skill include water levels (WLs) (e.g., Matgen et al. 2010), inundation volumes (Revilla-Romero et al. 2016), and inundated area (e.g., Hostache et al. 2018).

While the inundated area is directly observed by optical and microwave imaging sensors, satellite altimeters provide water levels on a large scale with typical footprint resolutions in the order of kilometres (Paiva et al. 2013; Hossain et al. 2014; Tourian et al. 2017; Huang et al. 2018b; Schneider et al. 2018). However, satellite altimetry is currently unable to resolve flood depth in narrow channels at local and catchment scales (Schumann et al. 2014c, 2018; Biancamaria et al. 2016; Grimaldi et al. 2016), which is problematic as almost 25% of global rivers have widths <120 m (Frasson et al. 2019). Some innovative research is ongoing in this direction which promises future improvements (Michailovsky et al. 2012; Bauer-Gottwein et al. 2015; Kim and Sharma 2019; Kim et al. 2019), especially with respect to the potential of swath altimetry for hydraulic flood inundation modelling applications (Domeneghetti et al. 2018b; Pitcher et al. 2018; Tuozzolo et al. 2019). The upcoming SWOT (Surface Water Ocean Topography) mission with an expected launch date of 2021, will use swath altimetry provide direct observations of water surface elevation, widths, and slopes simultaneously for rivers wider than 100 m (Baratelli et al. 2018). Inverse modelling can then be used to remotely derive discharge, bathymetry information (slope and geometry), and flow velocity values, within a hindcasting data assimilation framework (Durand et al. 2008; Yoon et al. 2012; Prigent et al. 2016; Baratelli et al. 2018; Brisset et al. 2018; Domeneghetti et al. 2018b, a; Oubanas et al. 2018a, b; Kim et al. 2019; Revel et al. 2019). Investigative studies preparing for the SWOT mission, have successfully demonstrated the value of combining virtual swath altimetry measurements with hydrodynamic models, for improved modeled depth and discharge estimates. Therefore, the SWOT satellite is expected to revolutionize hydraulic data assimilation research, as remote sensing-based spatially distributed water levels will be available to modellers for the very first time.

While multiple studies have noted that the processing of optical satellite imagery to extract inundated areas is relatively straightforward (Ogilvie et al. 2015; Lacava et al. 2019; Oliveira et al. 2019), cloud persistence during flood events hinders their systematic use in flood monitoring (Huang et al. 2018a). Per contra, Synthetic Aperture Radar (SAR) sensors which use active imaging techniques are capable of cloud penetration and capturing observations

through the day and night. SAR sensors are thus uniquely suited to flood mapping applications, especially for small to medium catchments, where flood events are quick and inundation often recedes before clouds have dissipated (Schumann and Moller 2015). Moreover, for narrow channels the surrounding topography can also significantly impact the echo shape returned to radar altimeters, adding uncertainties to the corresponding water surface elevation (Grimaldi et al. 2016). This makes SAR-derived WLs more suitable than radar altimetry for channels with sub-kilometre widths (Schumann et al. 2008c, d, 2010). Consequently, satellite-based flood extents can be overlaid on DEMs, to obtain shoreline WLs after a number of post-processing steps (Schumann et al. 2007a; Matgen et al. 2007a; Mason et al. 2012b). In fact, comparisons of WLs derived from LiDAR and topographic contours with those derived from the global SRTM DEM, have demonstrated that even coarse resolution DEMs have the potential to support flood modelling (Schumann et al. 2008b). Recent years have witnessed the development of a variety of data assimilation techniques, designed to enhance the exploitation of the available satellite-derived WL datasets for flood model improvements (Domeneghetti et al. 2019). However, research in this direction has also revealed a number of caveats in the WL derivation process, that must be effectively addressed to unlock the full potential of EO-data for flood disaster management (Schumann and Domeneghetti 2016).

This section provides a discussion on the issues related to the integration of EO-data with hydraulic flood forecasting models within a data assimilation framework, with an aim to harmonize and present the current progress in finding feasible solutions. Therefore, it is important to note that this section focuses exclusively on the use of EO-based hydraulic data assimilation to improve predictions of floodplain inundation extent, water level, and flow velocity. This implies that studies on the assimilation of satellite altimetry for improved discharge modelling, for example, are not within the scope of this thesis and accordingly have not been covered. Consequently, an overview of the state of the art of data assimilation techniques proposed for flood inundation model improvements is presented first. Next, the estimation of relevant hydraulic model states and fluxes using flood data assimilation is described. Furthermore, some selected case studies from recent literature which demonstrate the potential of EO-based data assimilation for improved flood inundation forecasting are presented. Finally, future work and open research questions in EO-based data assimilation for flood applications have been elucidated.

2.2.1 Assimilating Earth Observations into Hydraulic Flood Forecasting Models

Disentangling and dealing with multi source uncertainty in modelling is widely accepted as a scientific challenge in hydrological and hydraulic modelling (Blöschl et al. 2019). Accordingly, most hydraulic data assimilation studies have employed EO-data assimilation with the objective towards improved flood forecasting accuracy. As this requires dynamically updating the model trajectory at each assimilation time step, as and when a new observation becomes available, most of the studies have employed sequential filtering techniques (Andreadis 2018). A few studies have also investigated the potential of variational data assimilation approaches such as 4DVAR; although the main objective of these studies was not improving forecast accuracy in real time. These studies primarily focused on the selection of effective inputs and parameters in hindcasting scenarios, which implies that operational constraints of rapid processing times were largely ignored. For example, Lai & Monnier (2009) used EO-derived water level assimilation based on the 4DVAR algorithm for input updating; while a lumped friction parameter was estimated in Hostache et al. (2010) using the same approach. Within a similar 4DVAR-based assimilation framework, Lai et al. (2014) introduced the possibility of direct flood extent assimilation – eliminating the need for a water level processing step – for distributed channel roughness estimation. Analysis of the application of variational DA can also be found in Fletcher (2018a, 2018b), Lahoz et al. (2010), and Briggs et al. (2013). However, as this chapter is focused towards the use of EO-based assimilation for forecasting applications, variational data assimilation strategies have not been considered for a more detailed discussion. Accordingly, the following paragraphs focus on studies that used sequential filtering approaches with a clear objective towards improving forecast accuracy in real time.

2.2.1.1 Water Level Assimilation using the Kalman Filter and Variants

The earliest studies which investigated the potential of EO-based assimilation for flood variable estimation, typically used the Kalman Filter (KF) and its variants (Evensen 2003). The Kalman filter family of data assimilation approaches forecast the background error covariance matrix to calculate the Kalman gain, which determines how much of the observational information is taken up at each time step. This gain matrix is determined by the relative confidence in the model and the observations. The traditional Kalman filter approach achieves

this using a standard error propagation theory on the tangent linear model, while the extended Kalman filter (EKF) uses a Taylor's series expansion for linearization (Walker and Houser 2005). EKF allows the extension of KF to non-linear modelling problems, however, the computational cost exceeds model run times by a factor of one more than the number of assimilated observations (Habert et al. 2016). Here the computational cost is measured in terms of the number of binary matrix operations required at each time step to propagate the assimilation system forward. The Ensemble Kalman Filter or EnKF, computes the error covariance matrices from an ensemble of state forecasts using the Monte Carlo approach, significantly reducing computational cost and the required ensemble size (Madsen and Canizares 1999).

Although the studies discussed in this paragraph did not directly use EO-based flood variables for assimilation, they have been included here as they paved the way for future flood data assimilation studies. The first hydraulic data assimilation studies were by Madsen & Skotner (2005) and Neal et al. (2007), who assimilated ground gauge-based river level data at different points along river reaches. Madsen & Skotner (2005) developed a novel hybrid assimilation technique combining a simplified Kalman filter with an error forecast model, using gain functions with predefined shapes that reflect typical error correlation structures along the reach. Neal et al. (2007) used the EnKF to simultaneously update the states and inputs of a 1D-hydrodynamic model. One major structural assumption in both studies was that the boundary conditions were the sole source of uncertainty. Both studies sequentially updated an augmented state vector including discharge and stage (Madsen and Canizares 1999), as well as future state error covariance matrices which were also parameterised through the assimilation scheme (See Fig. 2.5). The error forecasting procedure allowed for an update of the inflow boundary conditions, which are known to strongly influence flood flow regimes between subsequent assimilation steps. As the errors introduced through the inflows at each time step, are pre-emptively estimated and accounted for in the assimilation framework, the forecast error reduction obtained is persistent. An autoregressive error model was used to synthetically generate and subsequently predict temporally correlated inflow errors. On testing a variety of temporal sampling intervals for field hydrometric observations of water stage, Neal et al. (2007) concluded that the state uncertainty was overestimated when the sampling rate was reduced. This indicated that the observations were assigned relatively lower weights than model predictions, when the temporal density of the measurements was low. Similarly, the spatially

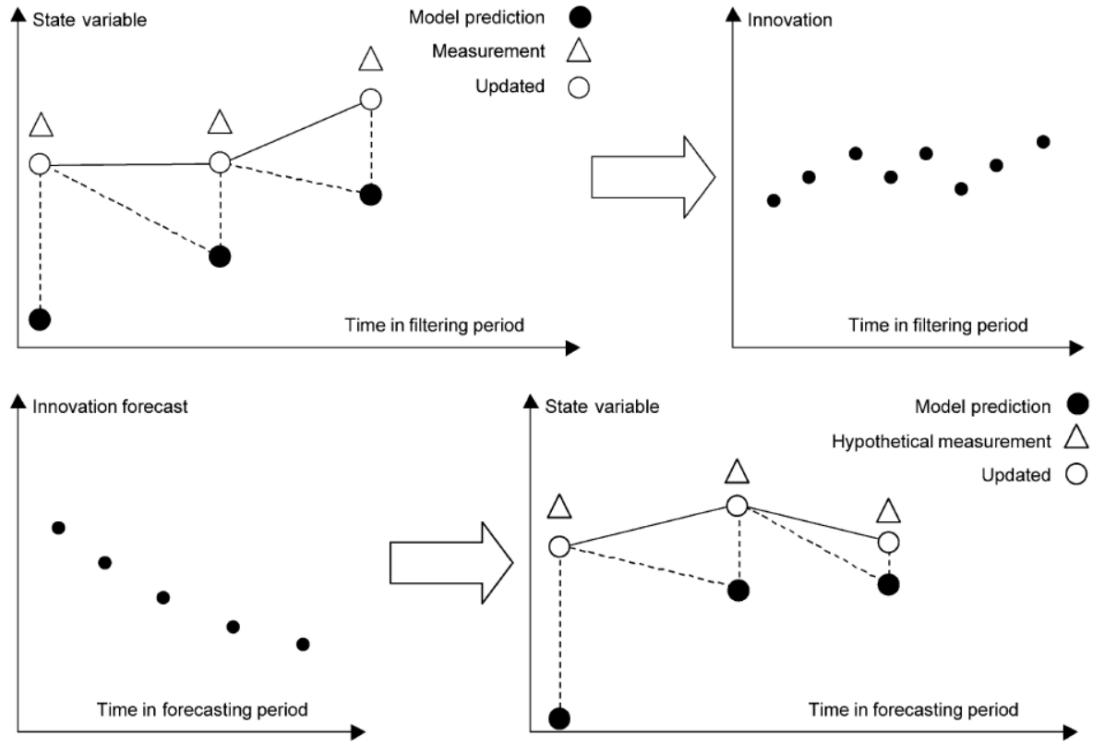

Fig. 2.5 Illustration of the combined filtering and error forecast procedure followed by Neal et al. (2007) and Madsen & Skotner (2005). Taken from Madsen & Skotner (2005), © Elsevier, 2005.

distributed systematic bias in the field sensor measurements led to an underestimation of the state uncertainty.

The following study on hydraulic data assimilation was actually geared towards the feasibility assessments of the proposed SWOT mission, then known as WatER - Water Elevation Recovery (Alsdorf et al. 2007). Through the assimilation of synthetic swath-altimetry data into the raster-based hydraulic model Lisflood-FP (Bates and De Roo 2000b) using the square root EnKF or EnSRF (Evensen 2004), Andreadis et al. (2007) investigated the prospects for estimating channel flow and water depth. Errors were propagated from boundary inflows, generated by corrupting outputs obtained from the Variable Infiltration Capacity (VIC) model. Given that this was a feasibility study, synthetic spatial fields of WL were generated using the NASA JPL SWOT simulator (Fu and Rodriguez 2004) by corrupting predictions from the "truth" model with Gaussian errors. A hydraulic data assimilation for EO-data was developed for the first time by Andreadis et al. (2007), to simultaneously update states and inputs by dynamically parameterizing the inflow error prediction model, following Madsen and Skotner (2005). The EnSRF-based approach was successfully able to retrieve channel water

depth and discharge from the corrupted open loop simulations. While the filter showed low sensitivity to assumed observation errors, it was highly sensitive to observation frequency; assimilation performance degraded substantially as the temporal frequency was halved.

The study by Matgen et al. (2007) considered real EO-derived WL observations, as opposed to the synthetic ones used by Andreadis et al. (2007), within a sequential data assimilation framework for the very first time. As state updating is the most straightforward way to utilize EO data to correct models in real time (Grimaldi et al. 2016), this proof-of-concept study implemented a similar strategy. Water stages were derived from ERS-2 SAR and ENVISAT ASAR images, by intersecting with a LiDAR DEM using the REFIX approach of Schumann et al. (2007). Considering all the sources of uncertainty in EO-based WL estimation, an interval of values is determined from the EO-based WL estimation approach and further refined by enforcing hydraulic coherence principles (Puech et al. 2007). Subsequently, the model generated water line is updated at the time steps of the EO data acquisition, to fall within the interval of EO-derived WLs at each model cross-section (Matgen et al., 2007). As the calibrated 1D HEC-RAS model already performed very well for the study reach, the reported improvements were minor and rapidly decayed within a few hours. For instance, the assimilation of ERS-2 based WLs improved forecast error by 23 cm at the assimilation time step, but this reduced to 3cm within 3 hours of the assimilation, after which the analysis trajectory completely decayed back to the background state values within ~4.5 hours. In contrast, the assimilation of ENVISAT-derived WLs increased the deviation of the assimilated state vector from the open loop from 3 cm to 6 cm, which was an almost 100% relative degradation. The decay to the original model trajectory in this case was faster (~2 hours). These results can primarily be attributed to the flashy catchment characteristics, where the flood wave peaked in a few hours and receded within three days. This study particularly highlighted that EO-based data assimilation can even degrade model predictions when observation errors are not appropriately represented. The implication here is that some *a priori* quality control might be necessary to effectively filter outliers which may negatively impact forecast errors.

Extending the work of Andreadis et al. (2007), Durand et al. (2008) used data assimilation to estimate bathymetric depth and slope, through an inverse modelling approach. This study used Lisflood-FP (Bates & De Roo 2000), and errors in the upstream inflow, DEM, bathymetry, and the channel roughness parameter for the first time. While there was no explicit objective towards flood forecasting, the realistic uncertainty scenario assumed in this study,

made it a rather significant contribution to flood data assimilation literature. Bathymetric slope, channel roughness, and upstream inflow uncertainties were modeled as multiplicative log-normal error (MLNE) distributions, with unit mean and different coefficients of variation. Topographic uncertainty in floodplain DEM is modeled as a zero mean additive normal distribution, while bathymetric depth was estimated from the slope ensemble previously described. A novel assimilation scheme which used Monte Carlo techniques to modify a linear parameter estimator for non-linear cases was designed. Synthetic SWOT WL fields were assimilated to obtain the reach-averaged bathymetric slope and point-based bathymetric depth. Using the assimilation scheme, the reach-averaged slope was estimated within 0.30 cm/km of the truth, while an improvement of 84% was observed at the downstream point estimate of channel depth. The experiments detailed in this study corroborated the low measurement error sensitivity observed by Andreadis et al. (2007). This phenomenon was primarily attributed to the higher-magnitude model errors, which likely dominate over measurement errors within the assimilation framework.

Similarly, Neal et al. (2009) built upon the earlier body of work to estimate discharge, using a combination of hydrodynamic modelling and EO-derived EnKF-based WL assimilation. This study considered the impact of incorrect bathymetry specification on the predicted flows by simulating a data limited scenario, where the channel cross-sections were simplified to a trapezoidal shape for the first time. The use of simple channel geometry led to an underestimation of hydraulic conveyance, resulting in higher predicted WLs for a given flow value. This is caused by an underrepresentation of river carrying capacity, which pushes the water surface elevations higher than they would actually be in the field. However, the assimilation of EO-based WLs was able to effectively improve the retrievals of observed discharge, even in the data scarce scenario where simplified channel geometries were considered. This study corroborated the findings of Matgen et al. (2007), who demonstrated that SAR images may not provide useful information across the entire model domain. It was also noted that forcing the model to reproduce observations at certain locations, may introduce measurement bias leading to model performance degradation.

The studies which exclusively used state updating or evaluated its impacts with respect to input updating, highlighted that the analysis vector rapidly decayed back to the background state trajectory immediately following the assimilation (Andreadis et al. 2007; Matgen et al. 2007b). Schumann et al. (2009) attributed this lack of persistence in accuracy to the dominating

effect of upstream boundary conditions in hydraulic modelling. They argued that as long as errors in the initial conditions and forcing data persisted, merely reinitializing the model with an updated state vector may not serve as an adequate solution. Based on these findings the simultaneous updating of states and inputs was recommended to obtain more persistent forecast error reductions.

Furthermore, the work of Schumann et al. (2008) showed that SAR-derived WL observations, mostly exhibit non-Gaussian probability density functions (PDF) at each cross-section. It therefore follows, that the distribution from all the cross sections considered for the full river reach, is also strictly non-normal in form (Kitagawa 1996). The EnKF algorithm assumes that the forecast and measurement error covariance matrices are jointly normal to make the data assimilation problem tractable; such that their respective PDFs can be sufficiently characterized using only the first and second moments of the distribution, i.e., mean and covariance (Moradkhani et al. 2005b). This raised questions about the theoretical justification of employing EnKF-based assimilation schemes for EO-based hydraulic data assimilation, where states are subject to unknown disturbances (Moradkhani et al. 2005a). The conditional probabilities of strongly non-linear model trajectories, can only be characterized sufficiently by tracking the higher order moments of the distribution (Moradkhani 2008). However, the spatial coverage of EO-data implies that some observations can be discarded, for example, those that fail a normality test as suggested by Neal et al. (2009). While simple and easy to implement, this approach might reject potentially useful information. Further studies then investigated possible ways to solve the problem of obtaining persistent improvements from hydraulic data assimilation while accounting for non-Gaussian uncertainties.

2.2.1.2 Water Level Assimilation using the Particle Filter and Variants

Particle Filters (PF) relax the assumption of Gaussianity regarding the forms of the posterior probability density functions of the models and observations, offering certain advantages over KF and variants (Moradkhani 2008). This enables PFs to easily manage the propagation of non-Gaussian distributions through highly non-linear hydraulic models, by tracking the temporal evolution of all the moments of the full probability density function (Plaza et al. 2012). This is achieved by updating the probabilities of any given model being true conditioned on the observation, rather than updating the state trajectory (Arulampalam et al. 2002). Exclusively updating the PDFs also provides the unique advantage of avoiding the

hydrostatic reinitialization shock, known to cause numerical instabilities in the hydraulic model domain, as fluxes drop to zero at the assimilation time step and momentum cannot be conserved (Hostache et al. 2018b).

Early implementations of the PF used the Sequential Importance Sampling (SIS) algorithm. The SIS algorithm used a Sequential Monte Carlo procedure to approximate the posterior true state PDF, through a number of independent random samples called particles, sampled directly from the state-space. Subsequently, the conditional probability of a given model realization being true given that the observation is true, is assigned as a weight to each particle and used to compute the weighted ensemble mean. Starting with a likely proposal distribution, weights and estimates are sequentially updated at every assimilation time step and the weighted mean vector is calculated, which is representative of the "analysis". PFs conserve mass for each particle unlike the EnKF, by selecting the most likely model runs rather than updating model states, which may cause discontinuities in the hydraulic model domain (Matgen et al. 2010).

The SIS algorithm, however, suffers from several significant limitations. Firstly, there is the problem of particle degeneracy, where most of the particles attain very low weights after a few assimilation steps, leading to an underrepresentation of the state-parameter space in the approximated posterior PDF. In extreme cases, only one particle will acquire the full weight of unity leading to filter collapse, implying that even the equifinality of hydraulic model states and parameters is no longer characterized. The second problem pertains to the ensemble size or particles required for an effective estimation of the posterior PDF. As the true state PDF approximation is dependent on discrete random sampling, it is understandable that the estimation will improve as the number of samples in increased. In theory this approximation process will work best as the number of particles tend to infinity; this is obvious, as discrete sampling is being used to characterize a continuous process. However, this is nearly impossible to achieve in practice, due to the huge computational demand of distributed hydraulic models.

Literature suggests several pragmatic solutions. Moradkhani et al. (2005a) suggest a variety of ensemble verification measures, to pre-emptively assess the skill and spread of the ensemble. Several resampling schemes have also been proposed, such as the Sequential Importance Resampling (SIR) algorithm. The SIR algorithm repopulates the particles by replicating the highly ranked particles, in proportion to their respective importance weights. In some cases simplistic solutions such as these may work, however, in most cases the resampling

leads to the problem of sample impoverishment. This is when the replicates represent only a few unique model realizations, yet the effective particle population remains spuriously high, again violating the principles of parameter equifinality. Techniques such as inflation which artificially scale the error covariance matrices, to account for the underrepresentation of variance due to small ensemble size, have also been proposed and can be used for sequential as well as variational assimilation (Evensen 2003; Slivinski et al. 2015; Browne 2016; Fletcher 2018b). Most studies use one or several of these in conjunction, to avoid the problems of filter collapse.

In this context, the study by Matgen et al. (2010) used the PF-SIR algorithm to assimilate synthetic EO-derived WLs into a coupled hydrological-hydraulic model. The PF-SIR algorithm was chosen primarily to relax the assumption of Gaussianity. However, as there were no alternative theoretical distributions yet identified in literature, only the sub-sample of observations that passed a normality test were used. Global weighting procedures such as the one used here, consider WLs simulated by a particular model realization along the entire channel reach as one particle. Weights at each cross-section were computed using a Gaussian PDF and aggregated through multiplication, by assuming (perhaps unrealistically) mutual statistical independence. The PF-based assimilation was able to retrieve the truth, even when errors of up to 5m were introduced in the synthetic WLs. This implied that for completely ungauged catchments, WLs extracted from a combination of coarse resolution satellite data and global DEMs, could also be useful to effectively constrain flood forecast errors (Schumann et al. 2008a, c). Consistently with previous findings (Andreadis et al. 2007; Matgen et al. 2007b), input updating was identified as a crucial aspect in retaining improvements to the model state trajectory. Interestingly, the study also found that a higher frequency of observations is required during the rising limb than during recession, as the errors in the precipitation are unpredictably compounded through the modelling cascade, resulting in increased model forecast errors.

In order to address the gap of defining a suitable spatially and temporally variable non-Gaussian distribution of observations, Giustarini et al. (2012) suggested using the full empirical distribution of WLs. Employing the empirical distribution with no assumptions about its form, finally allowed the optimal utilization of the advantages of PF over KF variants. At each cross-section, a histogram of the EO-derived WLs is computed, which is subsequently used to define the local likelihood of each particle. The final aggregation to global weights is achieved by following the approach proposed by Matgen et al. (2010). Two different case studies were

presented in Giustarini et al. (2012), using coarse and high-resolution EO datasets respectively; temporally persistent improvements in discharge estimation were obtained for both cases. Numerical experiments performed in this study, interestingly revealed a trade-off between WL and discharge estimation accuracy, particularly in regions where channel bathymetry is poorly specified.

2.2.1.3 Filter Localization and Flood Extent Assimilation

In a pioneering study which highlighted the possibility of degrading model forecasts through assimilation for the first time, Giustarini et al. (2011) demonstrated that the use of global weighting procedures always leads to compromise solutions. Using synthetic and real experiments, assimilating field data at the cross-sections where EO-derived WLs were available, this study argued that a single model run cannot perform equally well along the entire river reach. In fact the study showed that defining global weights as a product of local weights as proposed by Matgen et al. (2010), could even lead to the propagation of local scale systematic model errors over the whole domain. Global weights were found to favour acceptable solutions all over the domain, rather than well performing solutions locally. To avoid this problem, Giustarini et al. (2011) proposed the use of filter localization. Using a model setup identical to Matgen et al. (2010), the assimilation was implemented by considering the model simulated WL at each cross-section as a separate particle. Local weights were calculated for each particle, by comparing with the corresponding EO-WLs, using uniform as well as normal PDFs. While this study advocated the use of filter localization techniques going forward, it also warned of potentially introducing bias into the model predictions, as a function of local inconsistencies in the calibrated model. The use of local PF for diagnosing and correcting model errors was recommended for hindcasting applications, while inflow correction was identified as a major research gap for forecasting problems.

Filter localisation only updates the states close to the observation location, reducing the weight given to observations spatially disconnected from the estimated state variable (See Fig. 2.6). The impact of using small ensemble sizes can then be minimized, primarily by decreasing the state-parameter subspace in which the analysis is conducted (García-Pintado et al. 2013). The study by García-Pintado et al. (2015) also investigated this issue through a series of SAR-derived WL assimilation experiments, based on the Ensemble Transform Kalman Filter (ETKF) with and without localisation. ETKF (Hunt et al. 2007) is an adaptation of the square root form

of the EnKF. More specifically, the EnSRF implementation explicitly updates both the ensemble mean and the error covariance at each assimilation time step to theoretically match the Kalman filter. The ETKF additionally introduces a transform matrix to ensure a symmetric solution (García-Pintado et al. 2013). Results obtained by García-Pintado et al. (2015), corroborated the need for filter localisation to avoid the development of spurious correlations, within the forecast error covariance matrix (García-Pintado et al., 2014). The use of a flow distance based spatial domain localization metric was recommended, although further testing for channel networks with varied geometries is required. Simultaneous estimation of states and inputs proved necessary for persistent reductions in forecast error as in earlier studies (Neal et al. 2009). The use of localised ETKF for simultaneous parameter estimation (lumped channel roughness and distributed bathymetry) was also investigated, although the results proved to be inconclusive in terms of forecast improvements. Possible reasons could be the localization distance or the error covariance chosen for the stochastic generation of bathymetry estimates, as the standalone experiments on channel friction demonstrated adequate convergence. The findings highlighted that further experiments might be warranted to effectively tune the localization radii and the reach correlation length used in the bathymetry error generation model for improved forecast accuracies. Experimental results were corroborated by the findings of other studies, where empirical localization was found to significantly improve state estimation (Yamazaki et al. 2018; Revel et al. 2019).

As most studies observed and Schumann et al. (2009) incisively pointed out, the problem of local uncertainties stemming from the measurement bias of EO-derived WLs, required the development of more mature retrieval methods (García-Pintado et al. 2013; Andreadis and Schumann 2014). Moreover, the use of the same DEM both to retrieve the EO-derived WLs and for the implementation of the hydraulic model leads to adding the DEM uncertainty twice in the resulting predictions. While this might be an acceptable risk where highly accurate LiDAR-based elevation data are available, in the case of coarse resolution global topography this could severely impact the forecast error. Moreover, using the same elevation data additionally violates a basic principle of data assimilation, which mandates the use of model independent observations to improve the accuracy of subsequent predictions. Furthermore, even with the advent of sophisticated thinning algorithms for the automatic derivation and appropriate subsample of EO-WLs (Mason et al. 2012b), manual intervention is often still necessary (Hostache et al. 2018b).

Recent studies have, therefore, focused on the development of techniques capable of directly assimilating flood extents into flood forecasting modelling cascades, rather than water levels (Lai et al. 2014; Revilla-Romero et al. 2015, 2016; Hostache et al. 2018b; Shastry and Durand 2019). Lai *et al.* (2014) and Revilla-Romero *et al.* (2016) both interpret inundation extents as a function of the internal model states to develop the cost function for assimilation, i.e. water depth and discharge, respectively. The 4D-Var filter was used in the proof of concept study by Lai *et al.* (2014), for the assimilation of a MODIS-derived flood extent map (250m) to optimize a lumped friction parameter, with no particular aim towards forecasting. In contrast, Revilla-Romero *et al.* (2016) use the EnKF to assimilate dimensionless surface water extent observations (0.1°×0.1°) satellite-derived from the Global Flood Detection System (http://www.gdacs.org/flooddetection), to improve near real time (NRT) global flood forecasts. The EO-derived flood extents were interpreted as inundation volumes; the difference between the simulated and observed (interpreted) values, was used to update the models within the EnKF framework. The study was conducted on a global scale and gauge validation demonstrated improvements for a major portion (~60%) of those evaluated. Although the use of EnKF was perhaps not theoretically justified, significant improvements were noted in practice.

Adopting a different approach, Hostache *et al.* (2018) directly assimilated SAR-derived flood extents using a PF-based assimilation framework. Through a binomial comparison of modelled cell wet-dry status and satellite observed flood probabilities, improvements of up to 50% were obtained in forecast WLs at the assimilation time step. Inundation extents were retrieved from ENVISAT ASAR (resampled to 75m) into the hydraulic model LISFLOOD-FP. As modelled flood extents are derived based on simulated water depth, the variation in the number of wet-dry cells at most time steps is limited. Consequently, the development of an extent based cost function with enough sensitivity to isolate the best performing ensemble members and drive the assimilation is a scientific challenge (Lai et al. 2014). Hostache *et al.* (2018) used a number of pragmatic mathematical solutions to facilitate the direct assimilation of flood extents in a real case for the first time. For example, local weights were first computed on a pixel-by-pixel basis for each ensemble run or particle, by comparing the modelled vs. observed cell wet-dry status through the binomial distribution. The joint probability density of these pixel-wise weights is then used as the global weight of the particle, which is aggregated as a product of all pixel weights (Matgen et al. 2010; Giustarini et al. 2011). This method of calculating the joint probability density makes the assumption that the local pixel-wise weights are statistically independent (Hastie et al. 2009). As raster-based models numerically compute

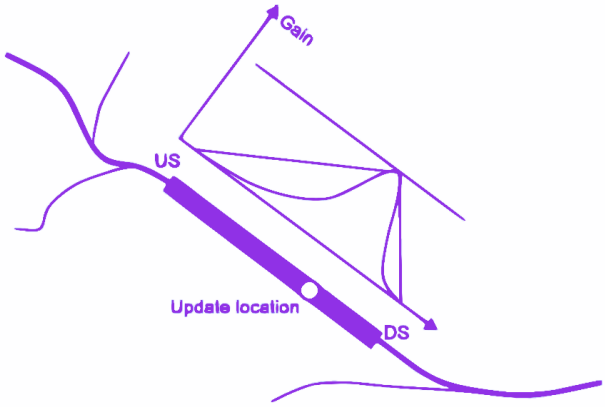

Fig. 2.6 Illustration of filter localisation for an example observation update location in 1D, adapted from Madsen and Skotner (2005). US = upstream and D S = downstream.

flow between cells, the values of simulated flow depths that are used to derive the wet-dry cell status cannot be considered independent. Moreover, flooding itself is a spatial phenomenon implying interdependence between modelled grid cells. Thus, the development of novel flood extent assimilation techniques to optimally combine flood inundation models with EO-flood extents still remains an open scientific challenge (Van Wesemael et al. 2019).

2.2.2 Observation Operators and Characteristics

One common issue highlighted by previous studies was the possible degradation of forecast skill, caused by the introduction of measurement bias by highly erroneous observations. Although the issue of observation timing (Matgen et al. 2010; Giustarini et al. 2011) and frequency (Andreadis et al. 2007; Neal et al. 2009) was briefly discussed in a few studies, the first explicit discussion and assessment was provided by García-Pintado et al., (2013). Similarly, Andreadis and Schumann (2014) evaluated assimilation performance sensitivity to the observation spatial location for the very first time. Further studies in this direction, evaluated the impact of domain length (Cooper et al. 2018), observation error correlations (Waller et al. 2018a), and observation operators used to calculate the innovation (Cooper et al. 2019).

In order to assess the impact of acquisition timing, García-Pintado et al., (2013) used an ETKF based assimilation strategy in conjunction with LISFLOOD-FP, with a synthetic experiment based on the July 2007 flood event in the widely researched Severn Catchment, UK (Mason et al. 2010b; Neal et al. 2011; Schumann et al. 2011). An error free model was

considered with no parametric uncertainty, to independently evaluate the impacts of observation timing on inflow error correction, through several satellite first visit and revisit scenarios. Results illustrated that frequent assimilation during the decreasing limb does not lead to sensible improvement in the forecast, implying that post peak overpass frequency could be reduced when considering budget limited scenarios. However, as satellite flood extents which are used to derive the WLs for assimilation can only be informative for out-of-bank flows, the first visit time should ideally closely follow channel overtopping. Corroborating the findings of Matgen et al. (2010), the results of Garcia-Pintado et al. (2013) demonstrated that multiple observations were necessary during the rising limb of the hydrograph, as additional errors are continuously introduced at the upstream boundary.

Subsequently, Andreadis and Schumann (2014) assessed the spatial observation impact on the forecast skill of hydraulic models, through the LETKF formulation of Hunt et al. (2007) within an ensemble sensitivity (ES) framework (Liu and Kalnay 2008). In the ES method, the LETKF cost function is modified such that each term can be calculated from the previously initialized ensemble forecasts. This eliminates the need for generating forecasts after the assimilation step. The cost function can be pre-emptively computed for each observation time and for different lead times, simply by selecting the appropriate time steps for the forecasts and observations (Andreadis and Schumann 2014). Using a continuous annual large-scale implementation of LISFLOOD-FP for the Ohio River, this seminal study showed for the very first time, that observations acquired during low flows consistently degrade forecast error. A synthetic study was conducted where model errors stemming from inflows, parameters, topography, and channel network delineation were considered. The observations were not explicitly assimilated, rather their impact was evaluated by just replacing the state and observation variables in the cost function equation. This allowed previously unused observations, such as river channel width and inundated area, to be considered alongside water surface elevation. On assimilating WLs, the largest improvements in forecast accuracy could be obtained during peak flows, although the error reductions decayed after 5–11 days in some locations and eventually led to forecast degradation. Interestingly, the findings indicated that the assimilation of channel top width mostly degraded forecast skill, while inundated area assimilation mostly resulted in improvements. Further examination revealed that the improvements possible from the assimilation of either observed variable, could be considered as a direct function of the variability (e.g., see Fig. 10 in Andreadis and Schumann (2014)). Authors also recommended the use of these methods to plan targeted satellite acquisitions, over

reaches that would either maximize the forecast error reduction or from the locations of fastest growing errors (Langland 2006). As this approach was tested with an almost real-world uncertainty scenario, it could potentially be operationally implemented to allow forecast improvements by incurring minimum possible costs.

The study by Cooper et al. (2018) showed that the hydrostatic reinitialization shock, caused by the fluxes dropping to zero just after the update step in hydraulic data assimilation, could be minimized by applying pre-assimilation velocities. This pragmatic solution was able to significantly reduce the root mean squared error values obtained just after the assimilation time. Using an idealized synthetic domain topography, the study also demonstrated that persistence of error reductions from WL assimilation could be improved by simply considering a longer domain length. The WL observations in different parts of the domain are usually strongly correlated, thereby allowing the ETKF to effectively update downstream WLs. State augmentation was used to simultaneously update model states and parameters, resulting in consistent improvements in forecast accuracy. This finding also suggests an interdependence between parametric and inflow errors, due to similar forecast error characteristics obtained from both. Furthermore, the joint-state parameter estimation using EO-derived WLs, can effectively detect and compensate for biases introduced in either or both.

In a study similar to Andreadis and Schumann (2014), the impact of observation quality on flood data assimilation was investigated by Waller et al. (2018). Through numerical experiments, the authors show that statistical averages of observation-minus-background and observation-minus-analysis residuals, can be useful for the estimation of error correlations in EO-derived WLs. Using the same flood event as García-Pintado et al. (2015), an analysis of observation error and spatiotemporal correlation was elucidated. Spatial analysis showed that the observed forecast error correlations were independent of observation errors. In terms of observation timing, the observations exhibit similar error standard deviations although the correlation length is fairly short. This is expected as this is the most dynamic part of the flood event, where ongoing precipitation is continuously adding errors to the domain while each tributary and sub-catchment respond differently. As the flood wave progresses, standard deviation decreases while the correlation length scale increases. Given that such observation error assessments are possible at a relatively low computational cost, the authors recommend its use in the identification of data anomalies.

Sensitivity to observation operators was examined by Cooper et al. (2019) using the same idealized test domain as Cooper et al. (2018). In this study, backscatter from a SAR image was directly assimilated to improve hydraulic model forecasts for the first time. Inflow and parametric uncertainty was considered, in a synthetic experiment based on the 2012 River Severn flood event, to compare the performance of a WL-based and a backscatter based observation operator. Synthetic SAR images were generated using a Gaussian mixture model, using the parameters empirically derived by Giustarini et al. (2016). These parameters were perturbed within 1% of the values to obtain an ensemble of observed SAR images. The number of backscatter observations used was limited to match the number of satellite WL observations that can typically be expected, to facilitate an intercomparison of the observation operators independent of observation frequency. The new backscatter operator generally performs better than the WL operator in all the test cases considered in this study. Using backscatter directly also eliminated the need for multiple image processing steps and the water level derivation, which has been identified as a major source of uncertainty (Schumann et al. 2009a). The authors acknowledge that the new operator can exclusively work well in conditions where a clear separation exists between the land-water distributions of backscatter response. As this is rarely the case in reality where backscatter is affected by multiple sources of uncertainty (Schumann 2019), the effectiveness of this operator has to be tested using a real case study.

2.2.3 Opportunities and Challenges

The high sensitivity of data assimilation algorithms to observation characteristics unlocks several innovative possibilities. Firstly, optimal acquisition strategies can be designed which allow maximum improvements in forecast skill while minimizing costs, based on the relationships between observation characteristics and assimilation performance (Andreadis 2018). Assimilation experiments can be used to inform acquisition planning, thereby optimizing the location, timing, and frequency of satellite-based flood observations to best support hydraulic forecasting of inundation. Similarly, assimilation can help to diagnose localized discrepancies, in model implementation or even parameterization.

In certain scenarios, integrating observations acquired at specific timings or locations, have no impact on the assimilation and can even lead to model degradation. This phenomenon is typically a function of localized dominating model errors, although in rare cases observation errors or the assimilation algorithm itself could also lead to similar problems. For example,

Schumann and Andreadis (2016) insightfully utilized this particular feature to objectively identify reaches which could especially benefit from high resolution topography. Using the LETKF-based ensemble sensitivity approach employed by Andreadis and Schumann (2014), optimum locations for LiDAR acquisitions along the Lower Zambezi River were identified, such that maximum improvements in accuracy could be obtained at minimum costs. The authors also demonstrate local improvements of up to 78% in risk estimates, which is substantial from an emergency management perspective.

In spite of the significant advances made towards hydraulic data assimilation of EO-data for improved operational food inundation forecasting, several areas still require extensive research. An objective estimation of the computational demand versus accuracy trade-off, considering the precision needs of the various stakeholders involved, is the first challenge which needs to be addressed. The representation of model uncertainties ideally requires a large number of simulations to effectively explore the state parameter space. However, in practice each hydraulic model run on a global or even continental-scale, may require several hours of computational time depending on the grid resolution. As Schumann et al. (2014) demonstrated, higher resolution models are required to efficiently resolve local scale flow pathways. Studies have shown that higher resolution modeling is now possible at much larger scales, as a consequence of the giant strides in scientific computing (Schumann et al. 2013; Dottori et al. 2016; Quinn et al. 2019; Uhe et al. 2019). Additionally, nested modelling approaches could be used to maximize computational efficiency. The work of Mason et al. (2015), for instance, shows that it might be possible to pre-emptively diagnose, areas of large vertical errors in the specified DEM. Subsequently, these highly uncertain subdomains of the hydraulic model, can serve as the focus of local scale high resolution modelling using LiDAR acquisitions, significantly improving flood forecast skill (Schumann and Andreadis 2016; Fleischmann et al. 2019).

Another significant challenge for effectively utilizing EO-based data assimilation for improved flood forecasts, is the adequate representation of anthropogenic factors like hydraulic infrastructure in models (Andreadis 2018). Large hydroelectric or irrigation dam reservoirs, often form a substantial portion of catchment storage capacity, yet data on reservoir operations is seldom available in the public domain (Do et al. 2010). Although alternative methods to determine factors like irrigation supply and demand are currently being explored in literature (Brocca et al. 2018; Zaussinger et al. 2018), significant uncertainties are remain in the accurate

estimation of flood regulation measures (Schumann et al. 2018). Moreover, the implementation of flood control structures in the DEM is absolutely vital, to capturing the actual inundation patterns observed on ground. In fact, misrepresentation of structures can be especially expensive in terms of risk management, in the rare case of a breach which can suddenly inundate large regions. Attempts are underway to develop global maps of control structures using EO-data, although several difficult challenges still remain (Wood et al. 2018).

The integration of other diverse datasets, for example point observations of water levels and velocities from crowdsourcing or depth Doppler maps of inundation from GNSS reflectometry, is another significant challenge that faces the flood data assimilation community. Using alternative datasets requires an objective estimation of the associated uncertainty and some data sources such as crowdsourcing are still rather young in this direction. A lot of research is ongoing, where the potential utility of crowd-sourced and citizen science observations for flood inundation modelling is being investigated (Assumpção et al. 2018; Zheng et al. 2018). Significant progress has been made in this direction, but the filtering and standardisation of the available data remain challenging (Mazzoleni 2016). Alternative observations of flow velocities are also becoming available (Fujita et al. 2007; Muste et al. 2011; Assumpção et al. 2018), which soon might be available to the flood data assimilation community. There is also a need to integrate the research progress in the assimilation of soil moisture (Patil and Ramsankaran 2018) and evapotranspiration (Hartanto et al. 2017) in hydrological models with hydraulic data assimilation frameworks, such that the utilization of EO-data for flood inundation forecasting can be optimally harmonized.

The final challenge is the translation of flood extent and depth estimates, into hazard and risk estimates, which are more relevant to the stakeholders (Ward et al. 2015). In fact open source/free models exist, which can integrate hydraulic model outputs with socioeconomic risk factors, e.g. the GLOFRIS (Global Flood Risk) framework (Winsemius et al. 2013) or HAZUS (https://msc.fema.gov/portal/resources/hazus) provided by FEMA (Federal Emergency Management Agency, US). Yet, the delivery of maps and inundation forecasts is often only in scientific terms. Open Street Maps (OSM) which use local knowledge of citizen scientists, to maintain a detailed global vector database of road/rail/waterways and buildings, is yet another rich resource which could potentially revolutionize flood emergency management. Maps of inundation could be intersected with the richly detailed OSM GIS layers, with structured crowd-sourced local information about assets at risk (Ward et al. 2015). The humanitarian benefits of

the progress in flood inundation modelling can only be quantified, through an evaluation of the impacts of improved forecasts on corresponding risk and hazard estimates (Pappenberger et al. 2015).

Schumann and Andreadis (2016) conducted such analyses for the first time in flood inundation modelling literature, although in streamflow forecasting this is a widely recommended practice (Cloke et al. 2013; Wetterhall et al. 2013; Pappenberger et al. 2015). In this context, the High-Resolution Settlement Layer (HRSL) developed by Facebook in association with the Centre for International Earth Science Information Network (CIESIN) can also prove to be a valuable resource. HRSL provides gridded population density estimates for 33 countries, at a ~30m resolution for the year 2015, based on recent census data and high-resolution 50cm DigitalGlobe imagery (https://www.ciesin.columbia.edu/data/hrsl/). This dataset could be readily used to quantify the change in risk estimates resulting from improved flood data assimilation techniques. Developing comprehensive forecasting frameworks which synergize the progress in Earth Observation, numerical modeling, data assimilation, and scientific computing to enhance global flood hazard estimates, could significantly increase flood resilience. In light of urbanization and climate change exacerbating the number and impacts of extreme weather events, developing effective techniques to minimize global flood hazard are urgently required.

2.3 Research Gaps

The current generation of high resolution SAR satellites like COSMO-Skymed, TerraSAR/TanDEM-X, Radarsat Constellation Mission, and Sentinel-1, and planned missions like NiSAR, SWOT (Surface Water Ocean Topography), TerraSAR/TanDEM-L, and CSK-2, have stimulated scientific research on the optimal use of this rich database for flood detection. The shift towards open data policies across space organizations could potentially lead to more innovative solutions to the persisting challenges in the field. Moreover, the continued investment in satellite SAR technology makes the development of model-data integration techniques quite essential for the future.

Identifying flooding under vegetated and urban land forms is a research topic of high interest with many open questions currently being examined. The use of ancillary datasets and complementary approaches like interferometric coherence can be helpful. However, diagnostic

assessments are required to evaluate the environmental conditions under which each of these approaches work and why. Furthermore, it's important to develop techniques that provide an associated measure of flood detection uncertainties, at least in regions exhibiting complex surface properties. A synthesis of all the relevant literature in this direction revealed the following gaps:

1. The shift towards fuzzy and probabilistic approaches in the field of SAR-based flood mapping necessitates the development of suitable validation techniques. Present strategies involve the use of binary validation maps for assessment, which seems counterintuitive, as the premise of this shift from deterministic approaches is that uncertainty cannot be eliminated. Investigative analysis of current mapping approaches for multi-sensor, multi-band data should be conducted across diverse geographical regions, to gain an in-depth understanding of why certain approaches favour a particular set of environmental conditions.
2. There is a clear need to develop advanced approaches which can detect flooding accurately from a single SAR image. In this context, it is worth investigating the potential of image features such as texture, which can automatically be derived from the image itself and reduce the dependence on ancillary data. Moreover, some statistical texture estimation approaches can even minimize wind and sensor geometry impacts, which could enhance the transferability of the mapping algorithms.
3. Crowd-sourced flood information has the potential to complement remote sensing based mapping of floods, as the advent of social media and the proliferation of smart phones has ushered in an era of citizen sensing. However, research in this direction is very new and many scientific challenges remain, especially in the objective estimation of uncertainties and the design of appropriate validation methodologies. The availability of data from diverse sources will continue to increase in the coming years and the flood mapping community needs to evolve rapidly to unlock the full potential of RS for disaster management.
4. Diagnostic assessments of localized flood flow behaviour need to be undertaken to identify which portions of the model domain could benefit the most from data assimilation. In fact, the possibility of forecast degradation as a consequence of assimilating in highly erroneous model sub-domains makes answering this question quite crucial.

5. The impact of assimilating highly uncertain observations, needs to be further investigated, especially with respect to completely ungauged catchments. In this regard, efficient methods for quality control of the assimilated observations while retaining the maximum amount of information possible, also need to be designed.
6. Filter localization methods which consider hydraulic relationships to objectively identify optimal regions of observation influence need to be developed, to support operational applications of hydraulic data assimilation. Similarly, advanced techniques to artificially inflate the error covariance need to be considered, which allow a comprehensive representation of model uncertainties even with limited ensemble members.
7. Effective spatial cost functions which are sensitive to subtle changes in the model state variable in question, need to be designed to better quantify the innovation and optimize the possible forecast improvements through assimilation. In this context, techniques to synthetically scale the objective function values such that the sensitivity to changes in state can be enhanced also need to be explored.

Research on hydraulic data assimilation of Earth Observations of floods, is relatively new and many scientific challenges remain. The availability of different flood observations from diverse sources, will only increase in future and the hydraulic data assimilation community needs to evolve rapidly to keep up with the pace of advancements in measurement techniques. Estimates of associated uncertainty will vary based on each observation technique and the measured variable in question. In order to unlock the full potential of Earth Observation data for hydraulic flood inundation forecasting, it is imperative to objectively estimate these uncertainties and remain cognizant of them during the assimilation process (Schumann et al. 2016). Harmonizing the progress in hydraulic modelling, data assimilation, and measurement techniques through the development of operational forecasting systems, is required to ensure the optimal utilization of EO-data and can finally result in tangible humanitarian benefits.

2.4 Chapter Summary

This chapter provided an overview of the state-of-the-art literature in SAR-based flood extent mapping and their integration with hydraulic flood inundation models through data assimilation. First, the challenges and opportunities associated with operational SAR-based

flood mapping at multiple spatial scales were critically discussed. The challenges of SAR image interpretation were presented next, specifically with reference to flood detection problems, along with the solutions proposed in literature to eliminate these effects. This was followed by a discussion on the representation of uncertainties in SAR-based flood mapping. An overview of the challenges and opportunities associated with the online integration of Earth Observation data with hydraulic flood inundation models using data assimilation techniques was critically examined thereafter. The progress in EO-based hydraulic data assimilation studies was critically examined here and the evolution of assimilation techniques with respect to the current state-of-the-art was presented. Sensitivity of hydraulic data assimilation algorithms to observation operators and characteristics like location, timing and frequency were then discussed. Finally, the challenges in hydraulic data assimilation were systematically illustrated and the opportunities with respect to improved flood forecasting were summarized.

PART II
DATA AND MODELS

CHAPTER THREE

“It is a capital mistake to theorize before one has data.”

- *Sherlock Holmes*, in the short story *A Scandal in Bohemia* by Sir Arthur Conan Doyle

3. Study Area and Data

In order to develop and comprehensively evaluate the methods outlined in Chapters 5 to 8, extensive state-of-the-art datasets including field and remote sensing observations were necessary. This chapter describes the data available to this study and the selected study area, the Clarence Catchment of New South Wales, Australia.

3.1 Introduction

The Clarence Catchment, Australia, was selected for this analysis as the New South Wales State Emergency Service expressed a strong interest in the development of a more accurate flood forecasting system for the Clarence Valley. Accordingly, detailed ancillary datasets were collected, to support the study of recent major flood events in the catchment, which had been heavily monitored. These datasets comprise of a LiDAR-based DEM, channel bathymetry, and inflows together with flood monitoring observations including gauged discharge and water levels (WLs), remotely sensed images capturing the flood events and crowd-sourced WLs. All the data were collected in the framework of the Bushfire and Natural Hazards Collaborative Research Centre grant ‘‘Improving flood forecast skill using remote sensing data". In particular, the support of The Clarence Valley Council, BMT-WBM Ltd., the Australian Bureau of Meteorology, Geoscience Australia, New South Wales State Emergency Services, New South Wales Land Registry Services, New South Wales Manly Hydraulics Laboratory, and e-GEOS, towards the collection and timely provision of the data is gratefully acknowledged. The following sections describe the test site in detail and the available data.

3.2 Study Area

The Clarence Catchment is situated in the far north coast of New South Wales. It is one of the largest river systems on the South-Eastern coast of Australia (Fig. 3.1), with a net drainage area of about 22,700 sq. kms. The Clarence Valley extends from 28°30' S to 30°25' S latitude and 152°4' E to 153°21' E longitude. The river mouth is located between the towns of Yamba and Iluka, with the estuary reaching 108 kilometres inland to Copmanhurst. The towns of Grafton, Ulmarra, Maclean, Yamba, and Iluka are the main centres of population along the estuary (https://www.industry.nsw.gov.au/).

The main stem of the river is approximately 394 km long and occupies the southern part of the Clarence-Moreton Basin in north-eastern New South Wales. The Clarence Catchment, which forms the Eastern part of the Great Artesian Basin, is markedly asymmetric in shape (Haworth and Ollier 1992). The main valley containing the main stem

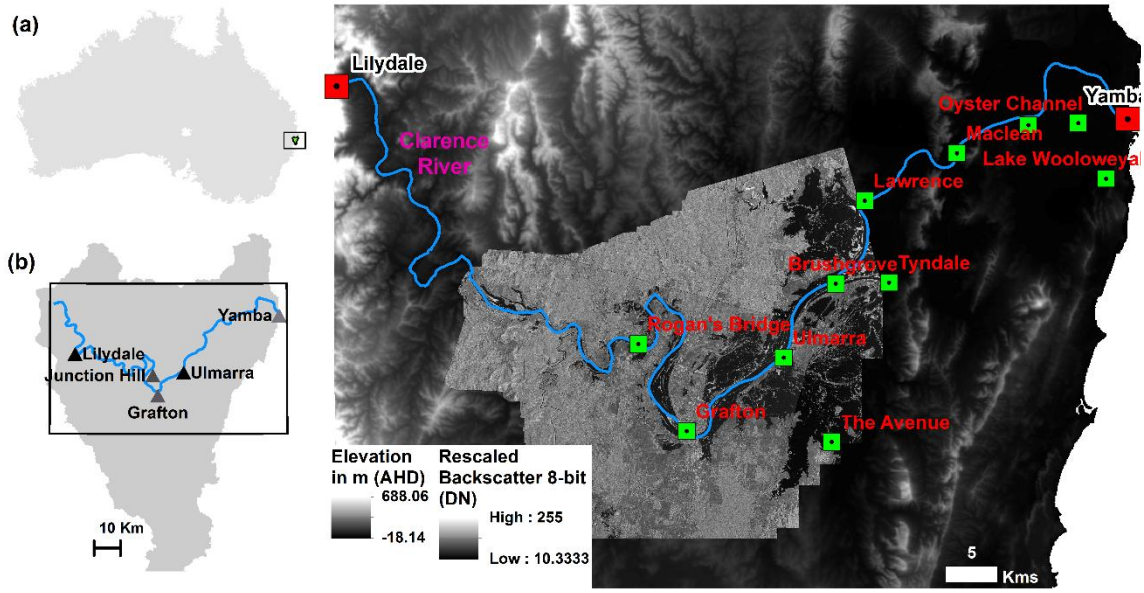

Fig. 3.1 Geographical location of the Clarence Catchment, in Australia shown in (a), with the Clarence River and nearby towns marked with respect to the Clarence River Catchment in (b). The extent of the model domain from Lilydale to Yamba is shown in (c), with upstream and downstream model boundary conditions marked in red squares while gauge locations are represented by green squares. The LiDAR DEM made available by Geoscience Australia is displayed as the base layer, with the spatial coverage of the two overlapping Cosmo-SkyMed SAR images covering the 2011 flood event, shown with respect to the model domain.

is almost 200 km long and aligned northwest-southeast, while most of its major tributaries flow southwest to northeast, with large gorges and deep valley systems on its western and south-western margins (Mcqueen 2016). The major tributaries to the Clarence River include the Maryland, Cataract, Timbarra, Mann, Guy Fawkes, Nymboida and Orara rivers. To the east from Coffs Harbour to Yamba, the catchment is bounded by the Coast Range, in the west by the Great Dividing Range (Northern Tablelands); by Baldblair, the Doughboy Ranges and the Dorrigo Plateau in the south; and by the MacPherson Ranges, in the north (NLWRA 2000). The lower floodplain is divided into two depositional basins by the Maclean-Tyndale Ridge which includes two large lakes, The Broadwater and Lake Wooloweyah (Mcqueen 2016). Generally, the catchment is characterised in its eastern extremities by tableland areas which fall away to the relatively large, flat coastal floodplain.

The land cover of the Clarence region is primarily dominated by grassland vegetation and agriculture, with some urban settlements around Grafton, Ulmarra, Maclean, Yamba, and Iluka. Most of the Clarence river basin is unregulated. Major water storages in this basin are associated with the Nymboida River and include the Shannon Creek Dam (capacity 30 GL), Karangi Dam (capacity 6 GL) and Rushford Road Reservoir (100 ML). A weir on the Nymboida River supplies water to the Karangi Dam. Non-irrigated agriculture dominates the Clarence Valley with only approximately 0.3% of the catchment area is under irrigation (https://www.bioregionalassessments.gov.au/).

The mean annual rainfall for the basin is 1,111 mm and mean annual actual evapotranspiration is 854 mm. The Clarence River is perennial with a mean annual flow of about ~181.6 cumecs and a runoff coefficient of about 0.23 (NLWRA 2000). The headwaters for the Clarence river basin represent some of the highest rainfall areas in NSW. As a result, flooding is a regular occurrence. There have been 73 major and moderate flood events since 1839, with the most recent major events recorded in 2009, 2011, and 2013 (Huxley and Beaman 2014). Dealing with this flood frequency, unquestionably requires improved flood modelling and forecasting solutions. The highest on record flood occurred in 2013, which reached water levels of 8.09m Australian Height Datum (AHD) at Grafton, Prince Street Gauge (Huxley and Beaman 2014). To put this in context, the bankfull height at this location is 1.1m (based on the 2 years ARI discharge) and the average low flow level varies between ~0m (low ebb tide level) and ~0.5m (high ebb tide level) (Stefania Grimaldi, *Personal Communication*).

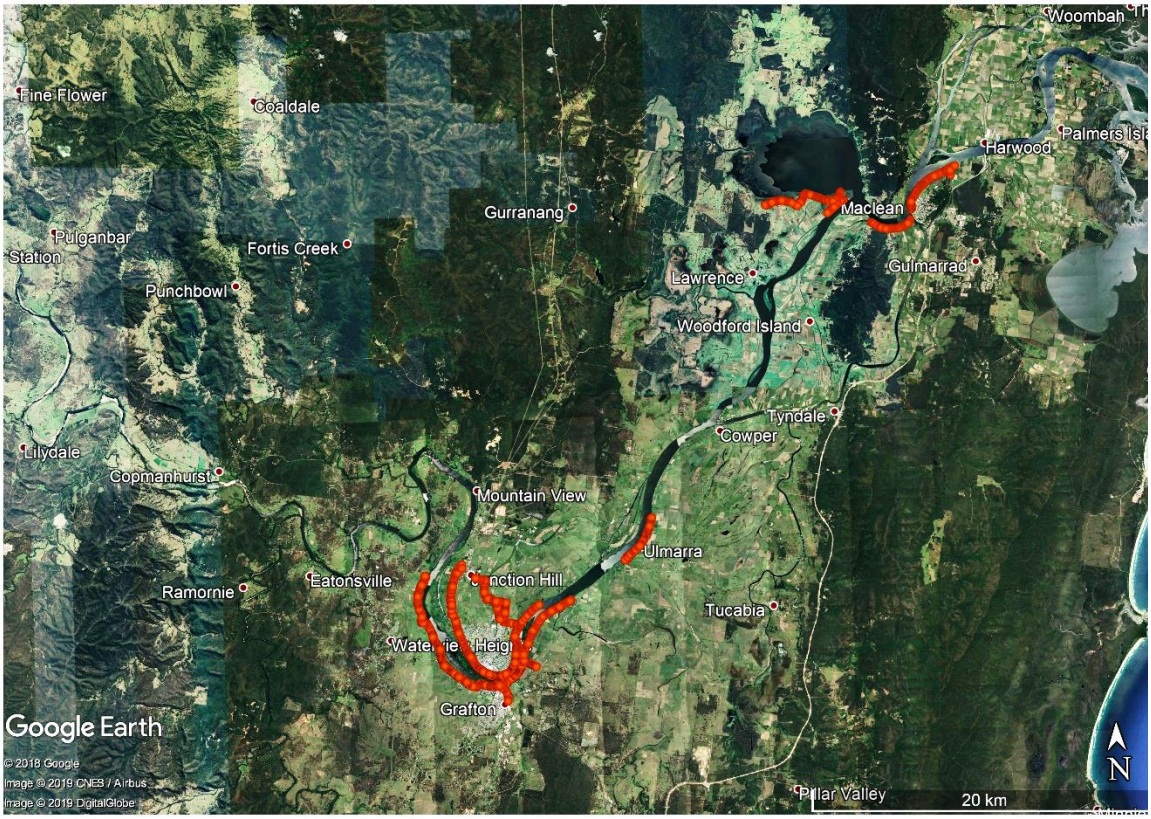

Fig. 3.2 Locations of the flood control levee system, constructed to protect the Towns of Grafton and Maclean from inundation. Source: Clarence Valley Council.

Floods in this catchment move fast, resulting in a flashy catchment response (Rogencamp 2004). For example, in 2011 the flood peak travelled the 125 km from Lilydale to Yamba in less than 30 h (Grimaldi et al. 2018). Low-intensity, long duration rainfall events are the dominant cause of flooding in the area, closely followed by the back propagation of ocean storm tides which control inundation dynamics as far upstream as Maclean (Ye et al. 1997). The catchment is characterized by largely variable flow velocities, ranging from 2-5 m/s in the channel and the levee system, to almost zero in the backwaters (Sinclair Knight Merz and Roads and Traffic Authority of NSW 2011). Extensive levee walls have been constructed to protect Grafton (both South and North), Ulmarra, and Maclean from flooding as shown in Fig. 3.2 (Rogencamp 2004).

3.3 Data Summary

The present study was benefited by the availability of remotely sensed data in the form of radar imagery and aerial photography, as well as LiDAR elevation data and field-recorded high water marks. Bathymetric data was collected during a field campaign in 2015 (Grimaldi et al., 2017) - described extensively in (Grimaldi et al. 2018) - and supplemented with pre-existing bathymetric datasets (Farr and Huxley 2013), forming a rich database. The data available to this study, compiled and kindly shared by Dr. Stefania Grimaldi (*Personal communication*, September, 2016), is summarized in Table 3.1.

Table 3.1 Summary of the relevant data available for hydraulic model implementation in the Clarence Catchment (Source: Stefania Grimaldi, *Personal Communication*).

Data	Provider/Source	Remarks
Hourly gauged water level and discharge	NSW Office of Water	2000 – 2014 (vary among gauges)
Hourly gauged water level	NSW Manly Hydraulics Laboratory	Spatial coverage: river branch downstream of Rogans Bridge
DEM	Geoscience Australia (Gallant et al. 2011)	Spatial resolution: 1 sec (InSAR)
	Clarence Valley Council	Spatial resolution: 1m (LiDAR)
Bathymetric data	BMT-WBM	Spatial coverage: from Mountain View to Yamba
	Port Authority of NSW	Spatial coverage: Yamba harbour
	Grimaldi et al., 2017	Spatial coverage: Copmanhurst to Mountain View
Geometric data of the levees	Clarence Valley Council	
Geometric data of the bridges	NSW Road's and Maritime Service	
	Clarence Valley Council	
Dynamic Land Cover	Geoscience Australia (Lymburner et al. 2011)	Spatial resolution: 250m
Debris marks and high water marks	Clarence Valley Council	Flood events in January 2011 and January-February 2013
Photos of flood events	Clarence Valley Council	Flood events in January 2011 and January-February 2013
	NSW State Emergency Service	
Airborne photography	Clarence Valley Council	Acquisition time: high and low flow days
	Land and Property Information NSW	
Satellite images	Geoscience Australia (SPOT6 image: AIRBUS) E-Geos (COSMO SkyMED images)	Optical and SAR images acquired during the flood events in January 2011 and January-February 2013

3.3.1 SAR Images

Two COSMO-SkyMed (CSK) X-band (9.6 GHz with a wavelength of 3.1 cm) HH-Polarized images, acquired by the CSK-3 satellite were available for this study. The data were acquired in Stripmap HIMAGE mode at 3 m resolution on 12 January, 2011 at 18:03 hours and 13 January, 2011 at 07:33 hours (AEDT). They were acquired at the peak of the 2011 flood event at Grafton and just after it on the falling limb, as depicted in Fig. 3.6. The CSK Level 1D Georeferenced Terrain Corrected (GTC) product delivered as an 8-bit image of digital numbers was used in this study. This was a consequence of project budgetary constraints, as at the time of ordering the SAR scenes from E-Geos, the project team did not intend to process the images. The objective was to obtain the 8-bit images along with interpretation from E-Geos for the Clarence Catchment, as the project objective was to

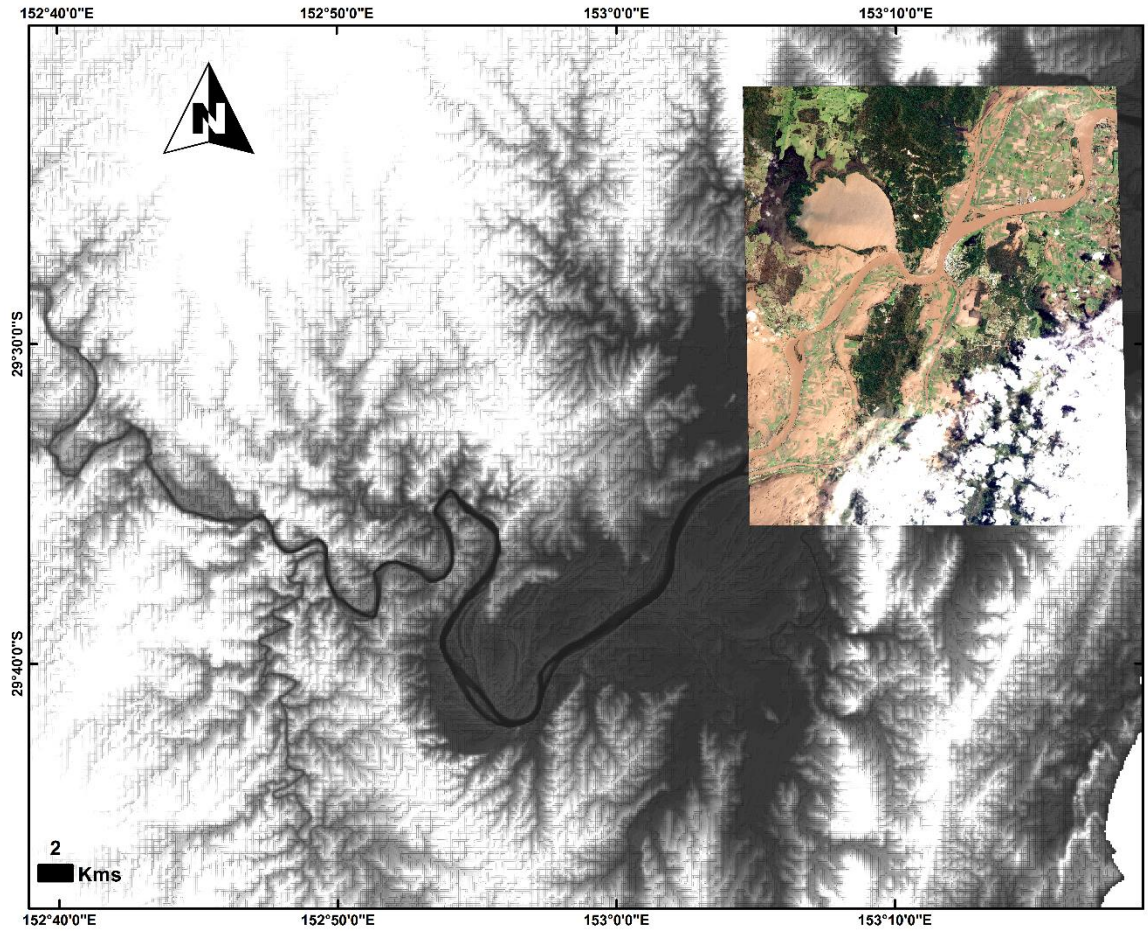

Fig. 3.3 Spatial extent of the SPOT-6 optical image covering the 2013 flood event in the Clarence, shown here with respect to the model domain. The LiDAR DEM available to this study is used as the base layer.

improve modelling using RS-data, rather than improve the RS-image interpretation itself. Moreover, the 8-bit images were made available to the study at a discounted price, compared to the native bit resolution backscatter products. The calibration process for the GTC product followed by E-Geos, corrects for local incidence angle impacts using a DEM, by normalising the backscatter to a 40° reference incidence angle (Italian Space Agency 2009). The domain comprised of 74,056,858 pixels each having an area of 9 m^2, bringing the total tile coverage to approximately 666.5 km^2. The spatial coverage of the SAR tiles with respect to the model domain, is illustrated in Fig. 3.1. The SAR images described here were used to develop and test the fuzzy flood mapping algorithm presented in Chapter 4 of this thesis. The fuzzy flood maps thereby generated, were subsequently assimilated into the flood forecasting model LISFLOOD-FP in Chapter 9.

3.3.2 Optical Imagery

An optical multi-spectral image from the Satellite Pour l'Observation de la Terre (SPOT) 6 satellite was available to this study, acquired at 31st Jan 2013 at 09:35 am (AEDT). The data were acquired at 6 m resolution and were delivered as an ortho-rectified, pan-sharpened multi-spectral (PMS) product at 1.5 m with four spectral bands, i.e. blue

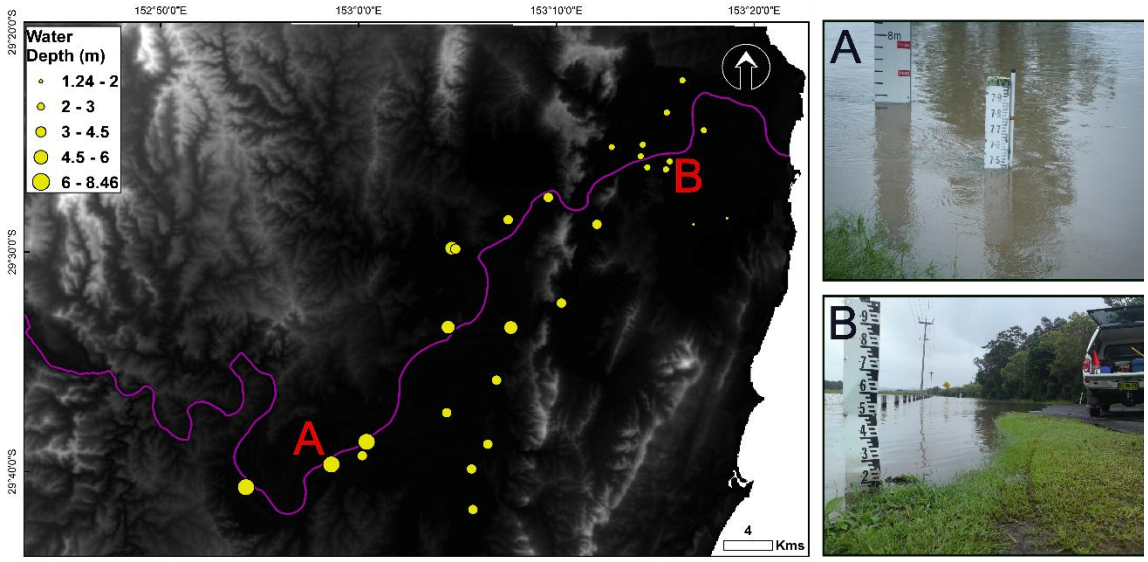

Fig. 3.4 Locations of the "crowd-sourced" water depth observations for the 2013 flood event in the Clarence Catchment. Sub-figures A and B show example images used for the depth calculation, by Clarence Valley Council.

(450-520 nm), green (530-590 nm), red (625-695 nm), and near infrared (760-890 nm). SPOT-6 PMS products have a radiometric resolution of 12 bits per pixel and the image was delivered in the JPEG 2000 raster format (Astrium Services 2013). The image comprised of a total of 255,072,110 pixels covering a total area of 573.91 km^2. About 20% of the tile was affected by cloud cover, obscuring the underlying inundated regions. In order to avoid the associated uncertainty, this portion of the image was removed from the analysis. Fig. 3.3 shows the spatial extent of the SPOT image with respect to the model domain, while Fig. 3.7 shows the temporal position with respect to the 2013 flood hydrographs. This image was used to delineate the flooded area and used in Chapter 4 for verification of selected model parameter values.

Two high-resolution optical aerial photographs covering the 2011 flood event in the Clarence Catchment, were provided by the NSW-Land and Property Information (LPI) Department for this study. As the NSW-LPI had only one airplane available to monitor the flooding, the spatial coverage was focused on urbanized areas. The images were captured on January 12, 2011, between 16:41 and 17:17 hours and 17:17 to 17:39 hours,

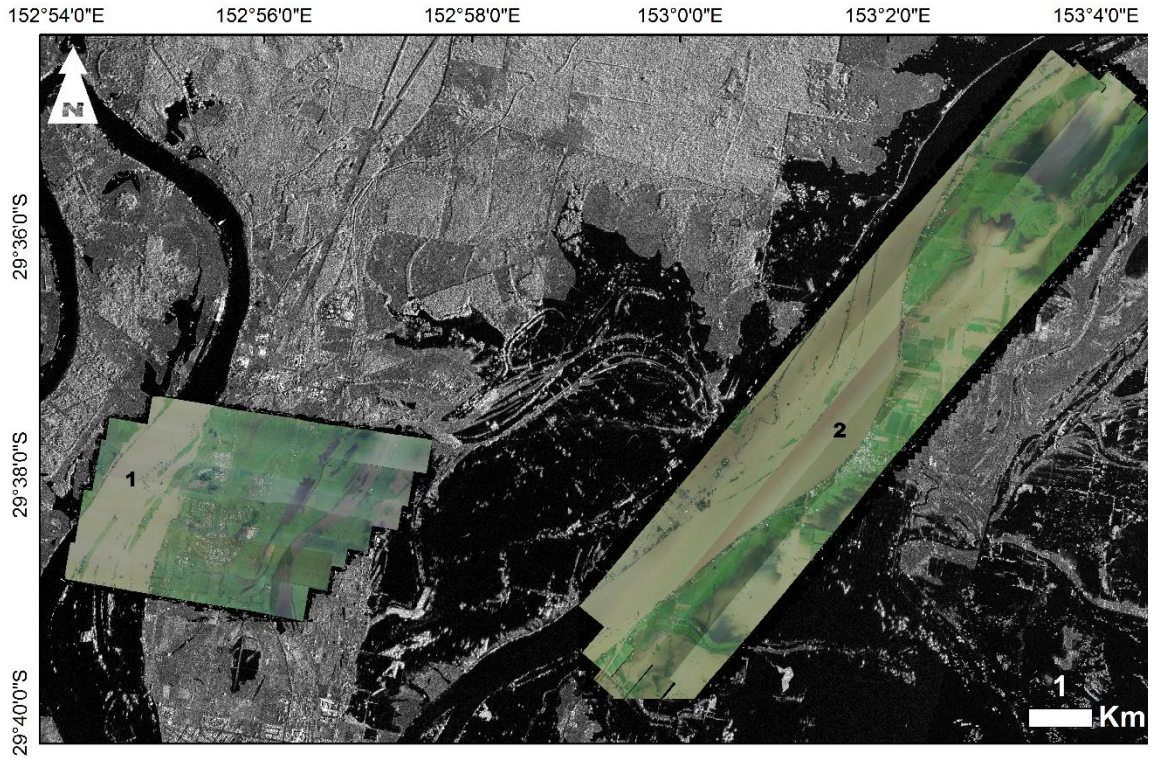

Fig. 3.5 Map illustrating the spatial coverage for the aerial photographs of Junction Hill (1) and Ulmarra (2), used in this study for validation of the flood mapping and data assimilation algorithms.

respectively. The images have a spatial resolution of 10 cm and cover Junction Hill and Ulmarra regions of the catchment, as shown in Fig. 3.5. These were used for validation of the proposed flood mapping approach in Chapter 4, and the real world application of the proposed data assimilation strategy in Chapter 9. Aerial photographs available from the region for non-flood scenarios were also used to derive distributed surface roughness values. First, the photographs were used to manually identify land-use classes at high resolution. Second, distributed roughness values were assigned based on standard prescriptions of each class available in literature (Arcement and Schneider 1989).

3.3.3 Topography

Topographic information was available in the form of a 1m Light Detection and Ranging (LiDAR) Bare Earth Digital Elevation Model (DEM), acquired in 2010 with a vertical accuracy of ±30 cm at 95% confidence (1.96 x RMSE) and horizontal accuracy of ±80 cm at 95% confidence (1.73 x RMSE) (New South Wales Land and Property Management Authority, 2010; Fig. 3.1). The vertical accuracy was assessed by comparing LiDAR point returns against survey check points as per the ICSM Guidelines for Digital Elevation Data (Intergovernmental Committee On Surveying & Mapping 2008). This dataset is freely available under a Creative Commons Attribution 4.0 license, for commercial and non-commercial applications at https://elevation.fsdf.org.au/, provided by Geoscience Australia. The channel bathymetry was reconstructed by interpolating between field observed cross-sections and stitched to the LiDAR DEM, for the part of the domain where it was available. The area upstream of Copmanhurst where LiDAR coverage was unavailable, was filled in with the SRTM-derived 30 m product enhanced by Geoscience Australia (DEM-H), which is described in greater detail in the next paragraph. For a detailed description of this combined topographic dataset, field data collection, and bathymetry generation, readers are referred to Grimaldi et al. (2018). The bathymetric dataset is freely available at https://figshare.com/articles/Bathymetric_survey_of_the_Upper_Clarence/5648251 (Grimaldi et al. 2017).

The SRTM-derived DEM-H (Gallant et al. 2011), a 1 arc second (~30 m) gridded DEM corrected for hydrological applications produced for the Australian continent by Geoscience Australia, was used in this thesis to represent the “best available” global

topography for the region (Jarihani et al. 2015). The DEM-H is considered superior to the globally available SRTM product for hydraulic modelling, as drainages were enforced using 1:250,000 scale watercourse lines and smoothed using ANUDEM software (Hutchinson 2011; Zheng et al. 2016), to ensure seamless hydraulic connectivity (Dowling et al. 2011). As coarse resolution SRTM products are often unable to accurately capture flow lines, which lead to erroneous inundation forecasts, the DEM-H with region specific hydrological enhancements was considered an appropriate choice (Mukherjee et al. 2013). The accuracy of DEM-H is similar to the raw SRTM 1-arc second product with geolocation errors <12.6 m and 90% of tested heights within 9.8 m for Australia (Gallant et al. 2011). This dataset is available for free, downloadable from http://elevation.fsdf.org.au/.

3.3.4 Land Cover Data

Land cover information was extracted from the National Dynamic Land Cover Dataset (DLCD) distributed by Geoscience Australia and the Australian Bureau of Agricultural and Resource Economics and Sciences (ABARES) at 250 m spatial resolution.

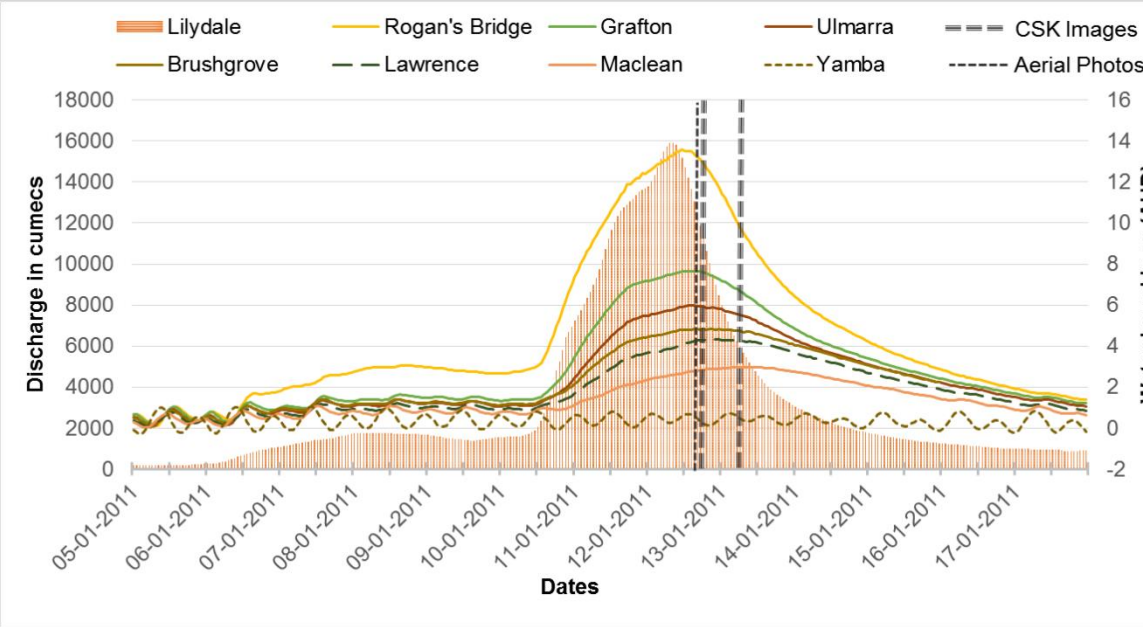

Fig. 3.6 Hydrographs recorded at the hydrometric gauges along the main stem of the Clarence River (locations shown in Fig. 3.1) for the 2011 flood event, shown together with the temporal acquisitions of available remote sensing data represented as the vertical black lines.

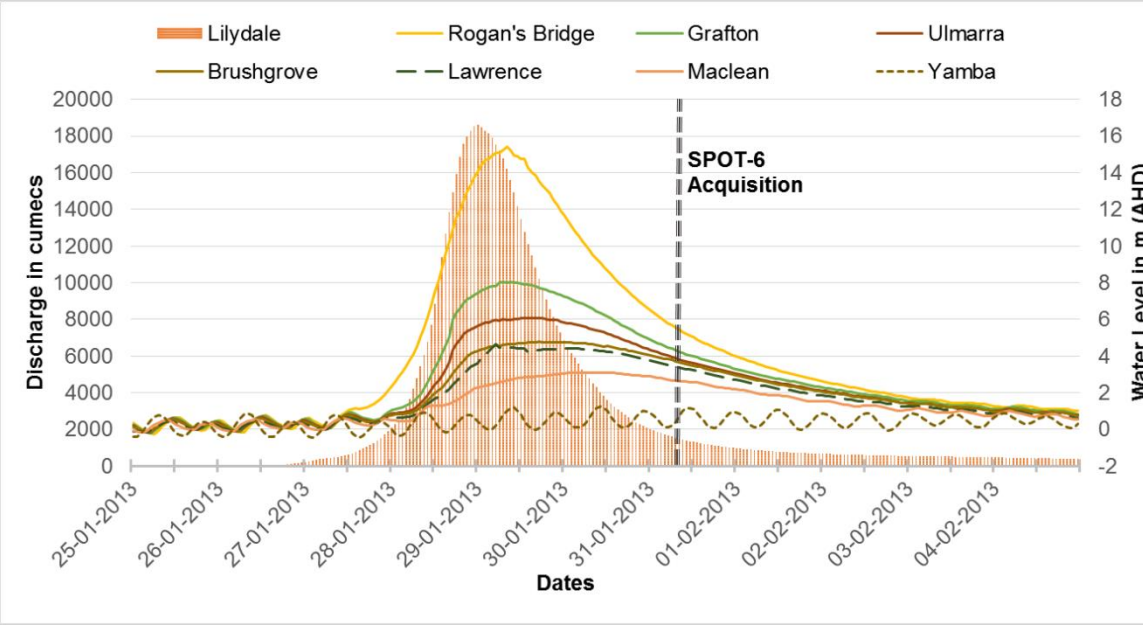

Fig. 3.7 As for Fig. 3.5 but for the 2013 flood event.

The base dataset for the DLCD is NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) 16-day Enhanced Vegetation Index (EVI) composite, collected between 2000 and 2008 (Lymburner et al. 2011). Each pixel was subject to time-series analysis to derive 12 coefficients representative of its statistical, phenological, and seasonal characteristics. These were subsequently classified and labelled using support vector machines (SVM). Readers are referred to Lymburner et al. (2011) for a more detailed description of this dataset. The DLCD dataset was used in this study to facilitate a land-use based performance assessment of the flood mapping algorithm developed in Chapter 4.

3.3.5 Crowdsourced Data

The Clarence Valley Council also provided field data in the form of photographs of wrack marks (debris deposited at the flood edge) and water marks (staining on the side of structures within the flooded area) available at http://clarenceriverfloods.blogspot.com.au/. Field observed water levels surveyed by the council just after the 2013 event were also provided. The water level observations presented here were derived based on field photographs and were therefore interpreted as crowdsourced, as field photographs form one of the most common citizen science data types that can be mined from social media for instance (Fohringer et al. 2015). These "crowdsourced" observations are used in Chapter 4 for hydraulic model calibration and shown in Fig. 3.4 alongside example photos.

3.3.6 Hydrometric Data

Hydrometric gauge information in the form of WL and discharge measurements were provided by the NSW Public Work's Manly Hydraulics Laboratory (MHL) and the Australian Bureau of Meteorology (BoM). The observations were recorded with a temporal frequency of fifteen minutes for the WL gauges and hourly for the discharge gauges. Missing data were interpolated using linear interpolation. Seasonality was not considered as the study required only event-based flood modelling. Discharge observations were available at Lilydale (BoM) and Grafton. Moreover, WL observations were available at Rogan's Bridge, Grafton, Ulmarra, Brushgrove, Lawrence, Maclean, Palmer's Island Bridge, and Yamba, from upstream to downstream along the main stem of the river. Gauge locations are shown in Fig. 3.1, while hydrographs recorded by gauges along the main stem

Table 3.2 Summary of the data utilised in this thesis, with the characteristics and usage for each listed.

Dataset	Main Characteristics	Usage
2 Cosmo-SkyMed Images	2011 event, SAR	Flood extent mapping (Chapter 4) and assimilation (Chapter 7 and 9)
1 SPOT-6 PMS Image	2013 event, Optical	Model verification (Chapter 5)
Aerial Photographs	2011 and 2013, Optical	Flood map validation (Chapter 4) and verification of assimilation performance – real case (Chapter 9)
LiDAR DEM	2001-2015, Optical	Flood inundation modelling (Chapters 5, 7, and 8)
DEM-H	2011, InSAR	Flood inundation modelling (Chapters 5 and 8)
LULC Data	2011, Optical	Flood map accuracy assessment (Chapter 5)
Hydrometric gauge data	1980-2017, Field	Model implementation (Chapter 5) and verification of assimilation performance – real case (Chapter 9)
High water marks and flood photos	2011 and 2013, Field	Model calibration (Chapter 5)

of the channel are shown in Fig. 3.6 and Fig. 3.7, for the 2011 and 2013 flood events respectively.

Additionally, WL observations were available at Tyndale, The Avenue, Oyster Channel, and Lake Wooloweyah. The WL values were recorded in metres with respect to AHD while the discharge observations were available in Mega Litres/day (ML/d) which were converted to cubic metres/second (cumecs). All of the gauges were used in Chapter 4 to calibrate the channel friction parameter in the hydraulic model, for the 2013 flood event in the Clarence. The gauges along the main stem were used for validation of the real world application of the assimilation algorithm in Chapter 9.

3.4 Chapter Summary

This chapter presented an overview of the study area and the data that were used to set up and validate the experiments outlined in this thesis. The study area description included details on the extent, drainage, geomorphology, topography, land cover, and flooding characteristics. Locations of flood control structures currently operative in the catchment were illustrated. The data summary comprised of reasons for choosing each dataset, detailed descriptions of each, and information on where and why they were used in the thesis, as summarized in Table 3.2. Positions of ground sensors and observations, as well as spatiotemporal coverage of the remote sensing data were also shown.

CHAPTER FOUR

"I claim that many patterns of Nature are so irregular and fragmented, that, compared with Euclid — a term used in this work to denote all of standard geometry — Nature exhibits not simply a higher degree but an altogether different level of complexity … The existence of these patterns challenges us to study these forms that Euclid leaves aside as being "formless", to investigate the morphology of the "amorphous"."

Benoit Mandelbrot, quoted in a review of *The Fractal Geometry of Nature* by J. W. Cannon in *The American Mathematical Monthly*, Vol. 91, No. 9 (November 1984), p. 594.

4. SAR-based Flood Inundation Mapping

The literature review revealed several gaps in the utilization of image texture in the field of SAR-based flood mapping. Consequently, a SAR texture optimization technique is proposed in this Chapter, to improve the utilization of texture in single image flood mapping and thus address these open research questions. Probabilistic flood maps developed using the methods detailed in this chapter, have subsequently been assimilated into LISFLOOD-FP in Chapter 7, using the framework developed in Chapter 5. The flood mapping algorithm presented in this chapter, was published as a peer-reviewed journal paper in Remote Sensing of Environment[3].

4.1 Introduction

Synthetic Aperture Radar (SAR) data are currently the most reliable resource for flood monitoring, though still subject to uncertainties which can only be objectively represented with probabilistic flood maps. Ideally, the inclusion of ancillary datasets could eliminate one or more sources of errors in SAR-based flood extraction (D'Addabbo et al. 2016). However, suitable supporting datasets are usually unavailable for the area of interest, especially for developing regions. Moreover, the present cohort of fuzzy flood classification algorithms utilize theoretical electromagnetic backscattering models for parameterization. Given that these are wavelength specific, they typically limit the transferability of fuzzy approaches across the range of SAR satellites. Therefore, this study introduced a texture-based image enhancement approach to

[3]**Dasgupta, A.,** Grimaldi, S., Ramsankaran, R., Pauwels, V. R. N., & Walker, J. P. (2018). Towards operational SAR-based flood mapping using neuro-fuzzy texture-based approaches. *Remote Sensing of Environment*, *215*(15 September 2018), pp. 313–329, http://doi.org/10.1016/j.rse.2018.06.019.

improve single image flood mapping, which can incorporate the spatial autocorrelation amongst pixel values to minimize the impact of sensor parameters.

Since texture can be derived from the SAR image, it also reduces the dependence on ancillary or complementary datasets. However, state-of-the-art texture based mapping approaches also struggle with the subjectivity in selecting application appropriate texture features, suitable window sizes, and optimal direction for identifying the feature of interest (Di Baldassarre et al. 2011). Optimized texture bands were considered alongside the SAR intensity image, within a neuro-fuzzy classifier to generate a fuzzy flood map. Moreover, Gaussian

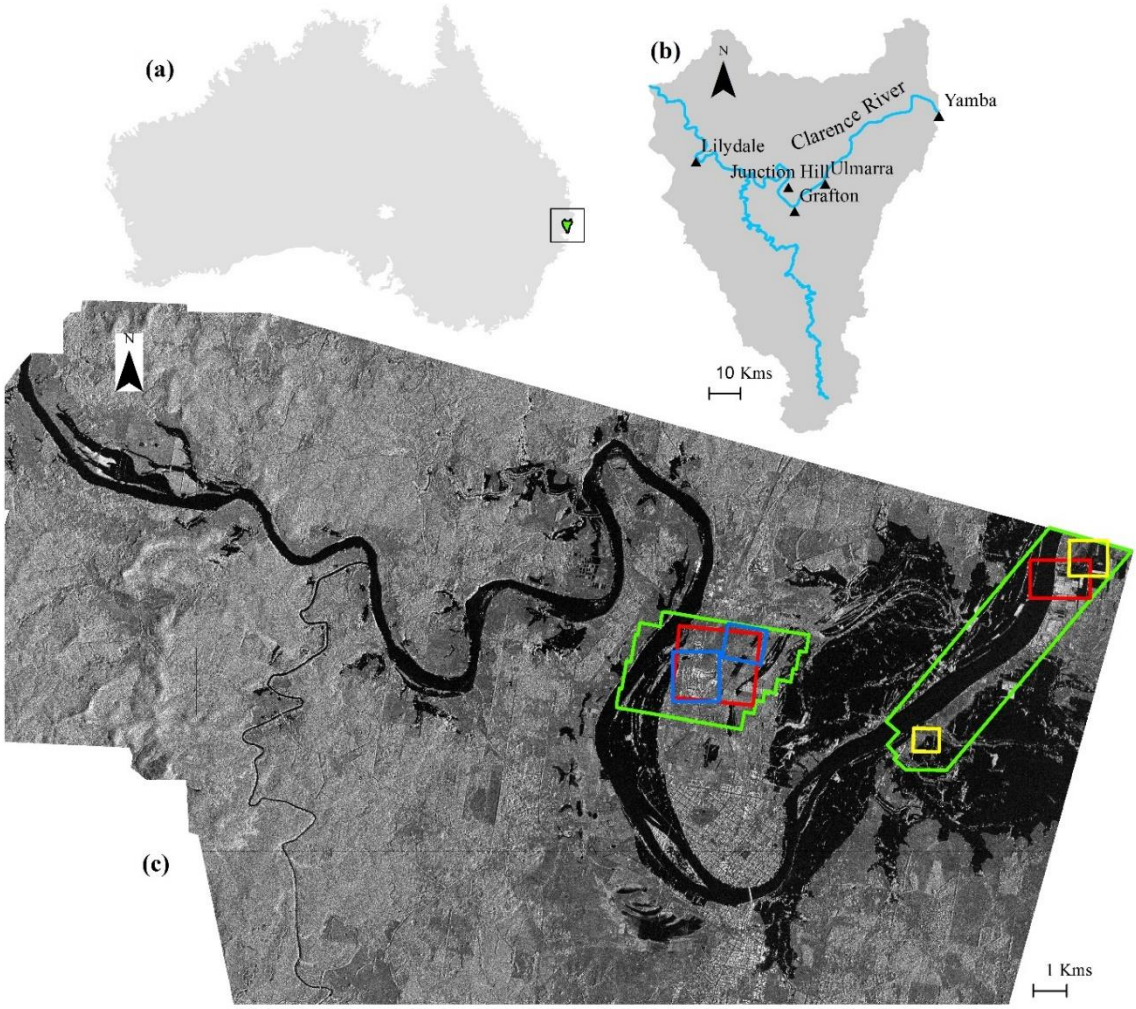

Fig. 4.1 Map displaying the location of the Clarence catchment (a), and the main drainage lines and towns (b). The COSMO-Skymed SAR image acquired on 12th Jan, 2011 is also shown (c), with the green polygons indicating the aerial photo coverage used for validation. The example subsets used in Fig. 4.5 are depicted in red, while those used in Fig. 4.8 and 4.13 are shown in blue and yellow respectively.

membership functions were chosen to represent the backscatter distribution of each class, based on the image histogram as in the probabilistic mapping approaches (Giustarini et al. 2016; Schlaffer et al. 2017). However, using the neural network for a data driven parameter estimation of these membership functions removes the need for identification of suitable prior probability distributions. Moreover, training the classifier on the image to be processed offers the additional advantage of accounting for image specific backscatter variability, caused by the reference incidence angle or wind effects.

Given a filtered SAR image, the ideal window size for texture estimation was first determined through semivariogram analysis. This was followed by an estimation of omnidirectional Grey Level Co-occurrence Matrices (GLCM) from which second-order statistical texture features were derived. An independent component analysis was then used to condense the maximum possible information into minimum bands, which were then added to the SAR image prior to classification. The class distributions were modelled as Gaussian functions within a fuzzy inference system, and parameterized using training data from the image itself. The resulting maps were evaluated using aerial photographs through reliability diagrams, as well as a fuzzy validation exercise novel to flood mapping literature. The fuzzy map comparison accounts for the uncertainties in manual shoreline extraction for validation data as well. The classification performance of the SAR image with added optimized texture bands was compared against a SAR image without any texture addition and a SAR image with some randomly selected texture features added. Finally, a land-use specific analysis was conducted to assess the spatial variability of classifier performance, to facilitate an area appropriate choice of classifiers for flood mapping. The following sections detail the general principles of SAR with respect to flood mapping, the proposed flood mapping framework, and the validation results for the Clarence Catchment (See Fig. 4.1) are subsequently presented.

4.2 Principles of SAR: Implications for Flood Mapping

Microwave interactions with the Earth's surface are governed by the SAR sensor configuration (wavelength, polarization, resolution and looking angle), surface characteristics (roughness and dielectric properties), and local slope. The high contrast in the radar image between flooded and non-flooded areas, is primarily the result of specular reflection from standing water. This phenomena makes water surfaces appear dark on SAR images, as opposed to the increased

scattering from rough land surfaces which makes terrain appear bright. As floods are mainly identified using the changes they cause in backscattering behaviour of land cover classes, flood mapping is then detecting the insurgence of water in three main land cover classes - bare soil, vegetation and urban areas. In order to understand these changes, one needs to be aware of the underlying mechanisms which drive the microwave interactions with the Earth's surface.

4.2.1 Geometric Distortion

The SAR side-looking sensing implies that two points will look closer in the slant-range image than they actually are on the ground. This effect gets more pronounced as we move closer to the antenna i.e. in the near range and is most noticeable at nadir (90°). The projection lines from the ground-range to slant-range are usually in the form of concentric circles around the antenna, due to the spherical divergence of radar pulses over such large distances (See Fig. 4.2). This leads to geometric distortion in the images, which needs to be accounted for in flood mapping problems.

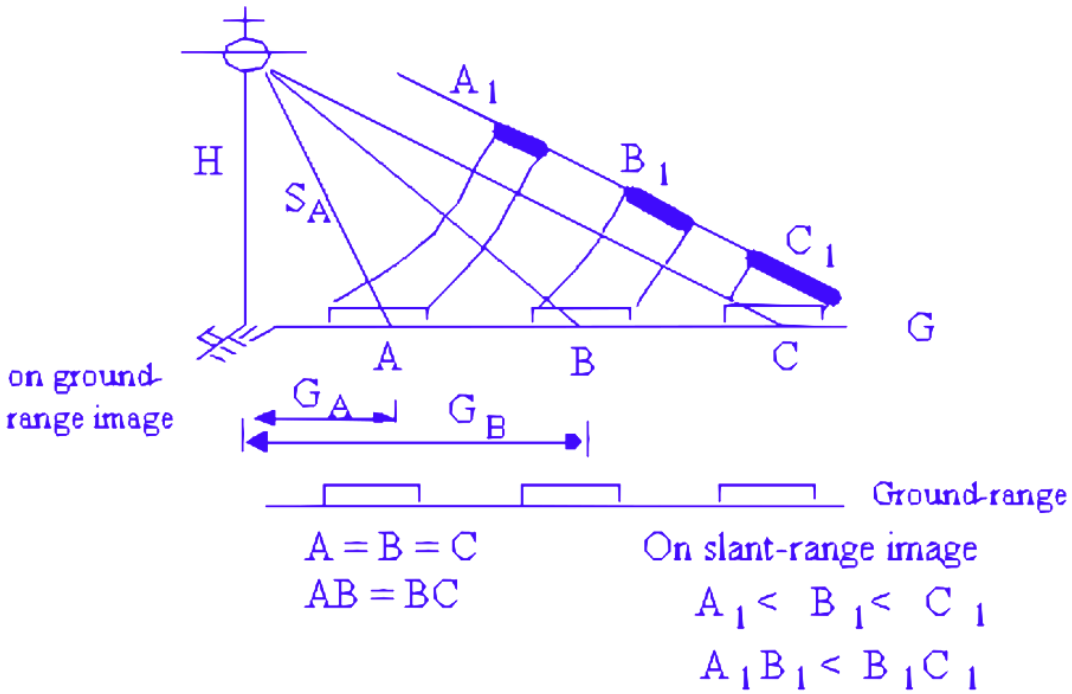

Fig. 4.2 The differences between ground-range and slant-range for side-looking radar geometries. Modified based on: http://nature.berkeley.edu/~penggong/textbook/chapter3/image/fig331.gif

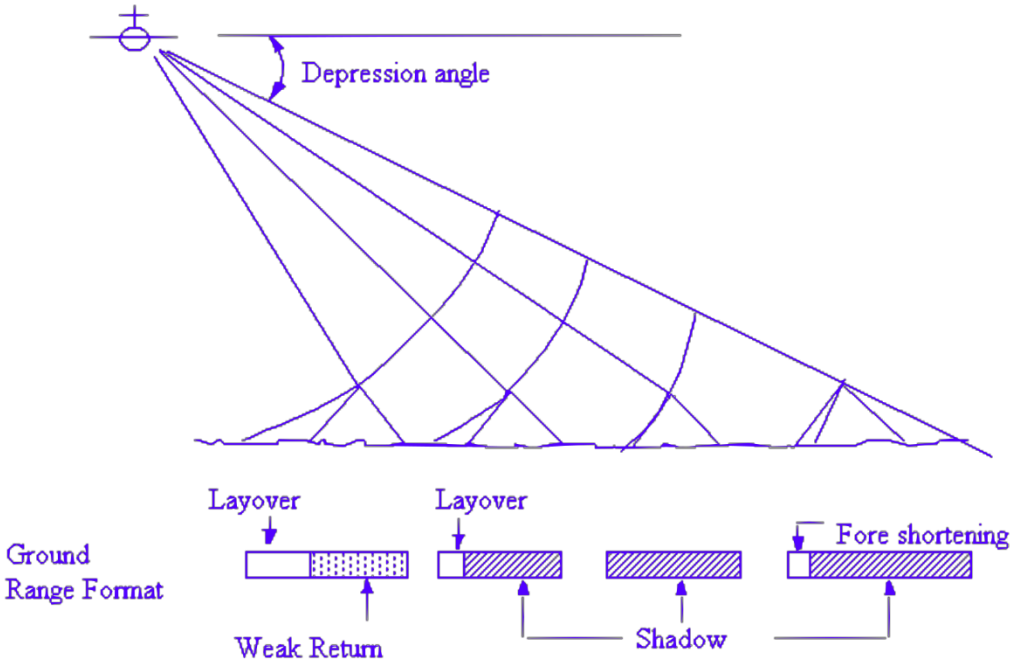

Fig. 4.3 Geometric distortions caused by the side looking radar viewing geometry. Modified based on: https://earth.esa.int/handbooks/asar/CNTR1-1-2.html

The imaging of steep vertical objects through SAR, results in a phenomenon known as relief displacement. The top of the structure shifts from its actual ground location, proportional to object height and radar look angles. In particular, these errors can be of three types, as shown in Fig. 4.3 and elucidated below.

1. Foreshortening - As radar imaging records distances from the antenna, the top of tall targets such as mountains would be encountered by the beam before the bottom is seen. This makes the object appear to "lean" towards the sensors with the foreslope (the slope facing the sensor) appearing shorter than in reality.

2. Layover - When slopes are steeper, targets in the valley region behind the mountains display a longer slant range. This causes objects in the slant range to be ordered in reverse of their actual ground positioning. Consequently, the front slope appears to overlay on the backslope on the SAR image.
3. Shadow - This effect is caused primarily by the side-looking imaging technique. As the radar views objects at an angle, it cannot "see" a certain region hidden by vertical structures, these appear as dark shadows in the SAR image.

4.2.2 Surface Roughness and Local Incidence Angle

Radar backscatter is a function of the surface roughness and local incidence angle of the surface from which the radar beam is scattered. The nature of this scattering determines the strength of the signal returned to the sensor and can be of several types as shown in Fig. 4.4.

4.2.2.1 Surface Scattering

The scale of the surface roughness can be represented by the root-mean-square deviation (h) from the mean height ($\bar{h}$) of the interface dividing the two homogeneous media, the atmosphere and the hypothetical perfectly smooth open water surface. Intuitively, one can imagine that when this deviation is comparable to the magnitude of the wavelength, the surface undulations return separate radar echoes to the antenna. The interference of these returned signals spuriously raises the recorded backscatter values, leading to a reduction in specular reflection effects.

According to the Rayleigh Criterion the phase change caused by h should be less than a quarter of the wavelength (λ). As this phase difference increases, scattering becomes more diffuse (i.e. energy reflected equally in all directions) and "smooth" surfaces would no longer be detectable by contrast. The local incidence angle θ_{loc_i} determines the path difference between the incident and the scattered path, and therefore can influence the maximum acceptable value of h for smooth surfaces. The Rayleigh Criterion describes smooth surfaces as those satisfying

$$h < \frac{\lambda}{8\cos\theta_{loc_i}}, \tag{4-1}$$

Equation 4.1 implies that h must be $\frac{1}{8}th$ of the wavelength to account for two-way travel of the radar path. In case of RS a stricter criterion for smoothness is needed, known as the Frauenhoefer Criterion which restricts h to $\frac{1}{32}th$ of λ, as the distance between the target and the sensor is usually several orders of magnitude larger than the wavelength of the incident beam. Based on the above, surface scattering mechanisms can broadly be classified into two categories:

1. *Specular scattering (or reflection):* Where the majority of the incident microwave energy is reflected off a smooth surface; the remaining energy is transmitted through the media in

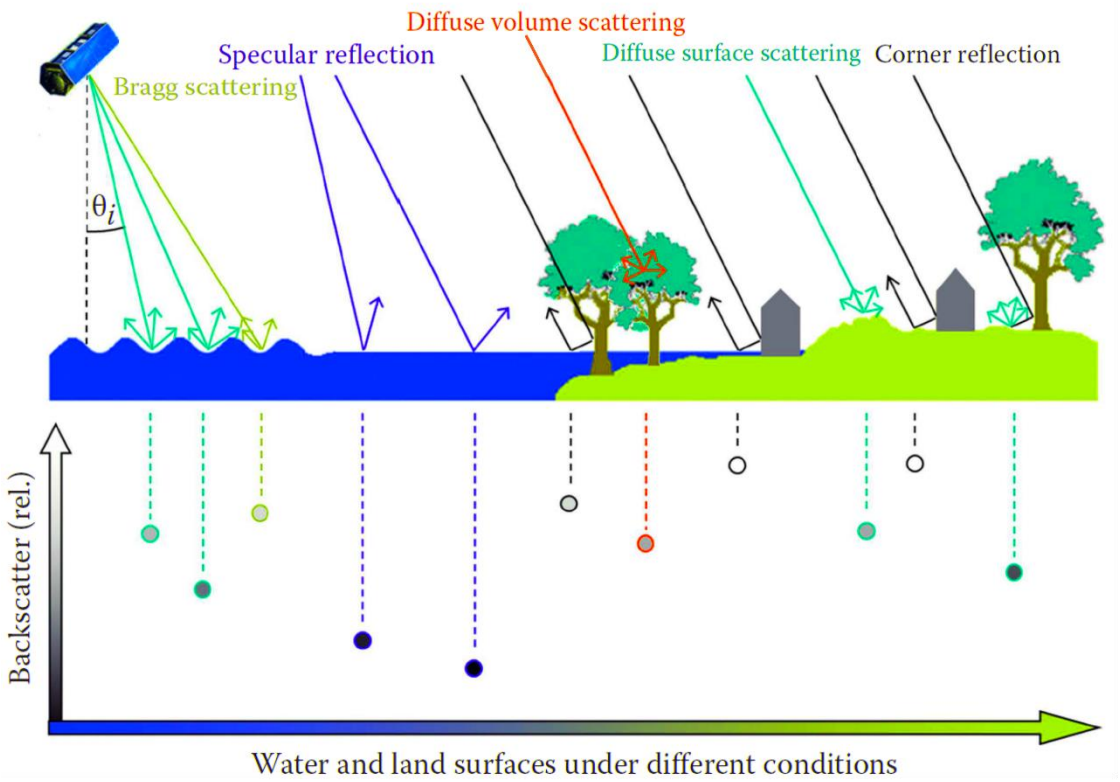

Fig. 4.4 Different scattering mechanisms displayed by radar interactions with water and land surfaces. Based on Martinis et al. (2015b).

accordance with Snell Law of refraction. Radar returns are negligible for a perfect specular reflector as the beam is reflected away from the receiver.

2. *Bragg scattering:* When the surface roughness of a slightly rough $\left(\frac{\lambda}{32} < h < \frac{\lambda}{8}\right)$ homogeneous medium exhibits periodicity, an identifiable and coherent pattern of backscatter is returned. Bragg surfaces are often an ordered collection of facets, where particular orientations occur at regular intervals. The facets oriented towards the sensor can often produce return signals which resonate with the incident beam resulting in bright image regions. This effect is more pronounced for steeper incidence angles as the sensor may record backscatter returned from both specular and Bragg scattering.
3. *Diffuse surface scattering:* As the surface roughness increases beyond $\frac{1}{8}th$ of the wavelength the scattering gets more random. An extremely rough surface should ideally scatter the radar beam in all directions with equal intensities, resulting in higher signal returns.

4. *Corner reflection:* The radar pulse is reflected back to the sensor when it encounters a smooth horizontal and smooth vertical surface orthogonal to each other, resulting in saturated pixels. Corner reflection can be caused by partially submerged vegetation or urban features such as roads, pavements and buildings, resulting in very strong returns.

4.2.2.2 Volume Scattering

Radar beams achieve measurable penetration depths directly proportional to the incident wavelength. Complex natural environments often include multiple media, having different dielectric constants, which results in volume scattering. When microwaves come in contact with a heterogeneous media exhibiting variations in electromagnetic properties, each wave is scattered in many different directions. The intensity of returned backscatter is a function of the density and heterogeneity of the traversed media. The shape, density, relative permittivity, and orientations of the various media encountered by the radar signal then determine the strength of the backscatter recorded at the antenna.

4.2.3 Dielectric Constant

The dielectric constant is defined as the ratio between the electrical permittivity of a homogeneous material with respect to that of a vacuum, given by a dimensionless quantity also known as relative permittivity. Natural materials do not respond to external electromagnetic fields instantaneously, unlike a vacuum. The response generally depends on the frequency of the applied field, which in this case is the radar signal.

Microwave frequencies are highly sensitive to changes in the dielectric constant, and can achieve considerable penetration depths as most dry natural materials exhibit values between 3 and 8, resulting in low reflectivity in this region of the EM spectrum. Conversely, liquid water has a dielectric constant of approximately 80 in the microwave region. Based on this, open water should return most of the signal right back to the sensor, exhibiting strong reflectivity. This doesn't happen in reality as the effects of specular reflection dominate and the beam is directed away from the sensor.

Since the dielectric constant is directly correlated to the moisture content per unit volume, higher vegetation/soil water content implies larger values of radar returns. Most of the sensible variation in backscatter recorded from natural surfaces can be explained by the moisture content of the underlying geographical features. This property is often exploited in the

derivation of remote sensing based soil moisture. As the dielectric constant depends on the frequency of the EM wave, the penetration depth is directly related to the wavelength. For instance, when detecting flooding under forests, using a SAR image operating in the longer wavelength region (lower frequency) allows the EM waves to penetrate the canopy to a greater extent than shorter wavelengths which experience surface scattering.

4.2.4 Polarization

The polarization of a SAR image refers to the orientation of the electric field intensity vector in the transmitted and received radar waveform as shown in Fig. 4.5. SAR sensors usually send and receive horizontally polarized waves, as the returns recorded are higher than in vertical polarization. When the transmitted and received radar wave polarizations are identical this is known as co-polarized; conversely cross-polarization refers to opposite send and receive polarizations. Fully polarimetric SAR images (antenna can record HH, HV, VH and VV returns simultaneously) can be useful for detecting flooding under vegetation or urban areas as illustrated in the following sections because they are able to highlight the double-bounce component of the surface scattering.

4.2.5 Speckle

SAR images are affected by conspicuous bright and dark spots, known as shot noise or speckle. This apparently random manifestation of light and dark pixels arises from constructive and destructive interference of scattered radar waves. A SAR pixel represents a few square metres of area on the ground, often presenting the incident microwaves with multiple scatterers. These waves which are scattered non-uniformly in all directions, interfere with one another before reaching the antenna. If this interference is constructive (the crests and troughs of the waveform superimpose), a strong signal return is recorded leading to a bright spot in the SAR image; conversely destructive interference causes the dark spots.

Speckle can be modelled as a random noise effect and several filtering techniques have been proposed to deal with this particular phenomenon over the years. Research suggests that the choice of the particular filtering technique chosen should be governed by the final application. Lee, Frost, Enhanced Lee, median and Gamma-Maximum-A-Posteriori are some of the speckle filters commonly applied as a pre-processors to SAR-based flood detection studies (Martinis et al. 2009; Voormansik et al. 2014).

4.3 Methodology

An overview of the proposed flood mapping approach is illustrated in Fig. 4.6.This methodology was applied to the two SAR images available to this study (See Section 3.3.1, Chapter 3), while high resolution aerial photographs acquired concurrently were used for validation (Section 3.3.5, Chapter 3). The reasons for choosing the processing steps and their subsequent data specific implementation are discussed at length in the following sections.

4.3.1 SAR Preprocessing

The COSMO images were preprocessed using the Gamma Maximum-A-Posteriori (GMAP) filter which suppresses speckle noise while preserving edges and image texture, a property conducive for flood detection (Senthilnath et al. 2013). A window size of 3 × 3 was used as higher resolution SAR images are more susceptible to speckle noise, due to backscatter interference from neighboring pixels adding to sub-pixel interference. The GMAP filtered image is then used for texture analysis, and hereafter referred to as the SAR image.

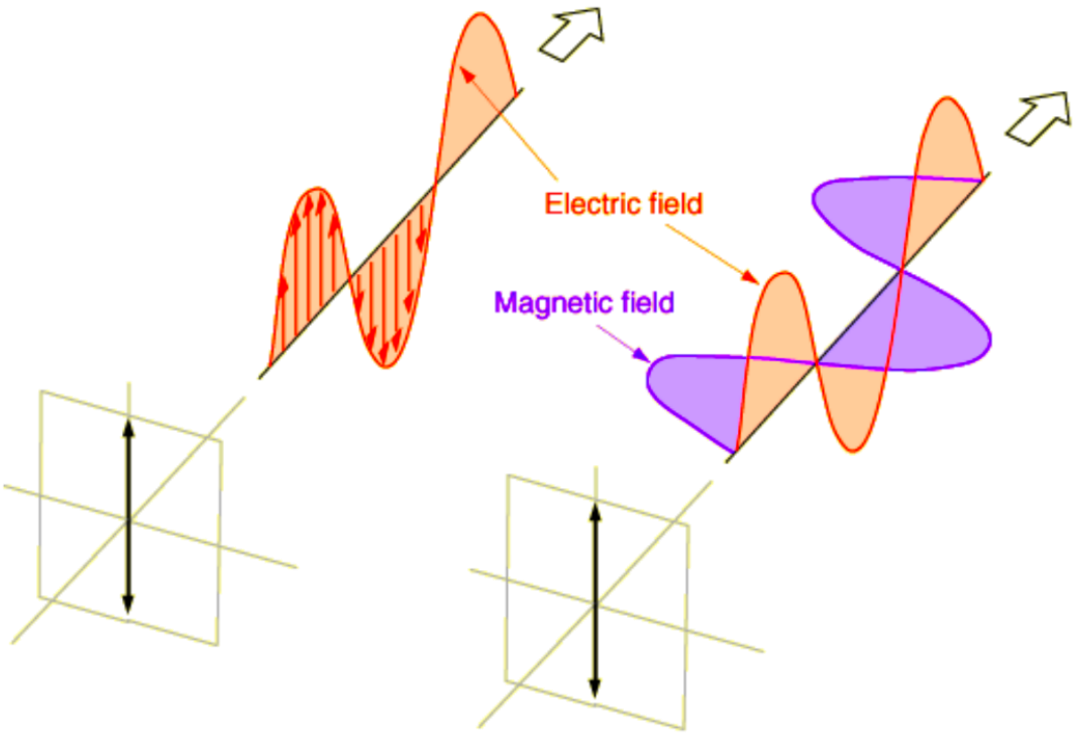

Fig. 4.5 Schematic of EM wave propagation, showing orthogonal electric and magnetic wave components (Source: http://hyperphysics.phy-astr.gsu.edu/hbase/phyopt/polclas.html).

4.3.2 Texture Analysis

Texture based approaches have the advantage that a single SAR image of the event can be used for their derivation and their extraction can be automated. Image texture can be defined as a measure of the transitional probabilities of pixel values, which can facilitate object identification in SAR data (He and Wang 1991; Haack and Bechdol 1999). As SAR images are rich in texture, it has often been utilized for flood identification (Schumann et al. 2009b; Senthilnath et al. 2013; Pradhan et al. 2014). However, most of the studies which utilize textural properties of SAR, have unfortunately failed to quantify their contribution to the overall improvement in flood mapping accuracy. Consequently, this study aims to explicitly assess the role of texture in flood delineation and quantify the maximum improvement possible through its inclusion.

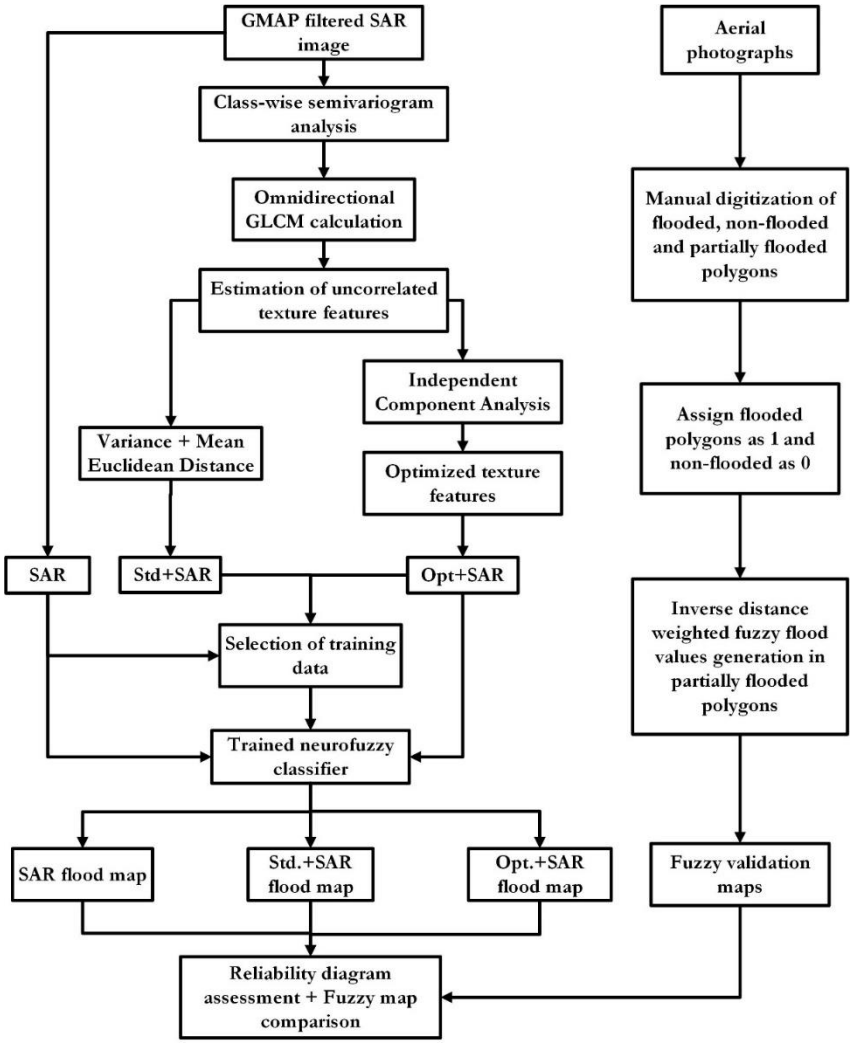

Fig. 4.6 Schematic of the overall classification framework used to generate the fuzzy flood maps.

Objectively selecting appropriate texture features for a particular case study is critical, as this can significantly impact the subsequent flood classification (Schumann et al. 2009a, 2012; Schumann and Di Baldassarre 2010b; Di Baldassarre et al. 2011). Moreover, SAR-based land cover mapping studies have demonstrated the merit of texture optimization methods to solve this problem (Carr 1996; Franklin et al. 1996; Carr and De Miranda 1998; Berberoglu et al. 2000, 2007; Haack and Bechdol 2000; De-yong et al. 2008; Balaguer et al. 2010). However, the impact of optimized texture still needs to be investigated in the context of SAR-based flood mapping as the implementation of texture based methods is fairly empirical (Amitrano et al. 2018; Ouled Sghaier et al. 2018). As texture is direction and scale dependent, these parameter choices can also influence notably the classification, and so must be explicitly considered in any texture based mapping approach (Franklin et al. 1996; Di Baldassarre et al. 2011).

Statistical texture estimation approaches were chosen for this analysis as they utilize non-deterministic properties, governing the distribution of pixel value pairs. Second-order image statistics are also useful for SAR-based flood extent mapping, as the range of spatial autocorrelation for speckle noise is limited to the image resolution in this space (Ulaby et al. 1986). Grey Level Co-occurrence Matrices (GLCM), which can be interpreted as joint gray level probability density distributions or 2-D image histograms, were used for this study due to their low sensitivity to image contrast (Kuplich et al. 2005). It then follows that the GLCM-based texture features which were subsequently derived and used to enhance flood identification in this study, were also relatively insensitive to the land-water backscatter contrast.

This implies that the effect of wind related surface roughening, which usually hampers accurate flood mapping from SAR by significantly increasing the backscatter of open water surfaces, can largely be mitigated. Adding contrast insensitive texture features as additional information layers reduces the dependence on backscatter, and therefore minimizes the impact of wind-induced backscatter variations. As the patterns in the backscatter rather than the backscatter itself were analysed, the probability of correctly classifying a wind roughened flood pixel was increased. This is a clear improvement over backscatter contrast dependent flood classification techniques such as histogram thresholding, which are completely unable to identify inundated pixels under windy conditions as the class distributions are no longer separable. Moreover, the use of a fuzzy mapping technique also allows for a clear expression

of the uncertainty in flood detection; especially in the overlap between the two class distributions.

Each entry in the $n \times n$ GLCM indicates the number of co-occurrences of pixel value pairs at a specific lag distance in a given direction, where *n* is the number of gray levels in the image. For example, the 45° GLCM with one pixel lag for a binary image would record the number of times each combination of grey level pairs ([0,0], [0,1], [1,0], [1,1]) appears in the image separated in the specified direction by one pixel distance. The optimum window size for GLCM calculation was estimated as the range of sensible intra-class variance through semivariogram analysis (Balaguer et al. 2010). Omnidirectional semivariogram curves were generated for both the flood and non-flood classes, by taking homogeneous subsets of size 400×400. This step can be automated in the future as the semivariograms for different sensors and resolutions can be precomputed and used as a look up table, based on archived satellite data. According to this scenario, when a new satellite image becomes available the algorithm would select an appropriate window size for texture estimation based on the specific sensor characteristics.

GLCM matrices obtained in the previous step were used to estimate the second-order textures proposed by Haralick et al., (1973). Mean, variance, homogeneity, contrast, dissimilarity, entropy, angular second moment, and correlation were the co-occurrence measures retained for further optimization as they were least correlated. Direction-invariant texture information was obtained prior to optimization, by averaging the texture values in all eight directions. An independent component transform (ICT) was used, to optimize the textural information and reduce the dimensionality.

As ICT assumes the errors to be of unit variance (white noise), the noise adjusted Principal Component Transform was used to whiten the noise and decorrelate it from the signal (Chica-Olmo and Abarca-Hernández 2000). From the noise whitened data, ICT extracts the direction with the least-Gaussian distribution, and removes the data explained by this variable. A cost function implying non-Gaussianity, such as skewness or kurtosis, is iteratively maximized until the remaining dataset can be explained by statistically independent variables. The first three independent components which contained approximately 99% of the GLCM texture information, characterized by significant Eigen values (Fig. 4.7), were added to the SAR image by layer stacking.

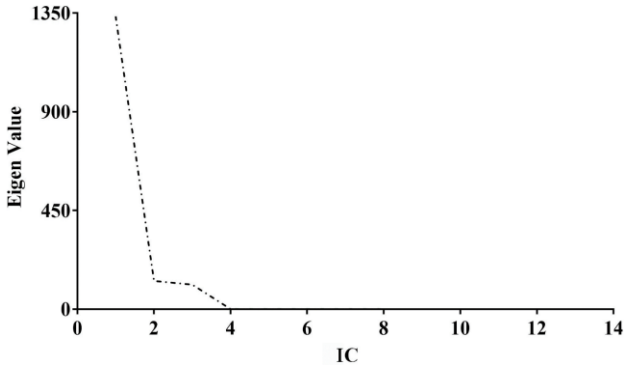

Fig. 4.7 Eigen values of the independent components obtained after analysing the texture bands.

4.3.3 The Adaptive Neuro-Fuzzy Inference System (ANFIS) classifier

4.3.3.1 Fuzzy Membership Function Definition

Fuzzy set theory is a probabilistic adaptation of the classical notion of crisp sets which provides an elegant solution to objectively dealing with the ambiguity of SAR-based flood mapping (Pulvirenti et al. 2011b). An element of fuzzy set is its representation of the degree of membership to a particular category or class, characterized by a predefined function with values ranging from 0 to1. In this study, the Takagi-Sugeno type fuzzy inference system was implemented for the classification, as it can handle non-linearities in the data distribution (Takagi and Sugeno 1985). Gaussian membership functions were chosen for both classes, as the histograms of flooded SAR images can be modelled as a mixture of two normal distributions (Giustarini et al. 2016). Studies have shown that the bimodality assumption fails if the observed flooded area is not significant compared to the tile size (Chini et al. 2017). Therefore, the image was first subset to extract the area of interest - including the flooded area, flood plains and nearby regions - through visual interpretation. This results in a comparable division of flooded and not flooded pixels, causing the image histogram to exhibit a clear bimodality.

4.3.3.2 Function Parameterization

Theoretical electromagnetic backscattering models have traditionally been used to define fuzzy membership functions and parameters (Pulvirenti et al. 2013). However, such approaches require detailed soil, vegetation, and land cover maps, to accurately estimate the expected

backscattering behaviour based on theoretical models, which are often unavailable. Moreover, as these models are wavelength specific, the parameterization is not applicable to data from other sensors or even other areas with a different distribution of scatterers on the ground. Furthermore, theoretical approaches are often unable to capture all practical considerations, for example, seasonal abscission in deciduous vegetation. The use of data driven models like artificial neural networks (ANN), which can learn data characteristics without prior process knowledge, is proposed to counteract this problem. Moreover, the variability in backscatter caused by wind and incidence angle effects - which alter the mean and standard deviation parameters of the class distribution - were implicitly accounted for in this approach, as they were estimated from the same image which needs to be classified.

The ANN was trained using polygons selected manually from the image by an expert, which were consistent throughout the analysis. The hybrid learning algorithm was used, which identifies parameters by iteratively minimizing errors using the gradient descent technique (Jang 1993). The number of epochs used for training were 100 for each case and the model error was estimated using cross-validation. In order to individually quantify the impact of optimized texture on classification performance, the classifier was tested with the following inputs:

1. The speckle filtered SAR image or the control without any addition of texture, called SAR hereafter.
2. The filtered SAR image enhanced with some arbitrarily selected common texture features – variance and mean Euclidean distance in this case – hereafter called Std+SAR.
3. The filtered SAR image with the optimized rather than arbitrary texture bands added, referred to as Opt+SAR.

4.3.3.3 Training Approach

The classifier was tested with three different sets of training data to ensure repeatability of results and to evaluate the sensitivity. The three training datasets are shown in Fig. 4.8 and the strategies for each training set are described below:

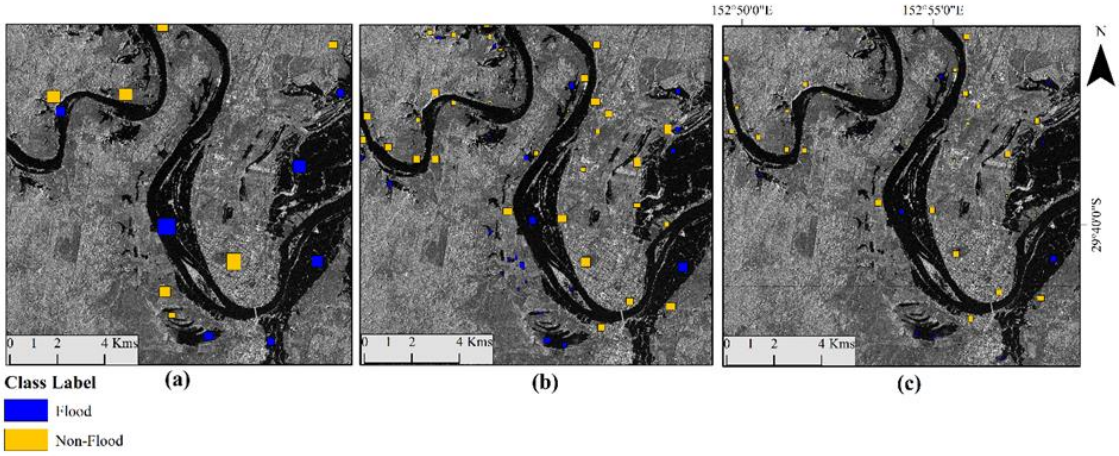

Fig. 4.8 The actual training polygons selected to train the neuro-fuzzy classifier are shown, where (a), (b), and (c), correspond to training sets 1, 2, and 3, respectively. The difference in the three trainings is primarily the size of individual polygons and their corresponding locations as shown by the coloured squares.

Set 1. Large area polygons were drawn for each class, encompassing the backscatter variability of the target classes, and spread out across the image. An equal number of polygons were chosen for the flood and non-flood classes.

Set 2. Smaller area polygons, each comprising of a nearly uniform subset of backscatter values representing one of the signatures, were selected for the training. Polygon selection was restricted to the flooded area and floodplains, where maximum classification accuracy was desired. More non-flood polygons were chosen than flood, to ensure sufficient representation of the entire spectrum of backscatter variability for training.

Set 3. Same as in training set 2, except that the size and number of polygons was further reduced to ensure that the classifier did not overfit the training data and to maximize computational efficiency.

As results of training Set 1 yielding large errors during the classifier assessment phase, it was not tested subsequently for the flood mapping or to assess the classification accuracy. However, a description of the training set has been provided here to highlight the problems with this set and inform readers about these limitations.

4.3.4 Validation Strategies

In order to ensure the reliability of this analysis, results were validated using two different methods. First, a fuzzy set approach was used to validate the SAR-based fuzzy maps against a fuzzy validation target. Second, reliability diagrams were used to assess the flood maps, being the most commonly used validation technique for probabilistic maps.

4.3.4.1 Fuzzy Set Validation Approach

Shoreline extraction from aerial photography can be quite ambiguous in densely vegetated and built up environments (Giustarini et al. 2013). The accuracy of the derived shoreline may vary from 10 to 100 m, depending largely on the skills of the photo interpreter (Mason et al. 2010a). In this particular case, the riparian vegetation at the edge of the storage areas in the flood plain made it rather challenging to "see" the underlying water edge. The illumination differences in the separate flight lines, combined with atmospheric effects, severely affected the clarity of boundaries in the area, as seen in Fig. 4.9. Studies have proposed marking ambiguous boundary regions as "no data" to remove the associated uncertainty (Giustarini et

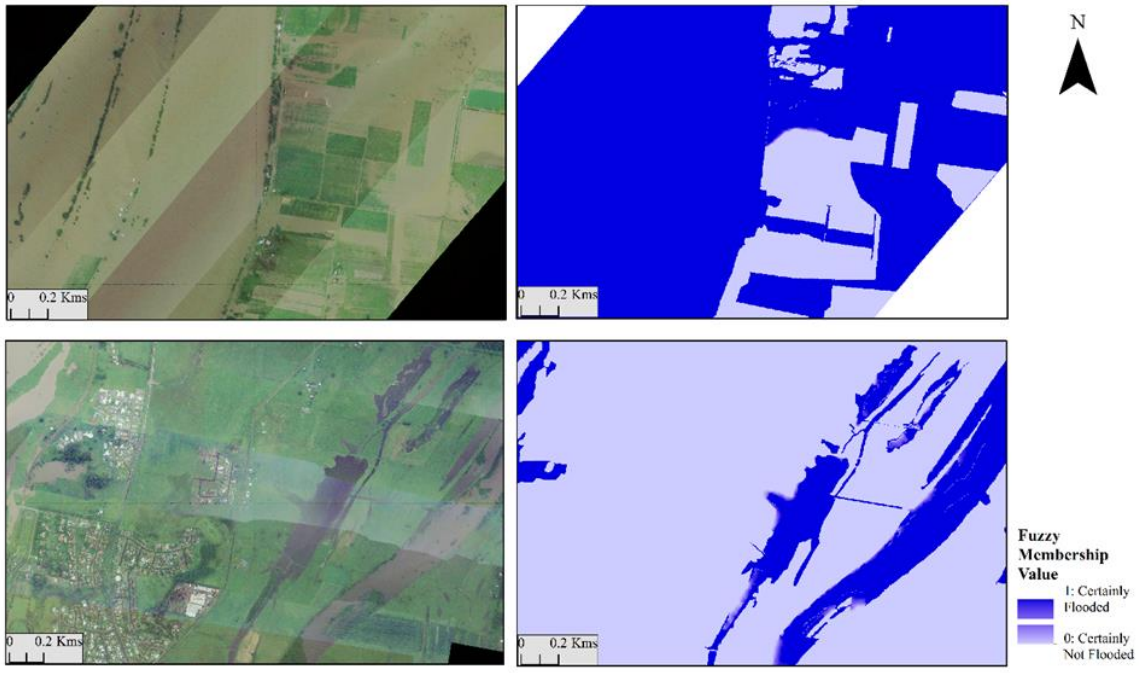

Fig. 4.9 Example subsets of the true colour aerial photographs (left) shown along with the corresponding manually fuzzified flood maps (right). Locations are shown in Figure 4.1.

al. 2016). However, it is more appropriate to use a performance measure which is capable of handling uncertainties, rather than discarding potentially valuable information.

A number of fuzzy performance measures capable of handling uncertainties in validation data have been developed for model calibration, especially those which use remote sensing data as targets (Pappenberger et al. 2007b). However, the utility of such metrics for the validation of SAR-based flood maps has not yet been tested. In order to facilitate this, the fuzzy map comparison method proposed by Hagen (2003) has been adopted in this study for the assessment of probabilistic flood maps for the first time. The proposed approach takes into account locational as well as categorical uncertainty in both the input as well as the validation data (Hagen-Zanker et al. 2005; Wealands et al. 2005). The resulting map comprises of pixel-wise similarity values ranging from 0 to 1, indicative of the local goodness of fit. As this evaluation approach requires an explicit representation of uncertainties in both datasets, a fuzzy flood map for validation was prepared as follows to reflect the errors of manual flood delineation:

1. The aerial photographs were manually digitized into three classes - clearly flooded, clearly non-flooded, and possibly flooded.
2. The flooded polygons were assigned a value of 1 and non-flooded polygons, a value of 0.
3. The partially flooded polygons were filled with intermediate values, interpolated using an inverse distance weighting (IDW) algorithm with an exponent of 2, to fuzzify the validation data.

4.3.4.2 Fuzzy Similarity Statistics

In order to characterize the similarity between the SAR-based flood map and the one derived from aerial photos, several statistics were calculated which could highlight the different aspects of classifier performance. First, deterministic grid-based statistics, such as Root Mean Squared Errors (RMSE) and Mean Absolute Errors (MAE), were calculated to assess classifier performance. This was followed by the map comparison method proposed by Hagen (2003), which allows the comparison of fuzzy maps while explicitly accounting for spatial and categorical uncertainties in both datasets.

For the two fuzzy vectors; $F_{SAR}$was obtained as the SAR-derived flood extent while F_{VAL} represents the validation data. The similarity measure *S* can therefore be computed for two pixels at the same location in both maps using:

$$S(F_{SAR}, F_{VAL}) = \left[\left| F_{SAR_{flood}}, F_{VAL_{flood}} \right|_{min}, \left| F_{SAR_{non-flood}}, F_{VAL_{non-flood}} \right|_{min}\right]_{max}, \quad (4\text{-}2)$$

where $F_{SAR_{flood}}$ and $F_{SAR_{non-flood}}$ denote the fuzzy values of the flood class and the non-flood class memberships at a given pixel in the SAR-based map, while $F_{VAL_{flood}}$ and $F_{VAL_{non-flood}}$ refer to corresponding values in the validation map. S is the maximum value within the set obtained by evaluating the minimum of the two fuzzy vectors F_{SAR} and F_{VAL} on a pixel by pixel basis (Pappenberger et al. 2007a), which can also be interpreted as the maximum grade of membership to the intersection of the fuzzy sets (Zadeh 1965). S takes values from 0 to 1 based on the degree of similarity, such that 0 was assigned to cells that were completely different and 1 was assigned to those which were identical. In order to account for fuzziness in location, the influence of neighborhood cells was also considered. The contribution of each neighborhood cell can be calculated using a distance decay function chosen based on the nature and magnitude of uncertainties, and the desired tolerance for spatial error (Hagen 2003).

In this study, the number and impact of neighbors included in the analysis were selected based on the semivariogram assessment results. The neighborhood influence was estimated based on the 3D exponential decay function with a halving distance of two pixels. This essentially means that the influence of the neighborhood reduced by half after the second pixel is encountered in any direction. The two way similarity between the fuzzy distance weighted neighborhood contributions $\left(F_{SAR_{nbh}}, F_{VAL_{nbh}}\right)$ and the fuzzy membership values for the central cells $\left(F_{SAR_{cc}}, F_{VAL_{cc}}\right)$ were calculated using the expression:

$$S(F_{SAR}, F_{VAL}) = \left|S\left(F_{SAR_{nbh}}, F_{VAL_{cc}}\right), S\left(F_{SAR_{cc}}, F_{VAL_{nbh}}\right)\right|_{min}, \quad (4\text{-}3)$$

For further information about the calculation of the neighborhood sets, the reader is referred to (Hagen-Zanker et al. 2005). Last, the fuzzy kappa K_{fuzzy} statistic is proposed as an overall performance measure, calculated using:

$$K_{fuzzy} = \frac{(S_{obs} - S_{exp})}{(1 - S_{exp})}, \quad (4\text{-}4)$$

where S_{obs} and S_{exp} are the observed and expected percentages of fuzzy agreement. Here the observed percentage of fuzzy agreement refers to the spatial similarity between the SAR-based

maps and validation data computed using Equation 2. Conversely, the expected agreement is a function of the number of classes and the image histogram calculated theoretically (Hagen 2003). K_{fuzzy} only differs from the traditional Cohen's kappa popularly used for map comparison, in the calculation of the expected percentage of agreement S_{exp} (Hagen-Zanker 2006).

The fuzzy Kappa statistic quantifies the improvement between the compared maps relative to a randomly generated categorical map with an identical histogram (Wealands et al. 2005). S_{exp} is estimated as the probability that a wet pixel observed in the validation data appears within a certain range of pixel distances or neighborhood in the corresponding SAR-based flood maps. This means that a flooded pixel in the validation data, which may have shifted in the SAR-derived map due a variety of factors such as sensor orientation or geo-location errors, was considered a match if it fell within reasonable bounds of uncertainty. S_{exp} can be viewed as a measure of the chance agreement that the two maps in consideration may exhibit, based on chosen neighborhood sizes and the number of classes in the dataset.

Neighborhood rings are defined as the set of cells located at an equal distance from the central cell. This implies that the chance agreement needs to be computed for each central cell and all possible neighborhood rings. Ideally this should be done for each pixel individually and for an infinite zone of influence, as edge pixels may have a different nature and number of possible neighborhood rings. However, on increasing the search radius for neighborhood identification from 10 cells to 500, the difference in values of S_{exp} was found to be insignificant, as the function used to model the influence decays rather rapidly. Furthermore, using a search radius of 500 increases the possible number of permutations and combinations, thereby drastically increasing the computational time. Finally, a radius of 10 pixels was used to calculate the expected agreement between the SAR-derived and aerial photo based fuzzy flood maps.

4.3.4.3 Reliability Diagram Assessment

The final performance assessment used reliability diagrams, being the currently accepted evaluation method for probabilistic maps in literature. As discussed previously, the limitation of this technique is that the validation data need to be binary. However, this analysis was conducted to facilitate the understanding of error characteristics with respect to state of the art mapping techniques and to understand the specific contributions of under or over

predictions. A threshold of 0.5 was chosen for defuzzification, to reflect the maximum uncertainty in the resulting deterministic flood map (Schlaffer et al. 2017). As the validation map was digitized on an aerial photo with a spatial resolution of 10 cm, the choice of this threshold has limited impact on the outcome of the analysis conducted at 3 m.

The agreement between the fuzzy membership values predicted by the neurofuzzy classifier, and the observed proportion of flooded pixels in the validation data can be characterized by a reliability diagram (Horritt 2006). The fuzzy membership values were binned into intervals of 0.1 and the ratio of wet cells to total number of cells in each bin, was plotted against the bin means. The ideal classifier assigns fuzzy membership values identical to the proportion of observed wet pixels in the validation data, with deviation from the 1:1 line representative of the classification error. As the distribution of pixels across the bins is non-uniform, a weighted RMSE (WRMSE) was calculated to objectively represent the uncertainty. The WRMSE is calculated by assigning weights to the bin errors, based on the bin population as in (Giustarini et al. 2016):

$$WRMSE = \sqrt{\frac{\sum_{i=1}^{N}(F_v - F_S)^2 n_i}{\sum_{i=1}^{N} n_i}}, \quad (4\text{-}5)$$

where n_i is the pixel count of each bin, F_v is the observed proportion of wet cells in the validation map, F_s is the fuzzy membership value predicted based on SAR analysis, and N is the total number of pixels in the validation domain. Reliability diagrams were also used to assess classifier performance for each land cover (LC) class in the study area through reliability diagrams. Such diagnostic analyses may lead to insights on when the proposed approach can be expected to perform well.

4.4 Results and Discussion

4.4.1 Window Size Selection

The choice of a window size is an essential step for texture estimation, to avoid inadvertently interpreting noise as a meaningful pattern. Therefore, the range of spatial autocorrelation observed through the flood and non-flood semivariograms was used as the window size for texture analysis. The semivariogram plots were estimated by visually selecting homogeneous subsets of flood and non-flood classes, to ensure that the observed range reflects

only the intra-class variance. The non-flood subsets were located clearly outside and away from the floodplain area, such that purely dry land pixels could be isolated.

Fig. 4.10 shows the semivariogram plots obtained for the flood and non-flood classes. The range of both semivariograms is consistent at 3 pixels as expected, as high-resolution data are more prone to noise and thus exhibit low spatial autocorrelation. The non-flood class exhibits significantly higher values of variance due to the diversity of surface scattering characteristics encountered on land.

4.4.2 Neuro-Fuzzy Classifier - Training, Testing, and Validation

The results for the neuro-fuzzy mapping are presented in the following manner. First, the general training procedure and results of the training are discussed, followed by an evaluation of the result maps obtained for each validation site. As the classifier uses a data driven estimation of the model parameters, it needs to be trained and subsequently subjected to rigorous testing. It is important to note that for statistical models, two types of errors need to be estimated:

1. Model selection error - which characterizes the prediction error of the selected model in the context of other available models. This analysis allows to select the one most suited to the data distribution;
2. Model assessment error - which evaluates the ability of the selected model to correctly classify new or previously “unseen” data points (Hastie et al. 2009).

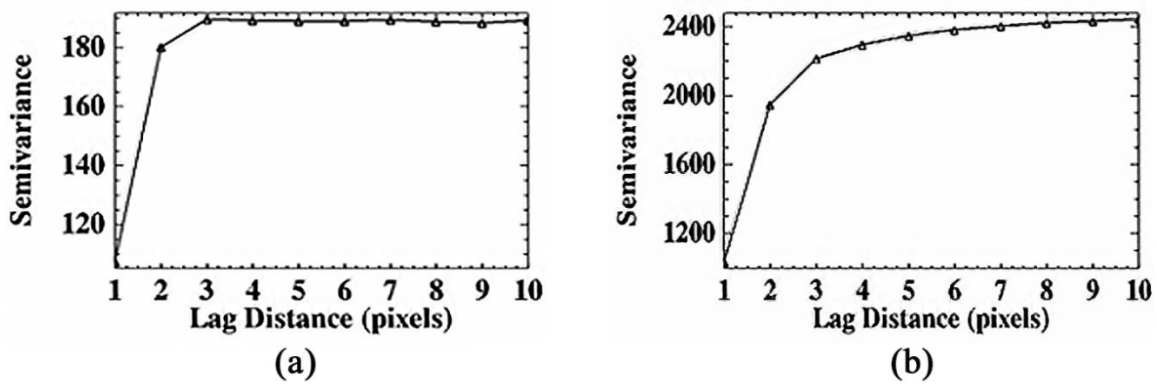

Fig. 4.10 Semivariograms showing spatial autocorrelation amongst backscatter values for the different classes in a COSMO-SkyMed 3m image.

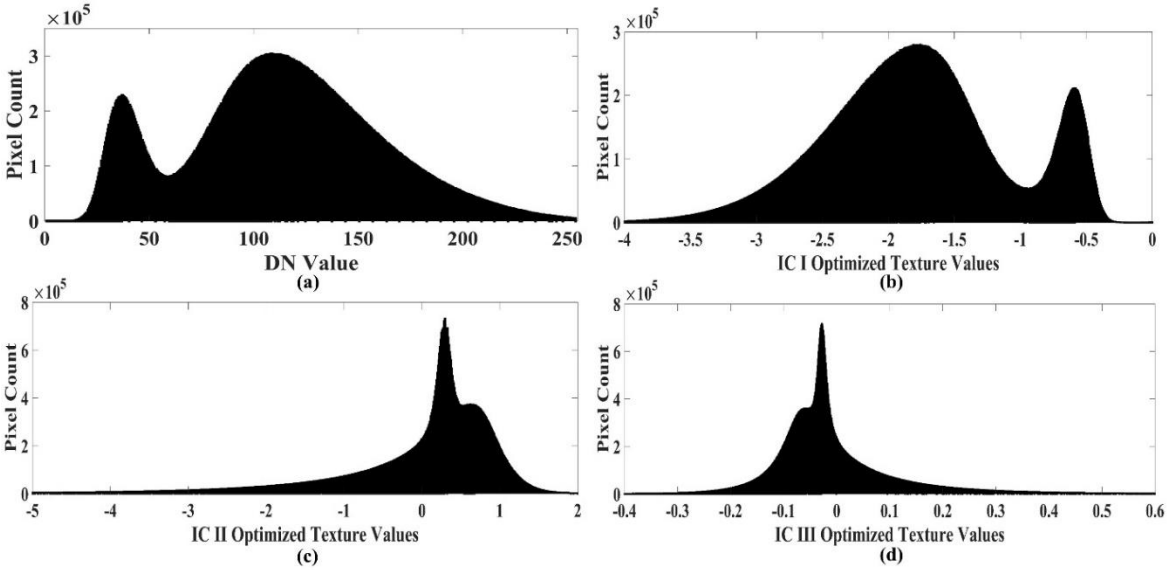

Fig. 4.11 Histograms depicting the bimodality in the distribution of pixel values for (a) the filtered SAR image in digital numbers, (b) Independent Component (IC) I texture values, (c) IC II texture values, and (d) IC III texture values.

The appropriate way to assess both, if sufficient data points are available, is to randomly divide them into three parts: a training set, a validation set, and a testing set. Based on the signal-to-noise ratio expected of a high resolution SAR image and the complexity of the Gaussian model, the split was chosen as 70% training, 15% validation and 15% testing (James et al. 2000). The training set includes the majority of data points by convention, and is used to fit the data driven model or distribution. Ideally, all distributions which are able to sufficiently explain the data characteristics should be examined. A separate dataset, hereafter referred to as the validation set, is used to compute the prediction error for all the competing models or the model selection error. The model with the lowest value of validation error is chosen for further analysis. Since in this case the Gaussian model was already selected based on the histogram bimodality, the validation error was used to select a suitable parameterization of the distribution.

Once a distribution has been trained and validated as the best fit model, another previously unseen batch of data, called the test set, is used for assessment. The test error provides a measure of the generalization capabilities of the chosen model and the ability to correctly classify new data points. If the validation set is reused for this evaluation, the true test error will be substantially underestimated. Ideally, multiple equally plausible model structures

should be tested to assign the test error. From the histograms of the SAR image and the optimized texture features presented in Fig. 4.11, it is apparent that the distributions were exhibiting slight deviations from Gaussianity. However, investigating different distributions with better fits was considered outside the scope of the present investigation, primarily because histograms of flooded SAR images are usually known to consist of a mixture of two Gaussian class distributions, which is also evident from the backscatter histogram shown in Fig. 4.11(a) (Chini et al., 2017). As the distributions of the texture classes were also nearly symmetrical and choosing a case-specific distribution function would limit transferability of the method, the Gaussian assumption was maintained throughout the analysis.

The error values followed the expected pattern of lower training set errors and larger test and validation set errors as summarized in Table 4.1. Training Set 3 led to a more generalized classifier as errors were nearly consistent across training, validation, and test sets. As elaborated earlier, minimizing the sample size can reduce overfitting which in turn can improve classifier performance, by reducing the bias-variance trade-off (James et al., 2000). This is in contrast to Set 2, where the validation errors were noticeably higher for both the texture based methods. The larger magnitude of test and validation errors can be due to overfitting to the training data in Set 2, which adversely affects model generalization. Generalization can be defined as the ability of a classifier to correctly identify previously "unseen" data points (Hastie et al., 2009). The pixel-wise comparison of deterministic statistics was carried out for both training Set 2 and Set 3, however, the reliability diagrams and fuzzy

Table 4.1 Mean Absolute Error statistics for classification model selection and predictive capability assessment.

	Training Set 2			Training Set 3		
Input	Training	Validation	Testing	Training	Validation	Testing
SAR	0.184	0.183	0.186	0.217	0.217	0.216
Std+SAR	0.155	0.178	0.159	0.174	0.175	0.175
Opt+SAR	0.121	0.136	0.125	0.117	0.12	0.12

statistics were only calculated for Set 3. Set 3 was chosen for the more detailed analysis as the classifier achieved a better generalization and was expected to perform better.

Note that the training, testing, and validation datasets defined in this section are subsets of the training data selected to "train" the neurofuzzy classifier, and the meanings of these terms are specific to the data driven modelling part of this study. After the model selection and training process was conducted, the trained model was used to generate fuzzy flood maps from SAR. These were validated against manually derived flood maps from aerial photography, hereafter referred to as the validation data.

4.4.3 Fuzzy Flood Maps - Accuracy Assessment

The fuzzy flood maps obtained post classification were assessed using two approaches - using the fuzzy map comparison and reliability assessment. For each validation site, the fuzzy flood maps, difference maps, spatial similarity maps and reliability diagrams were generated through the procedures detailed in Sub-section 4.3.4.

4.4.3.1 Validation Site 1: Junction Hill

The pixel-wise assessment results from the two training datasets were found to be somewhat inconsistent, as evident from Table 4.2. The proposed texture optimization approach succeeded in reducing the RMSE slightly (~2%) in Set 2. However, Set 3 exhibited a slightly higher RMSE, though increase of 10^{-3} can be considered to be negligible. It can easily be observed from the fuzzy flood maps illustrated in Fig. 4.12 that the optimized texture approach reduced the misdetection of linear and smooth urban features, like road networks, as flooded pixels in the urban land-use dominated (top row) subsets.

The proposed approach also reduced the fuzziness in the sparsely vegetated region which has varying backscatter but relatively homogenous texture, as seen in the second row of Fig. 4.12 where local improvements are visible. However, the pixel-wise assessment was unable to capture this improvement as both the maps had very different uncertainty characteristics. A deterministic differencing approach doesn't work with such continuous random variables as an exact match is nearly impossible. Furthermore, the difference maps depicted in Fig. 4.13 highlight the disparity between the error characteristics in the validation data and the SAR-based maps. In fact, the SAR-based maps provide more realistic estimates of uncertainty at each

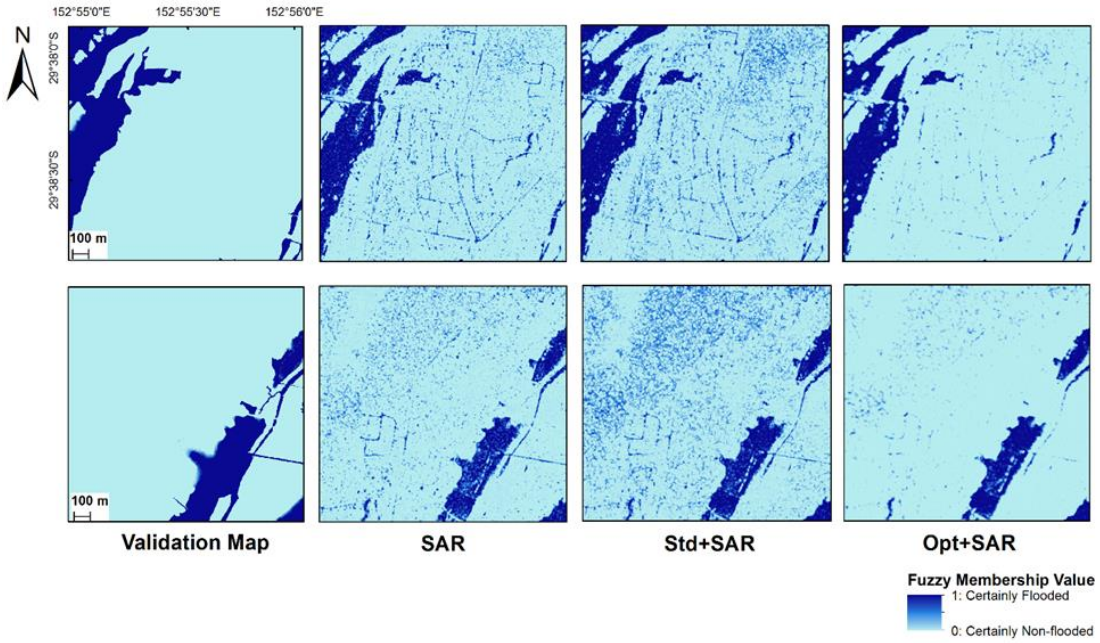

Fig. 4.12 Validation maps generated from aerial photographs are displayed in the first column, followed by flood maps derived by processing the following inputs through the ANFIS classifier; SAR alone in column two (SAR), arbitrarily selected textures with SAR in column three (Std+SAR), and optimized textures with SAR in column four (Opt+SAR). Areas depicting maximum reductions in uncertainty for the Junction Hill test site were chosen for illustration. The locations of the subsets used here are shown in Fig. 4.1.

pixel, as these are calculated objectively through the ANFIS classifier. Conversely, given that the chief contributor to the uncertainty in manually digitized flood maps is the skill of the analyst (Mason et al. 2010a), quantifying this value objectively is significantly more challenging.

The fuzzy set evaluation approach results in two sets of maps; one where only the fuzziness of the membership value or class assignment is considered and one which additionally includes both fuzziness of value as well as the fuzziness of location. The fuzzy similarity set presented in Fig. 4.14 shows a clear improvement but the transition zone uncertainty is

Table 4.2 Root Mean Squared Error statistics based on the pixel-wise deterministic difference operation.

	Training Set 2		Training Set 3	
	Junction Hill	**Ulmarra**	**Junction Hill**	**Ulmarra**
SAR	0.242	0.243	0.233	0.261
Std+SAR	0.239	0.246	0.241	0.257
Opt+SAR	0.222	0.24	0.236	0.24

spuriously inflated. As the validation data are manually digitized and fuzzified, the nature of uncertainties in the transition zone may be very different from that of SAR-based approaches. In contrast, the spatial similarity index illustrated in Fig. 4.15 reflects the improvements offered at the land-water boundary much better. A significant reduction in uncertainties is noticeable across the domain.

The overall performance (average spatial similarity index) values were 0.899, 0.905 and 0.915 for SAR, Std+SAR, and Opt+SAR, respectively (Table 4.3). The fuzzy kappa statistic, which corrects for the expected percentage of agreement and can be more informative, is identical to average similarity in this case due to low expected agreement. The K_{fuzzy} can provide a measure of the overall improvement, which is categorically required for many applications. As the expected value for similarity is directly related to the number of classes and the intra-class pixel distribution, the values obtained for this statistic can also be insightful (Hagen 2003).

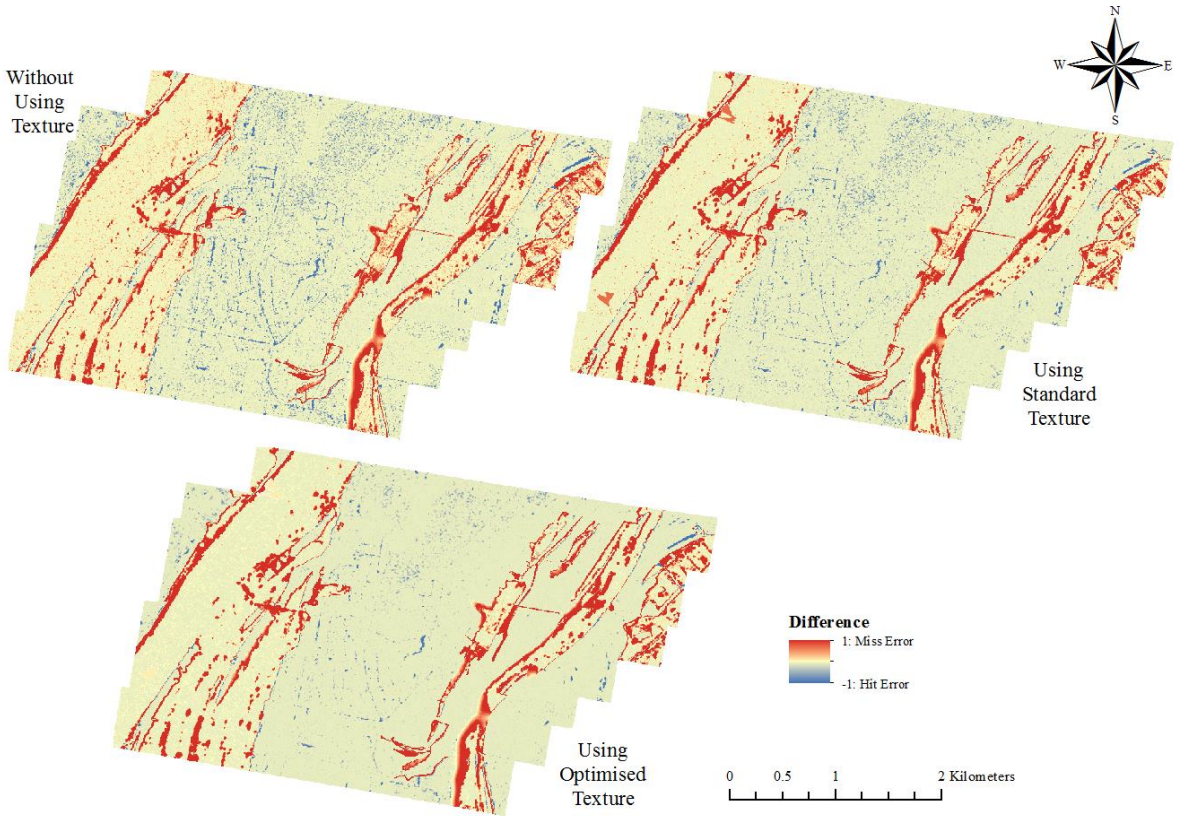

Fig. 4.13 Pixel-wise difference maps generated by subtracting the SAR-based flood maps, from the fuzzy validation map digitized from aerial photos, for entire Junction Hill region.

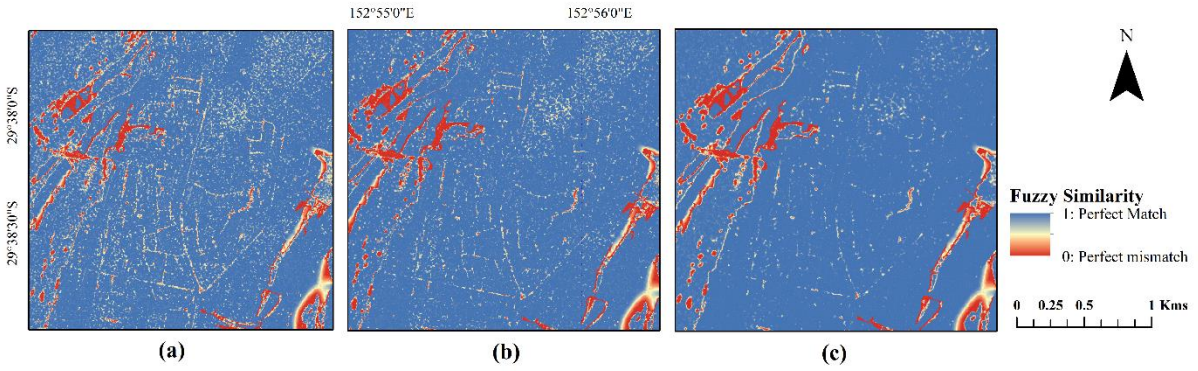

Fig. 4.14 Fuzzy similarity maps for the central cell comparison between the Junction Hill validation map and the SAR-based fuzzy maps, where (a) SAR represents the use of SAR alone as a classification input, (b) Std+SAR represents arbitrarily selected textures with SAR, and (c) Opt+SAR represents optimized textures with SAR.

Expected similarity can be understood as the probability of chance agreement between the two maps having identical image histograms. In the context of fuzzy spatial similarity, this value of possible chance agreement is computed for all the cells in the neighbourhood rings under consideration. One of the possible reasons for the low expected similarity values obtained might be the distance decay function chosen here, which was the 3-D exponential function with a halving distance of two pixels, reducing the neighbourhood influence drastically. This observation is in keeping with the expectation, as the nature of uncertainties in the validation data are very different from the SAR-based maps. Since the validation map is manually digitized and fuzzified, error variation is expected to differ substantially from the fuzzy flood maps from SAR, which objectively represent backscatter and classification uncertainty.

The spatial auto-correlation in high-resolution SAR data are also highly localized due to increased speckle noise, as already established through the variogram analysis. However, it is expected that K_{fuzzy} can help with assessing probabilistic flood maps and may add value to the average spatial similarity. The expected similarity S_{exp} would obviously vary with the spatial resolution of the maps, choice of the distance decay function, and the number of classes considered in the analysis. In cases like this where S_{exp} is nearly negligible, the average spatial similarity statistic S_{obs} may suffice as a test statistic for map comparison, as the K_{fuzzy} doesn't add any new information.

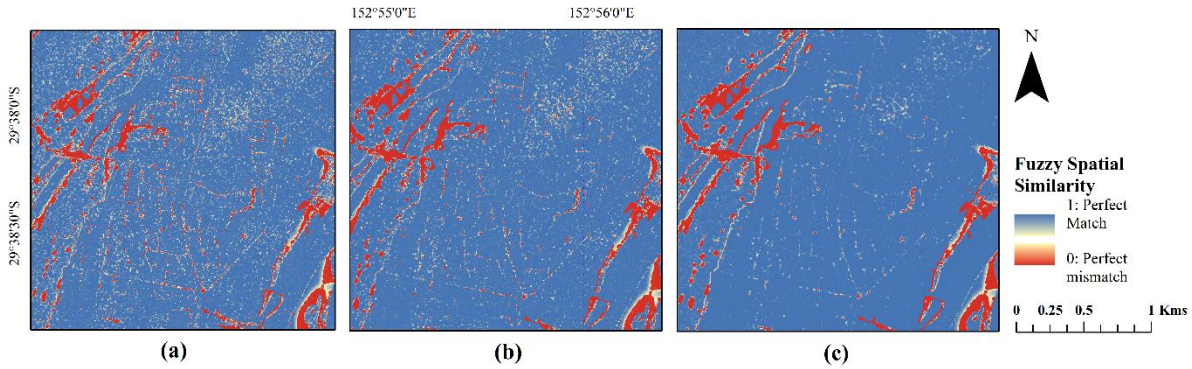

Fig. 4.15 As for Fig. 4.14 except with neighbourhood context included for the Junction Hill site with (a) SAR, (b) Std+SAR, and (c) Opt+SAR.

The reliability diagram for Junction Hill (Fig. 4.16) better reflects the improvement offered by Opt+SAR, which consistently gave predictions very close to the 1:1 line. Both the texture based approaches correctly classified nearly all the pixels for the last few bins, containing flood membership values ranging from 0.8-1. These are the certainly flooded pixels of the study area and the ones the algorithm primarily seeks to correctly identify. The addition of optimized texture, especially after optimization, seems to be conducive to this cause. Furthermore, omission errors seem to dominate over commission errors, for most of the uncertain bins. A closer examination reveals that the underprediction increases with the uncertainty, i.e. the Opt+SAR assigned lesser pixels to bins 4, 5, and 6 than observed in the validation data. The SAR image alone seemed to over predict in bin 9 and under predict in bin 10, both of which contained pixels almost certainly flooded, indicating a contribution of noise in the training. Consequently a reduction in the signal to noise ratio led to ambiguity in classification even within homogeneous image objects such as flood patches.

Texture estimation highlighted patterns and increased the separability between signal and noise. This reduced the uncertainty in feature extraction, offering some advantages for flood

Table 4.3 Summary of fuzzy statistics for the two validation sites.

	Junction Hill			**Ulmarra**		
Input	**Observed Similarity** (S_{obs})	**Expected Similarity** (S_{exp})	**Fuzzy Kappa** (K_{fuzzy})	**Observed Similarity** (S_{obs})	**Expected Similarity** (S_{exp})	**Fuzzy Kappa** (K_{fuzzy})
SAR	0.899	$3.341e^{-12}$	0.899	0.670	$4.843e^{-10}$	0.670
Std+SAR	0.905	$3.168e^{-12}$	0.905	0.689	$3.799e^{-11}$	0.689
Opt+SAR	0.915	$1.342e^{-12}$	0.915	0.713	$3.827e^{-10}$	0.713

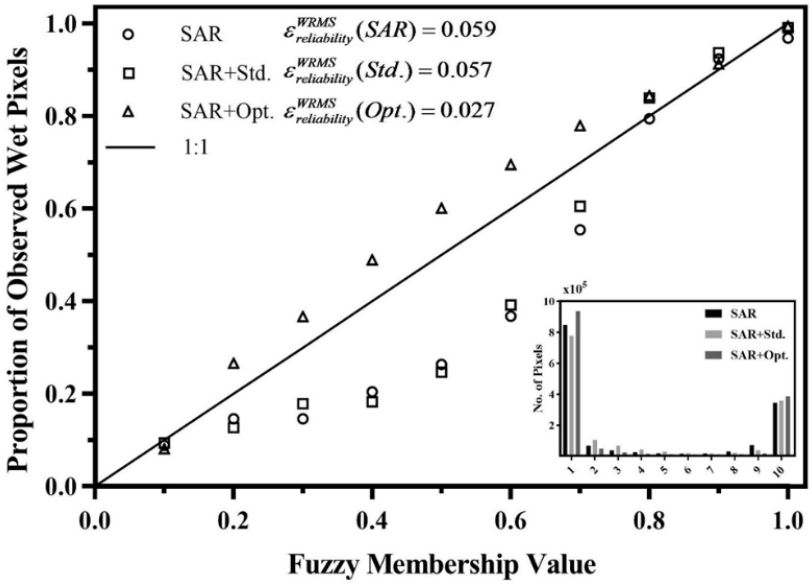

Fig. 4.16 Reliability diagram for the Junction Hill area with Weighted Root Mean Squared Error values and bin sizes represented in a sub-plot.

assessment. The change in the distribution of bin-sizes on the application of the Opt+SAR approach, evident from the subplot included in Fig. 4.16, corroborates this interpretation. The WRMSE values based on the deviation from the 1:1 line show a relative improvement of 54.2% over the use of SAR image without texture, and 52.6% over the use of standard textures when using the optimized texture approach. This clearly indicates the importance of choosing appropriate texture features, as the addition of arbitrary texture features reflects no significant improvement.

4.4.3.2 Validation Site 2: Ulmarra

At Ulmarra, the pixel based statistics show a reduction in RMSE for both training Set 2 and 3. However, the magnitude of this reduction was larger in Set 3. According to the flood maps in Fig. 4.17, this validation site had a larger proportion of flooded pixels. Since a significant percentage of the image is covered with a homogeneous land cover type, the impact of noise is more prominently visible. The optimized approach performed better in Set 3, due to the better model generalization achieved. The generalized model exhibited lower noise sensitivity. The Opt+SAR flood maps showed a noticeable reduction in fuzziness within the homogeneous regions and near the transition zone, in agreement with the findings from Junction Hill. While the pixel-wise RMSE values reduced with the addition of texture, it is important to

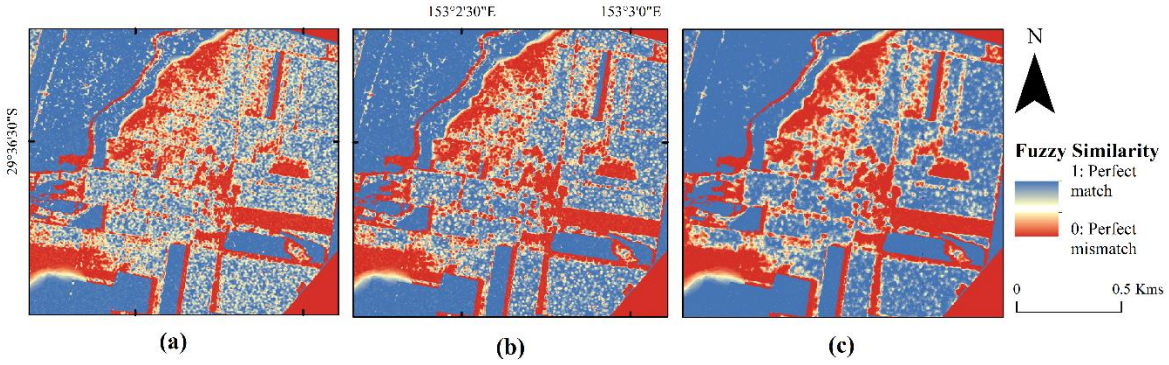

Fig. 4.18 As for Fig. 4.14 but for the Ulmarra region, where the agreement with validation data is shown for (a) SAR, (b) Std+SAR, and (c) Opt+SAR.

consider the failure of the deterministic differencing approach to capture improvements in mapping for the Junction Hill site. The problem of non-contiguous vegetation patches surrounded by flood pixels is observable even in Ulmarra. The deterministic difference maps for Ulmarra have not been included here for brevity as the inferences from the analysis were similar to those from Junction Hill.

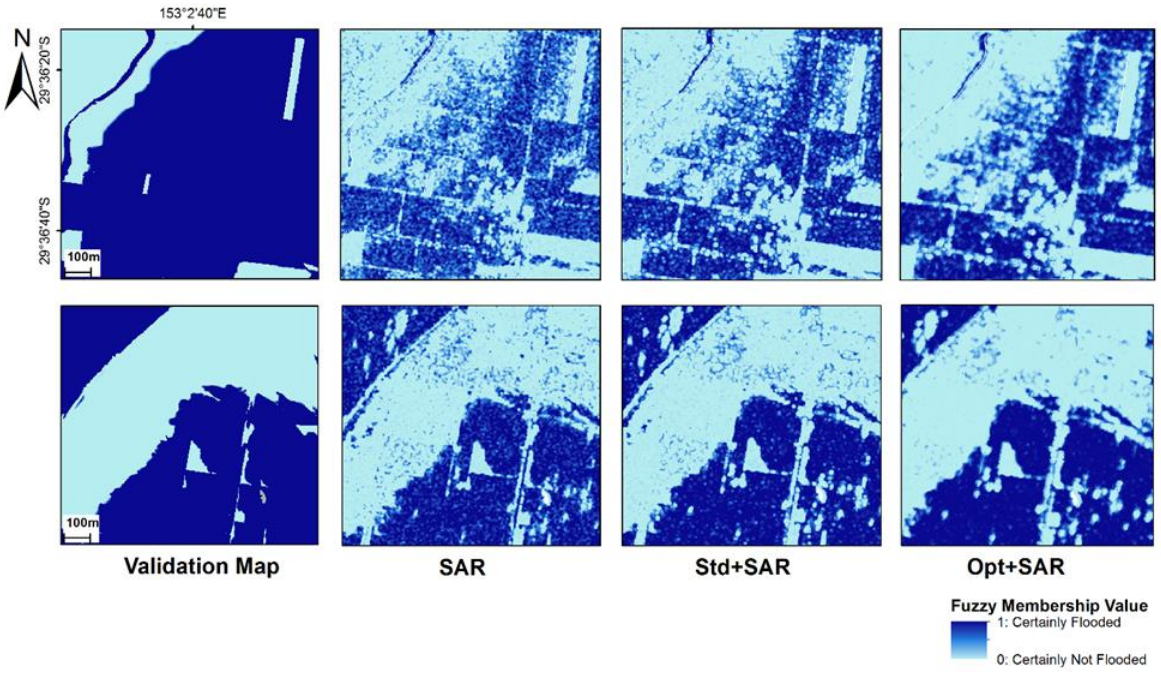

Fig. 4.17 As for Fig. 4.12 but for the Ulmarra test site, with locations of the chosen subsets highlighted in Fig. 4.1.

Additionally, some portion of the domain had agricultural fields with standing crops, highlighted in the top row of Fig. 4.17, making the underlying flood water very difficult to detect. A large portion of this area had been classified as non-flooded by all the SAR based approaches. As texture is a backscatter-derived property, the approach is unable to bring significant improvements in this area. Since the experiment was conducted using low-wavelength X-band data, which experiences severe scattering due to emergent vegetation, it was rather challenging to demonstrate the full potential of the proposed approach in this region. However, the RMSE still showed an improvement overall as the texture optimization seemed to reduce noise sensitivity. Since this particular test site was populated with homogeneous flood patches, where the effects of noise dominate, the overall errors were decreased.

The fuzzy similarity maps for Ulmarra (Figure 4.18) show large regions of uncertainty in the flood transition zone. Although the Opt+SAR map shows some improvements compared to the SAR and Std+SAR, it is noteworthy that the inclusion of texture itself offers limited enhancement in classification performance. This is evident from Fig. 4.18, which shows a notable reduction in uncertainties in the Opt+SAR compared to the Std+SAR approach. This implies that in areas where texture based approaches can be expected to work, such as distinguishing between water and water look-alike surfaces, the Opt+SAR technique will definitely provide better outputs than arbitrarily selected texture features.

Spatial similarity maps provide a suitable assessment measure for probabilistic maps, as a clearer reduction in the transition zone uncertainty is visible in Fig. 4.19. S_{obs} values were 0.67, 0.69, and 0.71 for SAR, Std+SAR, and Opt+SAR, respectively, as shown in Table 4.3. The K_{fuzzy} values were identical to S_{obs}, as in the case of Junction Hill, as values of S_{exp} were

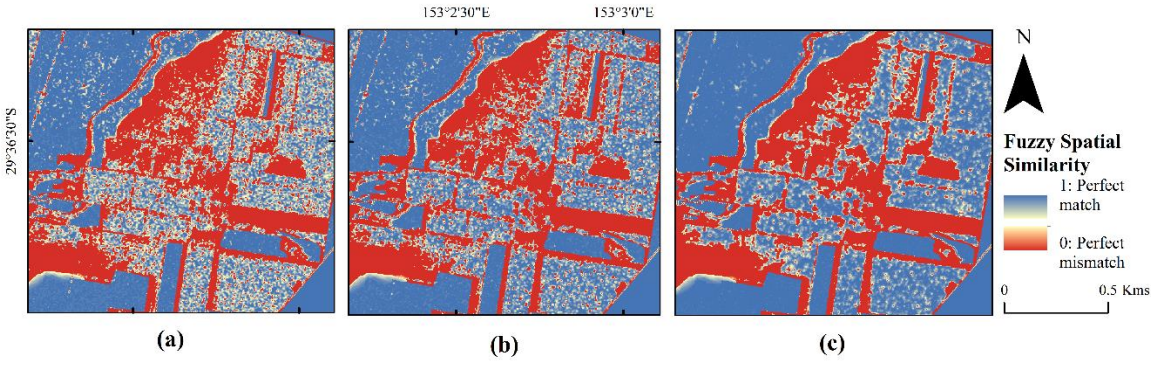

Fig. 4.19 As for Fig. 4.15 but for the Ulmarra test site with similarities shown as (a) SAR, (b) Std+SAR, and (c) Opt+SAR.

low for the same reasons. However, the K_{fuzzy} and S_{obs} both showed improvements when using the proposed technique at both test sites. This corroborates the visual observations from the fuzzy maps (Fig. 4.14 and Fig. 4.18), generated from each input image. The fuzzy map comparison technique used here can also be thought as something which mimics human interpretation of errors. The central cell comparison with the additional consideration of contextual neighbourhood influence is a powerful tool, which should further be imbibed for the assessment of probabilistic maps.

At Ulmarra the reliability diagram statistics are inconsistent with the results obtained from the fuzzy set analysis as illustrated in Fig. 4.20. It is important to observe that WRMSE increases with the addition of any texture in this case. One of the possible reasons for this could be the large number of certainly flooded pixels classified as not flooded in the SAR based maps. This large discrepancy is due to the emergent vegetation in the agricultural region. The deviation from the 1:1 line is large (-1), and as bin 1 has a large population size it consequently has a higher weightage, which amplifies the error statistic. Further, the WRMSE increases when adding optimized texture, as a large number of pixels shift from the uncertain bins to the certainly non flooded and flooded bins, increasing the weight associated with these errors. On closer examination, the reliability diagram reveals a notable shift from all the intermediate bins, containing the uncertain flood values towards the certain bins (1 and 10). This highlights that the optimization approach caused a shift in what backscatter value combinations are classified

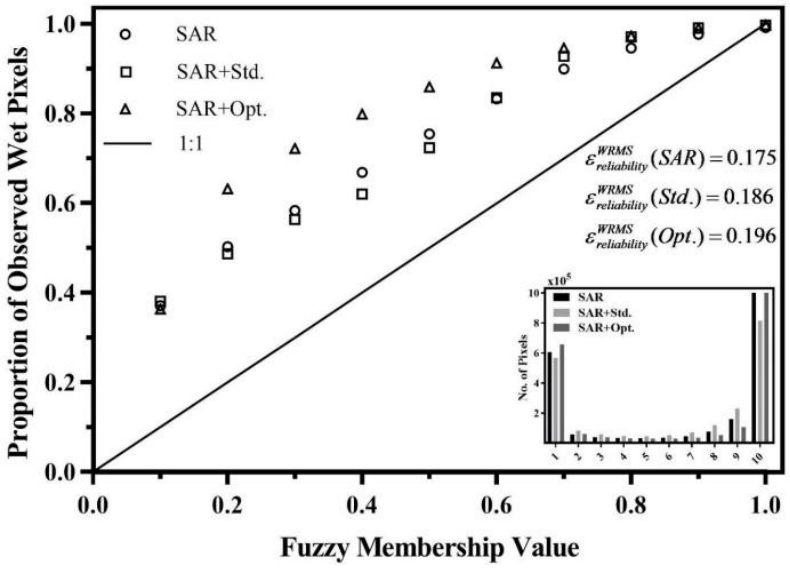

Fig. 4.20 As for Fig. 4.16 but for Ulmarra.

as uncertain, pushing mixed pixels or those with emergent vegetation towards the flood class. This observation is identical to the findings at Junction Hill, where the optimized texture addition resulted in clear error reductions.

Giving due consideration to the visibly flooded agricultural region (top row, Fig. 4.197), which is only partially flooded in the SAR-based maps, allows better understanding of the reduction in uncertainty through texture optimization. In the SAR-based maps the region consists of pixels mostly from bin 1 and 2, i.e. closer to the certainly non-flooded end on the fuzzy membership spectrum. However, as optimized texture enhances feature extraction and reduces noise effects, these patches were pushed towards the certainly non-flooded bins. This is consistent with the notable increase in the pixel count of bin 1 for Opt+SAR when compared to other inputs. Many of these pixels from bin 1 and 2, however, belong to the aforementioned error hotspots and were flooded in the validation data. Therefore, the proportion of wet pixels was increased in the validation data for the almost certainly non-flooded classes, especially in bin 2. As more pixels wrongly became "surer" of their non-flooded status, i.e. moving from uncertain bins towards bin 1 and 2, the error margin increased thus amplifying the overall classification error. It is clear from the above analysis that the use of reliability diagrams alone is insufficient for the assessment of fuzzy maps. The need to reconvert the continuous, fuzzy, SAR-based maps into discrete categorical maps with specified bin sizes as well as the need for a binary validation dataset for assessment, limits the applicability of reliability diagrams for an objective evaluation of probabilistic flood maps.

When comparing continuous spatial fields as in the fuzzy flood maps generated here, tolerance for locational or categorical errors is desirable (Pappenberger et al., 2007). Locational tolerance accounts for slight pixel shifts without denoting them in complete disagreement, while categorical tolerance allows to identify the higher similarity between "slightly wrong" values which are common in hydrological spatial fields (Wealands et al., 2005). As the fuzzy spatial similarity statistic accounts for both category and location fuzziness, the improvement brought about by Opt+SAR could be evaluated without the impact of spatial mismatches caused by the notably different nature of the uncertainties in the SAR and the validation data. In this case, addition of the fuzzy statistic to the accuracy assessment approach, proved to be rather informative.

The fuzzy map comparison added value to the reliability diagram analysis, with the local spatial improvement demonstrated through the maps and captured by the overall similarity

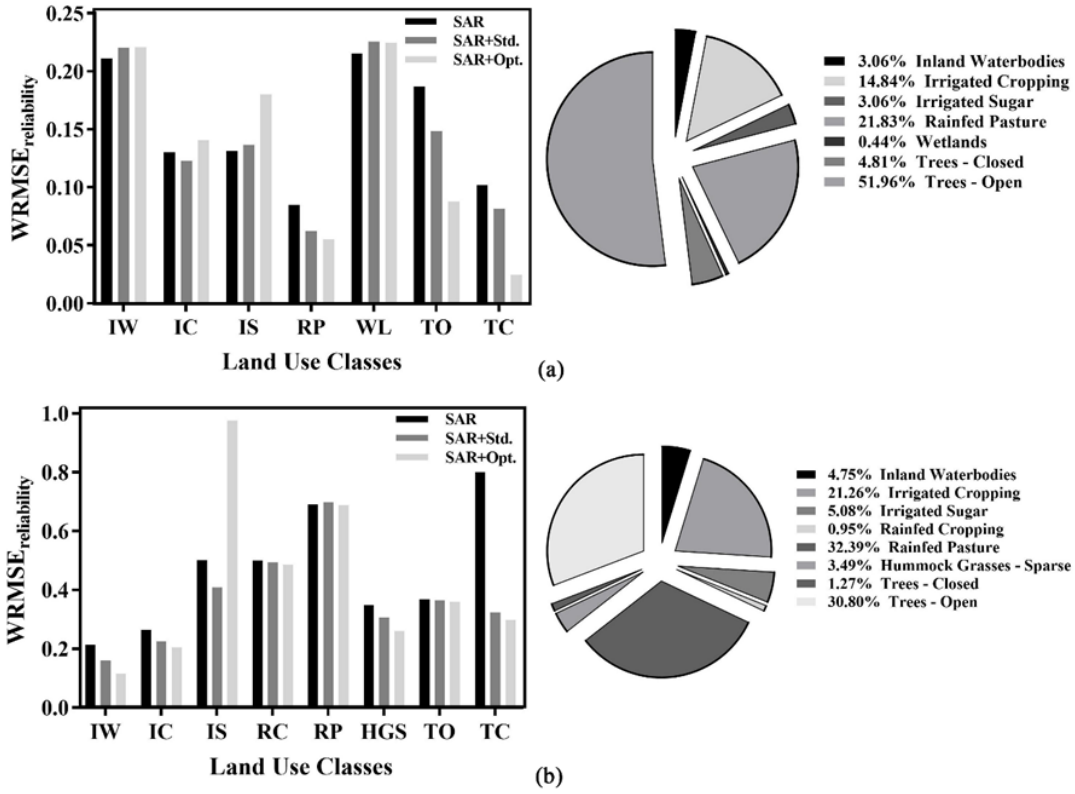

Fig. 4.21 Distribution of land cover classes and the corresponding Weighted Root Mean Squared Error values obtained for each of the SAR-based flood mapping techniques at (a) Junction Hill and (b) Ulmarra.

statistic (Hagen-Zanker et al., 2005). Even though the values of spatial agreement at Ulmarra were lower than those achieved at Junction Hill, which was expected due to the error hotspots critically discussed previously, a clear improvement was evident. All values were greater than 0.6 which has been considered a satisfactory measure of fit, given the expected uncertainties in shoreline locations (Pappenberger et al., 2007). By accounting for locational and attribute uncertainties in the computation of the local matching, chance agreement and image registration problems were also accounted for (Power et al., 2001). Furthermore, fuzzy map comparison provides a unique opportunity to assess the spatial characteristics of the classification errors, which can help to better diagnose their underlying cause. Correctly identifying the reasons and nature of uncertainties represents the first step in designing appropriate post-processing strategies, or even formulating better SAR-based flood mapping techniques for the future (Power et al., 2001).

4.4.4 Land Cover Based Performance Analysis

An investigative analysis of the classifier performance within different land cover classes was undertaken, to interpret the classifier response when exposed to pixels with varied surface characteristics. This can help to decide whether or not to use the proposed method based on the dominant land-use class in the region. In order to quantify the above, land cover maps of 250 m resolution were subset for the validation sites. The land cover maps were used to extract the underlying classes of the binned flood maps, which were subsequently overlaid on the binary validation data, to calculate the observed proportion of wet cells in each bin. Further, reliability diagrams were constructed individually for each land cover class and WRMSE values plotted for all the different approaches. Notably, the spatial average values of WRMSE in Fig. 4.16 and Fig. 4.20 differ from the WRMSE values obtained in Fig. 4.21, due to the calculation of case individual values for each land-use class. Based on the pixel count in each class, the values were found to be significantly different from the global spatial average. Although, if a weighted average based on the percentage distribution of the land-use classes was considered as shown in the pie diagrams of Fig. 4.21, the statistics were in agreement.

4.4.4.1 Validation Site 1: Junction Hill

The Junction Hill area is dominated by the "Trees - Open (TO)" land cover class, followed by the "Rain-fed Pasture (RP)" and "Irrigated Cropping (IC)" classes. According to Fig. 4.21 (a), WRMSE values had a significant reduction in the RP and TO classes, which might be the reason for the algorithm performing well in this region. The image region displaying a visible reduction in uncertainty through texture optimization, observable in the fuzzy flood maps illustrated in Fig. 4.12, is predominantly covered with the TO land cover class.

In the IC class however, classification performance was relatively poor due to the altered backscatter characteristics caused by emergent vegetation. The stems cause double bounce scattering of microwaves at the irrigated water surface making flood identification from SAR images in these particular land covers challenging. While the other classes may not contribute much to the overall RMSE values they provide crucial insight into classifier behaviour. The "Irrigated Sugar (IS)" shows a sharp increase in WRMSE with the proposed approach for the same reason as IC.

The water dominated classes like "Inland Water bodies (IW)" and "Wetlands (WL)" also showed a slight degradation in performance on texture addition. The contribution of the water based classes is considered limited in this context due to the land cover composition and

the small magnitude of the increase. However, this suggests that the slight increase in WRMSE for the Ulmarra region on texture addition may have been caused by the majority of pixels being flooded with 77.6 % wet cells. As the Junction Hill region has only 36.7 % coverage of flooded area in the aerial photo, the improvement is more noticeable in the mixed land cover classes. This implies that if permanent water can be masked prior to classification, the results may be further improved.

4.4.4.2 Validation Site 2: Ulmarra

Major classes in this region are TO, RP, and IC as evident from Fig. 4.21 (b), none of which seemed to benefit especially from the random addition of texture. IC seems to show a slight improvement over the poor performance at site 1, but as the crop species are unspecified, the difference in RMSE values can be assumed to be a function of plant morphological properties. TO and RP showed almost no change in WRMSE values for all three approaches. As most of the pixels at this site are inundated, the approach may not be able to demonstrate substantial improvement due to scattering of X-band from emergent vegetation. Conversely, the proposed algorithm was more effective in these land cover classes at Junction Hill, as the false alarm rate was reduced. Other classes in the region include “Hummock Grasses - Sparse (HGS)”, “Rainfed Cropping (RC)”, TC, and IS, of which all except IS exhibited a slight reduction in RMSE for the proposed approach, with TC showing the maximum decrease. IS showed an improvement in classification accuracy with texture addition but a degradation after texture optimization. However, the percentage coverage of the minority classes is insufficient to cause a quantifiable impact on the overall RMSE.

4.5 Chapter Summary

A novel single image SAR-based probabilistic flood mapping technique was developed and tested in this chapter. The literature review revealed that even though SAR images were rich in texture, the utilization of this information in flood mapping algorithms was fairly primitive, with appropriate texture feature selection, window size estimation, and directionality presenting the major challenges. Consequently, a framework was developed here to optimise textural information by choosing the window size based on spatial autocorrelation, removing directionality through an omnidirectional average, and condensing the information content into

the minimum possible number of bands. An adaptive neuro-fuzzy classification algorithm was then proposed to classify these enhanced SAR images into probabilistic flood maps. The backscatter distributions of flood and non-flood pixels in the image were modelled as Gaussian curves within a fuzzy inference system, parameterized by a neural network using training data selected from the input image. The fuzzy flood maps generated thus, were evaluated against fuzzy validation maps derived from aerial photographs, using a fuzzy map comparison strategy new to flood mapping literature and reliability diagrams. Overall and land-use based evaluations for two validation sites, show that the proposed mapping technique has the potential to improve operational flood mapping from single SAR images, when ancillary data are unavailable.

CHAPTER FIVE

"It is far better to foresee even without certainty than not to foresee at all."

- Henri Poincare, *The Foundations of Science: Science and Hypothesis, the Value of Science, Science and Method,* translated by Mélanie Frappier, Andrea Smith, and David J. Stump

5. Flood Inundation Modelling

5.1 Introduction

For any given flood forecasting problem, a model to describe the spatiotemporal evolution of the inundation is absolutely vital. Moreover, improving state estimation through assimilation of discrete, distributed observations of the system state, also necessarily requires a dynamic process model to numerically simulate the state of the system continuously across the space-time domain. While it is widely acknowledged that hydraulic models used to predict inundation suffer from a vast number of uncertainties (Beven et al. 2018), it is also common knowledge that a reliable flood forecasting system cannot be envisioned without a flood inundation model (Revilla-Romero et al. 2015; Ward et al. 2015; Alfieri et al. 2018; Hostache et al. 2018). It is therefore imperative to choose a suitable flood inundation model which can generate simulations fast enough to be useful in near real time and simultaneously allow updating through remote sensing data assimilation. It is also deemed necessary to achieve preliminary calibration of the model such that, the inundation forecasts generated by the open loop – ensemble of model runs with no data assimilation – are at least physically realistic.

Accordingly, this chapter first describes the principles of hydrodynamic modelling, the various uncertainties which affect it, and the guidelines followed for model selection. Thereafter, the model chosen for use in this thesis, LISFLOOD-FP is presented in detail. The model description section covers the governing equations, model assumptions, and limitations of the solver used in this thesis. This is followed by a detailed explanation of the particular flood model implementation used throughout this thesis and the chapter concludes with a discussion on model parameterization.

5.2 Principles of Hydrodynamic Modelling

Floods travel along river reaches as waves, with continuously varying velocity and depth through time and distance. Flood flow is characterized as gradually varied and unsteady, as flow properties gradually vary with time. Hydraulic models compute water level and velocity in the channel network and when the channel storage capacity is exceeded in the floodplain. Fully hydrodynamic models solve the complete Saint Venant or Shallow Water Equations, which are based on the conservation of mass and momentum (Horritt and Bates 2002). Complete de Saint Venant equations of continuity and momentum, are given below as Equations 5-1 and 5-2, respectively.

$$\frac{\partial \mathrm{Q}}{\partial \mathrm{x}} + \frac{\partial \mathrm{A}}{\partial \mathrm{t}} + \mathrm{q} = 0, \tag{5-1}$$

$$\left(\frac{1}{A}\right)\frac{\partial\left(Q^2/A\right)}{\partial t} + \left(\frac{1}{A}\right)\frac{\partial Q}{\partial t} + g\frac{\partial y}{\partial x} - g\left(S_0 - S_f\right) = 0, \tag{5-2}$$

where Q = discharge (m^3/sec); A = area (m^2); q = lateral flow per unit length (m^3/sec/m); x = along channel distance (m); y = depth of flow (m); g = acceleration due to gravity (m/sec^2); S_0 = bed slope; and S_f = friction slope.

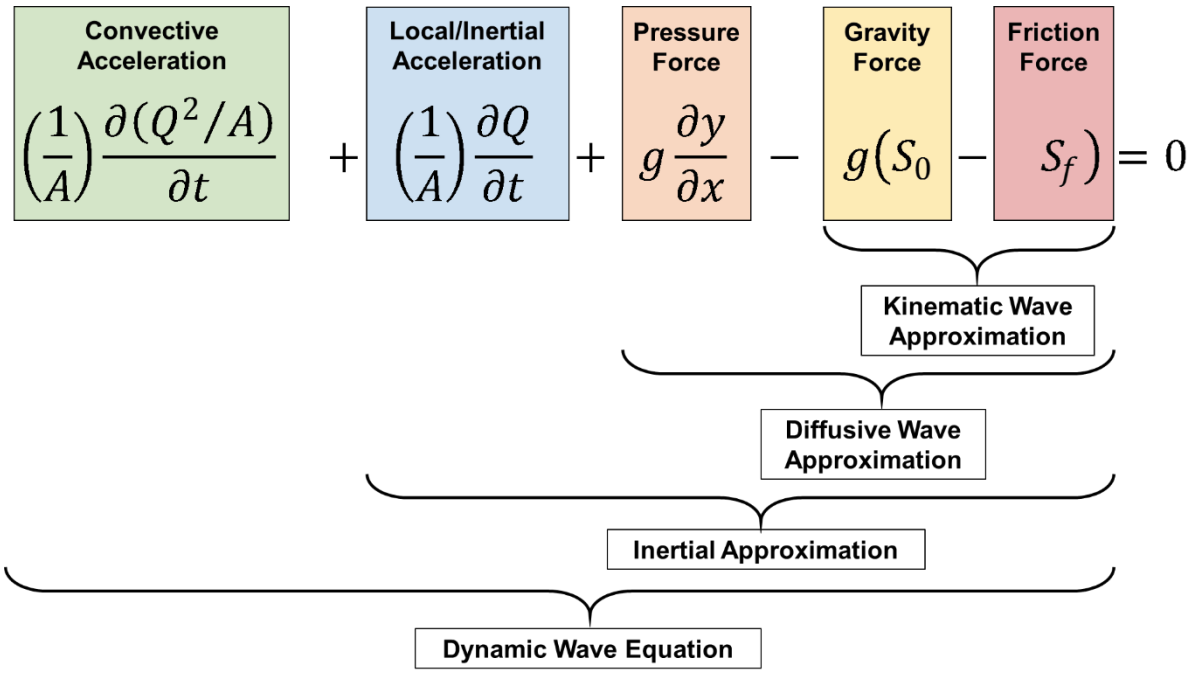

Fig. 5.1 Momentum equation of the Saint Venant equations, with definitions of all terms and various approximations used in literature.

It is not possible to solve the Saint Venant Equations analytically for complex real world applications (Mujumdar 2001). A wide variety of numerical models have been proposed in literature, each solving different approximations of the shallow water equations (SWE), thus accounting for varying degrees of physical complexity (Hunter et al. 2007). While all models usually solve the equation of continuity, most use an approximation of the momentum equation, to reduce the computational burden. Based on which terms are retained in the numerical solution, hydrodynamic models can be classified into kinematic wave, diffusive wave, inertial, and dynamic wave. Meanings of the various terms and the terms included in each approximation of Equation 5-2, are illustrated in Fig. 5.1.

Studies have shown that increasing physical complexity significantly increases the computational effort involved, while not necessarily translating into higher prediction accuracy (Neal et al. 2012). It is, therefore, vital to choose appropriate modelling tools for a particular flood simulation problem. For example, in steep slope channels with negligible backwater effects, gravity and friction forces may balance each other, resulting in uniform flow. Therefore, flood flows in such channels can be adequately represented using a kinematic wave approximation. Conversely, in mildly sloping channels with primarily downstream controls, when both inertial and pressure forces are important, the use of dynamic wave models is recommended. Hydrodynamic models are further classified on the basis of the spatial discretisation scheme used.

One dimensional models (1D) assume that flood flow only varies along the flow direction and lateral flows are negligible (Haile 2005). The model domain is discretised as a series of transverse cross-sections, perpendicular to the direction of flow. Governing equations are solved at each cross-section to derive water depth and velocity, assuming horizontal water levels (Schumann 2008). Channel geometry is assumed to be constant or vary linearly between two subsequent cross sections (Dasgupta 2015). The derivation of floodplain inundation and flood depth, requires further post-processing with a DEM. While such an approach is computationally efficient, it fails to depict the floodplain dynamics as the representation is not in the form of a surface but isolated cross-sections (Horritt and Bates 2002; Hunter et al. 2007). 1D models lack the potential to accurately represent the complex physical characteristics of the flood, leading to the proliferation of 2D, 1D2D and 3D models (Haile and Rientjes 2005; Tarekegn et al. 2010).

Two dimensional models (2D) consider transient flow characteristics in both the longitudinal and transverse directions. The Saint Venant Equations are solved in 2D, for estimates of water depth and velocity at each spatial unit with turbulence closure (Bates and De Roo 2000). Accurately depicting floodplain heterogeneity, especially in terms of roughness coefficient parameterization, is more challenging in this case (Haile and Rientjes 2007). The quality and reliability of 2D models, primarily depend on the accuracy of the topographic data. This includes channel bathymetry and the DEM used to represent floodplain elevations (Li and Wong 2010). 2D model domains can be discretised in the form of structured grids (finite difference method) or unstructured grids (finite volume and finite element methods) usually in the form of triangles or quadrilaterals (Bates and De Roo 2000). The computational resources required by 2D models, are highly sensitive to the choice of the grid size used (Carrivick 2006). Larger grid sizes significantly speed up the modelling, but microtopographic floodplain features are lost due to the averaging effect (Werner et al. 2005). Conversely, fine resolution grids allow for more detailed model outputs, however, computational times required scale almost exponentially (Schumann et al. 2007).

Coupled models which represent the channel in 1D while representing the floodplains in 2D have also been used to combine the advantages of both 1D and 2D model implementations (Haile and Rientjes 2005). In these hybrid modelling approaches; dominant main channel flows are modelled using a 1D element based on cross-sections covering the main channel only. Over-bank and floodplain flows are modelled in 2D typically using the grid-based discretization structure of DEMs, but in theory any spatial discretization is possible (Werner 2010). The coupling of the two domains at the end points of the channel section(s) to the adjoining floodplain cells can be achieved through a variety of methods depending on the model code (Moore 2011). For example, in the 1D/2D implementation of LISFLOOD-FP, water is routed onto adjacent floodplain cells once channel water depth reaches bankfull height, with only mass being conserved and not the momentum (Bates et al. 2013).

5.3 Uncertainties in Hydrodynamic Modelling

Hydrodynamic modelling is ridden with a variety of inherent uncertainties, which often cannot be completely removed through model calibration (Neal et al. 2009; Di Baldassarre et al. 2009; Mason et al. 2009; Stephens et al. 2012). The disparity between model calculated state

outputs e.g. water level and velocities computed at each computational grid cell (V_{sim}) and observed real world state data (V_{obs}) arises due to the following errors (Haile and Rientjes 2007).

- Random or systematic errors in the forcing data (ε_i), e.g. precipitation data
- Random or systematic errors in the recorded state data (ε_r), e.g. water levels
- Errors due to non-optimal parameter values (ε_{no})
- Errors due to incomplete or biased model structure (ε_s)
- Errors due to the time space model domain discretization (ε_d)
- Errors due to rounding off (ε_{ro})

Hence the total simulation error is given by the following relation:

$$V_{obs} - V_{sim} = \varepsilon_t = \varepsilon_i + \varepsilon_r + \varepsilon_{no} + \varepsilon_s + \varepsilon_d + \varepsilon_{ro}, \quad (5\text{-}3)$$

Note that these errors could have either positive or negative signs, which may compensate for each other when computing the total sum. In hydraulic modelling, channel geometry and roughness often compensate for each other, leading to different equally plausible parameter sets capable of mapping model predictions to the observed data (Neal et al. 2015; Wood et al. 2016; Grimaldi et al. 2018). This leads to the problem of state-parameter equifinality which has received significant attention in literature (Beven 2006). Accordingly, several statistical approaches have been developed to address these uncertainties, such as the Generalized Likelihood Uncertainty Estimation (GLUE) technique (Beven and Binley 1992). GLUE methods use Monte Carlo simulations to analyse a large number of model scenarios (Stedinger et al. 2008). Each competing model is assigned a prior probability according to the parameters used for the simulation. Finally, a likelihood measure is used to compare modelled and observed variables, such that the capabilities of each model to accurately represent observed conditions can be individually evaluated. Aronica et al. (2002), for example, extended the GLUE technique to estimate spatially distributed uncertainty in roughness values in hydraulic models through binary flood inundation maps. Similarly, Pappenberger et al. (2005) showed that cumulative distribution functions can be used to identify a range of equally well performing parameter sets within the GLUE framework. This also allowed the authors to produce event specific flood probability maps, which objectively represented parametric uncertainties and were more informative for stakeholders (Pappenberger et al. 2005). Similarly, Schumann et al.

(2008) extended the GLUE approach to quantify the uncertainty associated with water stages derived from a single SAR image.

In the context of integrating hydraulic models into operational flood forecasting chains, however, the consideration of uncertainties changes significantly. As Grimaldi et al. (2019) show, the errors could either increase in magnitude through the entire processing chain or compensate for each other. Flood forecasts, for example, require precipitation forecasts generated by numerical weather prediction models, which are then provided as inputs to a hydrological model. The streamflow computed by the hydrological model is then used as an input to the hydraulic model, to give distributed estimates of inundation extent and depth in the floodplains. This implies that hydraulic models running in forecast mode, often have to deal with highly uncertain inflow data (Matgen et al. 2010). Moreover, the lack of high-resolution and high accuracy topographic data, is also a major contributor of uncertainties.

Improved understanding of the uncertainty propagation through model cascades is essential to identify critical regions requiring further investigation, invest in data collection, improve prediction accuracy, and effectively communicate model uncertainty to end users (Ward et al. 2015; Birch et al. 2016). As Schumann & Bates (2018) show, the inherent errors in globally available DEMs are often to the order of several metres, rendering these unacceptable for many modelling applications. Inflow and topographic uncertainties are several orders of magnitude larger than any other sources of uncertainty in hydraulic modelling (Hostache et al. 2018). Therefore, in this study the impacts of these different errors on hydraulic model forecasts will be evaluated and the potential of the proposed assimilation technique to retrieve the true model state will be quantified through numerical experiments.

5.4 Model Selection

Selection of the appropriate modelling tool is primarily dependent on the purpose of the modelling (Beven 2012). Major factors that determine this decision include the required model outputs and inputs, model structure and process representation, and the availability of the model code (Koriche 2012). As the aim of this study was to improve the accuracy of operational flood forecasts through remote sensing data assimilation, the following features were considered necessary:

1. Ease of integration with remote sensing datasets.

2. Speed of simulation including parallelization of the model code.
3. Accessibility to the model code.

Based on the aforementioned considerations, the raster-based 2D-hydraulic model LISFLOOD-FP (Bates et al. 2010) was chosen for this study, as the executable was freely available from the University of Bristol for non-commercial applications. The primary objective of LISFLOOD-FP was to provide a simple raster-based model, with minimum process representation required to adequately simulate floodplain inundation, making it suitable for real-time application. A storage cell approach is used to speed-up the processing, whereby the floodplain is treated as a series of discrete cells. The flow between cells is calculated using an analytical flow equation such as that of Manning or Chezy, at a fraction of the computational cost required for numerical solutions of the St. Venant Equations (Horritt and Bates 2002; Coulthard et al. 2007; Hunter et al. 2008; Fewtrell et al. 2009; Gilles and Moore 2010; de Almeida et al. 2012; Neal et al. 2012; de Almeida and Bates 2013). Raster remote sensing data integration was also simplified by employing this gridded approach. This representation has proved to be sufficient for many flood studies even though inertia and supercritical effects were neglected (de Almeida and Bates 2013). LISFLOOD-FP is capable of rapidly providing distributed values of depths and flow velocities in 2-D, making it the ideal choice for operational flood forecasting and for this thesis.

LISFLOOD-FP includes a number of numerical solvers capable of simulating flood wave propagation within the channels and across the floodplains using various approximations of the Saint Venant equations (Bates et al. 2013). System characteristics, purpose of the modelling, and data availability should primarily determine the appropriateness of a given approximation (Schumann et al. 2013). As a novel data assimilation framework is developed in this thesis, only the impacts of inflow uncertainties were considered. Given the highly complex, non-linear relationships between floodplain inundation and depth, the impacts of assimilating flood extents needed to first be investigated neglecting all other sources of uncertainty (Hostache et al. 2018). For example, in low slope regions small changes in water depth can lead to large variations in flood extent, while the converse is true for steep confined valleys (Wood 2016). Moreover, even in data rich environments such as in developed nations, forecast inflow uncertainties always exist as a consequence of numerical model cascades necessary for forecasting. Since, observed bathymetry and a high resolution DEM were available, a full 2D specification of topography could be provided to the model, thereby limiting the uncertainty contribution from this source.

Accordingly, the full 2D implementation of the Lisflood-FP acceleration solver, which solves the inertial approximation to the SWEs to provide velocities and depths at each grid, was used for all the assimilation experiments detailed in this thesis. The following sections describe the model structure, implementation, and parameterization in greater detail.

5.5 LISFLOOD-FP: Model Equations and Implementation

Initial versions of LISFLOOD-FP used an explicit backward finite difference scheme to solve a 1D-kinematic wave representation of channel flow and a 2D-diffusive wave approximation of floodplain flow (Bates and De Roo 2000). The floodplain was discretized as a series of regular storage cells, for which the mass balance was updated at each time step based on water fluxes entering and leaving each cell. Fluxes between cells are calculated analytically using uniform flow formulae, reducing the associated computational cost per time step with respect to equivalent numerical solutions of the SWEs. All solvers assume subcritical, gradually varying flows only in the Cartesian coordinate directions and the momentum of the flows is not conserved at channel-floodplain connections. The following paragraphs describe the inertial acceleration solver in greater detail.

The acceleration solver uses the local inertial approximation of the 2D-momentum equation to simulate unsteady flow. Solutions are numerically computed through a backward finite difference method, explicit in time and first order in space (Bates et al. 2010), with a semi-implicit treatment for the friction term to aid stability (Bates et al. 2013). This approximation of the SWEs has proved to be numerically stable and has demonstrated accuracies comparable to analytical solutions and fully hydrodynamic models (Fewtrell et al. 2009, 2011), while increasing computational efficiency over diffusive models (de Almeida et al. 2012). The water flux (Q) passing through each cell face is individually estimated in the x and y directions; only the x-direction flow calculation equation has been included here for brevity, such that (Bates et al. 2010):

$$Q_{t+\Delta t} = \frac{q_t - g h_t \Delta t \frac{\delta(h_t+z)}{\delta x}}{\left(1+g\Delta t n^2 |q_t| / {h_t}^{7/3}\right)} \cdot \Delta x, \tag{5-4}$$

where Q is the flow ($m^3 s^{-1}$), q is the flow per unit width ($m^2 s^{-1}$), g is the acceleration due to gravity (9.81 ms^{-1}), h is the depth of flow (m), z is the river bed elevation (m), n is the Manning's roughness coefficient ($m^{-1/3}$ s), t is time (s), and Δx is the cell resolution (m). The

continuity equation is then applied to the full grid domain to update water depths inside each computational cell, thus achieving a 2-D solution. An adaptive time step is used in accordance with the Courant-Friedrich-Levy condition for small amplitude gravity waves, calculated using the following relationship (Bates et al. 2010).

$$\Delta t_{max} = \alpha \frac{\Delta x}{\sqrt{g h_t}}, \tag{5-5}$$

where a factor α is introduced to aid stability as the small wave amplitude assumption does not always hold for celerity calculations and as friction terms are included in the model. Values of α varying from 0.2–0.7 are mostly able to produce stable simulations for floodplain flow (Bates et al. 2010). Moreover, the stable time step given by Eqn. 5-5 is almost 1-3 times larger for typical cell sizes than a purely diffusive scheme and scales with $1/\Delta x$ instead of $(1/\Delta x)^2$, further increasing computational efficiency.

Hydraulic models require the specification of an inflow discharge hydrograph at the upstream boundary, initial conditions, detailed topographic and bathymetric data, surface roughness assessments, and outflow data at the downstream boundary. The discharge measurements available at the Lilydale gauging station were converted to flow per unit length by dividing by the grid size as required by LISFLOOD-FP and used as the upstream boundary (Neumann condition). Tidal water levels observed at Yamba were similarly used as the downstream boundary condition (Dirichlet condition), see Fig. 3.1 for the locations of Lilydale and Yamba, which form the boundaries of the study reach. This study used the inertial acceleration solver which was implemented in full-2D, with observed bathymetry, LiDAR topography, observed tidal levels, and calibrated channel roughness, only considering inflow errors. This implementation was chosen to develop and test the proposed flood extent assimilation framework, through both synthetic and real-world experiments, to prove the efficacy of the proposed approach independent of the multitude of uncertainties.

Initial conditions were computed by running a steady state simulation with the most commonly median flow value observed at Lilydale. The resulting simulated water depths were used as a start file for the model. Based on manual trial and error exercises, a spin-up period of six weeks was found to be enough to ensure numerical stability. The long spin-up period was required due to inaccuracies in the initial conditions, with the hydrographs showing large unrealistic oscillations for spin-ups shorter than one month. Distributed floodplain roughness values were assigned, according to land-use classes identified from aerial photographs, based

on recommendations made by Arcement & Schneider (1989). As the floodplain roughness is not a sensitive parameter for LISFLOOD-FP in most scenarios, this roughness map was kept constant through all the simulations (Wood et al. 2016). This choice allowed for a focused assessment of the impacts of channel roughness and inflows on the ensemble forecasts (Grimaldi et al. 2018).

5.6 Model Parameterization

5.6.1 Potential of Crowd-sourced Observations

Traditionally, hydraulic models have been calibrated with observations of channel flow and water depth, measured by hydrometric river gauges (Domeneghetti et al. 2014). For pluvial floods where the flooding could occur outside channels, however, gauges within the channel cannot provide useful information (Assumpção et al. 2018). Remote sensing forms part of the solution, however, some hurdles such as cost and frequency of acquisition have to be fully addressed to enable the routine use of RS data (Grimaldi et al. 2016). Therefore, as a complement to RS (or where RS data are not yet available), crowd sourced data can be utilized to supplement flood information. For example, for flash floods or fast moving floods in small catchments, the latency between satellite acquisition tasking and data delivery is often prohibitive, as the flood wave has already receded before an image can be captured. Consequently, novel sources of low-cost data are needed, which can be acquired frequently and in abundance. Citizen science (including citizen participation up to the scientist level) or crowd-sourcing (distributing a task among many agents), is an emerging concept in which citizens monitor the environment around them (See et al. 2016). In recent years, citizen science has provided distributed data on a variety of hydraulic variables, including water level (Kutija et al. 2014), flow velocity (Le Boursicaud et al. 2016; Le Coz et al. 2016), flood extent (Schnebele et al. 2014), topography (Shaad et al. 2016), and land-use land-cover (See et al. 2016).

The contributions and possibilities of citizen science for flood modelling were comprehensively reviewed in a recent paper by Assumpção et al. (2018). In their analysis of the current body of literature, they report the lack of appropriate techniques to utilize these data for model calibration and validation. The few studies which examine the impacts of including crowd-sourced water level data, either use qualitative approaches (Kutija et al. 2014; Yu et al. 2016) or focus on hydrological model validation with synthetic observations (Mazzoleni et al.

2015, 2018). Approaches to utilize crowd-sourced observations of water level for effective model parameterization still need to be developed. This study demonstrates the quantitative use of crowd-sourced flood observations to parameterize a hydraulic model for the first time. Here, crowd-sourced observations of floodplain water levels were used to identify the channel roughness parameter quantitatively. The primary objective of this experiment was to develop a simple framework to utilize water level observations which may be derived from crowd-sourced data for model calibration. Here calibration implies fine-tuning the model according to the available observations to arrive at best fit parameters (Assumpção et al. 2018). The parameter values identified using crowd-sourced data, were then compared with those derived from gauges, allowing verification of the parameter choice guided by crowd-sourced observations.

5.6.2 Framework for Flood Model Calibration using Crowdsourced Data

The overall methodology for this component of the research has been summarized in Fig. 5.2. LISFLOOD-FP Acceleration was set up in full 2D for the Clarence Catchment at 30

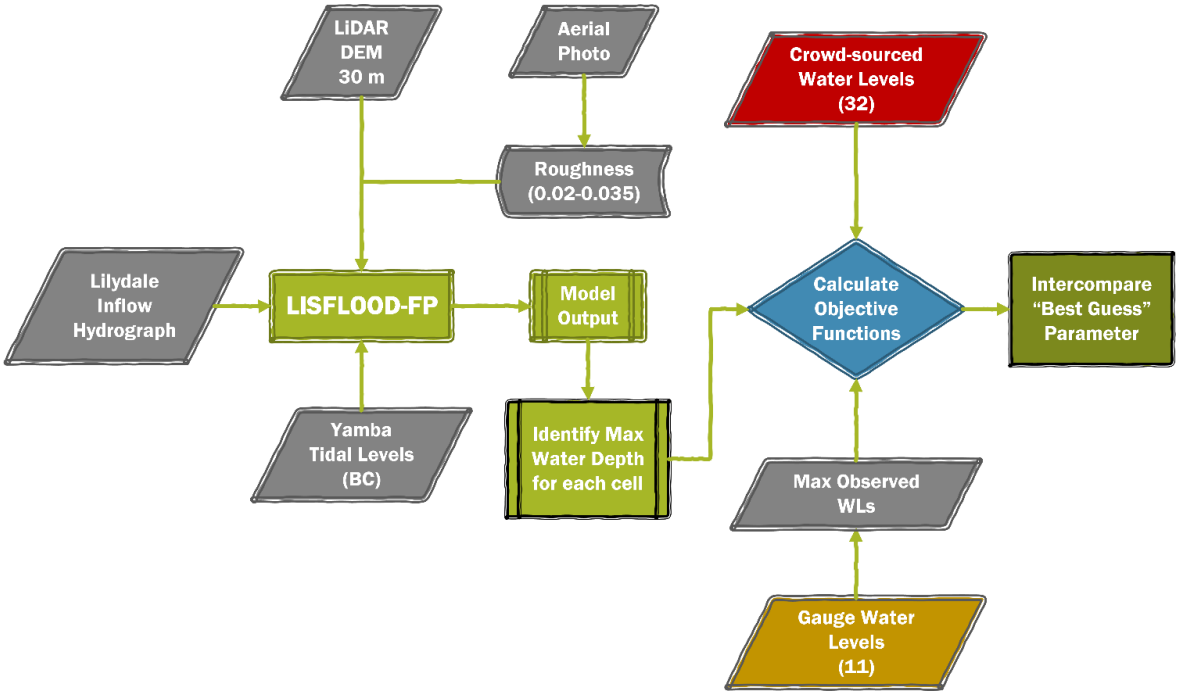

Fig. 5.2 Schematic of overall methodology used in this thesis for the parameterization of channel roughness in LISFLOOD-FP. The number of "crowd-sourced" and gauged water level locations have been included in the illustration, along with the range of roughness values considered for calibration.

m grid resolution, as Grimaldi et al. (2018) find it a cost-effective modelling solution for the Clarence Catchment. Implementation of this model requires a DEM, river geometry information, boundary conditions, and spatially distributed roughness coefficient values as previously discussed. Lateral tributaries and bridges were not included in the model setup, as they did not contribute significant water volumes during floods and the main objective of the thesis was towards operational forecasting of flood inundation (Rogencamp 2004). The LiDAR DEM available to this study was used to represent floodplain topography, while channel bathymetry was described by interpolating a continuous surface from field observations. The interested reader is referred to Grimaldi et al. (2018), where the generation of this bathymetric DEM is described in greater detail.

The 2013 flood event in the Clarence Catchment was used for this experiment. Observed inflows were introduced at Lilydale as hourly measurements of discharge, while tidal water levels were applied as the downstream boundary condition at Yamba. Using the tidal water levels as the downstream boundary, additionally allowed the evaluation of backwater effects on floodplain inundation for this catchment. Most hydraulic modelling studies choose to disaggregate spatially distributed coefficients of channel and floodplain roughness, into just one spatiotemporally invariant value for each (Werner et al. 2005). These are generally considered as effective parameters in hydraulic modelling, used to compensate for inadequate process and topographic representation (Horritt and Bates 2001; Jung et al. 2012). The floodplain roughness parameter is expected to be sensitive only during high velocity out-of-bank flows, as water shear will dominate resistance to flow once the floodplain is already inundated (Mason et al. 2003). As the events analysed in this thesis were not events of such magnitude, distributed time invariant values of floodplain roughness were assigned based on the land-use and kept constant for all runs.

Channel roughness is then, the only calibration parameter for this particular model implementation, which primarily controls the flood wave arrival time. Here, a lumped Manning's n value for the channel was optimized from 0.020 to 0.035 s/m1/3, which is the seasonal range of values for the Clarence River, by varying it in increments of 0.01 s/m1/3 (Grimaldi et al., 2018). Given that the objective of this thesis was towards improving flood forecasting through data assimilation of inundation extents, undertaking a distributed calibration exercise was deemed unnecessary. This is also the reason for not using a finer increment for the calibration. Moreover, the limits of the optimization were also fixed at the

bounds of physically reasonable values, because in an operational scenario, a fully calibrated model which has been "effectively" parameterized will mostly be unavailable (Andreadis and Schumann, 2014).

The observation set considered as "crowd-sourced" here, consisted of 32 distributed high water marks and wrack marks from the 2013 flood event ($m = 32$), whose timing of acquisition was unknown (See Chapter 3, Section 3.2.2). As these were observations of high water marks, it was reasonable to assume that these coincided with the peak flow. For each model grid cell where a corresponding crowd-sourced observation was available, the simulated maximum water depth (MWD) was first evaluated. Subsequently, two objective functions - Root Mean Squared Error ($RMSE_{MWD}$) and Mean Percentage Difference (MPD_{MWD}) - were calculated by comparing the model simulated maximum depth with each crowd-sourced/gauged value (i). RMSE was chosen to quantify absolute error in the simulation, while the MPD function allowed a relative assessment with respect to the observation values. The objective functions were computed as

$$RMSE_{MWD} = \sqrt{\frac{\sum_{i=1}^{m}(Sim_{MWD} - Obs_{MWD})^2}{m}}, \tag{5-6}$$

$$MPD_{MWD} = \frac{(Obs_{MWD} - Sim_{MWD})}{Obs_{MWD}} \times 100, \tag{5-7}$$

$$n_{opt} = \min_{i}(RMSE_i^{MWD} \times MPD_i^{MWD}), \tag{5-8}$$

The roughness value corresponding to the minima of the product of $RMSE_{MWD}$ and MPD_{MWD}, was selected as the best performing parameter n_{opt}. The product was considered as it is a simplified approach towards multi-objective optimization, as both the objective functions needed to be minimized. As the information content of the observations is distributed in space but limited in time, it was postulated that using more than one objective function with different priorities will allow for a more robust evaluation. Best fit parameters identified by using crowd-sourced and gauged water levels were inter-compared to assess the information content of the crowd-sourced data. The maximum water depth values given by the numerical model were finally compared with crowd-sourced and gauged water levels to arrive at the calibrated parameter value.

5.6.3 Results and Discussion

This section presents the results obtained from this novel calibration exercise based on crowd-sourced data and discusses the possible implications of this analysis. First, the model simulations of maximum water depth for different channel roughness values were compared with the crowd-sourced observations. Fig. 5.3 shows the distribution of RMSE and MPD values for the considered range of the channel roughness parameter. In this study, spatial variability in the roughness parameter was not considered, primarily as the objective was not to arrive at a fully calibrated model. It was assumed that in an operational scenario, complete and accurate knowledge of model parameters may not be available. Moreover, as previously discussed, hydraulic model uncertainties in the forecast mode are predominantly a function of topography and inflows (Andreadis and Schumann 2014). Consequently, the impacts of considered spatial heterogeneity in the roughness characterization, was not expected to yield significantly different results (Giustarini et al. 2011).

The maximum RMSE across all simulations was ~50cm and the maximum MPD ~40%, indicating that the model implementation was able to replicate ground conditions fairly closely.

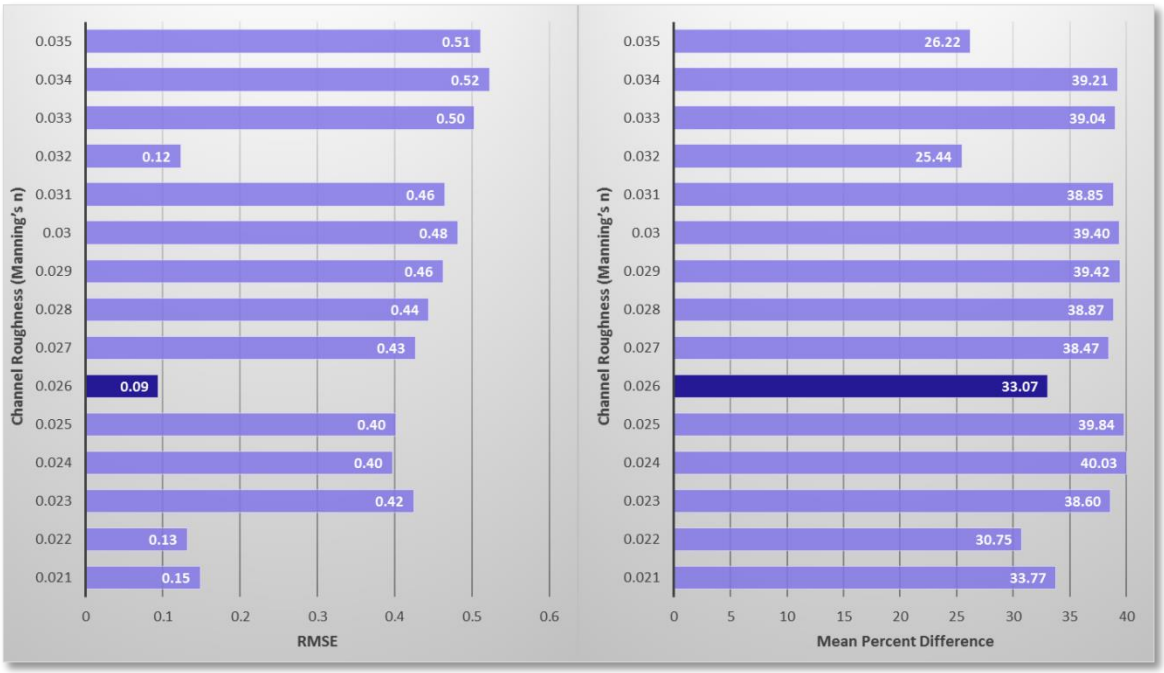

Fig. 5.3 Maximum water depths simulated by LISFLOOD-FP compared with crowd-sourced observations, with the plot on the left showing the root mean squared error values and the mean percent difference values on the right.

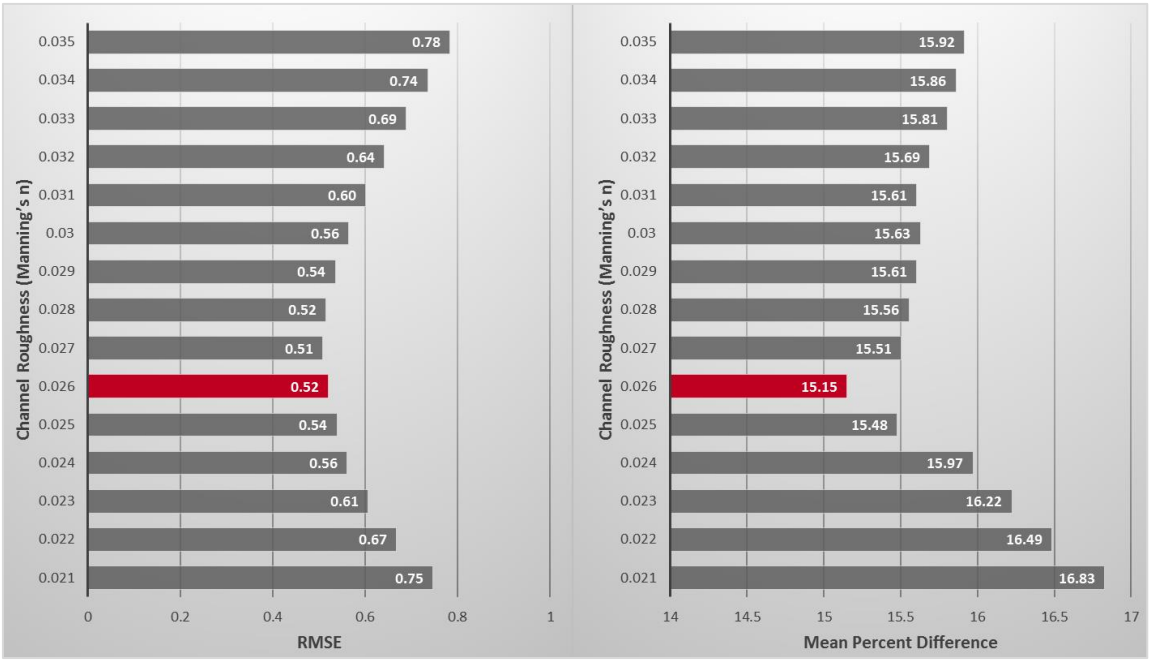

Fig. 5.4 As for Fig. 5.3, but for gauged maximum water levels.

Hostache et al. (2009) reported ±40cm RMSE through traditional calibration using a downstream limnigraph, where a LiDAR DEM with ±15cm and observed cross-sections with up to ±30cm uncertainty were used. They were further able to reduce the RMSE to ±23cm by including SAR-derived water levels in the calibration process. It is expected that the assimilation techniques developed in this thesis, will be able to achieve similar reductions in the RMSE values. The notable variation observed across the values of RMSE and MPD for the evaluated roughness range, implies high parameter sensitivity. There doesn't seem to have been any noticeable trend in the variation of the objective function values. Manning's n values of 0.026 and 0.032 seem to perform well across both objective functions, with RMSE values of 9 and 12cm, respectively, and MPD values of 33.07 and 25.44 percent, respectively. When the product is computed, $n = 0.026$ appears to be the better choice with a lower value for these error metrics, according to the crowd-sourced water level observations.

The second test involved a comparison between simulated and gauged maximum water depths, as shown in Fig. 5.4. In contrast to the previous comparison with crowd-sourced water levels, there is a clear trend in the objective function values in this experiment. The values of both error metrics first decreased with a corresponding increase in the magnitude of channel roughness, then increased after the optima. The maximum RMSE across all simulations was ~78cm and the maximum MPD ~17%, again indicating a suitable model setup. These findings

are aligned well with the expectations; as the water depth in the channel is greater, the corresponding RMSE is higher. Low values of MPD imply that the percentage error is actually lower than what was observed in the previous test against crowd-sourced water levels in the floodplain. In this experiment, Manning's n values between 0.025 and 0.028 seemed to perform well across both error metrics. On further examination of the product of the two objective functions, $n = 0.026$ appeared to be the clear choice here as well.

From this investigation, it was concluded that the best performing value for the channel roughness parameter was $n = 0.026$, which was chosen for further verification. Fig. 5.6 and Fig. 5.7 show plots of the simulated and observed water depths, for crowd-sourced and gauged

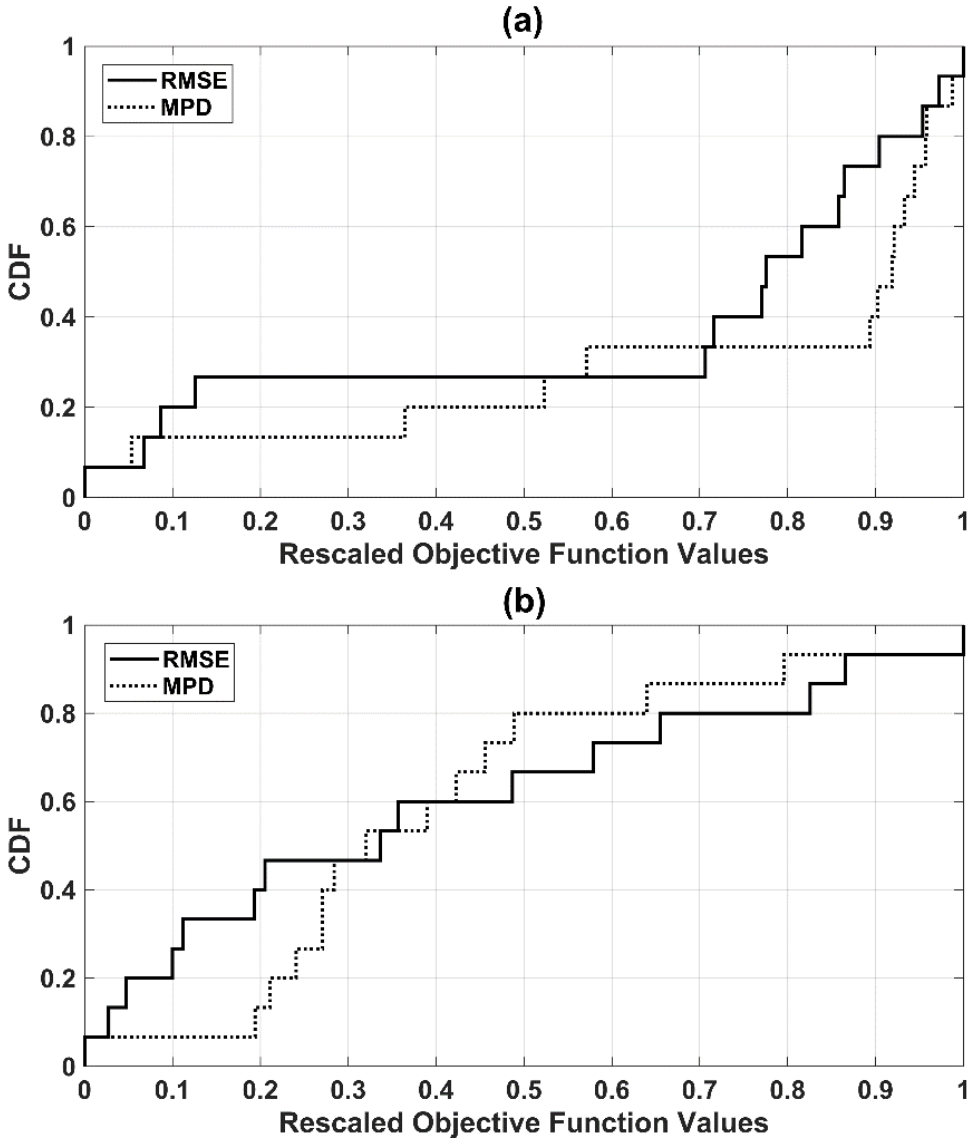

Fig. 5.5 Cumulative distribution functions showing the distribution characteristics of performance measures computed using (a) the crowd sourced water levels and (b) gauged water levels.

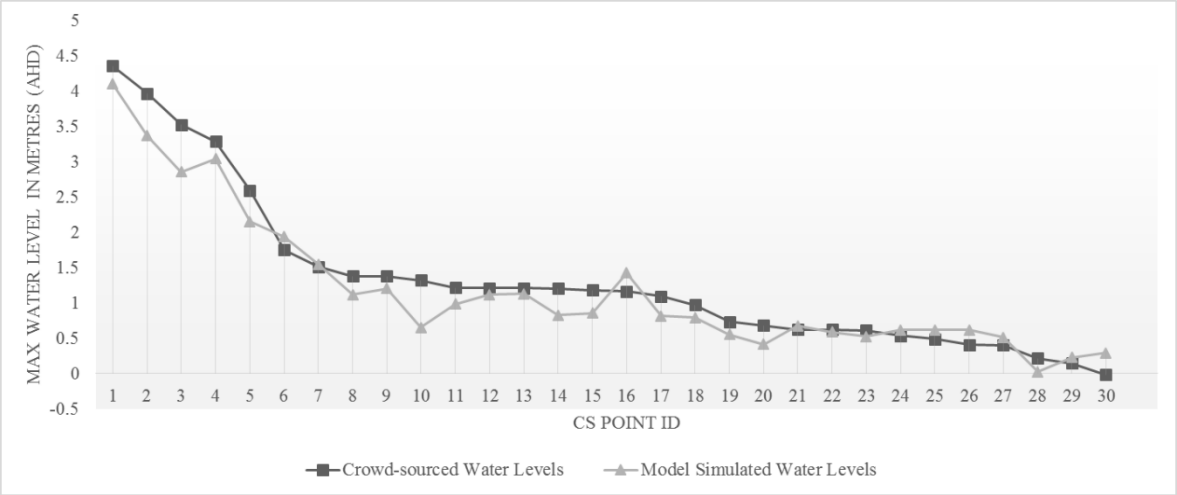

Fig. 5.6 Plot showing the maximum water levels simulated by the calibrated model using $n = 0.026$ and the crowd-sourced maximum water levels at all the available locations. Water level values have been sorted from the largest to the smallest according to the magnitude for illustration purposes.

data points, respectively. When examined in a distributed fashion there does not seem to be a clear trend of discrepancies between modelled and observed values from upstream to downstream (gauge locations are shown in Fig. 3.1), i.e. the model sometimes overestimates and sometimes underestimates the measurements. As expected, due to the relatively flat geomorphology of the region, higher values of water depth were observed within the channel associated with higher error magnitudes, conversely the error magnitudes were lower in the floodplain where elevation values are lower. The MPD is generally higher for the crowd-sourced points, as even low magnitude errors constitute a large percentage of the shallow observed water depth, while the opposite is true for the gauge-based assessment within the channel.

In the event of unbiased data, the relative and absolute error distributions should exhibit similar, if not identical trends (Schumann et al. 2008a). From Fig. 5.3 and Fig. 5.4, the similarity in the form of the error distributions is not immediately apparent, prohibiting an analysis of observation bias. Accordingly, cumulative distribution functions (CDF) were plotted in Fig. 5.5 separately for crowd-sourced and gauged maximum water depth observations for both objective functions, to evaluate the distribution of the errors. Values of the performance metrics were rescaled between 0 and 1, to ensure comparability of the plots and easy visibility on the axes. The CDF plots show nearly similar shapes, although some differences are evident in the higher error region for the crowd-sourced observations, while the lower errors exhibit some deviations

in the case of gauges. Similarities in the distribution shapes, reveal that the datasets under consideration are indeed unbiased, and can be reliably used for model performance evaluation.

These experiments illustrate that in the absence of gauge information, crowd-sourced water level observations might be able to provide sufficient information, to calibrate the channel roughness parameter in a hydraulic model. However, this might only be true for the present case study and in the floodplains, as in the presence of a levee system even an error of a few centimetres in water level predictions can cause false alarms/misses (Stefania Grimaldi, *personal communication*). In this context, there are still quite large discrepancies between gauged and modelled peak levels and perhaps the Nash-Sutcliffe efficiency (or other metrics) using the full hydrograph, would have allowed a more comprehensive evaluation of model accuracy. However, achieving a perfect calibration was not the objective of this thesis, as the parameters will be dynamically selected in real-time through data assimilation. The aim of this calibration exercise was to bring the magnitudes of errors are within acceptable ranges for assimilation, assuming a severely data limited scenario which may well be the case for most operational applications. Furthermore, the real world assimilation experiment presented in Chapter 9 should also reveal the quality of the calibration, as a full comparison with the hydrometric gauges available to this study is shown.

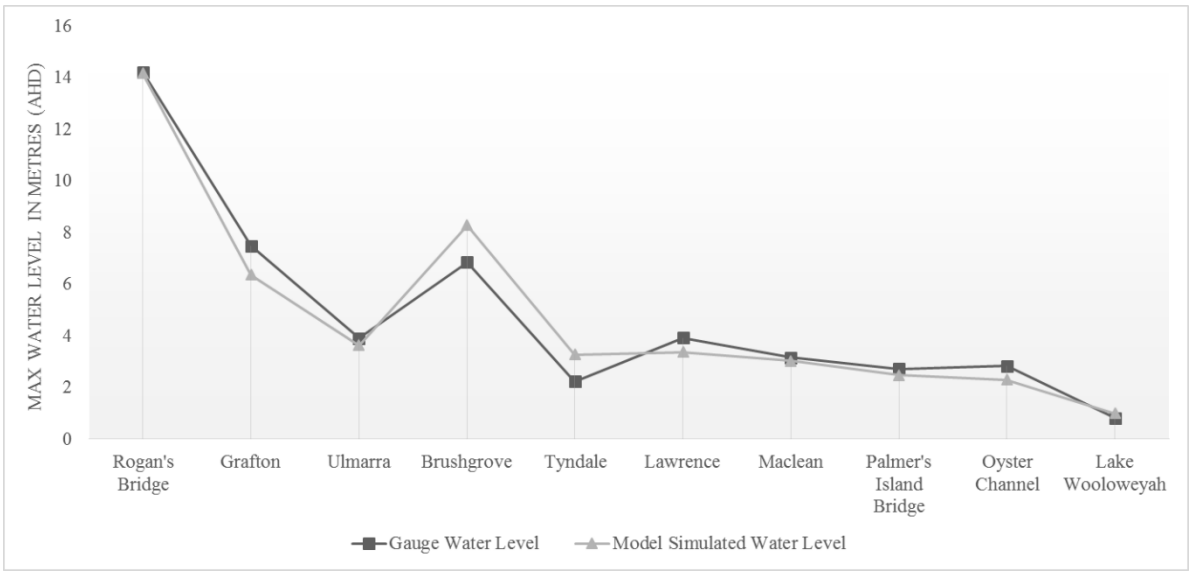

Fig. 5.7 Plot showing the maximum water levels simulated by the calibrated model using $n = 0.026$ and the gauged maximum water depths at all the available locations. Gauges are ordered from upstream to downstream. Note that Tyndale is not located along the main stem of the river and therefore has lower values than Lawrence.

Moreover, the number of crowd-sourced observations available to this study (32) was low as compared to the huge volumes of data expected from citizen science but highly accurate. As natural language processing and object extraction methods become more sophisticated, the processing of text and images/videos from social media for water level extraction are expected to be automated. If a large number of crowd-sourced water level observations with a time stamp were made available, the present methodology could be extended to accommodate those (Kutija et al. 2014), yielding further improvement in parameterization accuracy. The primary advantage of crowd-sourcing is that for the first time, calibration points can be in the floodplain where settlements usually exist rather than just in the channel, as it should not be assumed that a hydraulic model well calibrated in the channel will perform equally well in the floodplains (Pappenberger et al. 2007a). Crowd-sourced water levels therefore provide a unique opportunity to calibrate the model diagnostic variables in those areas where accurate estimates of flow and depth are required.

5.7 Model Evaluation

Mean and standard deviation values for the Gaussian parameter distributions (channel roughness, depth, and shape) used in this thesis, identified based on the aforementioned considerations, are summarized in Table 5.1. Interestingly, the calibrated channel roughness value seems quite physically meaningful, which is often not the case in model calibration studies. In fact, most studies have emphasized the need for effective friction parameters, to compensate for errors in the channel geometry and inflows, for example (Pappenberger et al. 2005). The first possible reason for this is, the use of a range of seasonal realistic values for the calibration. The second and perhaps more important one pertains to the representation of the

Table 5.1 Summary table of the parameter values used in this thesis.

Model Parameter	Mean (μ)	Standard Deviation (σ)
Channel roughness (n)	0.026	0.005
Depth multiplier (r)	1.05	0.005
Depth exponent (p)	0.36	0.005

bathymetry and topography (Grimaldi et al. 2018). It appears as though the hydraulic flow regimes are primarily controlled by channel geometry in this catchment, and therefore, as the bathymetric dataset used is highly accurate, a physically meaningful value could be identified as the best fit. Moreover, in the full 2D implementation used in this study, the only possible source of uncertainty was the inflow discharge, as the DEM, channel geometry, and the downstream gauged water levels were fairly accurate.

Parameter values chosen through the procedures outlined in the previous section, had to be verified using an independent dataset to ensure reliability of the estimates. This evaluation was performed against an optical image acquired in the falling limb of the 2013 flood hydrograph. This image was converted to Normalized Differential Water Index (NDWI) (McFeeters 1996) values to delineate the flood waters. The true colour composite of the SPOT image is juxtaposed against the derived NDWI image in Fig. 5.8. Problems of flood monitoring using optical data are also apparent from the Figure, as nearly 25% of the image is unusable due to cloud cover. Although the initial formula for the calculation of NDWI was developed for applications to the Landsat Multi Spectral Scanner (MSS) sensors, it has since been

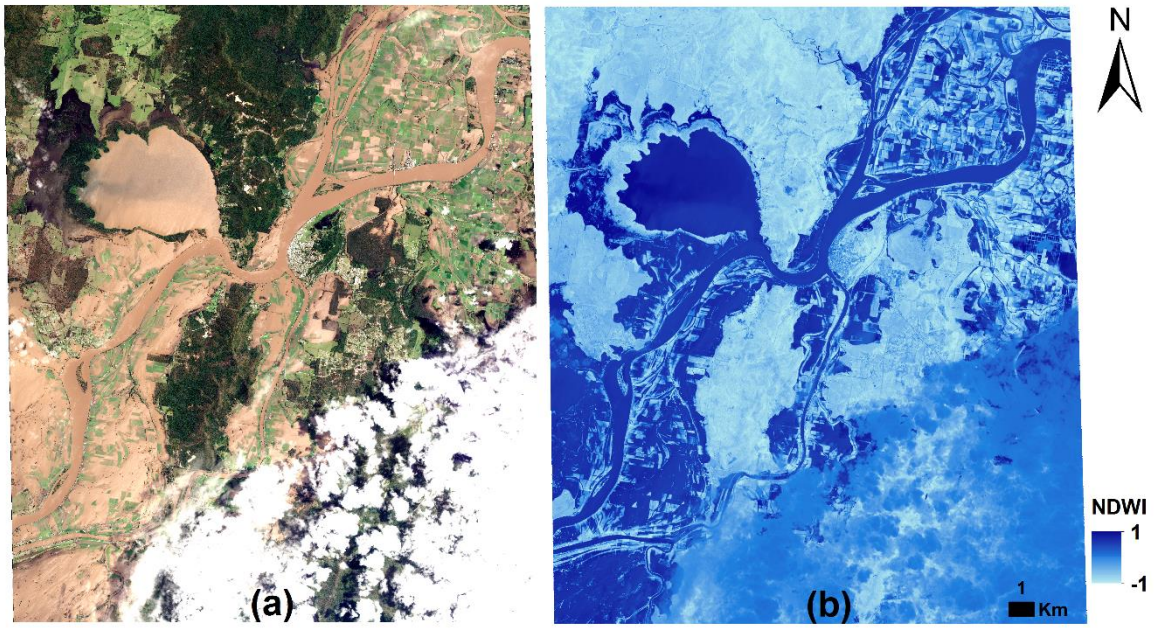

Fig. 5.8 Optical multispectral imagery from the SPOT-6 satellite, with (a) showing a true colour composite of the area, and (b) showing the Normalized Differential Water Index values derived from (a).

extended to almost all optical satellites (McFeeters 2013). The general equation used for the calculation of NDWI in this study is given as Equation 5-9.

$$NDWI = \frac{Green-NIR}{Green+NIR}, \quad (5\text{-}9)$$

where NIR refers to the Near-Infrared channel.

NDWI uses features of the water reflectance spectrum, i.e. maximum reflectance in the green region of the electromagnetic spectrum and minimum in the NIR region, to enhance the identifiability of water surfaces. It also exploits the high reflectance of terrestrial vegetation and soil in the NIR region, to aid the delineation of water bodies (McFeeters 1996). Using a band ratio approach for surface water detection, is of course, not devoid of uncertainties (Mukherjee and Samuel 2016); however, as the objective here was just to achieve an acceptable model set

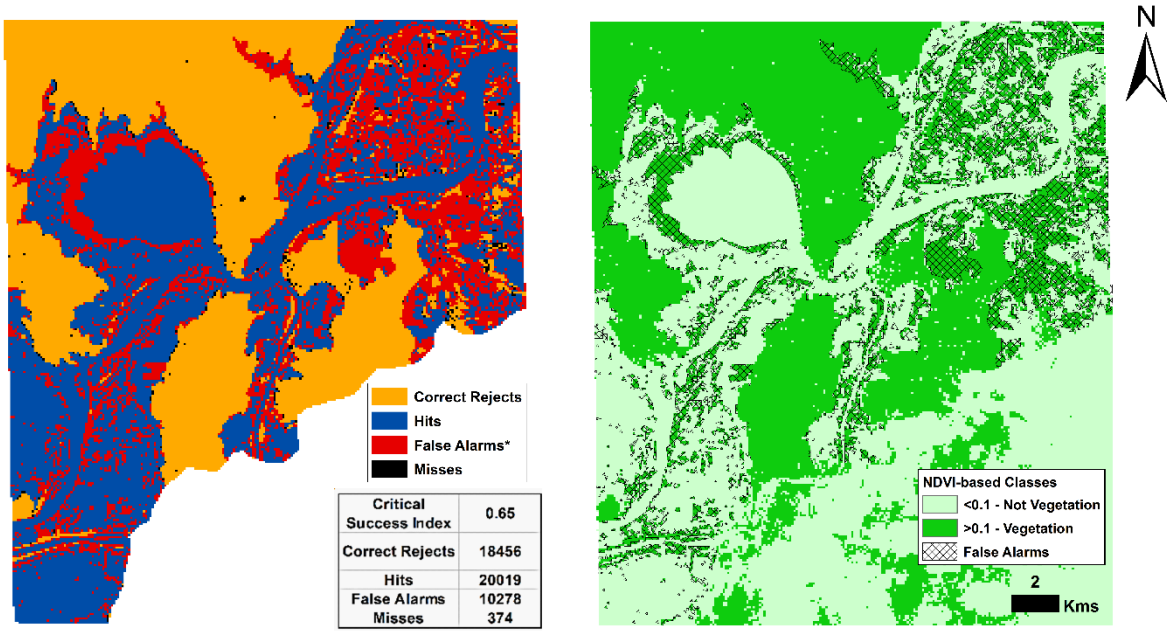

Fig. 5.9 Left panel shows the contingency map and statistics comparing the surface water extent map based on NDWI values derived from the SPOT-6 optical image against the inundation extents simulated by the LISFLOOD-FP acceleration solver in full 2D using the calibrated channel roughness parameter. False Alarms* indicates a lack of confidence in the inundation identified through the SPOT-6 image due to dense vegetation. Right panel shows the NDVI map showing area covered by vegetation and not vegetated regions, with respect to the extent of the False Alarms obtained.

up, this was considered sufficient to verify the parameter choices (Andreadis and Schumann 2014).

NDWI values greater than 0 are expected to represent water pixels, while negative values represent non-water land-use classes (Jain et al. 2005; Lu et al. 2011). Accordingly, the cloud-free portion of the SPOT image was processed to derive NDWI values, which was subsequently converted into a surface water map using a global threshold of 0 to retain positive values exclusively. NDWI values were derived from the SPOT image at the native resolution (1.5 m) of the pan-sharpened product (see Section 3.3.2), although these had to be upscaled to the model grid size of 90 m prior to making any comparisons. Model simulated water depths were extracted at the time of acquisition of the SPOT image and converted to inundation extent maps using a threshold of 1 cm. This depth threshold is used to derive flood extents throughout this thesis. Although some studies have justified the use of a 10 cm depth threshold due to reasons of uncertainty (Pappenberger et al. 2007b), it also means that for a pixel with 9 cm of water will not be considered inundated. This implies that 729 cubic metres of model simulated water volume per pixel, is ignored during the flood extent assimilation process. Consequently, a threshold of 1 cm was considered more suitable in this study (Hostache et al. 2018).

The inset table in the left panel of Fig. 5.9 shows a summary of the pixel statistics, also identical across both solvers. Number of correctly identified inundated pixels, is significantly larger than the misses and false alarms. The ratio of false alarms is quite high as flooding under vegetation cannot be identified in optical imagery, which observes only tree canopies. In order to corroborate this, the Normalized Differential Vegetation Index (Wang et al. 2011) was calculated to facilitate a qualitative comparison. The right panel of Fig. 5.9 shows the area identified as “False Alarms” drawn on a base layer of the SPOT-6 NDVI-based vegetation classes. As expected, most of the false alarms are perhaps flooded vegetation pixels which are not classified as water due to limitations of NDWI-based surface water extraction from optical images.

In spite of the limitations outlined earlier, a Critical Success Index (CSI) value of 0.65 is obtained, which is in the acceptable range for flood modelling and mapping exercises (Wood et al. 2016; Landuyt et al. 2018). The CSI score has been found to be slightly biased towards overprediction, catchment size, and event magnitude (Wealands et al. 2005; Stephens et al. 2014; Stephens and Bates 2015). However, as the aim was just to evaluate the model calibration in the Clarence Catchment for a single event, it was used here due to its ubiquity in flood science

literature. The model parameterization is, therefore, considered to be adequate based on this analysis.

5.8 Chapter Summary

This chapter reviewed the general principles of hydrodynamic modelling, model uncertainty, and model selection in brief. Thereafter, detailed descriptions of LISFLOOD-FP were provided, as this was selected to model flood inundation in this thesis. In particular, the inertial acceleration solver used in this study was comprehensively presented, alongside model equations and assumptions. This was followed by an overview of the flood inundation model implementation used in this study. Finally, the model parameterization process followed in this thesis was explained. A simple framework which was developed in this study to utilize crowd-sourced water level observations for calibrating the channel roughness parameter in hydraulic models, was detailed in this section. This generalized framework was designed such that it can be extended to any number of point water level observations that may become available irrespective of their source.

PART III
FLOOD EXTENT ASSIMILATION

CHAPTER SIX

“No man ever steps in the same river twice, for it’s not the same river and he’s not the same man.”

- Plato quoting the Greek philosopher Heraclitus, in *Cratylus, 402a*.

6. Flood Extent Assimilation: Framework Development

This chapter presents an overview of the data assimilation framework proposed in this thesis. This has been used to integrate satellite-based inundation extent with the hydraulic model LISFLOOD-FP implementations described in Chapter 5. The impacts of assimilating probabilistic flood maps, such as those developed in Chapter 4, were then evaluated through a series of synthetic experiments in Chapter 7 using this approach. The optimum spatiotemporal location for the observation used for assimilation was then identified through a numerical experiment in Chapter 8. Finally, Chapter 9 describes the real world application of the methods proposed in this thesis, where fuzzy flood maps generated through the mapping algorithm of Chapter 4 were assimilated using the algorithm developed in this Chapter. In this thesis, a standard particle filter based sequential importance sampling (SIS) assimilation scheme is used, with a unique cost function designed here to account for the uncertainty in flood extent observations. The following provides an overview of the mathematical concepts of data assimilation, the role of data assimilation in reducing model uncertainty, popular techniques in hydrology, and finally the proposed technique for this thesis and the reasons for its selection.

6.1 Introduction

The Heisenberg uncertainty principle states that the position and velocity of an object cannot be simultaneously measured with any amount of certainty, as they are continuously evolving with respect to time. Therefore, it follows that the likelihood of predicting the precise position and velocity of an object simultaneously, is rather low even in theory. Extending this argument to a flood wave, it is clear that any mathematical model attempting to describe the nature of flow in two dimensions, will remain a mere abstraction regardless of the level of

physical complexity incorporated in it (Bates et al. 2014b). Similarly, an observation technique will be unable to capture a continuous process without discretizing it; such is the nature of measurements and computations, unfortunately (Waller et al. 2018a). In order to then best describe a dynamic physical process like flooding, the two sources of information that typically exist - a mathematical model describing system dynamics and discrete observations of the system states - need to be considered (Nichols 2010). Optimal predictions need to incorporate both sources of information while acknowledging the inherent uncertainty in each, while data assimilation describes the mathematical framework for accomplishing this model data integration in an uncertain environment (Rodell et al. 2004).

Given the numerous inherent sources of uncertainty in hydraulic modelling, model data integration methods are rapidly becoming indispensable in hydrology, primarily to increase the reliability of flood predictions (Grimaldi et al. 2016). Moreover, the vast advances in satellite-based hydrological remote sensing and the increasing availability of such observations, urgently demands the development of efficient techniques for optimal utilization (Schumann et al. 2018). Data assimilation is ideal for combining the process knowledge embedded in dynamic models with real world observations, as it can use the information content in both weighted against the errors to improve the accuracy of state estimation (Walker and Houser 2005). System states are therefore recursively estimated by combining noisy measurements and model predictions synergistically. The error variance in the final state estimate should theoretically always be lower than the error variances of any of the contributing datasets. Accordingly, the data assimilation challenge can be viewed as an attempt to reduce the deviation of model predictions from real world observations while also taking into account the observational uncertainty.

The utilization of data assimilation methods has been gaining momentum in environmental modelling over the last two decades, with an increased number of data sources exploitable for modelling becoming available. Earliest applications of data assimilation were in the field of meteorology, but due to data deficiency the techniques could not easily be extended to hydrological applications for a long time (Walker and Houser 2005; Bates 2012). The growth of the satellite industry has aided the improvement in modelling techniques by making hydrology a relatively data rich environment (Bates 2012). Current trends suggest that the volume of hydrologic remote sensing data will grow exponentially in the years to come (Musa et al. 2015; Yan et al. 2015). However, the diversity of the disparate sources of information, due to the type of observations available, scale, sensor calibration, and accuracy will need to be

quantified and efficiently dealt with. The way forward then, seems to be in developing a systematic and comprehensive data assimilation framework which can account for the different kinds of uncertainties while improving the model performance. Data assimilation techniques can broadly be divided into sequential (Direct observer assimilation) and variational (Dynamic observer assimilation) approaches based on the method of model data integration used (Walker and Houser 2005). However, as this thesis focuses primarily on hydrological data assimilation, the classification of assimilation approaches described here is based on the source of the hydraulic model uncertainty that the assimilation seeks to reduce (Liu and Gupta 2007).

When this thesis was initiated there existed only two studies with any attempt towards flood extent assimilation, each of which converted the satellite observed flood extent into another variable before applying the cost function (Lai et al. 2014; Revilla-Romero et al. 2016). For example, Revilla-Romero et al. (2016) interpreted surface water extents from the global flood detection system (GFDS, http://www.gdacs.org/flooddetection/) as inundation volumes, which were subsequently used to find the innovation at each assimilation time step. Similarly, the study by Lai et al. (2014) converted satellite flood extents into water depths, using a series of assumptions which are often physically unrealistic although numerically feasible. The main reason why these studies opted for such conversions is that flood extents are not hydraulic model states and therefore, constructing an appropriate cost function is rather challenging. In this context, the work of Hostache et al. (2018a) can be considered as a seminal contribution, as they directly consider probabilistic flood extents derived from SAR images for the very first time. Cooper et al. (2018) suggest the use of backscatter values from SAR images directly, which would reduce processing time and as they argue increase information uptake per image. However, using backscatter directly and not a derived flood probability map, would also mean that the improvements in accuracy offered by the consideration of image texture (Dasgupta et al. 2018), ancillary data (D'Addabbo et al. 2016), interferometric coherence (Chini et al. 2019) or multiple frequencies (Plank et al. 2017) cannot be translated into improved observation quality. Accordingly, the flood extent assimilation (FEA) framework developed in this thesis improves upon the work of Hostache et al. (2018) and uses flood probability maps to implement the model data integration.

6.2 Principles of Data Assimilation

Mathematical models of environmental system dynamics can be used effectively to generate future predictions of the system behaviour, provided that the initial states of the system are known (Lakshmivarahan and Lewis 2010). Here, states refer to the condition of a given dynamic physical system at an instant in time. For example, water depth and flow velocities are hydraulic system states. However, observational data defining all the states of an environmental system are extremely rare. Moreover, the models as well as the initial states contain inherent inaccuracies, which can lead to significant discrepancies between the predicted and actual states of the system (Lahoz et al. 2010). Incorporating observations of the system into the model generated forecasts can therefore improve state estimates while simultaneously providing information on the associated uncertainties.

State estimation is an inverse problem which can be solved using filters operating based on feedback design techniques. In environmental data assimilation where the models are extremely complex and non-linear, with state variables on the order of ~10^8, system dynamics are often multi-scale, unstable, and chaotic (Nichols 2010). Typically a large number of observations are also available through Earth Observation satellites, although their uneven spatiotemporal distribution makes the data assimilation problem ill-posed, and the state estimates are consequently sensitive to observation errors.

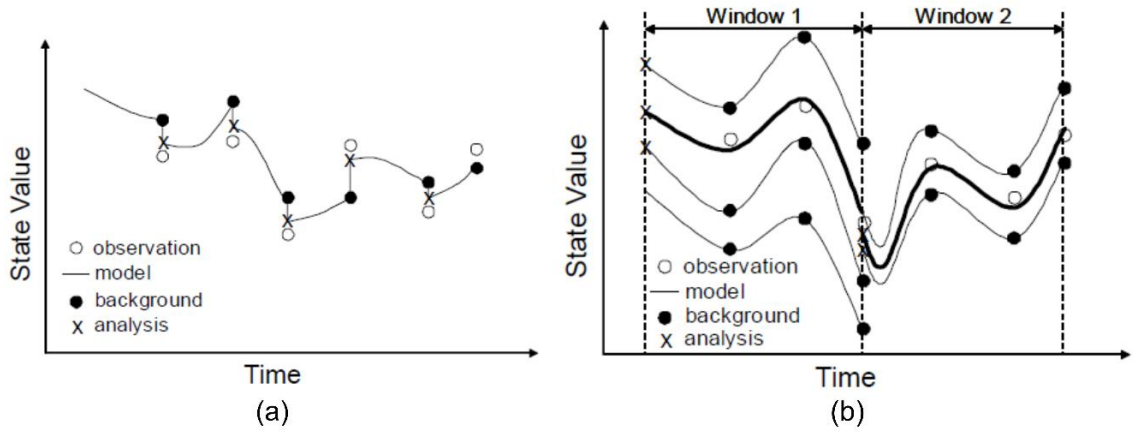

Fig. 6.1 Schematic of the a) sequential and b) variational data assimilation approaches, taken from Walker and Houser (2005).

This problem can be solved using one of two basic approaches. The first uses a sequential data assimilation scheme which dynamically feeds observations back into the model as they become available, while the second uses a variational data assimilation scheme where an optimal state trajectory is computed by fitting the observed data over a time window (See Fig. 6.1). The data assimilation problem can then be formulated as the task of finding the best estimates of system states **X** from (noisy) observations **Z,** given a (noisy) dynamic model of the system (Walker and Houser 2005). In order to demystify the large amount of jargon typically associated with data assimilation, a list of terminology adapted from Walker and Houser (2005) has been provided in Table 6.1.

6.2.1 Sequential Data Assimilation

The model forecast is sequentially updated using the difference between the observed value of the state $\mathbf{Z}_k$ and the model predicted value $\hat{\mathbf{Z}}_k$ known as the innovation, each time a new observation becomes available at time k. The predicted observation is computed from the background states $\mathbf{X}_k^b$ by applying the innovation (Andreadis and Schumann 2014). The gain matrix **K** (also known as the Kalman gain), which is a function of the relative confidence on the model and the observation, determines how much of the observation information needs to be incorporated in the forecast state estimate known as the analysis. When dealing with scalar quantities this value varies between 0 and 1, with 0 representing high observation uncertainty and 1 representing high model uncertainty. The analysis state vector $\hat{\mathbf{X}}_k^a$ calculated subsequently, then takes on the values of the actual observation if the errors in the background states are significantly larger than the observation errors. Conversely, if the actual observation uncertainty is relatively higher than the background, then the analysis remains unchanged from the original background value. The magnitude of the correction applied is then given by the product of the Kalman gain matrix and the innovation.

The analysis state vector is then defined as,

$$\hat{\mathbf{X}}_k^a = \mathbf{X}_k^b + \mathbf{K}\left(\mathbf{Z}_k - \hat{\mathbf{Z}}_k\right), \tag{6-1}$$

where the correction (second) term is the product of the gain matrix and the innovation.

Some commonly used sequential data assimilation methods include direct insertion (Mazzoleni et al. 2015), statistical/successive correction, optimal/statistical interpolation, analysis correction, nudging (Patil and Ramsankaran 2018), 3D-variational (Smith et al. 2013),

Kalman filter and variants (Evensen 2003), particle filter and variants (Arulampalam et al. 2002; Elvira et al. 2017), and evolutionary algorithm and variants (Dumedah 2012). While some of the approaches like direct insertion, nudging, and optimal interpolation offer ease of implementation, observational or background uncertainty is often ignored. The widely popular sequential algorithm known as the Kalman filter, allows the use of uncertainty estimates in models and observations. However, the linearization of model equations is necessarily required, which can cause instability in solutions and make the assimilation problem intractable for highly non-linear systems common to geophysics. The variants of the ensemble Kalman filter, particle filter, and evolutionary filters, while computationally more intensive, allow the full representation of model and observation uncertainties. These filters are therefore better suited for geophysical applications, as the models as well as Earth observation satellites are known to contain a variety of inherent uncertainties which must be acknowledged.

6.2.2 Variational Data Assimilation

Variational data assimilation techniques seek to minimize an objective function J over a pre-specified time window, to find the best fit between the model forecast and the observations, subject to the model uncertainty $\mathbf{\Sigma}$ and observation uncertainty $\mathbf{R}$. Accordingly,

$$J = {}^{1}/_{2}\left(\mathbf{X}_0 - \mathbf{X}_0^b\right)^{\mathrm{T}} \sum_0^{b^{-1}} \left(\mathbf{X}_0 - \mathbf{X}_0^b\right) + {}^{1}/_{2} \sum_0^{N-1} \left(\mathbf{Z}_k - \hat{\mathbf{Z}}_k\right)^{\mathrm{T}} \mathbf{R}_k^{-1} \left(\mathbf{Z}_k - \hat{\mathbf{Z}}_k\right), \quad (6\text{-}2)$$

where b refers to the background, the subscript represents time, and N is the length of the time window used. Typical implementations utilise an adjoint model to evaluate the derivatives of the objective function with respect to the initial model state vector $\mathbf{X}_0$ (Walker and Houser 2005). The adjoint model can be defined as a mathematical operator, which allows evaluation of the objective function sensitivity to changes in state equation solutions using a single forward and backward pass over the assimilation window.

While it is possible to numerically approximate the objective function derivatives with respect to each state using a larger number of passes, however, the use of an adjoint significantly improves computational feasibility. Variational data assimilation can then be interpreted as an optimization or calibration problem, where the state vector itself is “calibrated” to the observations over a given time window rather than the parameters, by modifying the initial values at the beginning of the window. Variational schemes can be formulated using strong (perfect model assumption) or weak constraints (model errors as white process noise). The

Table 6.1 Commonly used data assimilation terminology, after Walker and Houser (2005).

DA Term	Definition
State	Condition of a given physical system, e.g. water depth and flow velocities are hydraulic system states
State error	Deviation of the estimated state from the truth
Prognostic	A model state/flux required to propagate the model forward in time
Diagnostic	A model state/flux diagnosed from the prognostic states – not explicitly required to propagate the model
Observation	Measurement of a model diagnostic or prognostic variable
Model ensemble	Set of uniquely parameterized model realizations
Open loop	Model ensemble without any data assimilation
Error covariance matrix	Description of uncertainty in terms of standard deviations and correlations
State perturbation matrix	Matrix containing values of deviations of each individual state vector from the ensemble mean vector
Prediction	Model estimates of future states of a given system
Update	Correction to a model prediction using observations
Background	Model forecast, prediction, or state estimate prior to an update
Analysis	State estimate after an update
Innovation	Observation minus prediction, a priori residual between model prediction and observation of system state
Gain matrix	Correction factor applied to the innovation, calculated based on the level of confidence in the model and the observation
Tangent linear model	Linearized version of a non-linear model using the Taylor's series expansion, around a chosen equilibrium point
Adjoint	Inverse operator allowing the model to be run backwards in time

inclusion of the model error vector **w** and the model error variance-covariance matrix **Q** in Equation 6-2 then leads to,

$$J = \tfrac{1}{2}\left(\mathbf{X}_0-\mathbf{X}_0^b\right)^T \sum_0^{b-1}\left(\mathbf{X}_0-\mathbf{X}_0^b\right) + \tfrac{1}{2}\sum_0^{N-1}\left(\mathbf{Z}_k-\hat{\mathbf{Z}}_k\right)^T \mathbf{R}_k^{-1}\left(\mathbf{Z}_k-\hat{\mathbf{Z}}_k\right) + \tfrac{1}{2}\sum_0^{N-1}\mathbf{w}_k^T\,\mathbf{Q}_k^{-1}\mathbf{w}_k. \quad (6\text{-}3)$$

Variational methods are usually well suited for smoothing problems and provide rather accurate state estimates, albeit at considerable computational cost. As the backward integration of the model is essential over the entire time window at every assimilation time step, these

methods have limited use for real-time forecasting applications. Moreover, the development of robust adjoint models for the complex and non-linear equations of hydrological and hydraulic models, is a scientific challenge in itself and has impeded widespread applications of variational methods.

6.3 Reducing Model Uncertainty: The Role of Data Assimilation

Any given process model essentially uses physics of the real world to predict the true state of the system at any given point in time, which can subsequently be validated with observations. It is necessary, however, to consider the inherent errors in the model structure, linearization of the non-linear systems, spatiotemporal domain discretization, model initialization, forcings, and parameters (Haile and Rientjes 2007). Similarly, observation errors caused by sensor calibration issues, over-use, systematic bias, and rounding off, also need to be acknowledged (Waller et al. 2018a). Disentangling and dealing with uncertainty in hydrological and hydraulic modelling has widely been accepted as a necessity by the global hydrological research community (Blöschl et al. 2019). It is therefore relevant to understand the two different types of uncertainty in hydrological modelling and the characteristics of each as given below and illustrated in Fig. 6.2 (Gong et al. 2013).

1. *Aleatory uncertainty* or the uncertainty in repeatable events caused due to intrinsic randomness and unpredictability (O'Hagan 2006). This form of uncertainty is characterized by stationary statistical properties, which might be structured but can eventually be defined by a stationary random distribution (Beven 2016). For discrete variables, this randomness can then be quantified by individual probabilities of each possible outcome, while a probability density function can be used to characterize this for continuous variables (Beven 2014).
2. *Epistemic uncertainty* can be defined as the uncertainty in non-repeatable events caused primarily due to a lack of sufficient knowledge about the process, data, or parameters (O'Hagan 2006). It is challenging to objectively characterise epistemic uncertainties, precisely because they arise from a lack of knowledge about process representations, effective parameter values, input data or calibration data such as flood discharge estimates (Beven and Hall 2014). However, considering multiple competing alternative model structures, probability distributions, or probability density functions based on

Bayesian probability theory, are a few possible options to effectively characterize epistemic uncertainties (Woodhead 2007).

In flood management, epistemic uncertainties often lead to biased deterministic predictions even for calibrated flood inundation models, especially as appropriate calibration data for all model states - flow velocities for example - are rarely available (Beven 1993, 2014; Pappenberger et al. 2007a; Beven and Hall 2014; Sene et al. 2014; Beven et al. 2018a). Epistemic uncertainties in flood modelling can then be further classified by the source model component (Liu and Gupta 2007).

1. *Structural errors:* Models use various assumptions and approximations to characterize a complex real world system which introduces some inherent error in the system. The computational implementation which requires discretization over time and space also contributes to this error. If the model conceptualization is improper (i.e. physically significant processes are ignored) major errors may follow. As the system boundary is part of the model conceptualization, it contributes to the structural errors (Beven 2016).
2. *Parameter errors:* Parameters represent system properties continuously varying over space and time. Due to the spatiotemporal aggregation of parameters required for practical uses, the real world heterogeneity is inadequately represented quite often. Additionally, most parameter values cannot be directly measured on field or vary continuously making measurements expensive. For instance, channel geometry can be considered as one such parameter which is constantly in flux and thus very hard to monitor in field. The indirect estimation methods employed such as expert knowledge

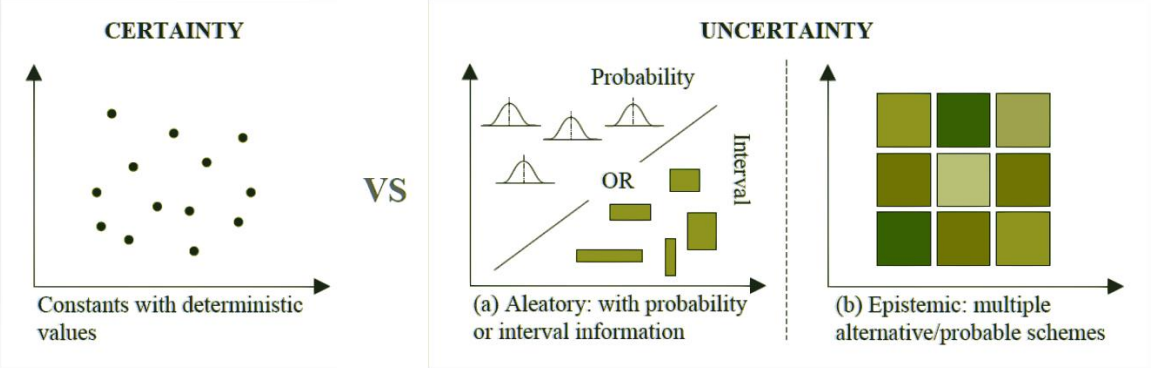

Fig. 6.2 Illustrative representation of the different facets of uncertainty, adapted from Hou, Li, and Liang (2019).

or model calibration, introduce uncertainty due to the ambiguity of the optimal parameter choice (Beven 2006).

3. *Data or measurement errors:* The uncertainty caused by model inputs and initial conditions fall into this bracket. They can typically be attributed to errors in the recording device or scale incompatibility between the sensor observation and corresponding model prediction which may require spatiotemporal aggregation or interpolation (Waller et al. 2018a).

In order to minimize these uncertainties, some characteristics which describe their probability distribution need to be known. However, such information may often be difficult to obtain in practice (Winsemius et al. 2013). For example, the frequentist approach to probability of observing a particular outcome of an event is defined based on occurrence frequency, and therefore, can only be used to quantify and reduce aleatory uncertainties (Oakley and O'Hagan 2004). In such situations assumptions of local linearity and gaussianity have often been made, perhaps incorrectly. In contrast, Bayesian probability theories evaluate the larger picture by linking the degree of belief in a given hypothesis before and after evidence is presented. In the context of hydrology for instance, this may imply one's confidence in the belief that a certain catchment specific calibrated parameter set will remain stationary across events of different magnitude (Oakley and O'Hagan 2004; Beven et al. 2018b). This interpretation of probability can be used to quantify both aleatory and epistemic uncertainties through probabilities (Woodhead 2007).

The recent cohort of Bayesian ensemble based methods, which use a prior distribution based on system knowledge and iteratively update it as observations become available, have proved most effective for uncertainty quantification in hydrology (Liu and Gupta 2007). Data assimilation allows the use of these Bayesian theories for correcting model state predictions in near real-time through online observations (Sene et al. 2014). In addition to assessing the overall forecast uncertainty, data assimilation techniques also aim to reduce the uncertainty using data on inputs, forcings, or the model state. Moreover, the different sources of forecast uncertainty can be disentangled and independently evaluated and corrected in an online manner for operational applications (Cloke et al. 2013). Furthermore, the available data can be supplemented by constraining the physically-based model equations with parsimonious observations, which can then be used to estimate unobserved quantities (Walker and Houser

2005). The next section examines the different applications of DA in hydrology and the resulting classification of techniques.

6.4 Hydrological Data Assimilation: The Systems Perspective

Hydrological data assimilation can broadly be classified into three categories based on the source of the error that one seeks to reduce. Each of these may be used in combination with one another or as standalone procedures (Walker and Houser 2005; Liu and Gupta 2007; Mazzoleni 2016).

1. *System identification:* The primary objective of these data assimilation approaches is to assign an appropriate "input to intermediate state to output" mapping process, which accurately represents the true conceptual relationships among them. Alternatively, system identification can be defined as the selection of the most conceptually appropriate model structure.
2. *Parameter estimation:* The best-fit parameter values are stochastically estimated using available data. This enables the model to make more accurate simulations. These techniques take into account the structural and measurement errors, thereby, offering some advantage over traditional deterministic calibration approaches.
3. *State estimation:* These methods combine all available state information derived from various data sources with that given by the process model optimally to give the best possible estimate of the "true" state of the system. Existing literature uses the term data assimilation synonymously with state estimation.

6.4.1 System Identification

Research has shown that for a given hydrological system, several parameter sets may exist which give equally acceptable real world predictions when used within the process model (Beven and Binley 1992; Beven 2006). It follows from this, that several equally plausible conceptual model structures may exist which can reproduce the characteristics of a particular hydrologic environment (Oakley and O'Hagan 2004; Woodhead 2007). However, real world hydrological systems are usually too complex and heterogeneous for all their unique behavioural properties to be adequately captured by a single process model (Neuman 2003). Therefore, by not considering equally probable model structures and allowing them to compete,

a systematic bias is introduced leading to an unconscious underestimation of the structural errors (Liu and Gupta 2007). Data assimilation techniques usually account for structural uncertainty by adding an unbiased error term to the model transition (input to output mapping) equation. However, to fully acknowledge the equifinality of model structures, various different model structures which work on diverse underlying assumptions need to be considered (Duan et al. 2007).

One such approach is the Bayesian Model Averaging (BMA) proposed by Hoeting et al. (1999) in which the prediction for the quantity of concern is approximated by a weighted sum of the predictions given by several independent or mutually exclusive model structures. The weights for the individual models are determined by using the Bayes theorem. There is a less computationally intensive interpretation called Maximum Likelihood BMA (MLBMA) where expert knowledge and field information about the system is used to assign likelihood values to the competing structures. MLBMA is based on the assumption that all models giving similar results may not be giving the output values as a function of their structure alone (Neuman 2003). Another popular approach is the Multi-Model Ensemble (MME) technique where the ensemble of model predictions is made up of smaller ensembles of independent models (Duan et al. 2007). The objective of MME techniques is to account for uncertainty propagation from various sources by sampling from the simulated output distributions of multiple plausible models (Butts et al. 2004). A well-chosen MME has given more consistent outputs than deterministic single model ensembles, primarily due to the broad range of solutions considered in the formulation (Georgakakos et al. 2004; Christensen and Lettenmaier 2007; Villarini et al. 2010).

6.4.2 Parameter Estimation

Hydrological parameter estimation techniques have evolved tremendously from elementary manual calibration to the high end variational assimilation schemes (Moradkhani et al. 2012). For example, Horritt (2006) used a basic trial and error method to assign parameter values and quantify the error in their selection using standard deviations, while Thiemann et al. (2001) used the Bayesian Recursive Estimation (BaRE) technique for model parameterization, which assimilates measurement data to recursively estimate state and parameter values, resulting in lower predictive uncertainty with each subsequent run. The latter technique, however, resulted in convergence to a single parameter value (deterministic) which did not

allow for effective uncertainty quantification (Beven and Young 2003; Gupta 2003). In contrast, the popular Generalized Likelihood Uncertainty Estimation (GLUE) technique, which uses a large number of Monte Carlo simulations, identified multiple equifinal parameter combinations which might be able to map inputs to outputs equally well (Beven and Binley 1992). While computationally intensive, the GLUE technique provided a major breakthrough towards probabilistic calibration methods. The Dynamic Identifiability Analysis (DynIA) used an approach similar to GLUE except multiple objectives were used to derive the equifinal parameter sets (Wagener et al. 2003). However, most of these approaches are no longer as popular, primarily due to the computational revolution which has allowed the use of more complex parameter optimization techniques.

Optimization techniques of higher complexity than the ones discussed above have also been found to perform well for parameter estimation. For example, Markov Chain Monte Carlo (MCMC) algorithms which search for optimal parameter combinations within the acceptable parameter space, and select values based on a predetermined likelihood measure, were adapted within a particle filter framework to allow real-time hydrological model improvement (Moradkhani et al. 2012). One such optimization scheme based on these methods is the Shuffled Complex Evolution Metropolis (SCEM) technique (Vrugt 2003). SCEM uses a large number of initial random samples to efficiently locate the true global optimum in the parameter space. Additionally, many parallel searches with different starting points are conducted enabling an unbiased exploration of the search space and identification of many equifinal parameter sets based on likelihood. Extending this approach to a multi objective analysis allows for an even more thorough check for optimal combinations of equally well performing parameters (Wöhling and Vrugt 2008; Chu et al. 2010). Particle swarm optimization (PSO) techniques based on the behaviour of bee/bird swarms have also been used for parameter estimation. Random particles fly about in a multi-dimensional acceptable parameter space, readjusting flight paths and trajectories based on their own and their neighbours' experience. Local and global best position estimate is updated periodically for each individual particle as demonstrated using the Muskingum routing model (Chu and Chang 2009). The same approach can also be extended for a multi-objective analysis by using the concepts of Pareto dominance (Gill et al. 2006). The advantages of using a dual state-parameter estimation approach using the Particle Filter and Simultaneous Optimization and Data Assimilation (SODA) with Ensemble Kalman Filtering (EnKF) have also been demonstrated through various numerical experiments (Moradkhani et al. 2005b; Vrugt et al. 2005).

Variational data assimilation techniques which essentially use the adjoint model equations (inverse problem) for assimilating state measurements, have also been used to solve for true parameter values (Lai et al. 2013). This is accomplished by minimizing a cost function representative of the differences between the simulated and observed state data. Chertok and Lardner (1996) used variational techniques for a parameter estimation problem but found that many instances of the measured state with good spatial spread are required for an accurate solution. As discussed in Chapter 2, a variational technique 4D-Var was also used to assimilate remote sensing derived flood extent data in a 2D flood model for estimation of parameters and improving the state estimation accuracy (Lai et al. 2013), while remote sensing derived water levels were assimilated using the same method for the estimation of initial conditions (Lai and Monnier 2009). However, prior studies have shown variational techniques to be difficult to implement and computationally unrewarding especially for forecasting applications (Wu et al. 2015). As the focus of this study is on flood forecasting, emphasis will be towards state estimation in near real-time. As variational data assimilation techniques cannot be implemented in an online manner for forecasting and operate over a fixed length time window propagating the algorithm backwards, they have not been considered for use in this thesis.

6.4.3 State Estimation

Sequential data assimilation techniques have widely been used in the last decade for state estimation and reduction of predictive uncertainty in hydrological models (Mazzoleni et al. 2017). Sequential assimilation schemes primarily work on the assumption that at any given time, prior estimates of the states known as the background values are known. The innovations (i.e. the differences between the sensor observed values of the true state and the model predicted observations) calculated using the background states at that point in time are used to correct the background state vector such that improved state estimates can be obtained, called analysis states (Walker and Houser 2005). The model equations are then propagated forward through time using the analysis states as the background values to the next point in time where observations are available, and the process is repeated sequentially (Fig. 6.1). This technique can be executed over single consecutive time steps, known as filtering, or over a window of several time steps, known as smoothing (Liu and Gupta 2007).

The Kalman filter was one of the first and most popular sequential data assimilation techniques to be used in hydrology (Moradkhani et al. 2012). Various extensions of the Kalman

filter method have been developed and used since. However, a major limitation of this approach is that the nature of the error distribution is assumed to be Gaussian which is almost certainly untrue for hydrodynamic cases (Wang and Cai 2008). In order to account for this non-Gaussianity, the ensemble technique was adapted. The Ensemble KF (EnKF) estimates error statistics from the sample member model forecasts, and thus no a priori assumptions need to be made about the distribution. The EnKF has been used in various studies for streamflow, soil moisture and water level assimilation (Devegowda et al. 2010; Xie and Zhang 2010; Mason et al. 2012b; Paiva et al. 2013; Abaza et al. 2014; Trudel et al. 2014). The EnKF technique has also evolved into more advanced filters like the Ensemble Square Root Filter (Chen et al. 2013), recursive EnKF (McMillan et al. 2013), Ensemble Kalman Smoother (Li et al., 2015, 2014; Li et al., 2013) and the Local Ensemble Transform Kalman Filter (García-Pintado et al. 2013, 2015; Andreadis and Schumann 2014; Cooper et al. 2018, 2019; Waller et al. 2018a) amongst others. Even with the ensemble approach, the evolution of Kalman-based methods are governed by second order characteristics and a linear state updating equation which cannot capture the system dynamics of highly non-linear models (Yan and Moradkhani 2016). Accordingly, the probability distributions of such systems can only be accurately computed by including higher-order moments in the analysis (Moradkhani 2008).

Particle filters (PF) based on Sequential Monte Carlo (SMC) methods represent the posterior probability distribution more comprehensively, thereby improving the capability of the assimilation scheme to cope with the propagation of non-gaussian noise within non-linear systems (Matgen et al. 2010). Independent random samples called particles, which are representative of the posterior probability are directly sampled discretely from the state space. As the number of these samples tends to infinity, the assumption is that the posterior probability distribution will tend towards the "true" non-Gaussian form. In the case of hydrodynamic models, the water levels generated through a single realization of the model at a time t is represented by a particle and the cross sections or computational grid points in the analysis domain provide the state variables (Giustarini et al. 2011). The particles are appropriately weighted and propagated by assimilating the state measurements into the model (Nakamura et al. 2009). The sampling is done using a Sequential or Bayesian Importance Sampling (SIS or BIS) procedure which chooses particles based on a known function called the proposal distribution or importance density (Fearnhead and Künsch 2017). The weights are subsequently assigned as the ratio of the target posterior probability density function (pdf) to the proposed

pdf. These weights are iteratively updated based on the probability of the particle having a value similar to the observed state at that time (Arulampalam et al. 2002).

The SIS algorithm is prone to degeneracy as particles with lower weights are continuously discarded. In order to efficiently mitigate this challenge, Sequential Importance Resampling was proposed (Moradkhani et al. 2005), which replicates particles in proportion to their weights and assigns a common weight (i.e. the reciprocal of the number of particles to all the new ones). The problem of sample impoverishment may also occur due to particles with higher weights getting selected multiple times, causing loss of sample diversity which leads to clustering of particles in small sub-regions of the state space. This problem has been overcome by perturbing the particles around the most likely values. Fig. 6.3 graphically shows the iterative fitting of the posterior pdf based on the SIS and SIR implementations, along with the problems associated with each. Many studies have established the relative supremacy of the particle filter over other state estimation techniques and the advantages of using a non-parametric assimilation algorithm, and thus it forms a major component of the data assimilation framework used in this thesis (Nakamura et al. 2009; Rémy et al. 2009; Briggs et al. 2013; Mattern et al. 2013; Noh et al. 2013; Thirel et al. 2013; Dumedah and Coulibaly 2013a, 2014; Li et al. 2014a; Kantas et al. 2015; Browne 2016; Penny and Miyoshi 2016; Yan and Moradkhani 2016; Jouin et al. 2016; van Leeuwen 2017; Elvira et al. 2017; Long and Hu 2017; Piazzi et al. 2018; Yan et al. 2018; Herbst and Schorfheide 2019).

Evolutionary Data Assimilation (EDA) techniques are relatively new to the field of hydrological data assimilation. The concepts of biological evolution and natural selection employed in population based sorting tools like genetic and evolutionary algorithms were adapted for data assimilation (Dumedah 2012). In genetic algorithms, prospective solutions (individuals) are made to compete and their performance assessed based on predefined objective functions (Bai and Li 2011). Each cycle of evolution is called a generation and the fittest individuals are naturally selected and made to reproduce using crossover and mutation, which ensure heredity and variation respectively (Ridolfi et al. 2014). The population size remains constant and the average health of the individuals in the population keeps improving with successive generations, but the number never converges to one solution in keeping with the principles of equifinality (Dumedah and Coulibaly 2013b). The individuals contain state and parameter values (genetic make-up – genotype) and the watershed or floodplain response (expressed behavior – phenotype) (Dumedah 2015). When the above concepts are extended for

multiple contrasting objectives, the vector having as its elements the relevant objective function values is optimized (Dumedah 2012).

The Pareto Optimal Set (POS) is the set of solutions remaining at the end of all the generations and the Pareto front is its mapping from parameter to objective space (Wöhling and Vrugt 2008; Zaji et al. 2018). Non-dominance is when all solutions perform equally well in all objectives being assessed. There are several algorithms which exist to rank non-dominant solutions like Strength Pareto Evolutionary Algorithm—II (SPEA-II), Pareto Archive Evolution Strategy (PAES), Pareto Envelope-based Selection Algorithm—II (PESA-II), Micro Genetic Algorithm and Non-dominated Sorting Genetic Algorithm – II (NSGA-II). The application of EDA proposed thus far, uses the NSGA-II, however, any of the listed ones may be adapted for the purpose. The POS is obtained by allowing the fittest solutions to compete with their elite parents (elitism); preserving stable populations in the neighbourhood of optimal solutions (niching); adjusting fitness of individuals and distribution to niches proportionally to their average fitness (fitness sharing); discouraging crowding anywhere in the search space by replacing many similar solutions with one (Samuel et al. 2014).

When adapted to data assimilation, each ensemble member is an individual having an associated parameter set and a hydrological response, and the sensitivity of one to the other can

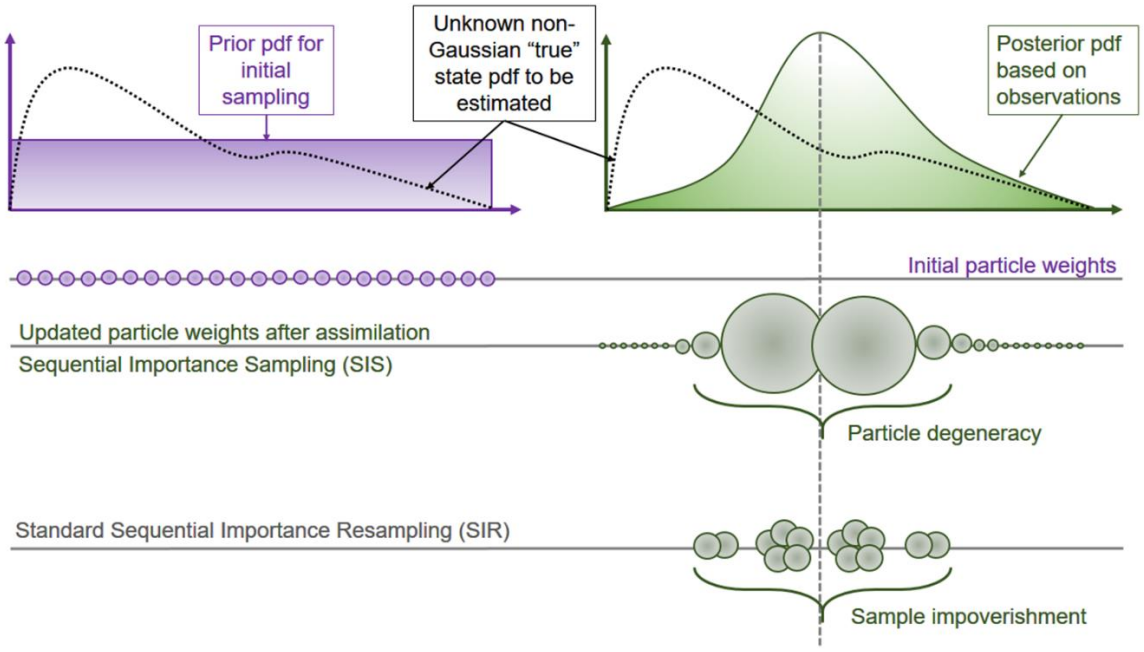

Fig. 6.3 Schematic representation of particle filter data assimilation, using the sequential importance sampling and the standard sequential importance resampling algorithms.

be assessed. These are made to compete based on various objective functions using Pareto dominance concepts at each time step; evolved and evaluated for subsequent time steps (Dumedah and Walker 2013). After the evaluation, only the Pareto optimal set is assimilated for the given time step and the procedure repeated for the next time step. The ensemble size is given by the product of the population and generations and remains constant throughout. The time variance of model states and parameters is acknowledged as they are re-estimated every time a new observation becomes available.

An assessment of the performance of EDA in comparison to the EnKF and the PF, showed that EDA gave the highest cluster convergence for parameter estimation. While EDA gave higher ensemble mean accuracy for short lead times, for longer lead times the PF performed better. As EDA is able to reduce the search space, smaller ensembles can be used after a while. In fact an adaptive ensemble size may be chosen to maximize computational efficiency. It is also worth mentioning that the assimilation of new information using the PF does not appear to override the background information completely, thus making it less sensitive to forcing variables. Previous studies have shown that if there is a large variation in the input or forcing data immediately after the assimilation time step, and the next observation is unavailable for some time after that, the model state rapidly decays to its original trajectory (Samuel et al. 2014). As the PF scheme is relatively more insensitive to extreme forcings, it may still retain assimilation effects. Moreover, as hydraulic models are computationally expensive and prone to numerical instabilities, restarting the model with updated states e.g. water depths, is usually ineffective as it can result in hydrostatic initialization shock. Consequently, the PF-SIS implementation remains the most suitable assimilation approach for flood modelling, as it does not require restarting the model and has successfully been applied to a real case study (Hostache et al. 2018b). Accordingly, in this thesis a new flood extent assimilation framework will be developed based on the PF-SIS algorithm.

6.5 Particle Filter-based Flood Extent Assimilation Framework

In the PF-SIS based flood extent assimilation algorithm developed in Hostache et al. (2018a) for the Severn Catchment, UK, local weights are first computed based on a pixel-wise comparison, as each model grid cell has an associated state vector. As the model state can only take two values of flood extent in this implementation, the binomial distribution was used to

approximate the posterior density (Renaud Hostache, *Personal communication*). The PF-SIS algorithm eventually requires a spatially lumped global weight to rank particles and calculate the weighted ensemble mean, the local weights must therefore eventually be aggregated to yield one global value. The joint probability density was used for this, computed as a simple product of all the local weights by assuming statistical independence of pixel-wise observation errors (Hostache et al. 2018b). However, in reality SAR backscatter errors are a function of speckle and the underlying land-use, which often exhibit spatial autocorrelation. In fact, for SAR-derived water level observations in the Severn Catchment, Waller et al. (2018b) showed that average observation error correlation existed for ~7km of separation distance. Assuming statistical independence for correlated observation errors during assimilation, this may lead to an underestimation of standard deviation and correlation length scale (Waller et al. 2016).

Additionally, as the product of a large number of low local weights values is infinitesimally small, an empirical rescaling factor α was introduced, such that even a model output in complete disagreement with the observation, had a strictly positive non-zero global weight. Flood probabilities obtained from SAR, were also truncated to fall within the range 0.001 and 0.999, to avoid the unreasonable assignment of a global weight of zero to a particle due to local mismatches. Pixels with flood probabilities lower than 0.5 in the observation, were masked out prior to assimilation. As these usually dominate SAR based flood maps, over-prediction is penalized much more than under-prediction, which is not conducive for operational applications. While these were all mathematically feasible and pragmatic solutions to the challenges of flood extent assimilation, there was significant scope for improvements which formed the basis of the algorithm development in this study.

For the model ensemble detailed in Chapter 7, LISFLOOD-FP was used to generate an ensemble of water depth maps, which were subsequently converted to binary flood extents using a 1cm threshold on the simulated water depth in each pixel. This implies that all grid cells with water depths >1cm were considered as flooded and all others were considered as non-flooded. While the selection of this threshold may seem arbitrary, it was chosen to account for the large grid size, where setting a 10cm threshold as commonly used (Pappenberger et al. 2005, 2007b, a), might ignore ~729 cubic meters (0.09×90^2) of water in the floodplain. The ensemble of binary flood maps simulated by the model, were then compared against the real or synthetically generated SAR-derived flood extents, using a novel cost function which directly considers the entire flooded area.

Particle filters are based on Sequential Monte Carlo simulations as discussed in the previous section, which enables a non-parametric representation of the continuous posterior probability distribution function. By randomly selecting a large number of discrete and independent samples from a proposal distribution, the true non-Gaussian posterior pdf is iteratively estimated (Arulampalam et al. 2002). Initially a uniform pdf is assumed, which is sequentially updated by assimilating observational information as and when it becomes available (Moradkhani et al. 2005a). Each ensemble member or a particular model realization, represents a distinct particle with unique associated inputs and parameters. At each assimilation time step, the Bayesian conditional probability of a particular model trajectory being "true" given that the observation is "true", is individually evaluated and used as the particle weight (Kantas et al. 2015). In the Sequential Importance Sampling algorithm, which is arguably the simplest implementation of the ensemble weighting procedure in particle filters, the weighted ensemble mean is subsequently calculated. This ensemble mean vector is then representative of the total information content of the model and the observations (Plaza et al. 2012).

The Sequential Importance Sampling algorithm initially generates particles based on a predetermined proposal distribution or importance density, as sampling from the complex, unknown, non-Gaussian posterior distribution is not possible (Moradkhani et al. 2005a). Although it is possible to use any proposal distribution given prior knowledge on the expected form of the posterior, a uniform distribution is often assumed when no other knowledge is available as in this case (Li et al., 2013). The importance weights are usually assigned as the reciprocal of the ensemble size, as in this thesis (Plaza Guingla et al. 2013). As new observations become available and importance weights are sequentially updated, a discrete approximation of the true continuous posterior is obtained (Plaza et al. 2012). The expectation of the state vector, computed as a weighted sum of all state vectors in the ensemble, will tend to the true value as the number of particles tends to infinity (Arulampalam et al. 2002). However, as a very large number of hydraulic model runs over a large model domain is computationally infeasible, the number of runs in practice needs to be determined based on empirical tests. For this study, 128 particles were used to represent the state-parameter distribution, although several remedial measures were used to avoid filter collapse.

Using the Sequential Importance Sampling (SIS) algorithm provides the unique advantage of avoiding hydrostatic initialization shock, which often occurs in state updating data assimilation filters applied to hydrodynamic models. As data assimilation is essentially a

statistical recombination procedure, the momentum across the domain is not conserved when states are updated and the model is restarted from the updated states (García-Pintado et al. 2015). This is because flow fluxes drop to zero when the model is stopped mid-simulation, and conserving the momentum of large volumes of water becomes impossible. As the Sequential Importance Sampling algorithm only alters the ensemble statistics without disturbing the delicate model dynamics, it perfectly suits the problem of hydraulic data assimilation. Although Cooper et al. (2018), demonstrate that this problem can be overcome by applying pre-assimilation flow velocities to the updated state vectors; this was only shown through application to a synthetic catchment topography, and therefore requires extensive testing before it can be used for real case studies with complex catchment topography (Cooper et al. 2018). However, the proposed data assimilation framework can be updated in the future, as novel methods to conserve the momentum of water in the domain during assimilation are developed and extensively tested.

In order to develop an appropriate cost function for the SIS-based particle filter implementation used here, a Critical Success Index (CSI) based ranking of particles was first computed to understand how the particle weights should look when ranked in comparison with the observed extent. Multiple cost functions were evaluated with respect to weight sensitivity and accurate particle ranking. The CSI could not be directly used as it is a binary pattern matching measure and does not allow for probabilistic representation of the observations. Several tests were conducted for the RMSE calculated from reliability diagrams, but as observed in Chapter 4, the lumped metric was uninformative for large flood extents near the peak. Mutual information (Shannon and Weaver 1964) was finally chosen to quantify the agreement between the observed probabilistic SAR-based flood map (O_f) and the modelled flooded area (M_f). At this step mutual information is used to compute the conditional probability of the observation being true given that each model is true. Mutual information can be defined as the measure of the amount of information one random variable contains about another (Cover and Thomas 1991).

In order to better interpret mutual information, the concept of entropy needs to be introduced. Entropy is defined as the measure of uncertainty of a random variable or the self-information contained in it (Hirschmüller 2008). Shannon and Weaver (1964) quantitatively defined entropy $H(X)$ of a random variable X as a continuous function of its probability distribution function $P_X(x)$, which fulfills the following conditions:

1. It should be maximal when $P_X(x)$ is uniform, and in this case it should increase with the number of possible values X can take;
2. It should remain the same if the probabilities assigned to different values of X are reordered; and
3. The uncertainty about two independent random variables should be the sum of the uncertainties about each of them.

It can be mathematically proven that the only description of uncertainty which fulfills all of the above conditions is given by,

$$H(X) = -\sum_{x\in X} P_X(x) \log P_X(x) = E_{P_X} \log \frac{1}{P_X(x)}, \quad (6\text{-}4)$$

where E_{P_X} is the expectation of the probability distribution function. Fig. 6.4 details the concept of entropy graphically for a single coin toss experiment, which is maximum for a fair coin with a probability of 0.5, where the uncertainty about the unbiased outcome or value that x can take is maximized. The conditional entropy $H(X|Y)$ between two random variables X and Y is then given by extending Eqn. 6-4 to

$$H(X|Y) = -\sum_{x\in X} \sum_{y\in \mathrm{Y}} P_{X,Y}(x,y) \log P_{X,Y}(x|y) = E_{P_{X,Y}} \log \frac{1}{P_{X,Y}(x,y)}, \quad (6\text{-}5)$$

where $P_{X,Y}$ is the joint probability distribution function. With these foundations, relative entropy can now be defined as the measure of the distance between two distributions, quantitatively expressed as the expected logarithm of the likelihood ratio. Relative entropy is also known as the Kullback-Leibler (KL) distance between two probability distributions p and q, which provides a measure of the error introduced by assuming that the distribution is q when the true distribution is p. Accordingly, the KL distance $D(p||q)$ or relative entropy is given by

$$D(p||q) = \sum_{x\in X} p(x) \log \frac{p(x)}{q(x)} = E_p \log \frac{p(x)}{q(x)}. \quad (6\text{-}6)$$

Relative entropy is always non-negative and zero only if $p = q$. However, it is not a true measure of distance as it is neither symmetric or satisfy the triangle inequality (Cover and Thomas 1991).

Mutual information can now be introduced as the measure of the reduction in the uncertainty of one random variable due to the knowledge of the other, calculated as the relative entropy between the joint and the product distributions of two random variables according to

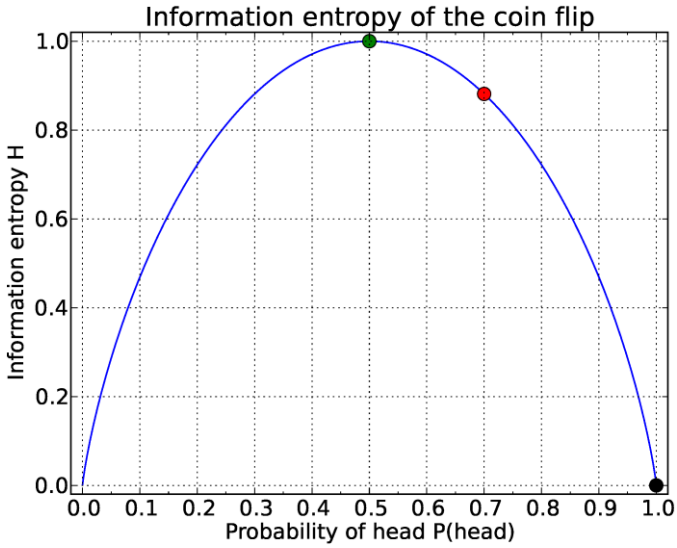

Fig. 6.4 Example plot of information entropy of a binary system, taken from Wellmann (2013). In the case of the fair coin with P(head)=P(tail)=0.5, the information entropy is maximal with a value of H(0:5)=1(green dot); in the case of the bent coin with P(head)=0.7, the uncertainty of the system is reduced, and the information entropy is accordingly lower H(0.7)≈0.88 (red dot). In the case of a double headed coin with P(head)=1, no uncertainty

$$\mathrm{I(X;Y)} = \mathrm{H(X)} - \mathrm{H(X|Y)} \tag{6-7}$$

$$= -\sum_{x\in X}\sum_{y\in \mathrm{Y}} P_{X,Y}(x,y)\log\frac{P_{X,Y}(x,y)}{P_X(x)P_Y(y)} \tag{6-8}$$

$$= E_{P_{X,Y}}\log\frac{P_{X,Y}(x,y)}{P_X(x)P_Y(y)}. \tag{6-9}$$

Mutual information can be expressed as a Venn diagram as in Fig. 6.5, showing the interrelationships between the various information measures used to define the uncertainty of two correlated random variables. While all of the above equations are defined for discrete random variables, they can be extended to continuous variables by replacing the summations with integrals. In order to compare discrete and continuous variables as in the present case, continuous variables can be quantized into a large but finite number of bins by assuming that the density is continuous within each bin. Here, 256 bins have been used after testing for sensitivity up to 1024 bins, as increasing beyond 256 yielded no notable impact on the calculation of mutual information. Mutual information is usually measured in bits when using log with a base of 2 as introduced by the computational information theorists. This was retained in the present case as the concept of bits to measure information also suits remote sensing images fairly well.

The use of mutual information has been rather ubiquitous in remote sensing, specifically intensity based image matching and registration studies (Chen et al. 2003; Hirschmüller 2008; Suri and Reinartz 2010; Horkaew and Puttinaovarat 2017; Liu et al. 2018). Moreover, it has also been used to establish causality of changes in flood occurrences with respect to mean precipitation patterns (Perdigão and Blöschl 2014). In the present context, mutual information was found to be the most suitable cost function as it allows an elegant comparison between discrete and continuous random variables. An additional advantage of using mutual information is that it quantifies the probability of correctly guessing the observation given each model, which forms the first step of particle filtering based on rules of Bayesian conditional probability (Godsill et al. 2004). Previous studies have highlighted the limitations of binary pattern matching measures currently in use, which tend to favor models which over predict and lose sensitivity when all the models exhibit positive bias (Wealands 2006; Stephens et al. 2014; Landuyt et al. 2018). This inhibits the cross-catchment comparison of binary pattern matching measures, which has prompted research towards more efficient evaluation measures. Here, mutual information is introduced for the first time to flood mapping and modelling literature with the expectation that the field could benefit from the rapid uptake of this metric, which overcomes the limitations of binary measures to a large extent.

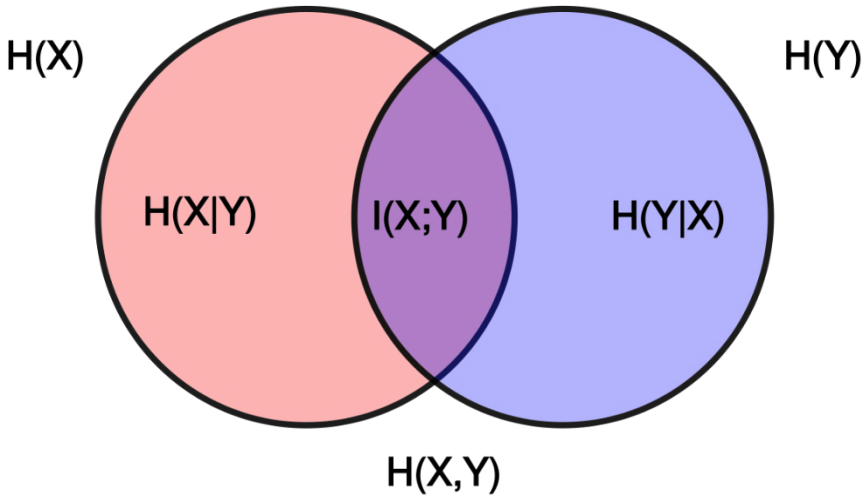

Fig. 6.5 Venn diagram showing additive and subtractive relationships various information measures associated with correlated random variables X and Y. The area contained by both circles is the joint entropy $H(X,Y)$. The circle on the left (red and violet) is the individual entropy $H(X)$, with the red being the conditional entropy $H(X|Y)$. The circle on the right (blue and violet) is $H(Y)$, with the blue being $H(Y|X)$. The violet is the mutual information $I(X;Y)$. Source: https://en.wikipedia.org/wiki/Mutual_information.

The mutual information based cost function allows for an efficient ranking of particles, however, an additional rescaling factor was introduced to further enhance particle weight variability. This rescaling factor based on the ratio of a given particle weight to the maximum weight obtained for all particles, was used for weight tempering and the exponent ϑ optimized through numerical experiments. In this thesis, the value of ϑ was fixed based on the consequent weight variability and consequent particle selection obtained. For example, a simple check that was used was whether a single particle had an associated weight of at least ~0.1. This indicates that the algorithm is able to distinguish between good and bad performing particles effectively. The ideal weight distribution has some variability, which is imperative to distinguish good models from bad, but the variance should not be so high as to tend towards degeneracy. The value of ϑ was optimized according to the aforementioned principles, and a value equal of $\vartheta = 4$ was identified through manual trial and error experiments. This value is specific to the particular catchment, event, and experimental setup used in this thesis, and therefore might need to be optimized for case specific implementations. In fact, from the values obtained here it appears that the optimum value of ϑ depends on the spatial coverage of the observation. This also makes sense intuitively as any spatial matching metric loses sensitivity as the area under evaluation increases. However, the value of $\vartheta = 4$ was found to produce positive impacts in most scenarios, as a general good fit parameter for most cases examined in this thesis. Accordingly, the final weight (w_i) for each particle at a given assimilation time step was calculated using

$$w_i = \frac{\left(Ag_i - \min_i Ag\right)}{\left(\max_i Ag - \min_i Ag\right)} \times \left(\frac{Ag_i}{\max_i Ag}\right)^{\vartheta} . \tag{6-10}$$

The use of a weight rescaling parameter is almost standard practice for particle filter applications and is known as tempering (Woodhead 2007; Fearnhead and Künsch 2017; Herbst and Schorfheide 2019). The ideal proposal distribution from which prior weights are sampled, must be different from the posterior to allow effective estimation (Godsill and Clapp 2001). However, if this difference is too large then the importance weights will be close to zero for frequent outcomes and close to one for rare values (Johansen 2015). Consequently, the state estimates obtained will be dominated by a small subset of the Markov chain (Woodhead 2007). Moreover, a full representation of the state-parameter space for a state vector size common to 2D flood inundation modelling, ideally requires a prohibitively large number of runs, impossible to consider in practice due to computational demands. Tempering allows the

inflation of weights such that the full state-parameter space can be spanned using a smaller number of model runs (Herbst and Schorfheide 2019). Finally, the global weights are normalized to calculate the conditional probabilities of a given particle being true given the observations, and to ensure that the posterior probability distribution function sums to unity according to

$$W_i = \frac{w_i}{\sum_{i=1}^{P} w_i}, \tag{6-11}$$

where P is the number of particles and W_i denotes the final global weight assigned to a particle for a given observation. Particle weights evolve over time, as new observations become available. When all the observations have been processed and their information content extracted in the form of the importance weights, the expectation of any state variable (S) gives the assimilated forecast. This can be computed as the weighted mean of the state vector (s_i) ensemble given by

$$E[S] = \sum_{i=1}^{P} W_i \times s_i \, . \tag{6-12}$$

This formula is used to compute the expectation of streamflow, flow velocities, and water depth. The expectation of water level and flood extent is derived from the expectation of water depth (Hostache et al. 2018b).

6.6 Chapter Summary

This chapter presented an overview of the data assimilation framework developed in this thesis for flood extent assimilation into hydraulic models. First, the concept of data assimilation was briefly introduced and subsequently contextualized based on the different types of uncertainties common to hydrological and hydraulic models. Second, the classification of data assimilation techniques from a hydrological systems perspective was described and an overview of current techniques presented. This section also justified the selection of a particle filter based framework for the flood extent assimilation in this thesis. Finally, the flood extent assimilation technique developed in this study was presented in detail. A novel cost function was formulated in this thesis based on the concept of mutual information, which allows the comparison of observed probabilistic flood maps with deterministic model simulated binary flood extents. Appendix A provides additional details of the experiments conducted in order to arrive at this metric, and the various different cost functions trialed before this. Weight

tempering was also introduced to overcome particle degeneracy that is common to sequential importance sampling implementations of the particle filter, through the use of an exponential rescaling factor. The catchment specific value for the exponent was fixed through numerical experiments to ensure optimal weight variability for efficient assimilation. The efficacy of the flood extent assimilation framework proposed in this chapter, is tested in Chapter 7 through synthetic and real world experiments, while the sensitivity to observation spatiotemporal characteristics is evaluated in Chapter 8 through synthetic experiments.

CHAPTER SEVEN

"A very small cause which escapes our notice determines a considerable effect that we cannot fail to see, and then we say that the effect is due to chance. If we knew exactly the laws of nature and the situation of the universe at the initial moment, we could predict exactly the situation of that same universe at a succeeding moment."

- Henri Poincaré, In *'Chance', Science et Méthode* (1908), Quoted in Richard Kautz, *Chaos: The Science of Predictable Random Motion* (2011), 167 as translated in *Science and Method* by F. Maitla.

7. Flood Extent Assimilation: Synthetic Study

The PF-based flood extent assimilation (FEA) scheme proposed in Chapter 6, which utilized an objective function based on mutual information, is evaluated in this chapter through synthetic experiments. As the primary objective of this thesis was to develop methods and evaluate the DA performance for a realistic operational scenario, a real flood event with real inflow measurements were used as the control for the synthetic experimental setup. However, this was undertaken in an identical twin experiment framework, where synthetic data were generated from the flood extents produced by the control run and subsequently assimilated using the same model, to assess filter behaviour and assimilation performance. Assimilation efficiency was quantified in terms of the simulated inundation extent, floodplain water depth, floodplain flow velocity, flood hazard estimates, and channel flow and depth in this chapter through synthetic experiments. The observations were considered together by multiplying weights forward in time and at the times at which actual SAR images were available to this study, to allow for comparisons across the synthetic experiment outlined in this chapter and the real world application of Chapter 9. As the synthetic and the real world experimental setups differed only in the spatial coverage of the observations, which for the synthetic case was the entire catchment while for the real case a small portion in the centre of the catchment, an assessment of the impacts of spatial coverage was also facilitated. The sensitivity of this algorithm to the observation spatial and temporal location with respect to the catchment morphology and the flood hydrograph are evaluated in Chapter 8.

7.1 Introduction

Unlike hydrological models, hydraulic models have only two diagnostic state variables: water depths and flow velocities, with information about flood extent and discharge being prognostic variables derived from these (Lai et al. 2014). This indeed makes the assimilation of

water levels significantly simpler than the assimilation of flood extents (Hostache et al. 2018b). Therefore, as evident from the comprehensive review of hydraulic data assimilation literature presented in Chapter 2, most studies have focused on assimilating synthetic (Garambois et al. 2019; Tuozzolo et al. 2019), in situ (Van Wesemael et al. 2019; Ziliani et al. 2019), or remote sensing-derived water levels (Lai and Monnier 2009; Giustarini et al. 2012b), rather than flood extents which are directly observed by satellites. However, a major limitation of water levels derived from remote sensing images is the dependence on the digital elevation model (DEM) accuracy and resolution (Schumann et al. 2008), which impedes the application of such approaches to data scarce regions. Even though the accuracy of Global DEMs is starting to improve (Gallant et al. 2011; Yamazaki et al. 2017), the spatial resolution is often insufficient for flood modelling at local scales (Grimaldi et al. 2018). Moreover, Hostache et al. (2018), argue that spatial flood information is lost during the interpretation of remote sensing-derived water levels, as only shoreline water heights can be reliably derived from the integration of flood extents and DEMs (Mason et al. 2012b, 2016; Shastry and Durand 2019). Furthermore, as the derivation of water levels from remotely sensed data is neither straightforward nor automatic, incorporating such assimilation frameworks into operational forecasting systems remains challenging (Hostache et al. 2009).

Recent studies have therefore focused on the development of techniques capable of directly assimilating flood extents into flood forecasting model cascades, rather than water levels (Lai et al. 2014; Revilla-Romero et al. 2015, 2016; Hostache et al. 2018b; Shastry and Durand 2019). As flood extents are derived based on simulated water depth, the variation in the number of wet-dry cells at most time steps is limited. Consequently, the development of an extent based cost function with enough sensitivity to isolate the best performing ensemble members and drive the assimilation has been a scientific challenge (Schumann et al. 2009a; Lai et al. 2014). Cooper et al. (2019) suggested converting modelled binary flood extents into synthetic SAR observations, based on the image statistics derived from flooded SAR images, then comparing simulated and observed backscatter values at the first dry and last wet model cell. The authors argue that using all the backscatter observations could potentially lead to overfitting the observation, especially for small ensemble sizes typical of hydraulic modelling studies. However, even though the backscatter operator improved the forecast for the idealized test domain used in the study, a substantial amount of useful information is discarded. Moreover, this technique also neglects the uncertainties in interpreting flooding based on backscatter alone, as Chapter 4 of this thesis as well as many prior studies have demonstrated

(Amitrano et al. 2016; D'Addabbo et al. 2016). Furthermore, the proposed operator also needs to be tested for a real case, in addition to the synthetic application currently illustrated, such that the applicability to real world situations can be assessed.

The work of Hostache *et al.* (2018) used several pragmatic mathematical solutions to facilitate the direct assimilation of flood extents in a real case for the first time. For example, local weights were first computed on a pixel-by-pixel basis for each ensemble run or particle, by comparing the modelled vs. observed cell wet-dry status through the binomial distribution. The joint probability density of these pixel-wise weights is then used as the global weight of the particle, which is aggregated as a product of all pixel weights. This method of calculating the joint probability density makes the assumption that the local pixel-wise weights are statistically independent (Hastie et al. 2009). As raster-based models numerically compute flow between cells, the values of simulated flow depths used to derive the wet-dry cell status cannot be considered independent. Moreover, flooding itself is a spatial phenomenon implying interdependence between modelled grid cells. Another limitation was that the observed flood probability values had to be truncated between 0.001 and 0.999, to ensure that the global weights computed as a product, did not decay to zero due to a complete mismatch at a few pixels. Moreover, the dry pixels were masked out prior to the assimilation, in order to remove the effects of asymmetric penalties which result from limited coverage of the inundated area within the domain. This study thus developed a novel flood extent assimilation framework to optimally combine flood inundation models with EO-flood extents by overcoming these limitations.

The mutual information-based (MIB) cost function used for the assimilation, provides a measure of the reduction in uncertainty about one random variable, given the complete knowledge of another. This makes the statistic sensitive to subtle changes in the inundation extent, and also uniquely suitable to the context of particle filters, rooted in the concepts of Bayesian conditional probabilities. Moreover, this process results in global weight values for each particle, eliminating the assumption of spatial independence required for the approach of Hostache et al. (2018a). The need for truncation of observed flood probabilities as well as the masking of dry pixels is also eliminated, reducing the number of pre-processing steps and making the assimilation process more straightforward. A tempering factor was also included to further enhance weight variability, as extent comparison metrics exhibit limited sensitivities when all models over predict (Schumann et al., 2008). Tempering techniques were subsequently

used to inflate the MIB-RMSE based weights artificially, to help the selection of a few well performing particles (Herbst and Schorfheide 2019). This step would not be necessary if a very large number of ensemble members (tending towards infinity) were used, covering the entire state-parameter space through generous sampling. However, as this is currently impossible for any operational hydraulic model implementation due to the computational expense, it was vital to introduce methods to counter the problem of under-representation of the combined error covariance of the modelling system. The next sections describe the experimental methods and results in greater detail.

7.2 Methods

This section outlines the methods used in this study, with Fig. 7.1 illustrating the overall workflow followed. The next sections describe the experimental design, ensemble generation,

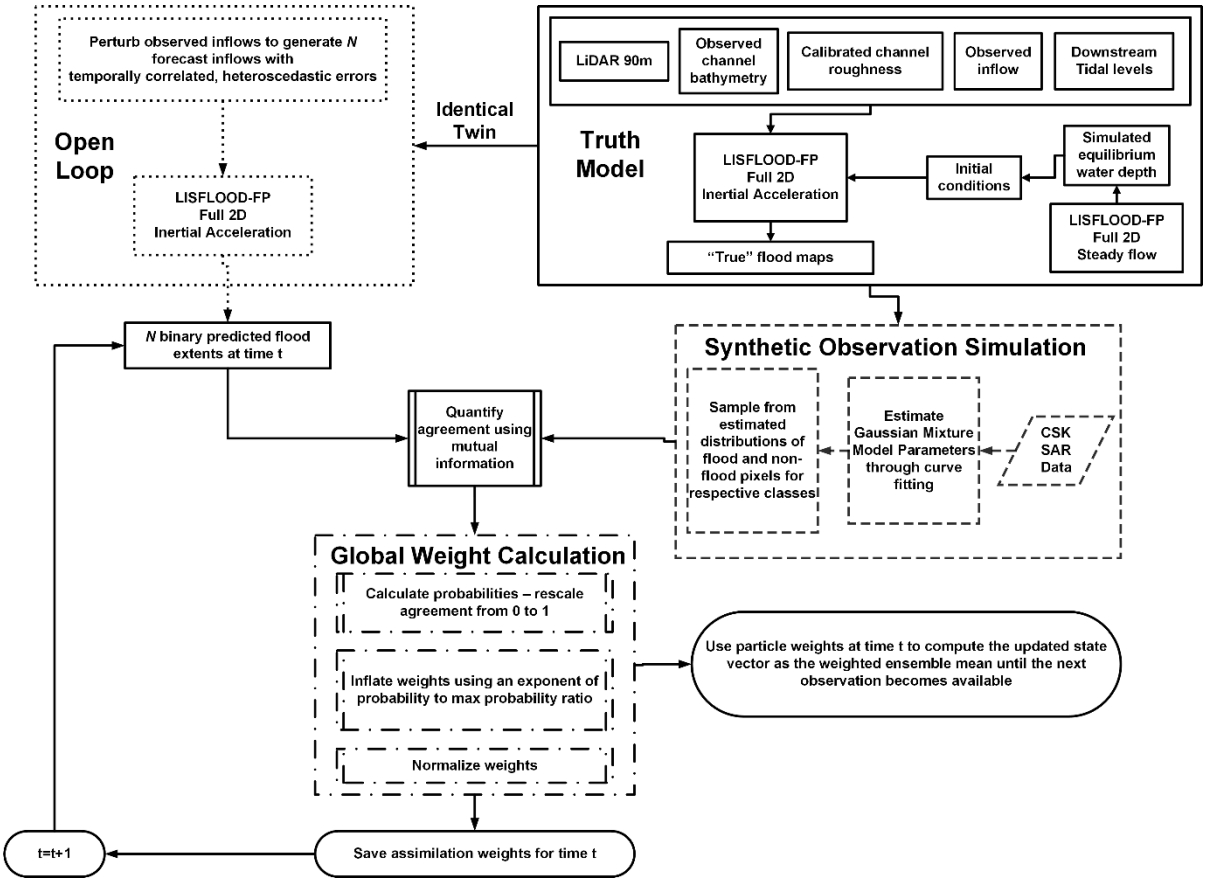

Fig. 7.1 Schematic of the synthetic assimilation experiment using an identical twin setup, where synthetic data were generated from the flood extents produced by the truth run and subsequently assimilated within the same model.

synthetic observation simulation, and performance metrics used to evaluate assimilation performance.

7.2.1 Experimental Design

Identical twin experiments common to data assimilation studies, were setup for the 2011 flood event in the Clarence Catchment, to assess the performance of the flood extent assimilation procedures outlined in Chapter 6. Uncertainties generated through incorrectly specified inflows were independently considered, to better understand the impact of flood extent assimilation. Observed inflows were used to initialise the "truth" model simulation, along with the calibrated channel friction parameter. Observed tidal levels were applied at the downstream boundary, and floodplain elevations were specified using the LiDAR DEM and the observed channel bathymetry. Considering the observed inflows as the truth, forecast inflow uncertainties were synthetically generated, representing outputs generated by hydrological models running in forecast mode (García-Pintado et al. 2013). The model implementation used in this chapter was described in detail in Chapter 5 of this thesis, with the simulation of the inflow ensembles presented in the following section.

Flood extents simulated by the "truth" model were used to generate probabilistic satellite-based flood observations for the synthetic experiments, corresponding to the actual SAR image acquisition timings. The observation simulation technique is described in subsequent sections. Assimilation results were evaluated against the benchmark "truth" model in a distributed fashion for the simulated floodplain states, and at point gauge locations for the channel. For the real world application of the proposed assimilation technique, the actual SAR-based observations of flood extent derived using the methods presented in Chapter 4 were used. Gauge observations were used for channel performance evaluation, while aerial photographs were utilised to quantify the accuracy of simulated flood extents after the assimilation.

7.2.2 Ensemble Generation

The success of any assimilation experiment is driven by the perturbation technique chosen and the sensibility of the ensemble spread (De Lannoy et al. 2006). In this study, the upstream boundary condition, supplied in the form of hourly observed discharge values, was assumed to be the only source of uncertainty. In order to emulate discharge forecasts produced by hydrological models, temporally correlated, heteroscedastic errors were synthetically

generated with a positive mean bias (Gobeyn et al. 2017). As the uncertainty of the rainfall forecasts from weather prediction models is propagated down the modelling chain, the nature of hydrological streamflow forecast errors is often rather complex. Consequently, the error model used to generate these forecast inflow errors is described next.

The forecast discharge errors were simulated by considering a positive multiplicative mean bias of 20% (Matgen et al. 2010). Further, a temporal correlation and time-varying variance is also imposed on the errors (García-Pintado et al. 2013). The temporal evolution was modelled according to (Evensen, 2003)

$$q_k = \alpha_t q_{k-1} + \sqrt{1-{\alpha_t}^2} w_{k-1}\,, \tag{7-1}$$

where $w_k \in \mathcal{N}(0,1)$ is white noise, $\alpha_t \in [0,1)$ is a time decorrelation factor, k is the simulation time, and the white noise component of the model errors is represented by $q_k \in \mathcal{N}(0,0.15)$ (García-Pintado et al. 2015). The time decorrelation of the stochastic forcing variable is controlled by the factor α_t, where low values generate a series nearly white in time, and higher values lead to highly correlated errors. This implies that α_t is determined by the decay time of the temporal correlation (τ) and the time interval between subsequent flow observations (Δt), computed according to

$$\alpha_t = 1 - \frac{\Delta t}{\tau}\,. \tag{7-2}$$

As operational global flood forecasting systems currently provide discharge forecasts at a daily time step, a decay time of 3 days was assumed following the work of García-Pintado et al. (2013). Finally, the forecast discharge time series was calculated based on,

$$Q_k = Q_{truth} + \sqrt{\Delta t}\sigma\rho q_k + 0.2 * Q_{truth}\,, \tag{7-3}$$

where σ is the variance of model errors and ρ is a variance growth scaling factor.

The use of a scaling factor ensures that variance growth becomes independent of α and Δt over the simulation time period (Evensen 2003), computed through,

$$\rho = \sqrt{\frac{1}{\Delta t}\frac{(1-{\alpha_t}^2)}{n-2\alpha_t-n{\alpha_t}^2+2{\alpha_t}^{n+1}}}\,. \tag{7-4}$$

In order to implement heteroscedastic errors the value of σ in Eq. 7-3 was replaced by the time varying σ_k using,

$$\sigma_k = c.v.* Q_k\,, \tag{7-5}$$

where c.v. is the coefficient of variation, the value for which was estimated from the historical rating curves at Grafton, Prince Street gauge (Farr and Huxley 2013). As the average value of c.v. was observed to be close to ~0.2, a slightly higher value (c.v.=0.25) was used as recommended by García-Pintado et al. (2013). For a more detailed description of the inflow ensemble generation techniques described here and for the derivation of the equations used, readers are referred to Evensen (2003) and García-Pintado et al. (2013).

An ensemble of 128 simulations was considered for each of the cases examined in this study, which is the most commonly used ensemble size for particle filters in hydraulic data assimilation literature and represent an optimum computational burden versus accuracy trade-off (e.g., Hostache et al., 2018a, 2010). While the ensemble size is rather small with respect to the domain size and the size of the state vector, given the computational requirements of each run this was considered a sufficient number (Hostache et al. 2018b). Even though studies show that increasing the ensemble size may result in improved assimilation performance, there is no guarantee that this will occur (Ziliani et al. 2019). In fact for a state vector of the size of the one in this thesis (~63,500 wet cells), the ensembles would need to exceed the length of the state vector at least by a few orders of magnitude (>100,000 at least) (Banister and Nichols 2012). This was not attempted as it was computationally infeasible and slight increases in ensemble size (e.g., 256 or 528) have been shown to result in only marginal performance improvements for particle filters (Plaza Guingla et al. 2013). Moreover, as the objective of this thesis was to prepare flood extent assimilation methods for operational applications, increasing the ensemble members arbitrarily was not considered a viable option.

Even though larger ensemble sizes are possible within a research environment, their operational implementation remains significantly challenging. As an example, consider the Global Flood Awareness System or GloFAS, the only global flood forecasting system currently in existence hosted by the European Centre for Medium Range Weather Forecasting, which uses an ensemble of 51 hydrological model runs (Alfieri et al. 2013; Dottori et al. 2016; Hirpa et al. 2018b). In contrast, hydrological data assimilation studies have considered even up to a million ensemble members as the main objective is towards research (e.g., Moradkhani et al., 2005b), while in an operational scenario the trade-off between computational time and accuracy becomes critical. Accordingly, the experiments in this thesis were executed with the current best estimate for particle filters in literature of 128 model runs (Plaza Guingla et al. 2013).

As mentioned earlier, the success of any synthetic data assimilation experiment is largely dependent on the skill of the generated ensembles and the sensibility of the ensemble spread (Nester et al. 2012) Ensemble skill is in turn a function of the ensemble size and the perturbation technique chosen. Accordingly, the ensemble generation parameters were tuned to ensure sensible spread as suggested by Moradkhani et al. (2005), especially as the ensemble size could not be increased further due to operational applicability considerations. Forecast skill was evaluated using the Normalized RMSE Ratio (NRR) proposed by Moradkhani et al. (2005), where an NRR value of ~1 is ideal, while values of $NRR>>1$ indicate too little spread and $NRR<<1$ represents too much spread (Matgen et al. 2010). NRR is estimated according to,

$$NRR = \frac{Ra}{E[Ra]}, \tag{7-6}$$

where Ra is the ratio of the time-averaged RMSE of the ensemble mean, and the mean RMSE of the ensemble members. If the observation is statistically indistinguishable from the ensemble, the expectation of Ra is given by,

$$E[Ra] = \sqrt{\frac{(n+1)}{2n}}, \tag{7-7}$$

where n is the ensemble size considered. For an in-depth description of the forecast evaluation methods used in this study, readers are referred to Moradkhani et al. (2005) and Matgen et al. (2010). The value of NRR for the forecast inflow ensemble was ~0.99 which is almost equal to the ideal value of unity, indicating sufficient ensemble spread and skill (Moradkhani et al. 2005).

7.2.3 Synthetic Satellite Observation Simulation

In order to assimilate a SAR-derived flood extent, a probabilistic map objectively representing the sum of uncertainties from instrument error, flood classification, and orthorectification errors was necessary. Therefore, the neuro-fuzzy flood mapping algorithm developed in Chapter 4 was used to derive probabilistic flood maps from the Cosmo-Skymed SAR images. Synthetic observations were also simulated at the corresponding time steps, using an approach similar to Cooper et al. (2019). Backscatter distributions of flood and non-flood classes were assumed to follow the form of Gaussian Mixture Models (GMM). Parameters for each Gaussian curve in the observed SAR image histograms, were estimated using non-linear curve fitting. Values were identified using both SAR images available to this thesis and

averaged to incorporate information from both images. The parameters finally used for the synthetic observation simulation are summarized in Table 7.1.

The water depth rasters simulated by the truth model at the assimilation time steps were then converted to binary flood extents using a minimum water depth threshold of 1cm. For each pixel in the modelled flood and non-flood classes, a backscatter value was sampled from the corresponding normal distribution defined based on the observed mean and standard deviation, to generate a synthetic SAR image corresponding to the actual SAR image. The synthetic and real SAR images for each assimilation time step are illustrated in Fig. 7.2. Subsequently, the synthetic SAR images were translated into probabilistic flood maps using the Bayesian approach proposed by Giustarini et al. (2016). The pixel-wise conditional probabilities of flooding given a particular backscatter value $p(F|\sigma^0)$ were calculated according to

$$p(F|\sigma^0) = \frac{p(\sigma^0|F)p(\bar{F})}{p(\sigma^0)} \tag{7-8}$$

$$p(\sigma^0) = p(\sigma^0|F)p(\bar{F}) + p(\sigma^0|NF)p(\overline{NF}) \tag{7-9}$$

$$p(F|\sigma^0) = \frac{\left(\frac{1}{\sqrt{2\pi}s_{\sigma^0(F)}} exp\left[-\frac{1}{2}\frac{\left(\sigma^0-\mu_{\sigma^0(F)}\right)^2}{s_{\sigma^0(F)}{}^2}\right]\right)p(\bar{F})}{\left(\alpha_F\, exp\left[-\frac{1}{2}\frac{\left(\sigma^0-\mu_{\sigma^0(F)}\right)^2}{s_{\sigma^0(F)}{}^2}\right]p(\bar{F}) + \alpha_{NF}\, exp\left[-\frac{1}{2}\frac{\left(\sigma^0-\mu_{\sigma^0(NF)}\right)^2}{s_{\sigma^0(NF)}{}^2}\right]p(\overline{NF})\right)}, \tag{7-10}$$

where $p(\sigma^0|F)$ is the conditional probability of observing a specific backscatter value given that the pixel is flooded, and $p(\sigma^0)$ is the marginal probability distribution of backscatter values at a given pixel. The flood and non-flood classes are denoted by *F* and *NF* respectively and subscripted symbols represent the class specific Gaussian mixture model parameters defined in Table 7.1. The terms $p(\bar{F})$ and $p(\overline{NF})$ denote the prior probabilities of observing each class, which were assumed to be 0.5 as recommended by the authors in the absence of any ancillary information (Schlaffer et al. 2017).

As the flood mapping algorithm developed in Chapter 4 of this thesis was specifically designed to reduce uncertainties in flood classification from a single SAR image in operational settings, it was not used to process the synthetic images. The objective here was to just generate flood probabilities from synthetic backscatter, for which the optimized texture based enhancement procedure was deemed unnecessary. As the requirement of observed flood

Table 7.1 Summary table of the backscatter distribution parameters used in this thesis, estimated from 8-bit Cosmo-Skymed SAR flood observations.

Class	*Mixing proportion* (α_F/α_{NF})	*Mean* (μ)	*Standard deviation* (s)
Flood (F)	781.512	39.495	6.491
Non-flood (NF)	1391.190	125.994	29.458

probabilities for the synthetic assimilation experiment was met through the simple approach of Giustarini et al. (2016), it was considered adequate and thus there was no need to improve upon some benchmark flood mapping accuracy.

7.2.4 Performance Metrics

Model performance metrics are defined as mathematical measures designed to quantify the fit between simulated behaviour and real world observations (Krause and Boyle 2005). The nature of the observed data can be considered as one of the key determinant in designing a

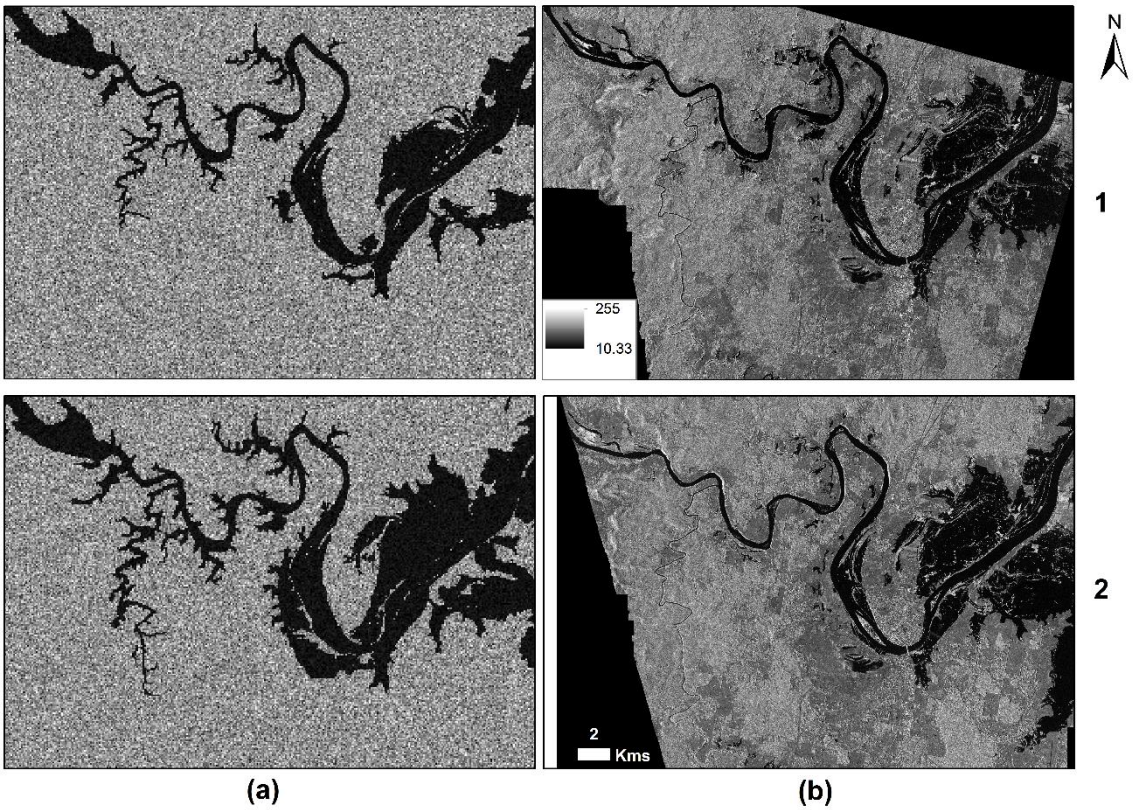

Fig. 7.2 Synthetic and real SAR images juxtaposed in columns (a) and (b), respectively, for assimilation time steps 1 and 2.

suitable model performance evaluation strategy (Beven 2012). The choice of an efficiency criterion can significantly impact the final performance ranking of model realizations (Stephens et al. 2014). As different likelihood measures prioritize different behavioural errors, it follows that the choice of a particular criterion should then be guided by the downstream purpose of the modelling exercise (Pappenberger et al. 2007). However, due to unavoidable subjectivity in the choice of an appropriate performance statistic, hydrodynamic model evaluation should ideally be treated as a multi-objective optimization problem (Gupta and Kling 2011). Note that the results were interpreted by comparing objective function values obtained by evaluating the open loop and the assimilated ensemble, against the truth model.

In this study, the focus is on improving operational flood forecasts through flood extent assimilation, and performance metrics were carefully selected to reflect model accuracy in simulating relevant inundation dynamics correctly. For the synthetic case, the impact on simulated flood extents was quantified through contingency maps and contingency matrix based statistics. Simulated water levels were evaluated spatially through water depth difference maps, and a spatial mean Root Mean Squared Error (RMSE) statistic was used to quantify overall performance. The spatially distributed impact of the assimilation on x- and y-direction flow velocities was then evaluated through flow velocity difference maps, while the percentage improvement in RMSE was used to compare the relative impact of the two assimilated images. The relative performance of the assimilation with respect to the open loop was examined through the Brier Skill Score (BSS), using hydrographs from the truth model at river gauging sites. Absolute errors in the channel performance were evaluated using RMSE and hydrograph plots. Finally, the impact on inflow uncertainties was quantified through hydrograph comparisons.

The flood extents were evaluated using the Critical Success Index (CSI) and the Kappa statistic, both of which were calculated based on the contingency matrix illustrated in Table 7.2. Choices of the metrics were guided by sensitivity to accurate flooded area prediction and comparability across literature. The following equations were used to calculate

$$CSI = \frac{A}{A+B+C}, \tag{7-11}$$

$$Kappa = \frac{P_o - P_e}{1 - P_e}, \tag{7-12}$$

where P_o and P_e refer to observed and expected agreement, respectively, and the meanings of A, B, C, and D are as per the contingency table given in Table 7.2.

$$P_e = \frac{((A+B)*(A+C)+(C+D)*(B+D))}{(A+B+C+D)}, \quad (7\text{-}13)$$

$$P_o = A + D, \quad (7\text{-}14)$$

As evident from Equations 7-11, the CSI measure ignores the correctly simulated non-flooded areas to eliminate the disproportionate impacts of the typically larger non-flooded areas in the model domain on the value of the metric. Similarly, the Kappa statistic allows for the evaluation of flood simulation capabilities, independent of the impacts of random chance agreements between the modelled and observed image classes. Spatial comparisons were also included through the contingency maps, which illustrate the locations of the classes in Table 7.2. For the synthetic case, these comparisons were made against flood extents derived from the truth model, while aerial photographs were used to evaluate simulated flood extents for the real world application.

For the modelled water depth assessment, a global measure of error was required to quantify the overall model performance in absolute terms. The RMSE statistic was preferred for this purpose, due to its ubiquity in hydraulic data assimilation literature. Therefore, it can facilitate benchmarking vis-à-vis other studies (García-Pintado et al. 2013, 2015; Cooper et al. 2018). In the context of the present research, the RMSE was calculated between the weighted ensemble mean $E[WD_k]$, and the water depths simulated by the truth model WD_k^{truth}, at select time steps according to

$$WD_{RMSE} = \sqrt{\frac{\sum_{i=1}^{P}\left(E[WD_k]-WD_k^{truth}\right)^2}{P}}, \quad (7\text{-}15)$$

Table 7.2 Contingency matrix used for the calculation of binary pattern matching based flood extent performance measures. The green colour specifies where the model and observation are in agreement, while the red and blue, refer to under and overprediction respectively.

Contingency Matrix		**Modelled**	
		Flooded	Non-flooded
Observed	Flooded	**A) Correct Flooded (Hits)**	**C) Underprediction (Misses)**
	Non-flooded	**B) Overprediction (False Alarms)**	**D) Correct Non-flooded (Correct Rejects)**

where P is the total number of model grid cells. Similarly, water depth difference (WDD) maps were chosen to understand the spatial patterns in forecast error. As Cooper *et al.* (2018) show, these can provide valuable insights into spatiotemporal error propagation through the model domain, which lumped statistics like the RMSE are unable to detect. For each pixel, the WDD map contains the deviation of the ensemble mean from the true WD at the corresponding location, given by

$$WDD_k = E[WD_k] - WD_k^{truth}. \tag{7-16}$$

Flow velocities were evaluated in the exact same way as water depths. Both absolute errors and their spatial distribution were considered, through RMSE and flow velocity difference maps (FVD). Metrics were computed for the x- and y-directions, by changing the variable from water depth to flow velocity in Equations 7-15 and 7-16. These assessments were only possible for the synthetic case, as spatially distributed observations of water depth and flow velocity were unavailable.

The model performance within the channel was evaluated using hydrometric gauges. River gauges recording water depth and discharge were considered at the locations where real gauges along the Clarence River already exist. Synthetic gauges were also added at a few locations to enable a more comprehensive gauge based evaluation. The relative skill of the assimilated ensemble with respect to the open loop was quantified through the Brier Skill Score (BSS), calculated according to

$$BSS = 1 - \frac{\overline{(\boldsymbol{Assim.} - \boldsymbol{Truth})^2}}{\overline{(\boldsymbol{OL} - \boldsymbol{Truth})^2}}, \tag{7-17}$$

where the variables in bold denote state vectors while the overline denotes an average. Values of $BSS \in (-\infty, 1]$, where $BSS = 0$ indicates no change in forecast skill with respect to the open loop while $BSS = 1$ is the ideal score. RMSE and Kling Gupta Efficiency plots were also considered for the gauge evaluation, to quantify the absolute errors in addition to a comparison of the truth, open loop, and assimilated water level and discharge hydrographs. The RMSE metric allowed for a simple lumped measure of the absolute error in the forecasts, while the KGE was chosen for a more holistic evaluation of the forecast hydrograph. The KGE metric was calculated according to

$$KGE = 1 - ED, \tag{7-18}$$

$$ED = \sqrt{(r-1)^2 + (\alpha_r - 1)^2 + (\beta - 1)^2}, \tag{7-19}$$

where ED is the Euclidean Distance from the ideal point in objective space, $r = \alpha_r = \beta = 1$, and r is the Pearson's correlation coefficient between observed and simulated hydrometric time series, while α_r is a measure of relative variability in the simulated and observed values using a ratio of the standard deviations, and β is the ratio of the mean of the simulated and observed values. KGE quantifies the errors in the simulated versus observed flows from a multi-objective perspective, by simultaneously focusing on the correlation, variability error, and bias error as separate criteria to be optimised (Gupta et al. 2009). Each of these criteria are equally important from a hydrological perspective, as α and β relate to the ability of the model to reproduce the first two moments of the observation distribution, while r relates to the model's ability to reproduce the timing and shape (of the hydrograph, for example). The KGE metric was preferred over the more popular Nash Sutcliffe Efficiency (NSE), as it is known to be biased towards models which underestimate the variability in flows (Gupta et al. 2009). The various metrics used to assess the model performance in terms of the different simulated flood variables are summarized in Table 7.3.

Table 7.3 Summary table of the performance evaluation metrics used in this thesis to assess the accuracy of different flood variables.

Modelled Variable	**Evaluation Metrics**
Flood extent	Critical Success Index, Cohen's Kappa
Floodplain Water Depth	Root Mean Squared Errors, Water Depth Difference Maps
Floodplain Flow Velocities	Root Mean Squared Errors, Flow Velocity Difference Maps
Flood Hazard ($D \times V$)	Root Mean Squared Errors, Hazard Difference Maps
Channel Discharge and Water Levels	Root Mean Squared Errors, Kling Gupta Efficiency, Brier Skill Scores

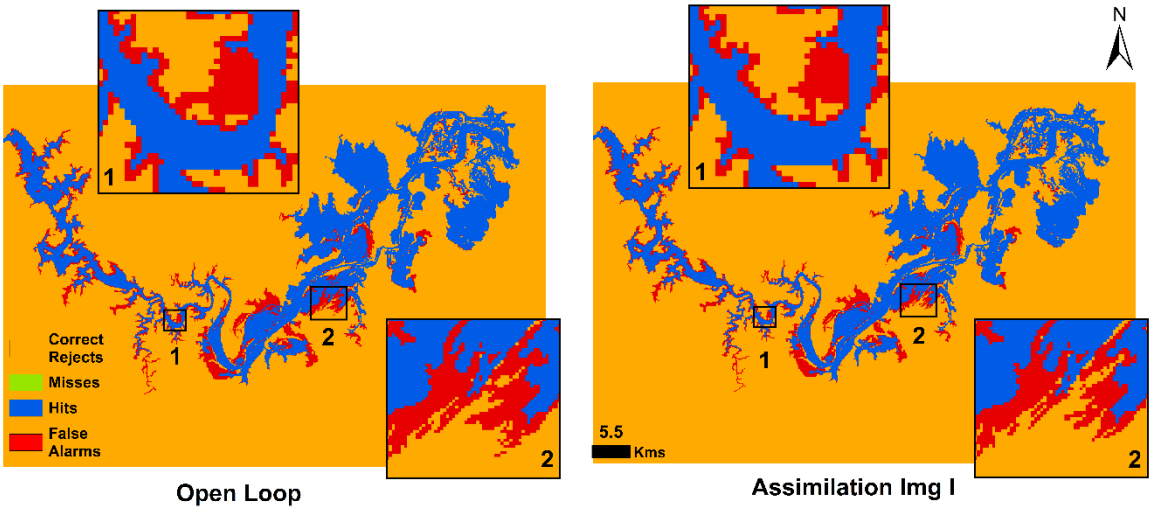

Fig. 7.3 Contingency maps comparing the forecast versus true flood extents, for the open loop and assimilation runs, at the first assimilation time step.

7.3 Results and Discussion

The impacts of assimilating synthetic SAR observations on the open loop ensemble generated by propagating erroneous simulated forecast inflow errors through Lisflood-FP are summarized here. The key difference between the synthetic and the real-world experiments implemented in this thesis was the spatial coverage of the observations, which in the synthetic case extended to the entire catchment. The timing of the synthetic images was assumed to be

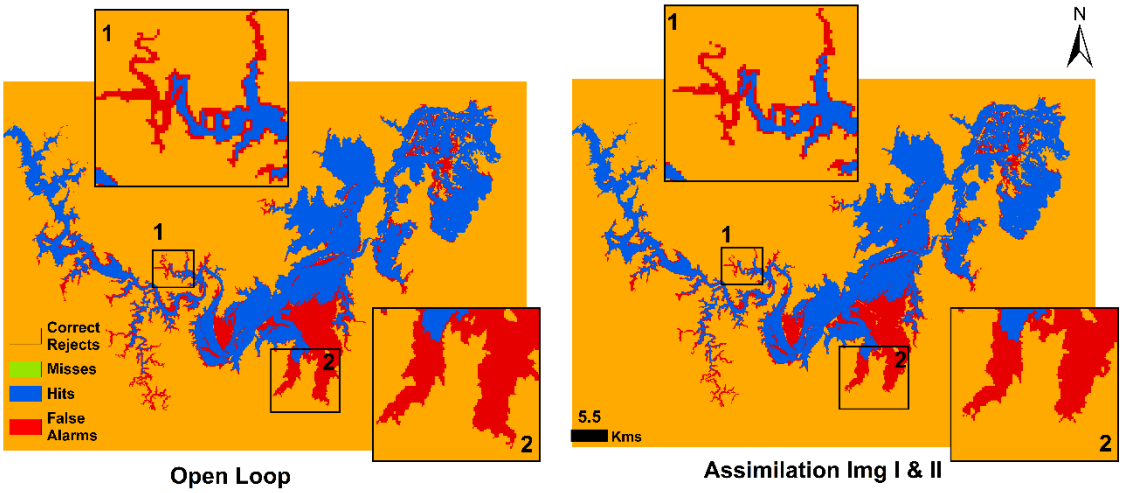

Fig. 7.4 As for Fig. 7.3 but at the second assimilation time step after considering both images together.

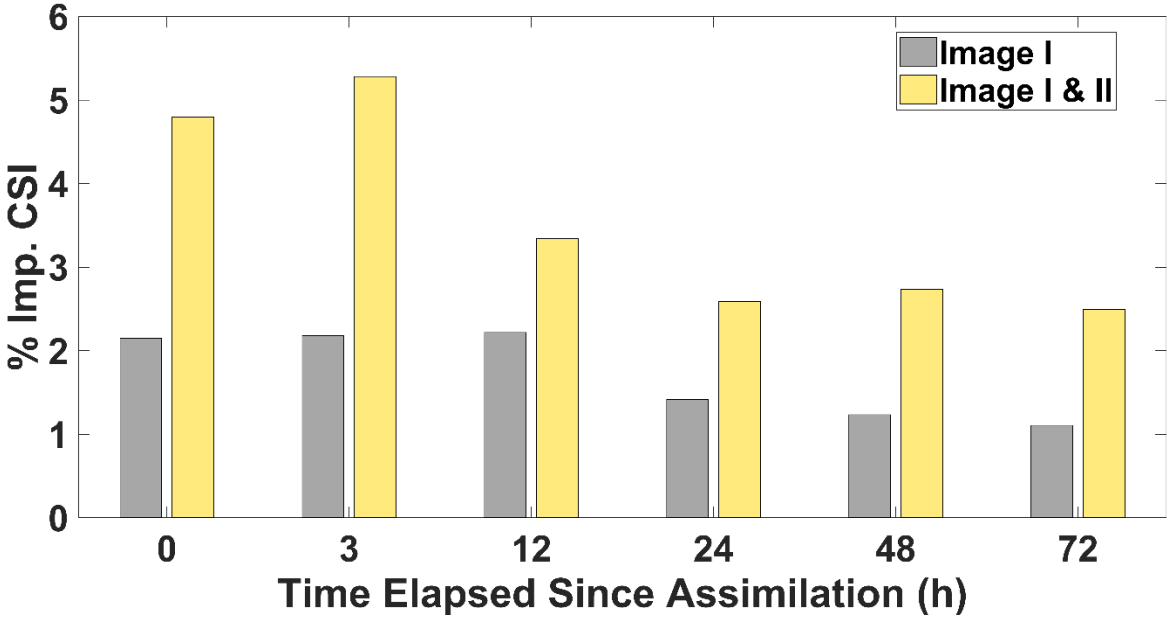

Fig. 7.5 Percentage improvement in Critical Success Index before and after the assimilation, for different lead times from the time of assimilation.

identical to the actual SAR acquisitions. This chapter allowed an examination of the impacts of spatial coverage of the observation. Forecast performance was first evaluated in terms of flood extent, followed by a spatial analysis of water depth errors. The distributed impacts on floodplain flow velocity and flood hazard ($D \times V$), were then assessed by this study for the first time in assimilation literature. Finally, the impacts on in-channel performance were quantified in terms of discharge and water level estimation capabilities.

7.3.1 Impact on Simulated Inundation Extent

Contingency maps were used to quantify the differences between the true flood extent versus the open loop and assimilated forecasts at the assimilation time steps. Maps resulting from assimilating the first image are displayed in Fig. 7.3, while Fig. 7.4 shows the results from assimilating both images together with forward weight multiplication. It is important to note that all references to the assimilation of the second image in this Chapter inherently include information from the first image as well, according to the multiple image assimilation design where weights are multiplied forward.

From the figures, it is evident that the open loop consistently overestimated the true inundated area and there were no missed pixels, due to the positively biased inflow errors used

in this experiment. This overprediction was expected, as there was a twenty percent multiplicative bias imposed on the model domain through the inflow error model used to generate the open loop ensemble. The closed loop, or otherwise known as the assimilated ensembles, exhibited lower numbers of false alarm pixels as a consequence of the assimilation.

The assimilation of synthetic "observed" flood extents generated from the true flood extents should pull the simulation towards the truth, with more observations helping to constrain the forecast and resulting in fewer false alarms. The largest reductions in overprediction (Boxes 1 and 2 in both contingency maps) seem to be centred within the catchment in this study. A limited number of pixels changed their wet-dry status in the topographically constrained upstream part of the reach, between Lilydale to Rogan's Bridge (See Fig. 7.8 for locations). This was in line with expectations as the assimilation of extent can only have limited impact in regions where flood extents are not sensitive to subtle changes in the water depth.

Similarly, the assimilation was not expected to produce large impacts near the downstream boundary between Brushgrove to Yamba, where the inundation is dominated by tidal backwater effects which are identical for the truth and the open loop ensembles used in this experiment. The main shift observed in pixels after the assimilation was from the false alarm class to correct rejects, while the number of hits remained fairly consistent. This is not unusual as the open loop constantly overestimated the extent with most of the true extent being correctly simulated and the performance of the ensemble members differing primarily in terms

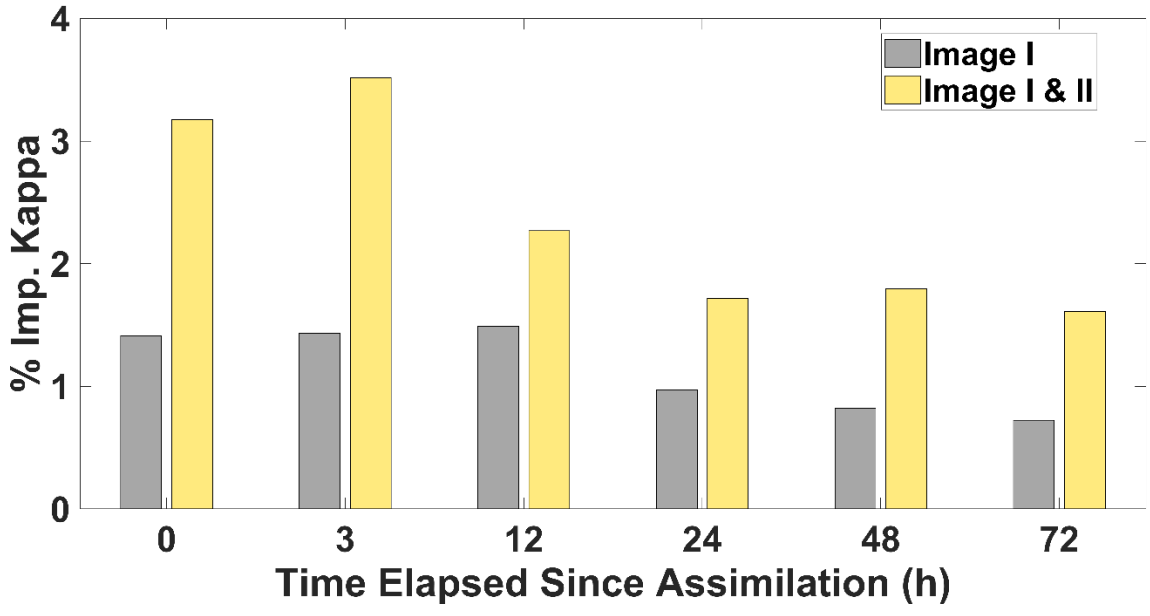

Fig. 7.6 As for Fig. 7.5 but for Cohen's Kappa.

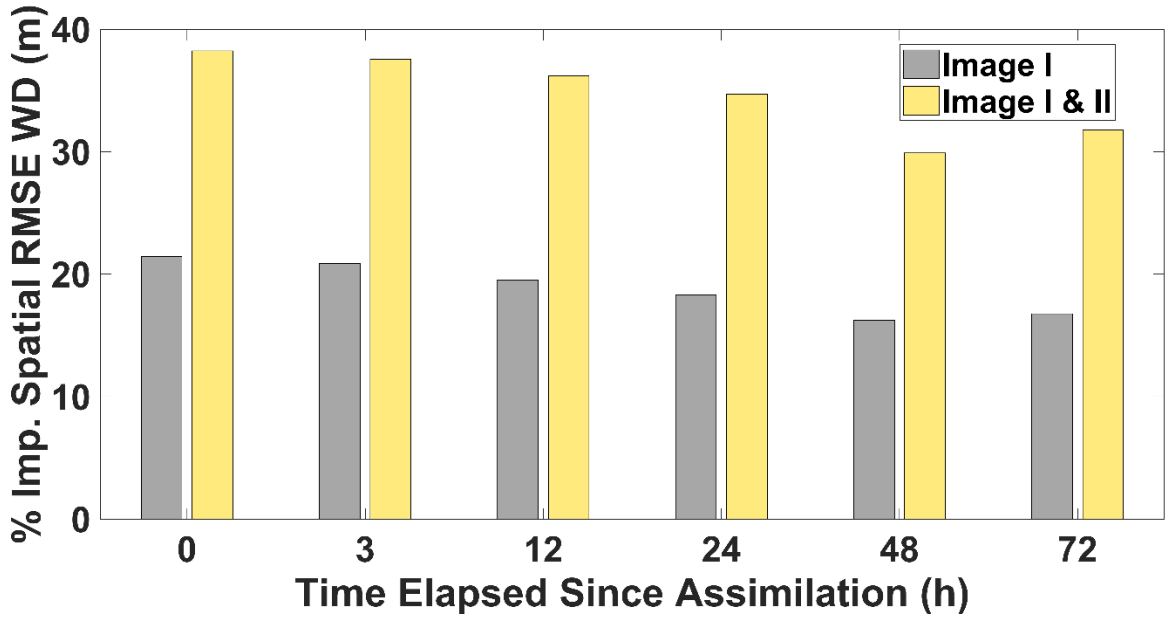

Fig. 7.7 As for Fig. 7.5 but for spatially averaged RMSE in water depth.

of false alarms. The percentage improvement in the CSI (Fig. 7.5) and the Kappa (Fig. 7.6) were therefore expected to primarily be a function of the reduction in false alarms.

In terms of the temporal positions of the images with respect to the hydrograph, assimilating the first image ~6h after the peak reduced false alarms by ~2%, while adding the second image ~18h post peak reduced overestimation by ~5% at the assimilation time steps. The percentage improvement in the Kappa as well as the CSI values with increasing lead time was fairly consistent and exhibited similar slightly decreasing trends. As the assimilated images were acquired in the falling limb, the multiplicative error also started to reduce with time after the peak with decreasing input inflow values. At this time during the event, observations can be more informative and thus better constrain the forecast as the correlation between time steps is relatively higher. Adding the second image nearly doubled the percentage improvement for both the objective functions for all the lead times examined here, indicating that more observations could potentially further improve the forecast. Improvements from the assimilation were retained at least up to the examined lead time of 3 days.

7.3.2 Impact on Floodplain Water Depth Simulation

On comparing the lumped spatial RMSE values between the assimilation and the open loop forecasts, the errors in simulated floodplain water depth consistently reduced over time. In fact, the percentage improvement in spatial mean RMSE (Fig. 7.7), revealed that the assimilation achieved average improvements ~30% for up to 72 hours. Assimilating Image I led to an average improvement of ~20% in the RMSE statistic, while adding Image II nearly doubled it to ~37%. It is possible that the images available to this study were at favourable temporal positions with respect to the flood hydrograph, and consequently proved to be informative for the inundation forecast. As post peak images are usually characterised by transitioning flows, the inundation patterns are typically dominated by flood propagation mechanisms during this time. In contrast, during the rising limb inflow errors dominate and therefore result in inconsistent improvements due to the continuous addition of more unpredictable errors to the domain. This finding is in keeping with those of earlier studies which have found flood extent observations acquired just after the peak flow to have higher information content for hydraulic model parameterization (Gobeyn et al., 2017; Wood, 2016; Wood et al., 2016a).

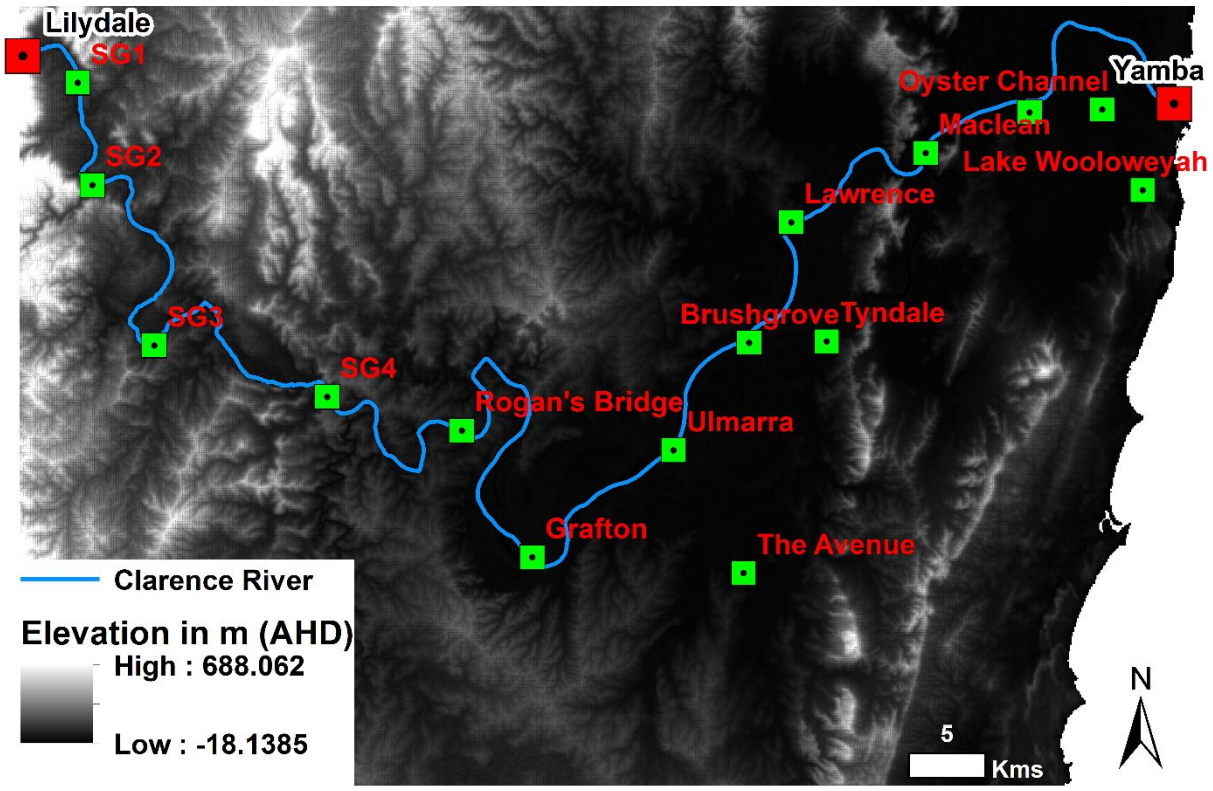

Fig. 7.8 The location of the gauges used for the discussion on the WDD maps.

.

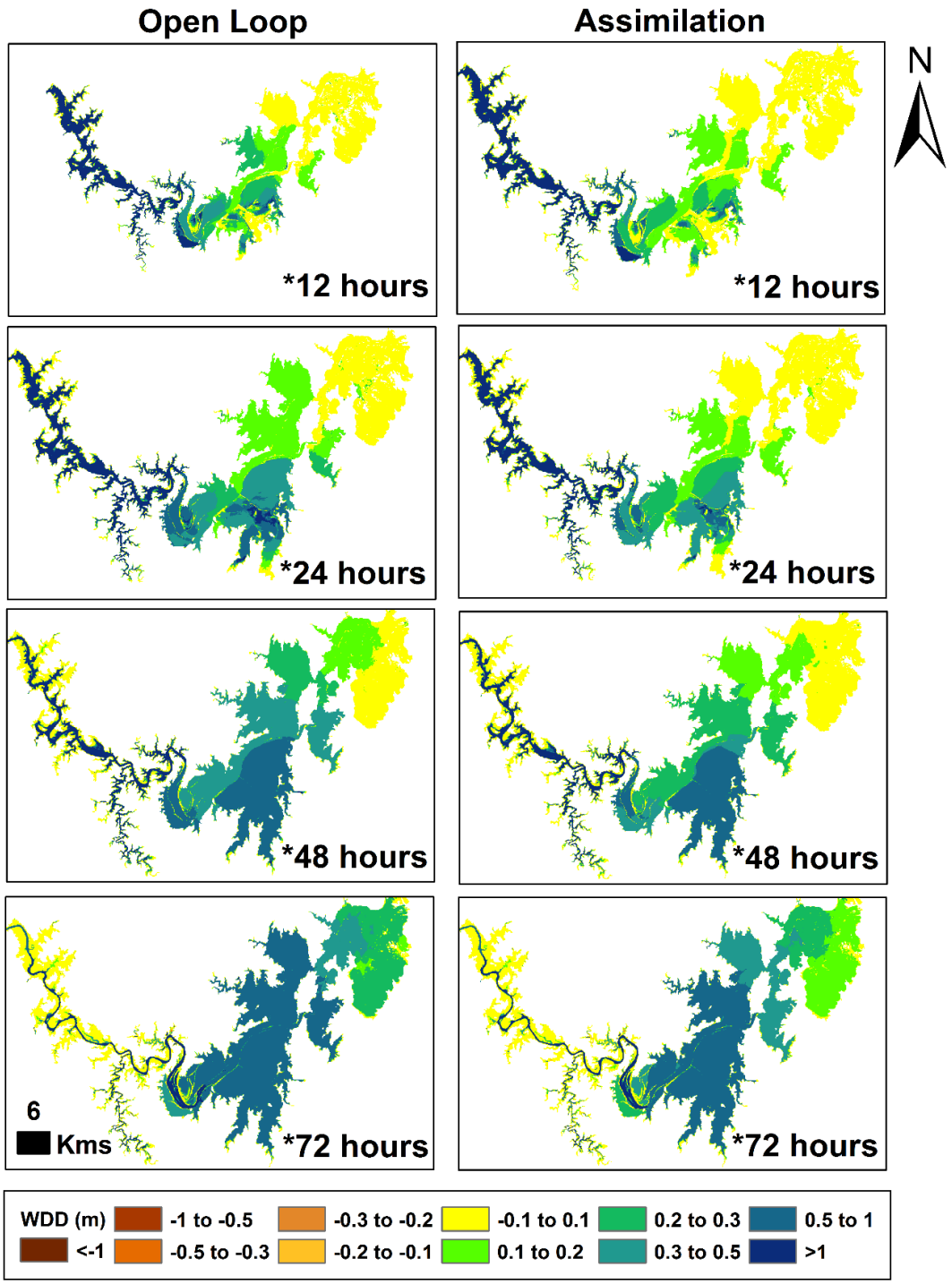

Fig. 7.9 Forecast ensemble mean minus true water depth for different *lead times measured from the first assimilation time step after assimilating only the first image.

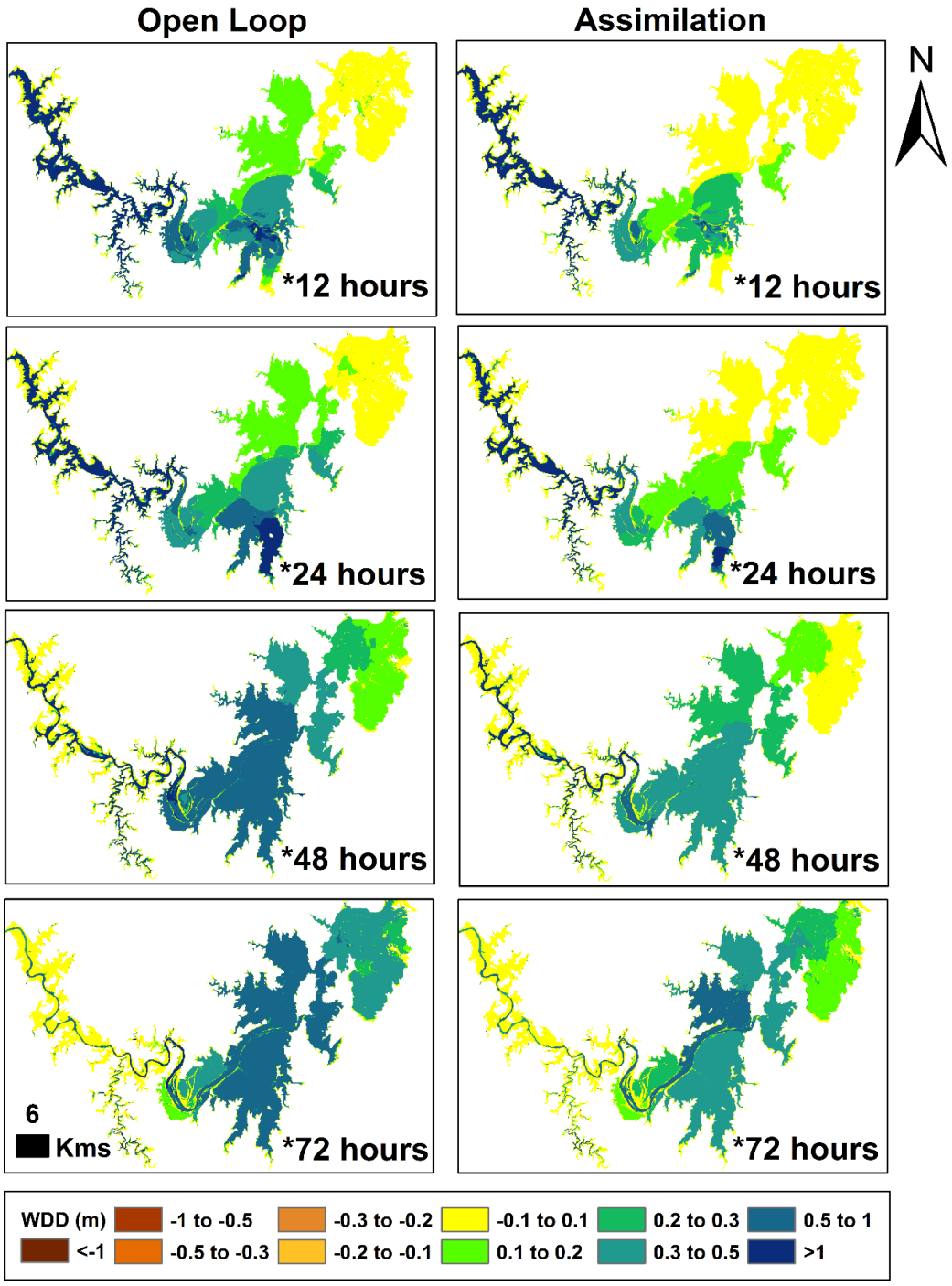

Fig. 7.10 As for

Fig. 7.9 but for time after the second assimilation time step when both images were assimilated. Note that the time steps shown here are different from the ones shown in the previous figure.

In order to better understand the spatial evolution of forecast errors, the water depth difference between the forecast ensemble mean and the truth at each grid was calculated, and presented here for an assimilation experiment based on real topography for the first time. The forecast minus true water depth difference (WDD) is shown in Fig. 7.9 and Fig. 7.10 for the first and second assimilation time steps, respectively. Therefore, the positive errors imply that the forecast overestimates the true water depth, while the converse is true for the negative errors where the forecast underestimates the truth. The colours ranging from rust to indigo refer to the direction of the deviation from truth; i.e. yellow is within ±10cm of the truth, the different shades of green through to indigo represent excess water, and the shades going from orange to rust indicate a lack of water. The intensity of colours in each direction is representative of the magnitude of the deviation from the truth. WDD maps were computed for lead times of 12, 24, 48, and 72h to assess the spatiotemporal evolution of errors in the domain after the assimilation.

The first observation that can be made from these figures is that the open loop was positively biased for most of the domain throughout; again this was expected due to the nature of the inflow errors used here. There seemed to be a significant reduction in the magnitude of errors after the assimilation of both images, in comparison to the assimilation of only one image, as expected due to increased information content of the forecast with the integration of more observations.

For each lead time the assimilated forecast was able to reduce the errors by at least one (~20cm) and up to several orders of magnitude. As an example, if the central portion of the catchment is considered, it is easy to observe that a large number of pixels moved to lower intensity hues between the left and the right columns - showing the open loop and the assimilated forecast, respectively - implying a definitive nudge towards the truth. As expected, the assimilation had limited impact in the upstream part of the catchment. Interestingly, the downstream part of the catchment exhibited notable reductions in uncertainty even after the assimilation of only one image, although the effect was more pronounced when both images were considered together.

If the temporal locations of the images and the individual impact of each considered, it is evident that the magnitude of change in WDD was greater for the second image with noticeable improvements retained for ~3 days. For instance, on assimilating the first image, the positive errors in the region starting from Rogan’s Bridge down to the outlet at Yamba were consistently reduced by at least one order of magnitude for all lead times examined up to 48h.

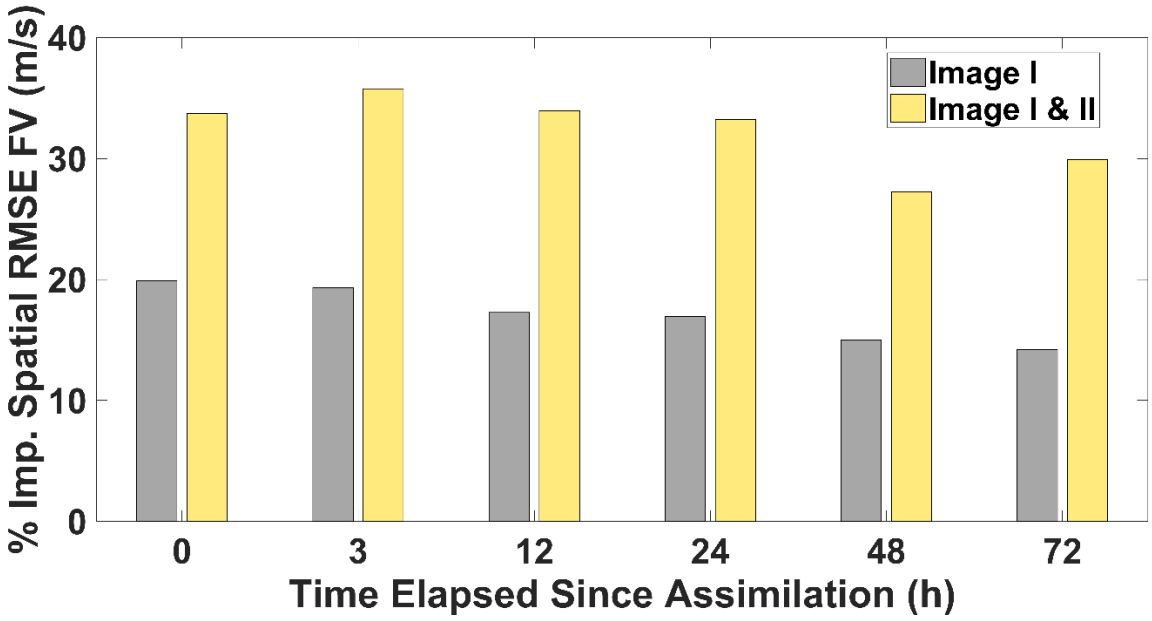

Fig. 7.11 As for Fig. 7.7 but for RMSE in flow velocities.

In the 72h case, the error reduction was only noticeable downstream of Maclean, but this apparent lack of improvement could also be due to the colour discretization used here, which lumped larger error values within a single class. However, when the second image was assimilated, there were consistent improvements up to lead times of 72h. Even after 72h from the second assimilation time step, the WDD map for the assimilated forecast reduced the errors by one order of magnitude in all areas downstream of Grafton. In general, there seemed to be a greater impact when both images were considered, in keeping with the general expectation and the observations from the lumped spatial RMSE plot, although the closely spaced post-peak temporal position of these images may have been a contributing factor. For the assimilation of single as well as multiple images, clear reductions in uncertainty were visible for large areas, highlighting the usefulness of the proposed algorithm for improving flood inundation forecasts

7.3.3 Impact on Floodplain Flow Velocity Simulation

In this section, the distributed impact of flood extent assimilation on flow velocities is estimated through flow velocity difference (FVD) maps, which were computed as the pixelwise difference between the forecast mean and the truth at different lead times from the assimilation similar to the WDD maps. Lumped impact on simulated floodplain flow velocities was also estimated through an examination of the spatially averaged RMSE in flow velocities, as in the previous section.

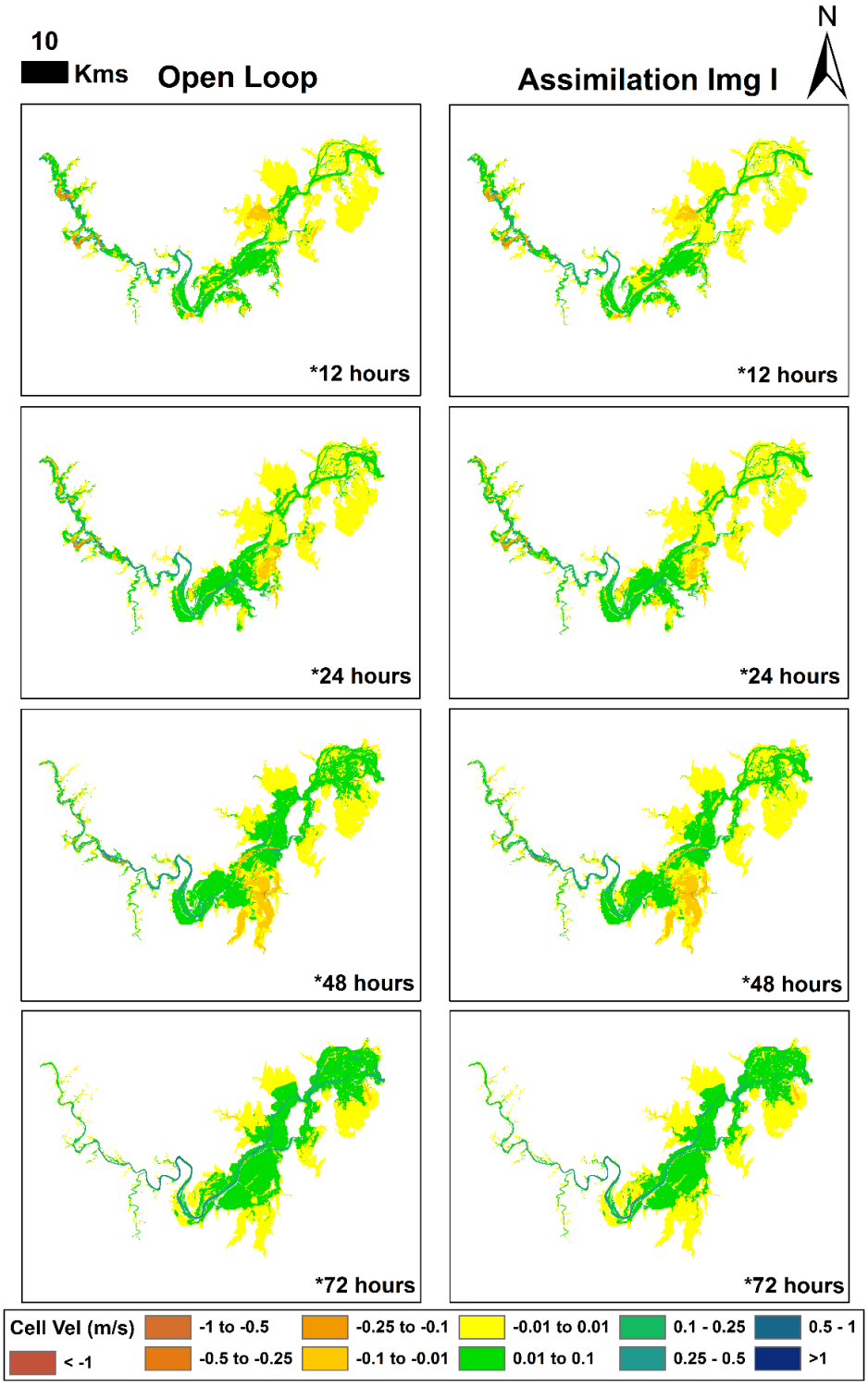

Fig. 7.12 Difference between the forecast mean flow velocities and the truth computed at each grid cell, shown here for different lead times following the first assimilation time step. The difference is calculated as forecast minus truth, so positive errors represent overestimation while negative errors show underestimation.

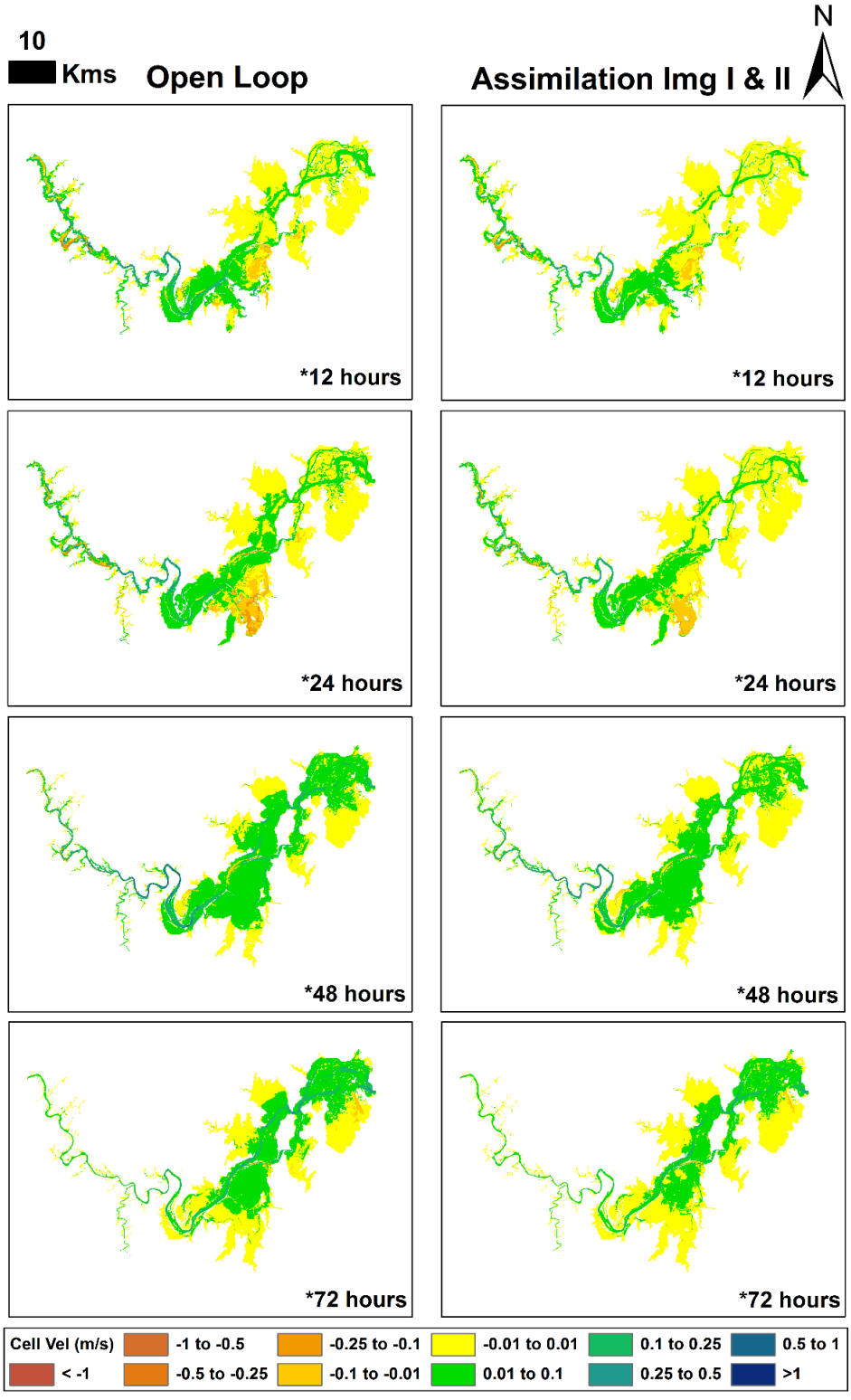

Fig. 7.13 As for Fig. 7.12, but for the assimilation of both images together at the second assimilation time step.

The cell velocities were computed by combining the cell edge velocities disaggregated in the x- and y- directions, the intermediate model states output by Lisflood-FP. The colour scheme used here is identical to the previous section, although it is important to note that the discretisation of classes differs slightly based on the range of bias values observed. The minimum error class of ±0.01m/s is represented by the colour yellow in this section as well.

It could be argued that the simulation of flow velocities is not as accurate in hydraulic models as it should be, primarily as momentum is not conserved in channel floodplain connections, for example. However, as this was a synthetic experiment and the flow velocities used as benchmark for the evaluation were also simulated using the same model, this comparison was expected to yield useful information about the impact of assimilation on floodplain flow velocities, which has not yet been studied. Channel velocities are implicitly considered when evaluating against gauge discharge, but floodplain flow velocities have never been spatially examined in the past, primarily due to a lack of appropriate observations to facilitate such a comparison. Since most high value assets are typically located in the floodplain, such analyses are invaluable to a comprehensive spatiotemporal characterization of the hydrodynamic model errors and the impact of assimilation on their subsequent propagation.

The first and most obvious observation from the FVD maps illustrated in Fig. 7.12 and Fig. 7.13 is that the assimilation improved the forecast for all examined lead times, evident from the increase in the number of yellow cells. Although overall error reductions were observed, it is interesting to note that the assimilation was able to correct for both under- and over-estimation, which was locally introduced in some regions by the inflow errors in the open loop. There was a general trend of overestimation in velocities in the open loop which reduced after the assimilation, although the extent of the overestimation was much lesser in magnitude than the water depths. This is due to the fact that errors were only introduced in the magnitude of the hydrograph and not explicitly in the timing of the flows, implying that the errors in velocity were primarily a consequence of trying to hydraulically move larger quantities of water across the floodplain. The impact of the assimilation remained consistent across all examined lead times, after the assimilation of one as well as both images, as evident from the percentage improvement in spatial mean RMSE in flow velocity illustrated in Fig. 7.11. In fact, after assimilating both images together at the second assimilation time step, >30% improvements were observed consistently for at least 72 hours, which is significant for better estimating flood hazard.

Assimilating the first image did not reveal large spatial impacts although the lumped RMSE metric suggested ~20% improvements over the open loop. After assimilating the second image, the assimilation consistently improved velocity estimates over the entire domain, visibly bringing more pixels to lower uncertainty classes. One possible reason for this could be the temporal position of Image II, which was almost a day after the observed peak. At this time during the flood, the flows are out of bank and rapidly transitioning, leading to model propagation driven inundation patterns. As the error from the inflows is not dominant at this time, the assimilation of a "true" flood extent observation was able to effectively constrain the model trajectory for larger durations. Moreover, the errors in velocity are primarily a function of the propagation, rather than the synthetically introduced uncertainties as previously noted.

It therefore naturally follows that the relative improvements in water depth simulations would exceed the velocity error reductions as a consequence of the assimilation. However, this assessment illustrated that the assimilation improved all model states and did not negatively impact the delicate equilibrium between them by pulling the model trajectory in a certain direction. This implies that the proposed flood extent assimilation algorithm has the potential to reduce flood inundation forecasting uncertainties with significant implications for the

Table 7.4 Flood hazard classification according to human safety, given by the Australian Rainfall Runoff Revision conducted by Engineers Australia in 2010 (Cox and Shand, T.D.Blacka 2010). The green colour refers to the safe category, with colours going towards red signifying increasing values of flood hazard

<table>
<tr><th>D.V m^2/s</th><th>Infants/young children (H.M < 25 m.kg) and frail old persons</th><th>Children (H.M = 25 to 50 m.kg)</th><th>Adults (H.M > 50 m.kg)</th></tr>
<tr><td>0-0.4</td><td rowspan="5">Extreme Hazard; Dangerous to all</td><td>Low Hazard</td><td rowspan="2">Low Hazard</td></tr>
<tr><td>0.4-0.6</td><td>Significant Hazard; Dangerous to most</td></tr>
<tr><td>0.6-0.8</td><td rowspan="3">Extreme Hazard; Dangerous to all</td><td>Moderate Hazard; Dangerous to some</td></tr>
<tr><td>0.8-1.2</td><td>Significant Hazard; Dangerous to most</td></tr>
<tr><td>>1.2</td><td>Extreme Hazard; Dangerous to all</td></tr>
</table>

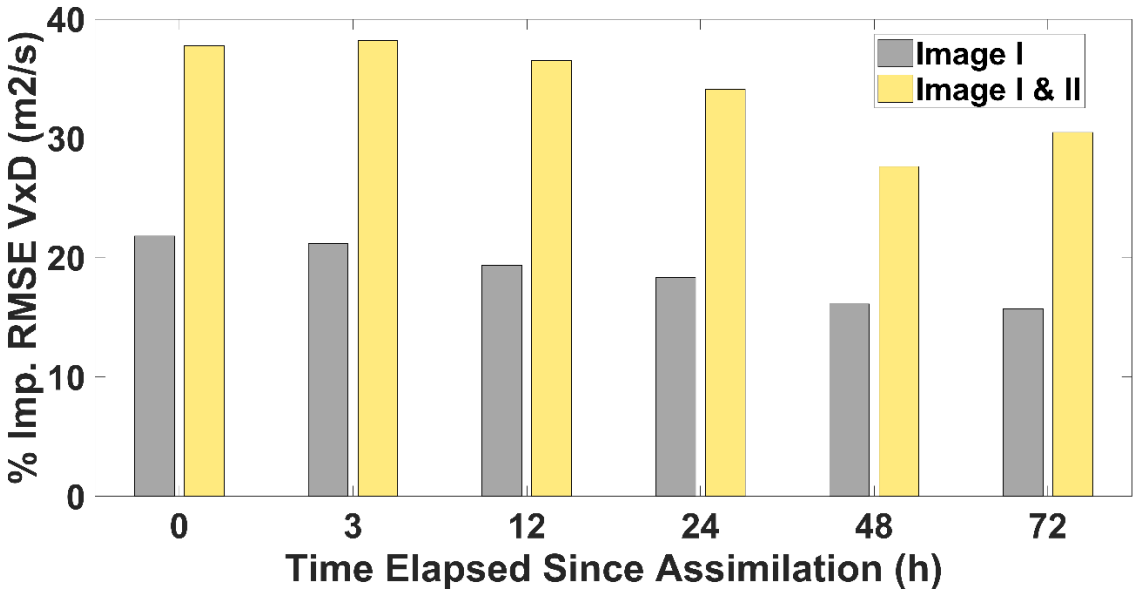

Fig. 7.14 As for Fig. 7.11 but for flood hazard estimates.

estimation of flood hazard. Flood hazard is a function of flood depth and velocity, and better hazard estimates could potentially better inform preparedness and help to create more resilient societies.

7.3.4 Impact on Flood Hazard Simulation

Anthropogenic activity in floodplains is inevitable and therefore, the hazard posed by floods needs to be accurately quantified in order to minimise risk. Human safety is compromised when exposed to flows exceeding their capability to remain standing (National Flood Risk Advisory Group and Australia-New Zealand Emergeny Management Committee 2014). Australian Rainfall and Runoff (ARR) guidelines stipulate that "to prevent pedestrians being swept along streets and other drainage paths during major storm events, the product of velocities (V) and depths (D) in streets and major flow paths generally should not exceed $D \times V = 0.4$ m^2/s" (Cox and Shand, T.D.Blacka 2010). Relationships between the height and mass of a person (H.M; m.kg) and the tolerable flow value (D.V; m^2/s), have been defined according to general guidelines, which classifies adults (H.M > 50 m.kg) and children (H.M = 25 to 50 m.kg). Infants and very young children (H.M < 25 m.kg) are considered unsafe in any flow without adult support. The safety guidelines provided by Engineers Australia in accordance with ARR 2015 have been summarized in Table 7.4.

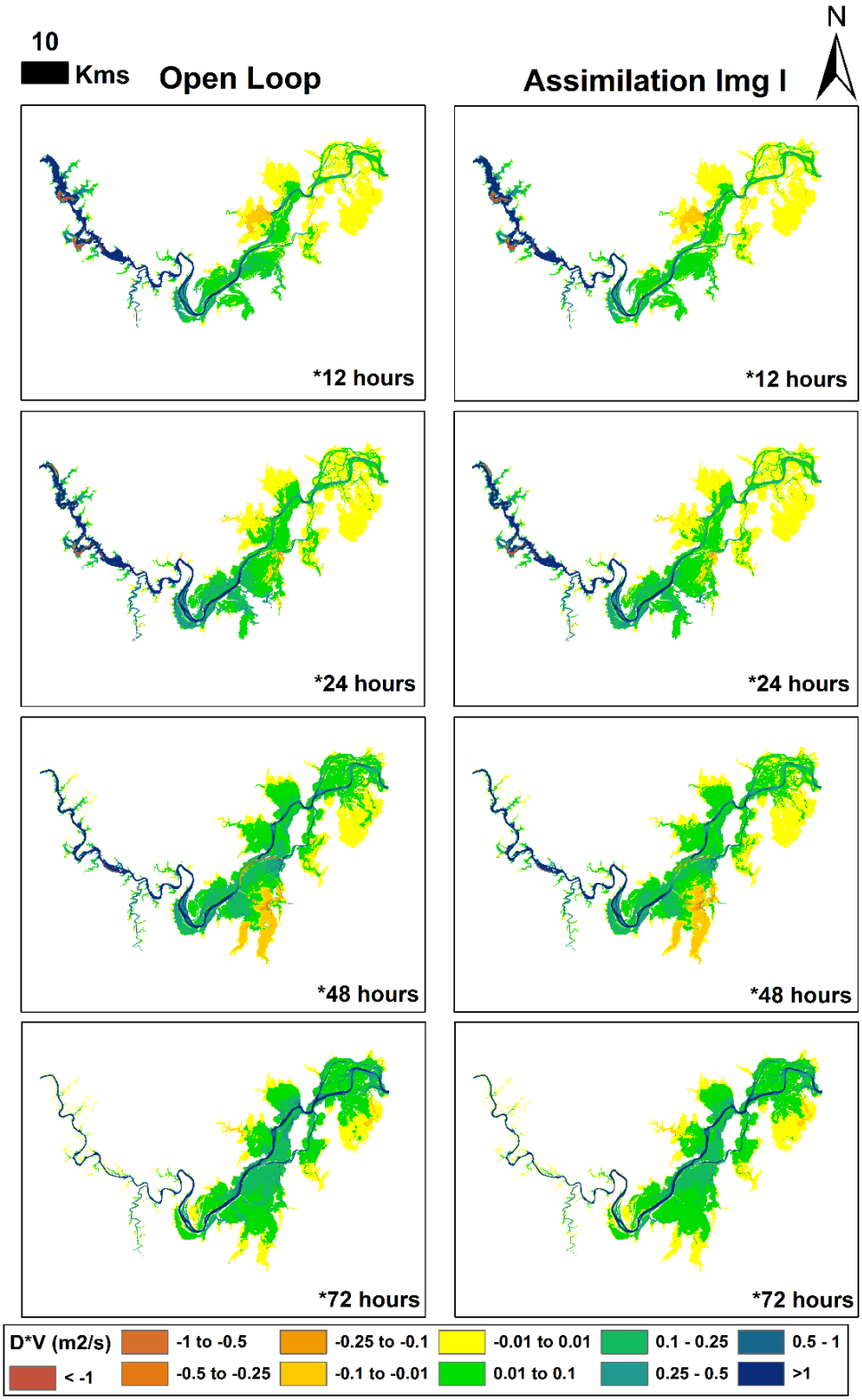

Fig. 7.15 As for Fig. 7.12 but for the product of flow velocities and depth for different lead times from the first assimilation time step.

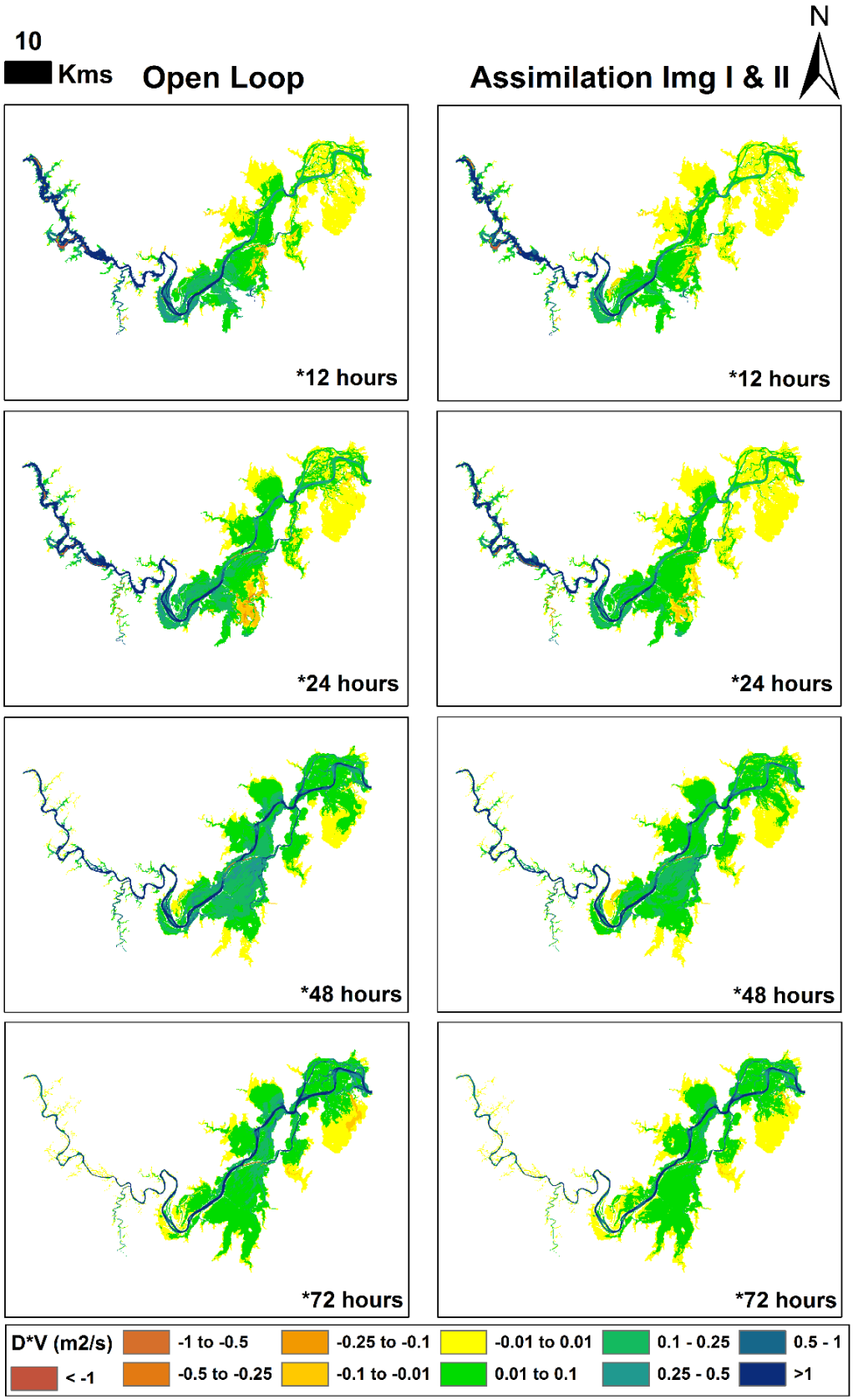

Fig. 7.16 As for Fig. 7.15 but for the second assimilation time step.

Although it should also be noted that loss of stability could occur in milder flow conditions when problems like uneven, slippery, or obstacles in bottom conditions, or flow conditions like floating debris, low temperature, poor visibility, unsteady flow and strong winds exist. Moreover, the training and musculature of the human subject might also be relevant, in addition to disability and psychological factors.

The lumped improvement in the spatial mean RMSE in hazard estimates is quantified as percentages in Fig. 7.14. Additionally, Fig. 7.15 and Fig. 7.16 show the hazard difference maps, which have been calculated exactly as the maps from the preceding sections, as forecast hazard (product of velocity and depth) minus truth. Accordingly, positive values signify overestimation of flood hazard, while the converse is true for the negative values. The colour scale and discretisation used for the figures is identical to the previous section on flow velocities. Both figures show that on average the open loop forecast slightly overestimated the hazard, which was expected due the positively biased inflows used for the simulation. However, the assimilation was able to reduce the difference between forecast and true hazard estimates, especially in the floodplains, for all lead times but with decreasing impact through time.

Assimilating the first image was able to correct for both positive and negative errors in hazard estimates, but the improvements were marginal and not immediately apparent in the spatially distributed error maps. On assimilating both images, the overall positive impacts were larger and the hazard estimates consistently improved around the Grafton area, which is significant from a flood management perspective. The assimilation exhibited positive effects at least up to the last examined lead time of 72h, illustrated by the increase in the number of yellow pixels (minimum error category) in the right hand side panels of Fig. 7.15 and Fig. 7.16. Importantly, the estimation of flood hazard in the areas with human settlements in the floodplain such as Grafton, Ulmarra, and Maclean, was considerably improved after the assimilation. These results suggest that flood preparedness and planning in this catchment could significantly benefit from the proposed assimilation strategy.

7.3.5 Impact on Channel Flow and Water Level Simulation

The assimilation results for the channel discharge and water levels, at eleven gauging stations (four synthetic and seven real gauges; Locations in Fig. 7.8) along the main stem of the river are shown in Fig. 7.17 and Fig. 7.18, respectively. The open loop is always shown in red, the truth in green and the assimilation of image one and that of both images is depicted by blue

and a dashed magenta line, respectively. The gauges have been arranged upstream to downstream in the different rows of the figures, with the final subplot illustrating the gauge locations as a reference. A cursory examination of both figures reveals that the assimilation always improved the estimation of flow and depth in the channel, with two images exhibiting stronger impacts than the assimilation of only image one as expected.

From the discharge plots of Fig. 7.17, it appears as though the overall impacts were not significant at the upstream gauges, with the impacts increasing downstream of Rogan's Bridge to Palmer's Island Bridge, especially following the the assimilation of both images. Another

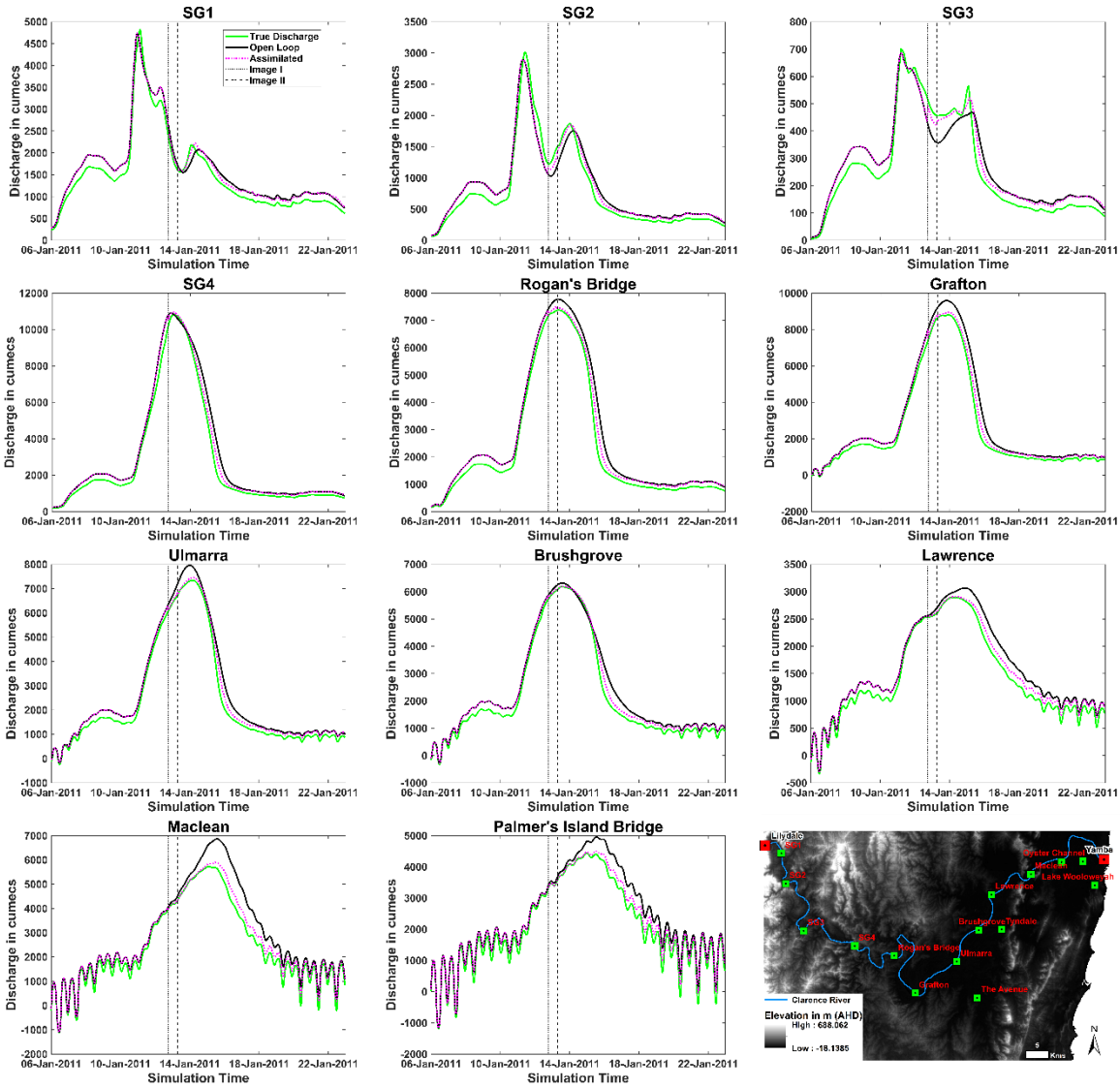

Fig. 7.17 Channel discharge time series at the synthetic and real gauge locations along the main stem of the Clarence River, with the open-loop expectation (red), assimilation experiment expectation (blue-I and dashed magenta-I+II), and the truth (green).

observation that can be made from the figures is that in spite of the positively biased inflow errors introduced, there are some portions of the hydrograph where the open loop under predicted the flow at a few upstream gauges. The gauges SG1, SG2, and SG3 where this phenomenon occurred, exhibited limited sensitivity to the assimilation. However, the assimilation was able to pull the state trajectory upwards towards the truth by increasing the discharge during the update. This is interesting, primarily as the truth was consistently lower than the open loop for all the other gauges, where the assimilation resulted in lower discharge values, showcasing the ability to deal with positive as well as negative errors quite effectively.

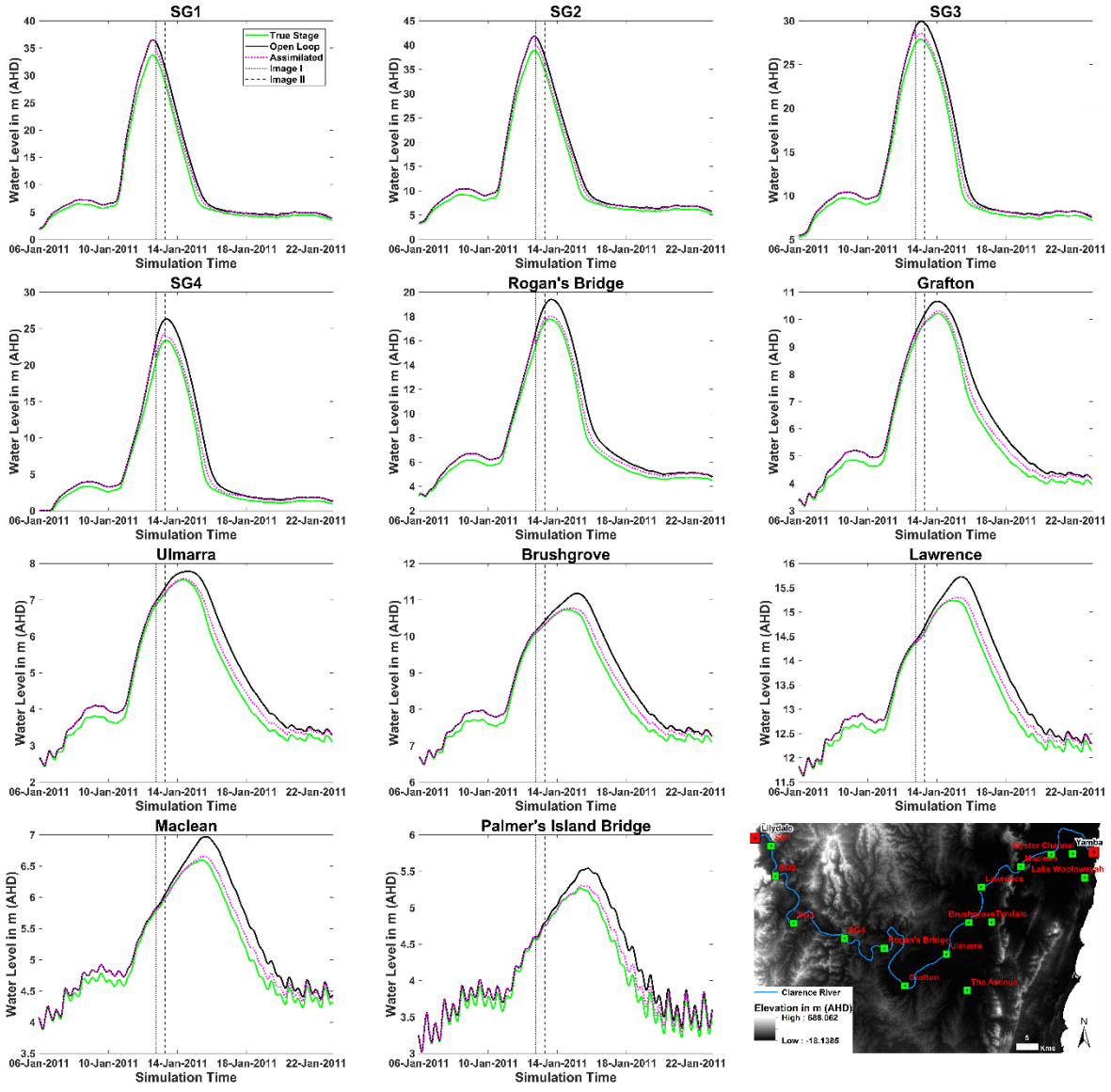

Fig. 7.18 As for Fig. 7.17 but for simulated channel water levels.

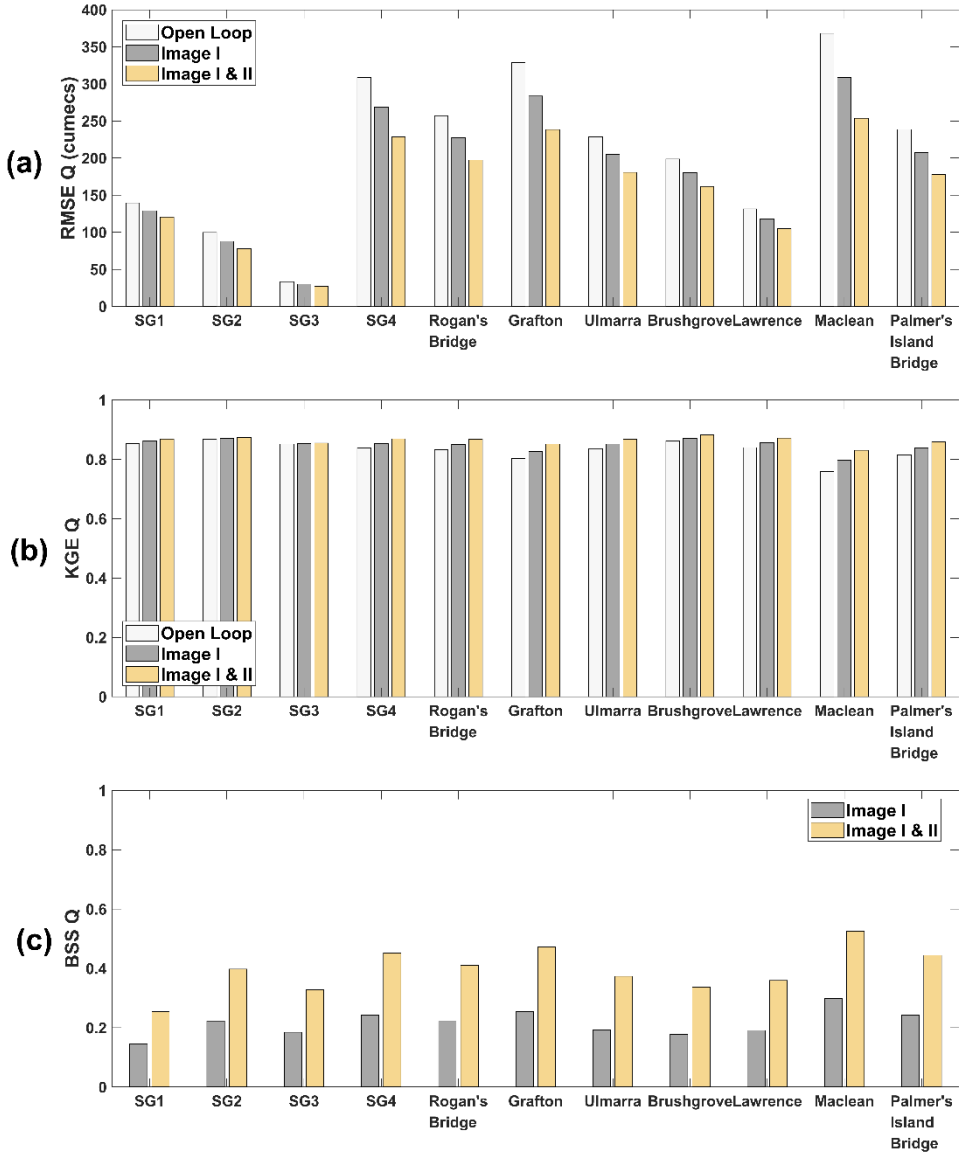

Fig. 7.19 Illustration of the channel evaluation efficiency metrics, (a) RMSE, (b) KGE, and (c) BSS for the simulated discharge, at the gauges along the main stem of the river (Locations shown in Fig. 7.8).

For the water levels, this effect is observable from the first gauge at the upstream end i.e. SG1. The water level hydrograph of the open loop simulation was consistently higher than the truth, even for gauges SG1, SG2, and SG3, suggesting that the negative bias observed in the discharge at these locations could be dominated by the contribution of the velocity. One possible reason for this is the placement of the synthetic gauges, as the shape of these particular synthetic gauge hydrographs also differs from the others. The two peaks exhibited by the hydrographs observed at SG1, SG2, and SG3, suggest that the placement might be close to a

tributary, and the rise and fall of the water in the inlet may be responsible for the second peak observed.

Since, Lisflood-FP calculates discharge as a function of the water crossing over a given cross-section in a specified direction, i.e. a gauge location needs to be specified and the cross-section needs to be defined from this point in a particular direction (N, E, W, or S only). As channel flows cannot be effectively separated in the four Cartesian directions alone, therefore, placing the gauge at a tributary inlet could explain this two peak effect. This hypothesis was

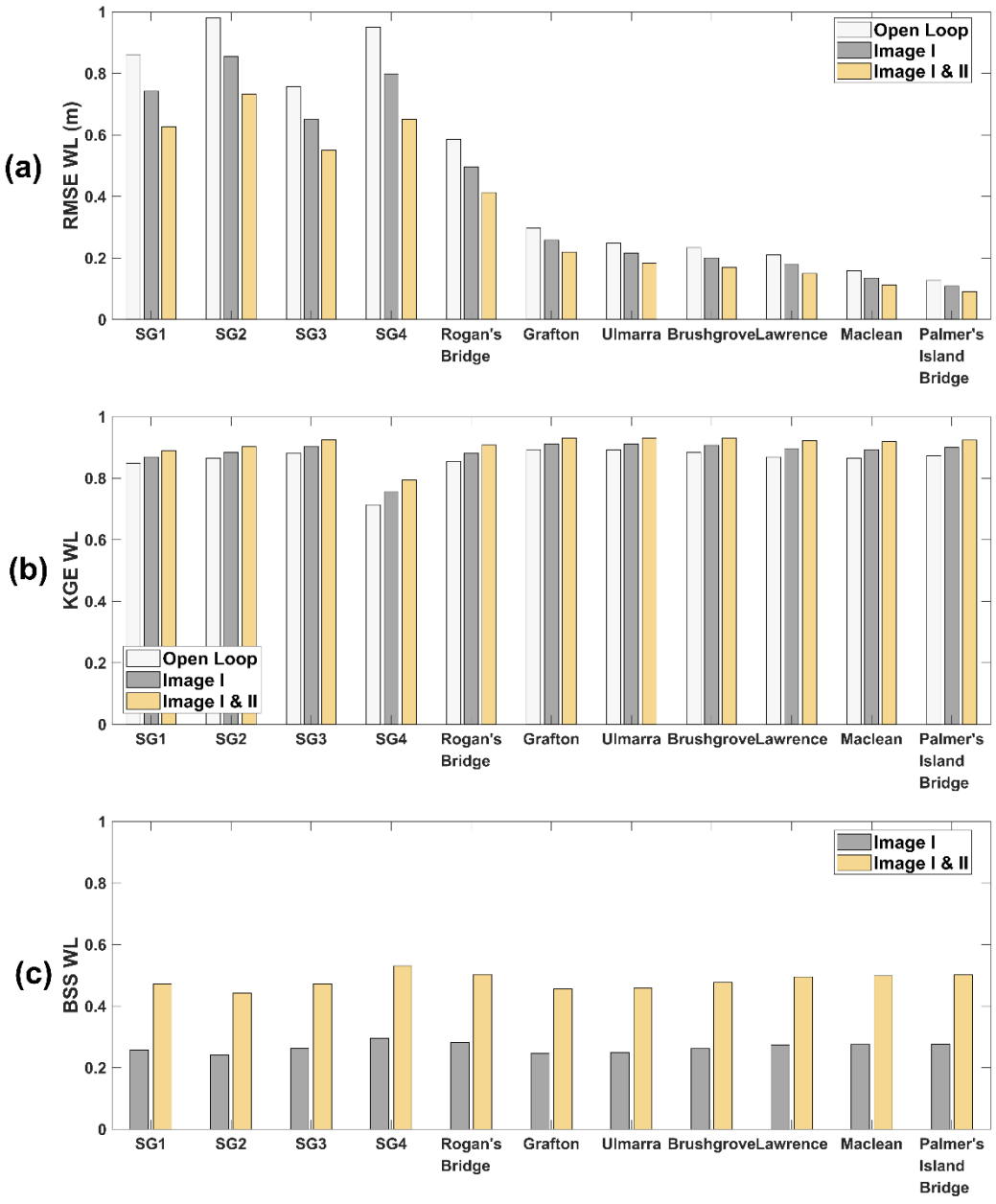

Fig. 7.20 As for Fig. 7.19, but for the simulated channel water levels.

also corroborated by the volume of the hydrograph at SG3, which was significantly lower than all the other hydrographs, implying that it might primarily be illustrating the flow within a small tributary rather than the main channel. Another interesting thing to note from the channel hydrographs is that the improvements persist for several days (~7) after the assimilation, especially downstream of Grafton where the flood peak actually arrived after both images had been assimilated, which can be crucial from an emergency management perspective.

Next, the channel flow and water level time series was statistically evaluated using the RMSE, KGE, and the BSS statistics, shown in Fig. 7.19 and Fig. 7.20, as sub-figures (a), (b), and (c), respectively. The BSS implicitly compared the forecast skill of the assimilated ensemble with the open loop, with values ranging from -∞ to an ideal value of 1.

The RMSE values for the open loop were high for the water level time series at the upstream synthetic gauges, but the real gauges located downstream from Rogan's Bridge exhibited lower errors with respect to the truth. The converse was true of the RMSE values for the channel discharge, which were low at the first three upstream synthetic gauges and then consistently high for the rest of the channel. This seems like a function of the simulated velocity, especially as the trends are noticeably contrary to those observed in the channel water level assessment. Lisflood-FP can effectively compute velocities in the upstream topographic reaches but perhaps the simulation of velocities in the more hydrodynamic reaches varies with changing water volumes introduced in the domain. The errors as well as the corresponding percentage improvement as a consequence of the assimilation were consistent from upstream to downstream. This is typically true for errors propagated upstream to downstream, since one uncertain value is used to calculate the next, purely as a function of the numerical model domain discretisation. Since, the errors are introduced as inflows at the upstream boundary, the improvements from assimilation can be sustained for longer durations downstream (Cooper et al. 2018).

The KGE metric displays almost similar trends for the discharge and water level time series and is generally quite high. Improvements seem consistent from upstream to downstream with only a marginal increase noticeable in the case of water depths. The simultaneous impact of assimilating both images was nearly double in comparison to the assimilation of only the first image for the gauge evaluation, as evident from the Brier Skill Score plots for the discharge and water levels. An assessment of the BSS plots also revealed that the forecast always strongly benefited from the assimilation, with more consistent improvements visible in the water depth

hydrographs relative to the discharge. There was a general increasing trend in assimilation benefits from upstream to downstream in the discharge statistics, although the magnitude of the trend was marginal. This is expected to be a function of the forecast velocity improvements achieved due to the assimilation, as the water levels do not exhibit this trend. Moreover, velocity errors amplify across the domain during flood propagation as a consequence of numerical modelling and the grid-wise domain discretisation. It therefore follows that the impacts of correcting the forecast would be more evident as the errors increase from upstream to downstream. The assimilation of images one and that of one and two together with weight multiplication, resulted in average improvements of about ~20% and ~40%, respectively, for both simulated discharge and water depth. Overall the proposed flood extent assimilation algorithm had a positive impact, demonstrating the potential for improving flood inundation forecasting capabilities, especially in ungauged catchments.

7.4 Chapter Summary

This chapter evaluated the performance of the flood extent assimilation algorithm proposed in this thesis through a synthetic experiment. First, the experimental design and ensemble generation were described, followed by a description of the methods used to generate the synthetic satellite SAR observations and the performance metrics used for the evaluation of the forecasts. Assimilation impacts were evaluated in terms of the simulated flood extent, floodplain water depth, floodplain flow velocities, and the channel flow and depth for the synthetic case. Overall the effects of the assimilation were strongly positive, with the magnitude of improvements generally increasing on assimilating additional observations and from upstream to downstream. An evaluation of the channel performance revealed that the assimilation is able to handle unsystematic bias, as the assimilation is able to pull the model trajectory towards the truth when the true state is higher or lower than the forecast. Subsequently, Chapter 8 evaluates algorithm sensitivity towards observation spatiotemporal characteristics, while Chapter 9 describes the impacts of implementing the proposed assimilation algorithm for a real world application.

CHAPTER EIGHT

"In one word, to draw the rule from experience, one must generalize; this is a necessity that imposes itself on the most circumspect observer."

- Henri Poincaré, *La valeur de la science.* In *Anton Bovier, Statistical Mechanics of Disordered Systems (2006), 186.*

8. Sensitivity to Observation Characteristics

The experimental results obtained in Chapter 7 demonstrated the potential of the flood extent assimilation framework proposed in Chapter 6 of this thesis. In this chapter, the sensitivity of the proposed assimilation technique to the spatiotemporal characteristics of the satellite-based flood observation is examined. The primary objective of this investigation is to find the most informative spatiotemporal SAR-acquisition scenario for flood forecasting to facilitate the design of a targeted observation system. Those sub-regions of the domain where the model is unable to predict the real-world flow dynamics are expected to enable maximum improvements in forecast accuracy, if independent satellite flood extent observations are integrated through assimilation at the appropriate time and location.

8.1 Introduction

Multiple studies in literature have examined the relationship between SAR-acquisition timing and the usefulness of information relevant to flood modelling applications (García-Pintado et al. 2013; Wood et al. 2016; Gobeyn et al. 2017). However, there has been little attention given to the spatial location of the Earth observations, with only two studies that have explicitly investigated this issue (Andreadis and Schumann 2014; Schumann and Andreadis 2016). The seminal study by Andreadis and Schumann (2014) which highlighted the possibility of forecast degradation through assimilation for the first time, used a Localised Ensemble Transform Kalman Filtering approach, to pre-emptively estimate the spatiotemporal impacts of assimilating observations of water surface elevation, channel top width, and inundated area. Specifically, lumped values of each variable (for example, inundated area in sq. kms.) were compared using a cost function, to quantify the impact of observations that were not spatially coincident with the forecast reach.

This investigation was of interest as the spatial coverage of satellite observations might be partial relative to the model domain depending on the specific sensor, orbit, and imaging mode chosen for image acquisition. This assumption typically holds true for high resolution satellite SAR observations (Boni et al. 2016). In fact, the experiments of Andreadis and Schumann (2014) specifically examined the impacts of updating the entire model trajectory, based on observations located at a single reach. Through a sequential implementation of this assimilation framework to localized 5km sub-reaches of the Ohio River, the most valuable observation locations for improved forecast skill were identified. In a similar study Schumann and Andreadis (2016) were the first to investigate the potential of assimilating improved topography, using a targeted observation system in reaches characterised by hydraulic behaviour. This study was novel in quantitatively evaluating the humanitarian and socioeconomic impact of the improved forecasts as a consequence of the assimilation. However, the study again examined impacts on inundated area in a lumped fashion, limiting insights into the spatial evolution of forecast errors after the assimilation.

Although these studies presented the first steps towards an investigation of the impacts of spatiotemporal observation characteristics on flood extent assimilation performance, the use of a lumped value for inundated area did not allow for spatial comparison. Specifically, as the locations of the inundation of the models and observations were not compared in the particular implementation of assimilation employed by the aforementioned studies, the quantified disagreement cannot be considered representative of spatial patterns of flooding. The assimilation framework of Chapter 6 used in this experiment does utilize a lumped statistic for comparison of extents, however, the calculation of this lumped value is rooted in pixel-wise spatial comparisons of flood inundation. Multiple studies have only examined the spatiotemporal impact of assimilating SAR-derived water levels, while the spatiotemporal sensitivity of direct flood extent assimilation has largely been ignored (Lai et al. 2014). Inundation extents can be good proxies for correcting forecast errors, but characteristics of flood extent assimilation, especially in terms of model response, can be starkly different to water level assimilation as it is not an actual model state. Through spatially lumped assessments of inundated area, Andreadis and Schumann (2014) discovered that the assimilation impact might be a function of temporal variability. As the recent spate of high-resolution SAR sensors can only observe a small part of the catchment, acquisition planning can significantly benefit from knowledge on which part of the catchment can lead to maximum reductions in forecast uncertainty (Boni et al. 2016).

SAR acquisitions are fairly expensive; it is therefore, imperative that the utilization and downstream impact of each image be optimized through studies on designing targeted observation strategies (García-Pintado et al. 2013). Accordingly, the spatiotemporal sensitivity of flood extent assimilation was critically examined in this chapter, through a series of synthetic numerical experiments. In order to ensure transferability of results to other catchments, a reach-wise spatial domain discretization was chosen following Schumann and Andreadis (2016), allowing findings that could be generalised to other hydraulically similar reaches. Synthetic observations generated in reach-specific model sub-domains using the approaches outlined in Chapter 7 were assimilated using the DA framework of Chapter 6, into the forecast inflow open loop ensemble described in Chapter 7. The following sections exclusively describe those components of the experimental design and performance metrics, which differ from the previous chapter.

8.2 Methods

Designing a targeted observation system first requires a preliminary investigation of reach hydraulic behaviour, to understand which parts of the model domain may benefit most from the assimilation. This section describes the methodology used to formulate the experimental design and the performance criterion used to evaluate the spatiotemporal observation sensitivity.

8.2.1 Experimental Design

One approach for identifying hydraulically similar reaches is to simulate the channel flow and observe the resulting water surface profile, whereby hydraulically uniform reaches will exhibit consistent water surface slopes. However, this requires running the hydraulic model once pre-emptively to evaluate the reach hydraulic behaviour, which adds to the computational expense of the modelling exercise, even though the resulting sub-reach characterization would be quite reliable. In this chapter, the reach hydraulic characterization approach of Schumann and Andreadis (2016) was employed, and the results evaluated against a water surface profile simulated using the Full-2D Lisflood-FP inertial acceleration model implementation at 30m. Using the bank height variations of the SRTM DEM, Schumann and Andreadis (2016) identified hydraulically similar sub-reaches of the Zambezi River, by computing the deviations

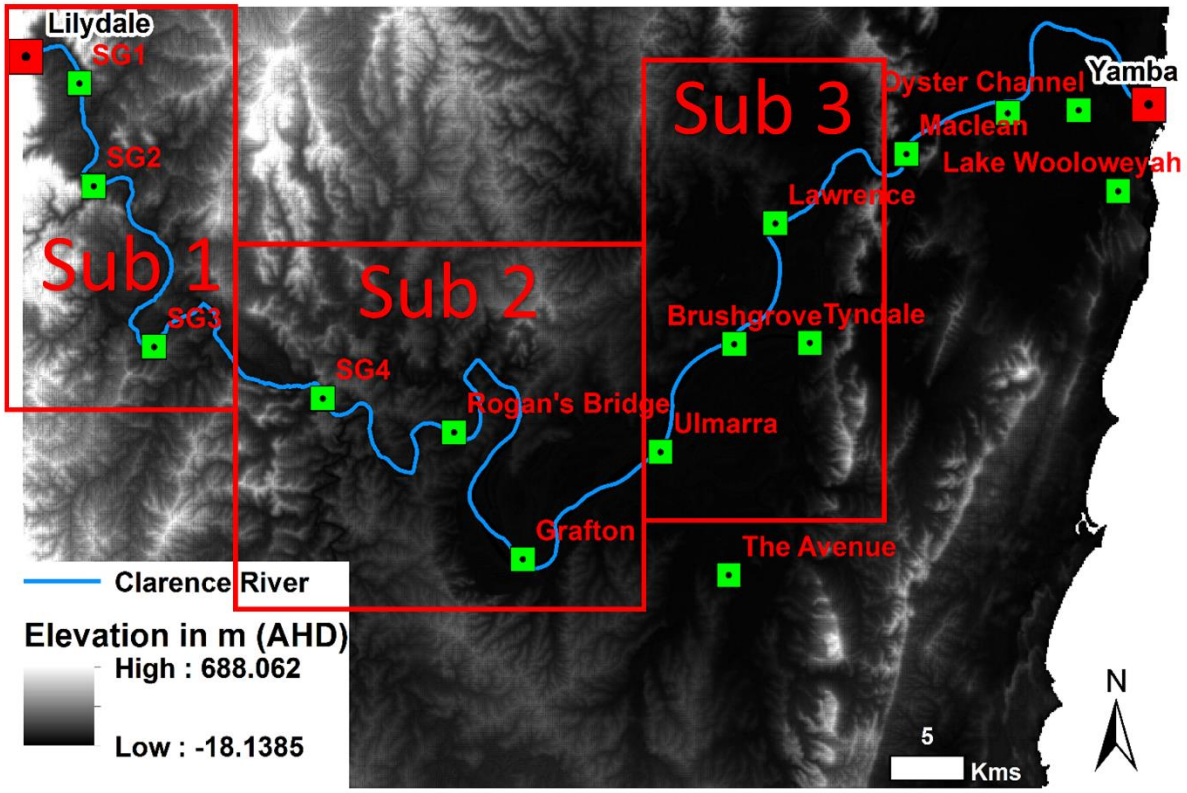

Fig. 8.1 Spatial locations of the three model sub-domains covering the three identified sub-reaches are shown in red squares, along with the locations of the real and synthetic gauges considered for performance assessment.

of the piecewise linear regression trends and from the reach linear trend. The reach linear trend was assumed to match the kinematic wave profile and the differences from this linearity were expected to represent hydrodynamic flow behaviour. In keeping with the hypothesis, the study revealed that the impacts of assimilating observations in reaches exhibiting more diffusive flow behaviour was indeed greater than in those which could be characterized by a linear approximation. In fact, those sub-reaches which departed from the kinematic wave by more than two standard deviations were found to be ideal spots for targeted observation of better floodplain topography.

The approach of Schumann and Andreadis (2016) was slightly modified in this chapter, by replacing the SRTM-derived derived bank heights with surveyed and interpolated bathymetric data for this first-order in-channel hydraulic analysis. Elevations were sampled at every 1-km along the thalweg line, being the point of lowest elevation in a channel cross-section, for the entire modelled reach of the Clarence River. The elevations were plotted

upstream to downstream with respect to chainage, and piecewise linear regression and ordinary linear regression used to evaluate the sub-reach and reach linear trends respectively. Results of this assessment were verified against a model simulated water surface profile, and used to define the model sub-regions for the targeted observation assimilation experiment. At each assimilation time step, synthetic SAR images generated using the methods detailed in Chapter 7 were assimilated into the hydraulic model forecasts, using the flood extent assimilation algorithm of Chapter 6. Images were individually assimilated at each of the previously identified sub-reaches, and impacts evaluated for the entire domain, at the sub-domains where the images were assimilated, and at the point gauge locations.

The experimental design of this study was identical to Chapter 7, with variations only in the spatial location, timing, and interval of the observations. Identifying the optimum spatiotemporal image acquisition scenario in each sub-reach was the primary objective here. Accordingly, the particle weights were computed through local comparisons of the model simulated and synthetic observed flood extents but applied to the model as global weights as in the synthetic experiments and real case study of Chapter 7. For this analysis, observations were assimilated in all the model sub-domains identified previously, with the maximum possible frequency of 12 h starting at the rising limb of the flood hydrograph. The first test considered all the images independently, i.e. observation information was not carried over from one time step to the next and each image was treated as a single available observation of the event. This experiment allowed answering the question, where and when should a flood observation be acquired, if the satellite can be tasked with only a single image acquisition. All the images were simultaneously considered, carrying weights forward by multiplication in the multiple image assimilation case. This implied that when examining the impact of a particular image, it was assumed that all the images acquired before this had been assimilated. The reason for establishing this experimental design was to answer the question, if a richly detailed temporal coverage of a flood event was available or to be curated, what would be the optimum combination of spatial location, first visit time, and revisit interval for improved flood forecast quality.

8.2.2 Performance Metrics

The sensitivity of channel flow forecasts to the first visit and revisit frequency was evaluated through Brier Skill Scores. Plots of BSS values obtained for different first visit and

revisit combinations for all the gauges were evaluated, to identify the optimum temporal satellite acquisition scenario for each sub-reach, and also for the single image assimilation scenario. The spatiotemporal mean RMSE in water depth, (i.e. the RMSE value averaged across all the model cells and the time window under consideration) was used to quantify algorithm sensitivity to spatial location both locally and globally. More specifically, the global analysis was similar to the previous chapters where the weights assigned using the real world observation location, were used to calculate the weighted ensemble mean and subsequently evaluated across the entire domain. In contrast, the local analysis assessed the impacts only within the observed area using the same RMSE metric. This metric, albeit lumped over space and time, was chosen to illustrate the mean error trends in the spatial domains with time, which could be simultaneously plotted and inter-compared. Previous studies such as the one by García-Pintado et al. (2013), have looked at the impact of observation timing for water level assimilation on channel flow and water level forecasts. The impacts on the spatial simulation of water depths, has largely been ignored in this regard. This study therefore provides a first insight into the optimum selection of image location, timing, and imaging interval, with respect to the accurate simulation of floodplain water depths.

8.3 Results and Discussion

This section summarizes the results obtained from the observation spatiotemporal sensitivity analysis. Specifically, each section was designed to address the following research questions:

1. Is a purely DEM-based approach adequate for the characterization of reach hydraulic behaviour?
2. If only one image can be acquired for assimilation, when and where should it be acquired to obtain maximum improvements in the consequent forecast?
3. How does the first visit time and the revisit interval, with respect to reach hydraulic characteristics, impact the efficiency of flood extent assimilation?
4. How much reduction in forecast error is possible through flood extent assimilation, if observation spatiotemporal characteristics are optimized?

The following sections present quantitative results to answer each of the abovementioned questions.

8.3.1 Sub-reach Hydraulic Characterization

This section addresses the first research question, with the results from the sub-reach hydraulic characterization illustrated in Fig. 8.2. The plots show the thalweg bathymetric elevations and the water surface elevations at the corresponding locations, each 1km along the river chainage. Based on the deviation from reach linear/kinematic behaviour, as suggested by Schumann and Andreadis (2016), three sub-reaches could clearly be recognized through the DEM-based approach presented in Fig. 8.2 (a). The plot of the model simulated water surface elevations shown in Fig. 8.2 (b) corroborated the identified reaches with the identified sub-reaches retained for the further experiments presented in this chapter.

In terms of reach hydraulic behaviour, the first sub-reach (0 to ~40 km) exhibits nearly linear flow, while the second (~40 to ~90 km) and third (~90 to ~150 km) display more complex flow behaviour. Note that all distances are expressed in terms of chainage measured from the upstream inflow boundary at Lilydale. The model analysis showed considerable backwater effects in the downstream part of the reach which the DEM analysis also implied, suggesting that sub-reach identification is possible using bathymetric elevations. Moreover, the sub-reaches that were identified through the DEM- and model-based approaches were characterized by nearly similar chainage distances. This implied that the DEM-based methods outlined here have the potential to guide future assimilation studies and facilitate targeted satellite acquisition planning for improved flood forecasts.

8.3.2 Impact of Assimilating a Single Image on Forecast Accuracy

This experiment was designed to advise the timing and location of a single SAR image during a flood event for maximum assimilation impact, in terms of the catchment morphology and hydrograph dynamics. Accordingly, single SAR images were assimilated at different times across the hydrograph, with images at 12h intervals starting from 6th Jan 2011. Impacts were evaluated for assimilation in different sub-domains of the catchment in terms of Brier skill scores (BSS) for channel water level simulations at synthetic as well as real gauge locations (shown in Fig. 8.1), and in terms of the spatiotemporal mean RMSE in water depth in Fig. 8.4. The main difference between the BSS and the spatiotemporal mean RMSE both computed for the simulated water depth, is that the latter computes a spatial average over the entire domain in the global case and in the smaller assimilation sub-domains in the local case. As a

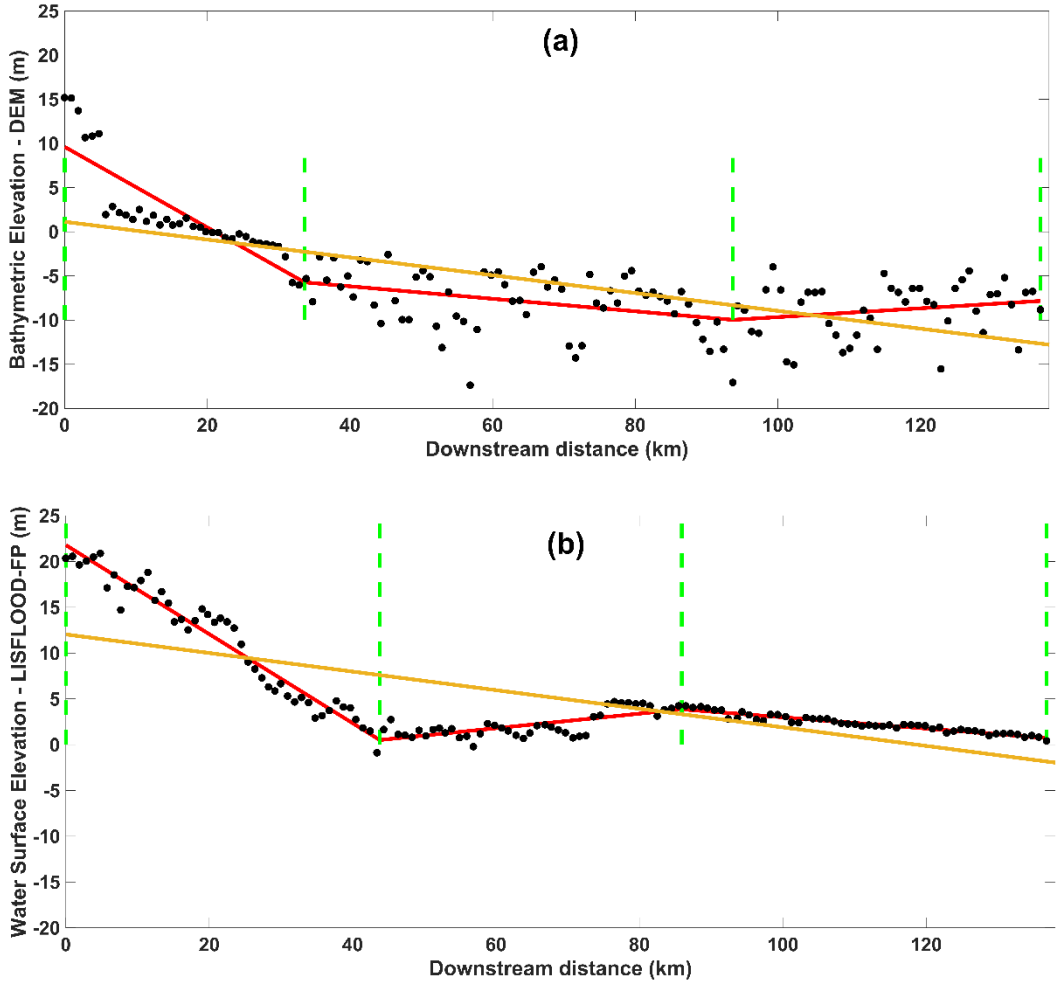

Fig. 8.2 Shows (a) thalweg bathymetric elevations extracted from a LiDAR DEM (30m) and (b) maximum water surface elevations simulated by the hydrodynamic model LISFLOOD-FP, at each 1km of downstream flow distance, along with their deviation from a sub-reach linear approximation (red lines) and the kinematic wave approximation plotted along the entire main stem of the Clarence River (mustard line).

consequence, the spatiotemporal impact of the assimilation on floodplain water level simulations for different observational configurations, could be evaluated to identify an optimum targeted observation design for the catchment which had never been previously investigated.

The BSS values for all single image assimilation time steps considered are shown in Fig. 8.3, for each of the three gauge locations falling within each sub-domain. In all the BSS plots shown in this chapter, the scores are calculated based on a discharge hydrograph comparison, which implicitly includes information on flow velocities. The water level BSS, which was also examined but not included here for brevity, exhibited nearly identical trends in terms of assimilation impacts to the discharge plots shown here. As the statistic is representative

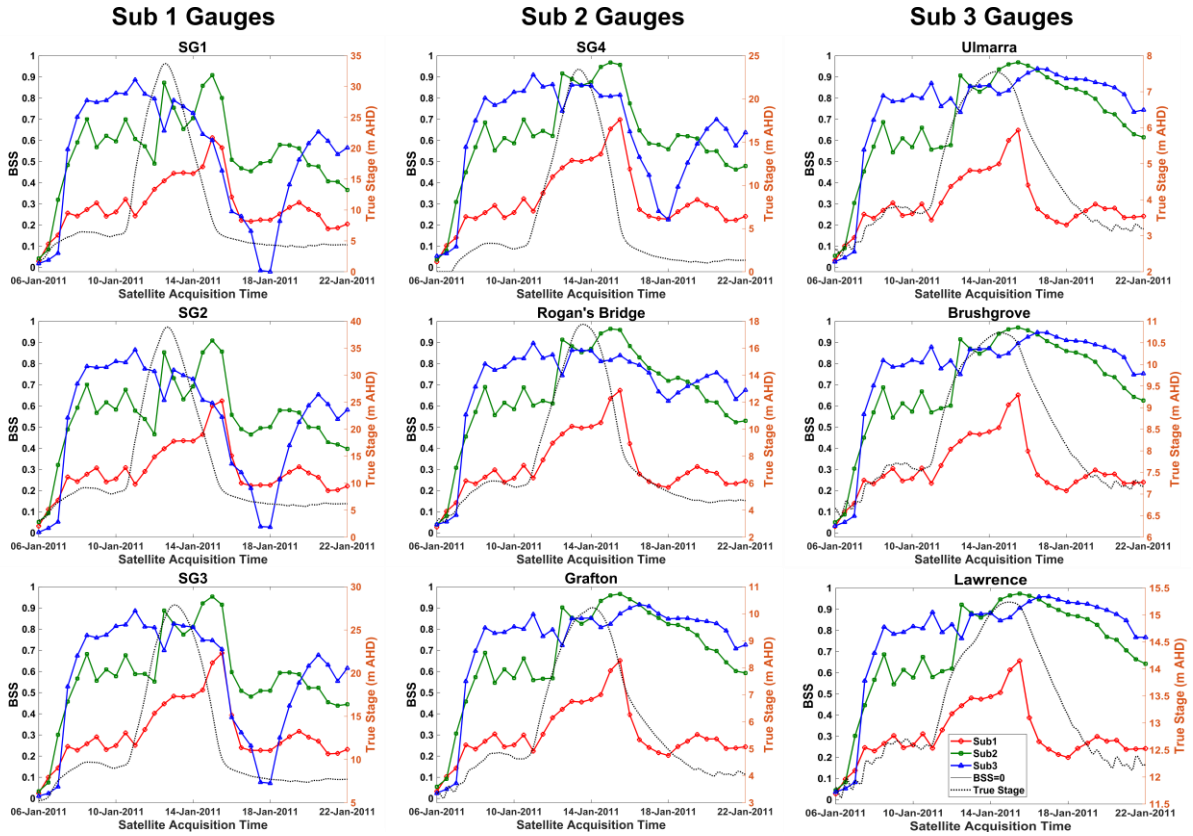

Fig. 8.3 Brier Skill Scores (BSS) obtained for single image assimilation in each sub-reach, from the time of the satellite overpass to the end of the forecast. Observations were independently considered each 12h starting from the 6th of Jan with BSS calculated at nine water level gauges along the channel (three in each sub-domain); the true stage at the location is shown in all subplots as a reference. Positive values of BSS imply forecast improvements, while negative values imply degradation and 0 implies no change from the open loop. Each point on each curve is representative of the satellite acquisition time and the corresponding BSS obtained from the time of the satellite overpass to the end of the forecast.

of the improvement in the mean squared error of the forecast water levels with respect to the open loop, the patterns observable in the relative improvement as a consequence of the flood extent assimilation at point gauge locations were quite similar to discharge as expected.

The first thing that can be observed from Fig. 8.3 is that the optimum timing of a single image acquisition varied for the three different sub-reaches. This confirmed the hypothesis that the assimilation performance is sensitive to the location of the observation. Overall, maximum improvements were observed when images were assimilated in sub-reach 2 and 3, confirming the hypothesis that the extent assimilation would exhibit larger impacts in reaches exhibiting more dynamic flow behaviour. In sub-reach 1, the flood extent assimilation resulted in limited improvement in comparison to the other sub-reaches due to a narrow constrained valley where

the extents exhibited low sensitivity to the simulated water depth, and therefore were uninformative about model performance once the valley was full. Assimilation efficiency of Sub-reach 1 remained almost constant upstream to downstream as the impacts of the assimilation were propagated through the domain by the numerical model. In contrast, when images were assimilated in sub-reach 2 and 3, the efficiency significantly declined, specifically at the upstream gauges located within Sub-reach 1. This effect was most pronounced for the images acquired at and around the channel peak in Sub-reach 3. Assimilating single images using the proposed algorithm demonstrated a large positive impact as indicated by positive BSS values. Except the images acquired in the rising limb before the flood peak, all images assimilated in sub-reach 2 and 3 after the 14th of Jan improved the water level forecast by >70%.

The impact of flood extent assimilation increased in general as the flooded area increased in the domain, as evidenced by the increasing BSS with time for sub-reach 2 and 3. In Sub-reach 1, the assimilation efficiency increased with time until just after the flood wave traversed the reach and then declined rapidly as the valley filling occurred. The overall assimilation improvement for each image increased from upstream to downstream in sub-reach 2 and 3, while the converse was true for sub-reach 1. In terms of the timing, images acquired at and around the peak seemed to have the largest improvement on the forecasts, with this effect increasing from upstream to downstream. For instance, the assimilation of the images at the peak in Sub-reach 2 produced larger improvement at the gauges located downstream in Sub-reach 3 than for those located within the sub-domain or in Sub-reach 1. The maximum improvement (~95%) using a single image was observed in Sub-reach 2 when an image was assimilated ~24h after the peak, which was expected due to the dynamic flow behaviour of the reach while being independent of backwater effects. Images at and after the peak consistently produced large positive impact in line with expectations, as the inflows and the error added at each consecutive time step is also diminishing. This allowed the assimilation to a have a more consistent positive impact, as the errors became dominated by the flood propagation mechanism of the model thereby allowing the selection of well-performing particles across the domain.

The spatial water depth RMSE evaluation shown in Fig. 8.4 revealed a slightly different pattern. The open loop errors peaked when the flood peak entered the domain at Lilydale, and the images assimilated in Sub-reaches 2 and 3 when the flood peak was traversing Sub-reach 1 had the largest improvement in water depth errors across the entire domain (Fig. 8.4 (a)). Between 7th and 13th of Jan, the assimilation of flood extents in Sub-reach 3 had the largest

improvement followed by Sub-reach 2 and 1, respectively. After Jan 13th relative differences in the assimilation impacts of Sub-reach 2 and 3 became nearly negligible, while still remaining notably larger than those observed in Sub-reach 1. This finding corroborated the initial hypothesis and the findings from the gauge assessment, which highlighted that flood extent assimilation impacts were more significant in reaches exhibiting more diffuse flow behaviour. As the floodplain neared valley filling towards the end of the flood event, the impact of the assimilation on the spatial simulation of water depths also decayed rapidly, due to reduced sensitivity of extents to the model diagnostic states. This is in contrast to the observations from the point gauge locations in Fig. 8.3, where images after the peak towards the end of the flood event continued to exhibit large improvements in Sub-reach 2 and 3 in particular.

For the local case shown in Fig. 8.4 (b), where the water depth RMSE was locally averaged across the specific model sub-domains used for the assimilation, the findings remained consistent with the global case. The magnitude of local improvements obtained through assimilation in Sub-reach 1 were lower than in the global case. Moreover, for Sub-reach 1 and 2 the local impacts became negligible after Jan 15th. It is interesting to note that Sub-reach 2 continued to demonstrate a global spatial improvement on the water depth RMSE, even as the local improvement decayed to zero, matching the global performance of Sub-reach 3. Assimilation in Sub-reach 3 led to both local and global spatial improvements in RMSE consistently, starting from the very early images. In fact, the early images seem to be able to better constrain the spatial errors than those acquired after the peak. The largest improvements

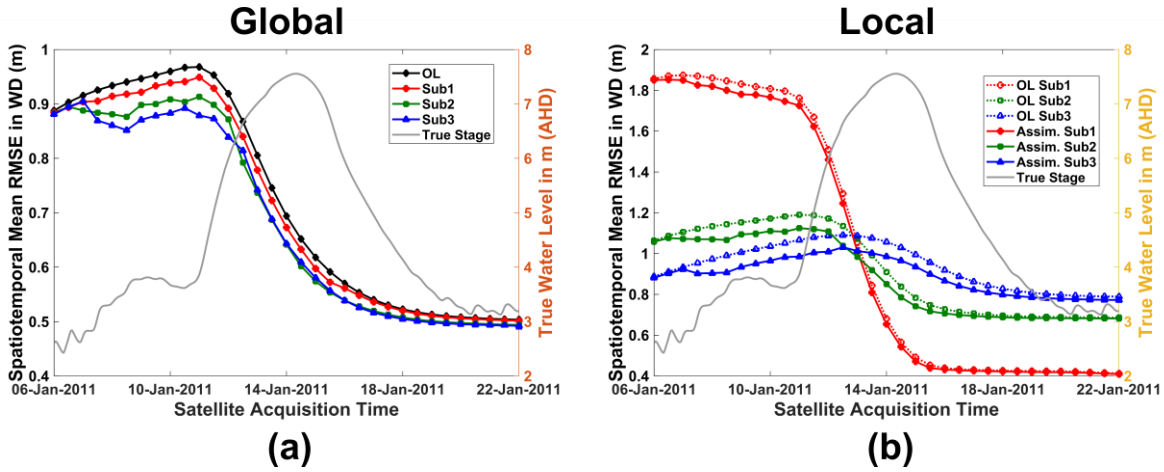

Fig. 8.4 As for Fig. 8.3 but for the spatiotemporal mean RMSE in water depth shown in (a) for the global RMSE averaged across the entire model domain and (b) for the local RMSE averaged within the model sub-domains used for the assimilation.

were observed in Sub-reach 3, followed by Sub-reach 2 and 1, where the improvements were only marginal in the latter case. A maximum reduction of up to 20cm in the water depth RMSE was obtained through early assimilation in Sub-reach 3, which could be quite significant from an emergency management perspective.

8.3.3 Optimizing Multiple Image Assimilation

This section identified the optimum observation design for this catchment in order to achieve maximum improvement in forecast accuracy through the assimilation of observed flood extents. Specifically, assessments were made with respect to reach flow behaviour to allow for generalization across reaches exhibiting similar flow dynamics. The following research questions were investigated:

1. Which images have the largest positive impact until the next assimilation time step, both spatially for the floodplain and for the channel?
2. How does the assimilation efficiency (the improvements resulting from the assimilation) vary with respect to revisit intervals when weights are carried forward by multiplication?
3. What is the temporal correlation length for observations starting from different revisit times i.e. for how long into the future does the observation continue to have a positive impact?
4. When does the consideration of multiple images produce the largest positive impact?

8.3.3.1 Maximum Improvements between Assimilation Time Steps

This section addresses the first research question, regarding the spatial and temporal location of the best performing image until the next assimilation time step. Only the impacts on the forecast until the next assimilation time step has been computed here for the 12h revisit case. As the previous section already demonstrated possible improvements for longer lead times by calculating the evaluation metrics until the end of the flood event, this section focused on the performance of single images for a multiple image assimilation scenario. Metrics used for the assessment were also identical to the previous section, where within channel flow rate performance was evaluated through Brier Skill Scores at point gauge locations shown in Fig. 8.5. Spatial performance was quantified through spatiotemporal mean RMSE water depth computed globally (across the entire domain) and locally (within each assimilation sub-domain), as in the previous section and throughout this chapter, illustrated in Fig. 8.6 (a) and (b), respectively.

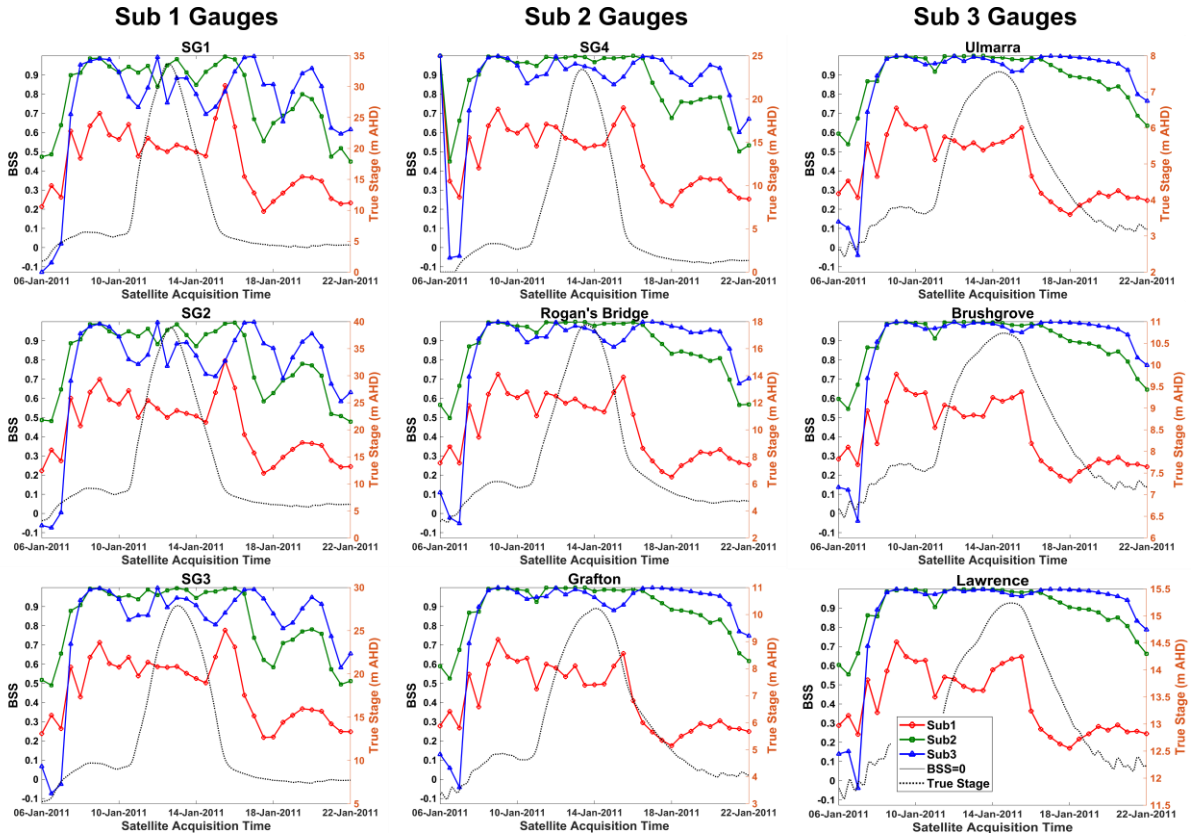

Fig. 8.5 As for Fig. 8.3 except for the time window used for the BSS calculation. Here BSS is calculated from the assimilation time until the next image becomes available. As images are considered every 12 hours, this time window is restricted to 12h after each assimilation time step.

The first observation from Fig. 8.5 is that the magnitude of improvements was significantly larger for a lead time of 12h in comparison to the full forecast impacts, as expected. The difference obtained in BSS values when assimilating images in Sub-reach 2 and 3 decreased to almost zero near the channel peak and then increased again as the floodplain inundation neared the maxima in Sub-reach 3. Assimilation in Sub-reach 1 consistently exhibited lower efficiency than in the other Sub-reaches, which was also expected, due to the dominant kinematic flow controls in the region. In other words, the open loop model forecast already performed quite well in this region and therefore it was difficult to obtain relative improvements. Moreover, the prognostic extents quickly lost sensitivity to the diagnostic water depth variable in this region, as the valley filled with the threshold amount of water used for the model flood extent calculation. As in the previous case the impacts of assimilation were propagated from upstream to downstream, thus explaining the nearly identical BSS curves obtained for Sub-reach 1 at all the gauges. In contrast, the efficiency of assimilating in Sub-reach 2 and 3 reduced at the gauges upstream of the assimilation locations. For instance, the

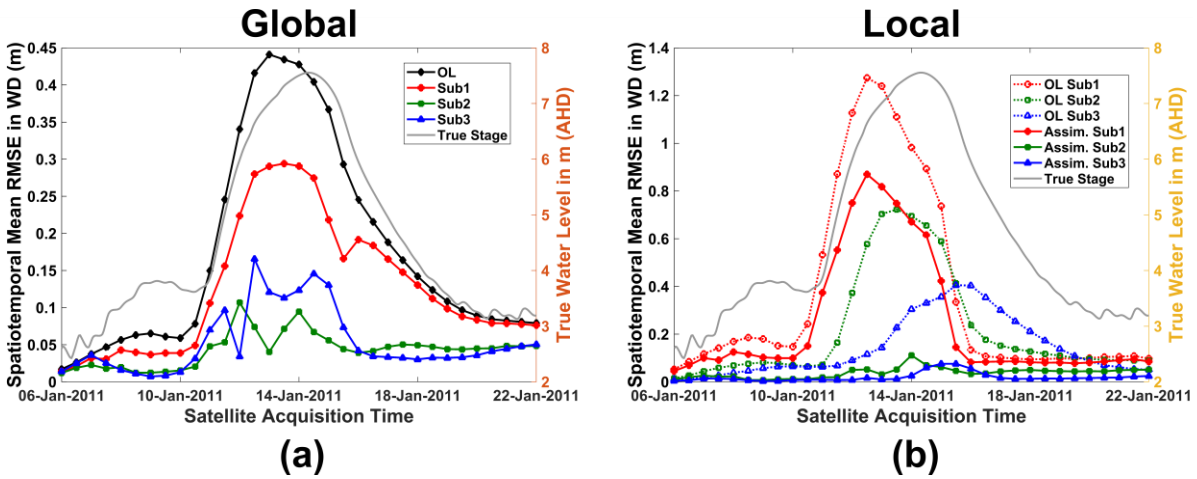

Fig. 8.6 As for Fig. 8.4 but for the time window used for the calculation of the spatiotemporal mean RMSE. Here, the time window used is the same as in Fig. 8.5, i.e. the 12h between one assimilation time step to the next.

impact of assimilating in Sub-reach 2 remained consistent between the gauges located in Sub-reach 2 and further downstream in Sub-reach 3, while the improvement on upstream gauges was reduced. Most of the images considered in Sub-reach 2 and 3 between 9th to 20th January, produced improvements in the mean squared errors of the assimilated forecast of >90 over the open loop (BSS>=0.9, see Fig. 8.5), indicating the feasibility of the proposed methods to reduce flood inundation forecast errors.

The plots of the spatiotemporal mean RMSE in water depth, illustrated in Fig. 8.6, revealed that the largest global improvements were produced when assimilating in Sub-reach 2, followed by Sub-reach 3 and 1. On 13th Jan, just before the flood peak arrived in Sub-reach 2, the global spatial RMSE in WD reduced by ~40cm on averaged across the ~550,000 modelled cells. In fact, even for assimilation in Sub-reach 1 the maximum error reduction was ~20cm, highlighting that the assimilation of flood extent can effectively constrain the forecast. Maximum relative improvements were observed when the open loop exhibited higher errors, being from the time that the flood peak entered the domain, as also evident from Fig. 8.4. Local improvements were again maximum for Sub-reach 2 with a maximum reduction of ~70cm of error observed at the peak. Initial assimilation in Sub-reach 2 and 3 reduced the local average RMSE in water depth to almost zero until the next assimilation time step, suggesting that the 12h lead time considered here falls within the window of system memory. Images considered after the peak in Sub-reach 1 had negligible impact on both, the global and the local water depth RMSE values. The optimum timing of images for assimilation that emerged from a spatial evaluation of the forecast, was different from what was identified from the BSS plots at the

point gauge locations. Improvement in the spatial objective function was maximum when the flood peak entered the domain and the distributed errors were the largest, i.e. images temporally located before and at the peak with respect to the inflow hydrograph. In contrast, those images the assimilation of which resulted in maximum improvements for the water level forecast at the gauges, were temporally located post peak with respect to the location specific hydrographs within the assimilation sub-domains.

Images assimilated between 12th and 16th of January demonstrated a strong positive impact irrespective of the spatial location of the assimilation, although the impact was notably greater in the reaches exhibiting more hydrodynamic flow behaviour. In fact the images that performed best were those in which the extent varied the most between time steps, being when flows are beginning to transition from the channel to the floodplain. The results also support the hypothesis that the optimum temporal acquisition scenario would be different for each sub-domain; specifically, in relation to the location of the assimilated image with respect to the catchment and reach hydraulic behaviour. Hydrodynamic sub-reaches with highly variable flows which are dominated by the errors of flood propagation but not significantly influenced by backwater effects, stand to benefit the most from flood extent assimilation efforts and should be targeted for observations.

8.3.3.2 Impact of Revisit Interval on Multi-image Assimilation

In this section multiple images were simultaneously considered, starting from different first-visit times across the flood event by carrying weights forward through multiplication. Moreover, the first-visit times were sequentially shifted forward by one revisit interval. For example, for the 12h revisit case first visit times started from 6th Jan 2011, with 33 images assimilated simultaneously, to 22nd Jan 2011 where only one image was assimilated. This experimental design was chosen to investigate the optimum combination of spatial location, first visit time, and revisit interval for improved flood forecast quality, when a rich satellite image database is available. The following sections examine the impacts on channel performance through Brier Skill Scores and on floodplain water depth simulation through the spatiotemporal mean RMSE water depth as in the previous sections.

The BSS plots for the multiple image assimilation scenarios, with different first visit and revisit intervals of 12h, 24h, and 48h are illustrated in this section as Fig. 8.7, Fig. 8.8, and Fig. 8.9, respectively. It is evident from these plots that the assimilation of flood extent was

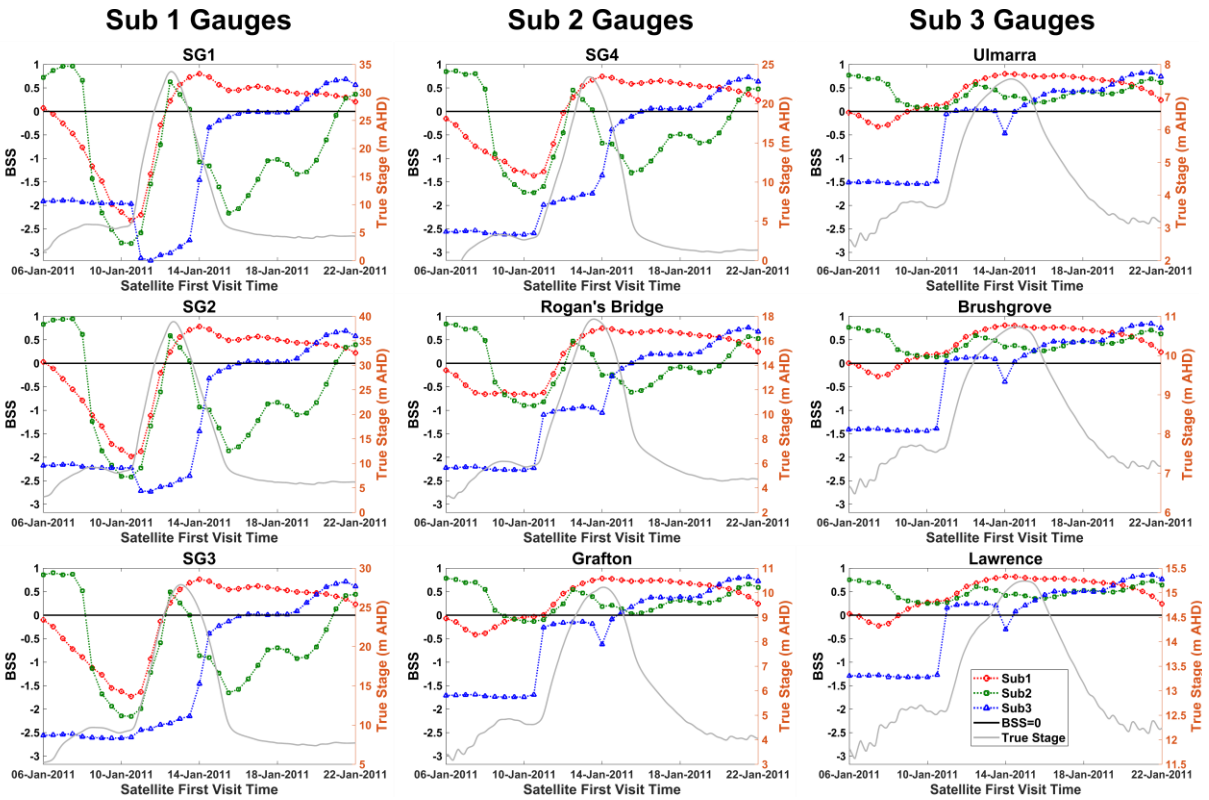

Fig. 8.7 As for Fig. 8.3 but for the multiple image assimilation case with a revisit interval of 12h and weights carried forward by multiplication. Each point on each curve corresponds to the first visit time and the BSS obtained

extremely sensitive to first visit time and to revisit intervals in this multiple image scenario, as particle weights are carried forward through multiplication. The plots represent the BSS values obtained when all images for the particular revisit frequency up to 22nd Jan were assimilated; the abscissa of each point corresponds to the first visit time, starting from which multiple images were considered. This implies that for each consecutive point on the curves the number of observations is one less than the predecessor. For the 12h revisit scenario in Fig. 8.7, a total of 33 images were considered simultaneously at the first point of the curves on 6th Jan in each sub-reach, although the first improvements were only observed when images were assimilated starting from of 9th Jan. The positive impacts of the assimilation increase from upstream to downstream. For instance, at the gauges of Sub-reach 1, the first positive impacts were observed only when images starting from and after the peak on 13th Jan were assimilated.

Interestingly, the impacts of the assimilation were most consistently positive for assimilation in Sub-reach 1, implying a longer correlation window between extents in that region. This seems likely as the shape of the valley in that particular sub-domain does not allow for high variability in simulated extents over time. At Sub-reach 1, improvements from the

assimilation seemed to increase because of using multiple images in comparison to the single image case. In contrast, the single image assimilation worked better in Sub-reach 2 and 3, where the flood extent is more variable and uncorrelated between time steps, with different particles performing well at different times. Simply put, the forward multiplication of weights does not work well as different particles perform better at different times throughout the flood event. Moreover, these results establish that assimilating images at all times and locations of the model domain does not guarantee a better result, and can even be significantly detrimental.

The results for the 24h and 48h assimilation frequency are shown in Fig. 8.8 and Fig. 8.9, respectively. Interestingly, the reduction in assimilation frequency from 12h to 24h and 48h did not have a large impact on the timing and magnitude of the maximum efficiency achievable through the assimilation. While the general pattern of the curves obtained for all of the sub-reaches remained generally the same, the magnitude of maximum BSS for all the three Sub-reaches declined. Moreover, the difference in the magnitude of BSS observed for different imaging frequencies was greatest for first visit times before the peak and decreased in the post peak images. The limited sensitivity to revisit frequency after the channel peak evident here, also corroborates the findings of previous studies investigating the optimum timing for SAR-

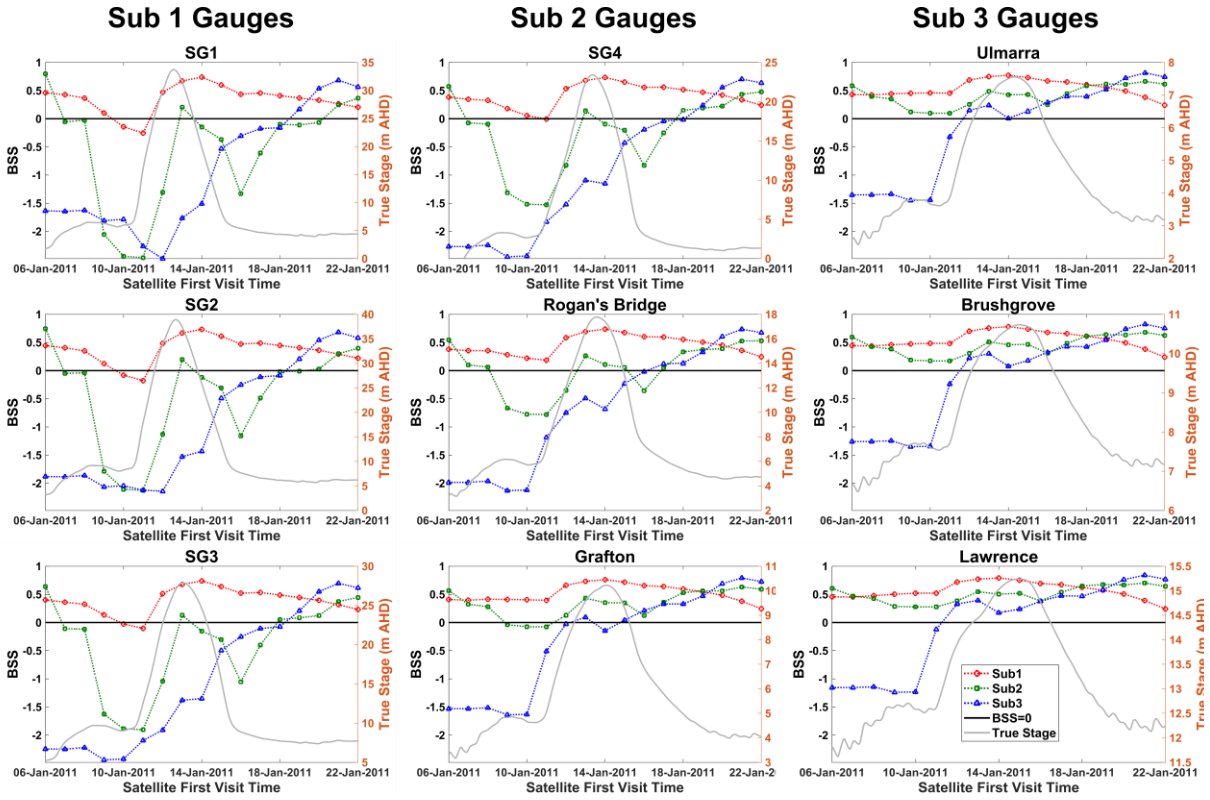

Fig. 8.8 As for Fig. 8.7 but for a revisit interval of 24h.

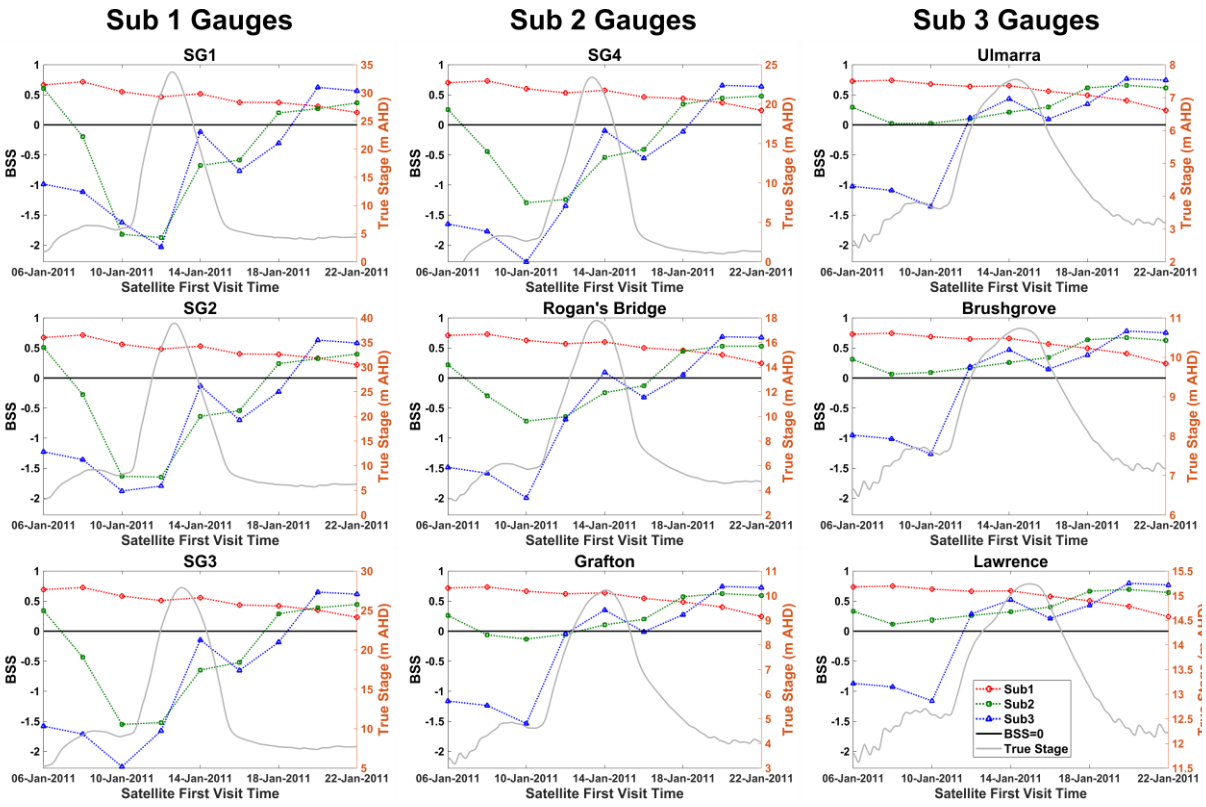

Fig. 8.9 As for Fig. 8.7 but for a revisit interval of 48h.

derived water level assimilation (García-Pintado et al. 2013). Assimilation impact for images considered within Sub-reach 1 was fairly consistent across all gauges in the domain. Images assimilated in Sub-reach 2 generally had a positive impact at the downstream gauges. However, at the gauges within the sub-domain or those located upstream, the impact of assimilating multiple images together at Sub-reach 2 and 3 was almost consistently negative. This is primarily due to the fact that the particles that are selected based on assimilation in these hydrodynamic sub-domains, do not perform well in the topography controlled sub-reach, implying that different particles perform better at reaches exhibiting different flow behaviour.

8.3.3.3 Observation Correlation Length

This section investigated the maximum observation correlation length for flood extent assimilation. Here correlation length is defined as the number of different observations which continue to select the same particles from the ensemble, implying that the observational information is correlated forward in time as some particles consistently perform well. This results in a larger positive impact on the forecast accuracy when the images are considered together through forward weight multiplication, as the weights of the well performing particles

consistently increase. The maximum correlation length is then the number of observations which produce the maximum improvements in forecast accuracy when weights are multiplied forward. In other words, maximum correlation length is defined as the set of observations which produce the maximum improvement when the same well performing particles are repeatedly selected through the assimilation and assigned higher weights as a consequence of the forward weight multiplication. The difference between the observation correlation length and the maximum observation correlation length is simply that the latter case refers to the maxima observed in the former case. More simply, if the correlation between images from one time step to the next was temporally plotted, the number of observations which resulted in the largest value obtained for the correlation would be the maximum observation correlation length.

Fig. 8.10 (a) and (b) illustrate the observation correlation length and maximum observation correlation length, respectively, with respect to the different first visit times considered throughout this chapter. For the purpose of this investigation, forecast improvements were quantified through Brier Skill Scores, from the time of the assimilation until the next time step. The duration of the time step following the assimilation varied according to the revisit intervals considered (ie. 12, 24, and 48h). Observation correlation lengths shown here thus quantify the maximum number of images that demonstrate a positive Brier Skill Score over the time window following the assimilation. Likewise, the maximum correlation length corresponds to the number of images where BSS is maximized. Note that the y-axis of the plots are defined in terms of the revisit intervals, therefore, observation correlation in terms of number of days can be directly interpreted from the second row of plots or by dividing/multiplying by the revisit interval. For example, for revisit intervals of 12h and 48h, the correlation length in terms of number of observations needs to be divided or multiplied by two respectively, in order to arrive at the time in number of days.

The first thing that becomes clear from Fig. 8.10 is that the observation correlation length differs significantly from maximum correlation length for most times examined. With the exception of very early first visit times, the images being assimilated continued to have positive impacts for all the images considered together. However, as this plot does not highlight the magnitude of the BSS, it is reasonable to assume from the maximum correlation plots that the assimilation efficiency increased to a local maxima and then started to decline again.

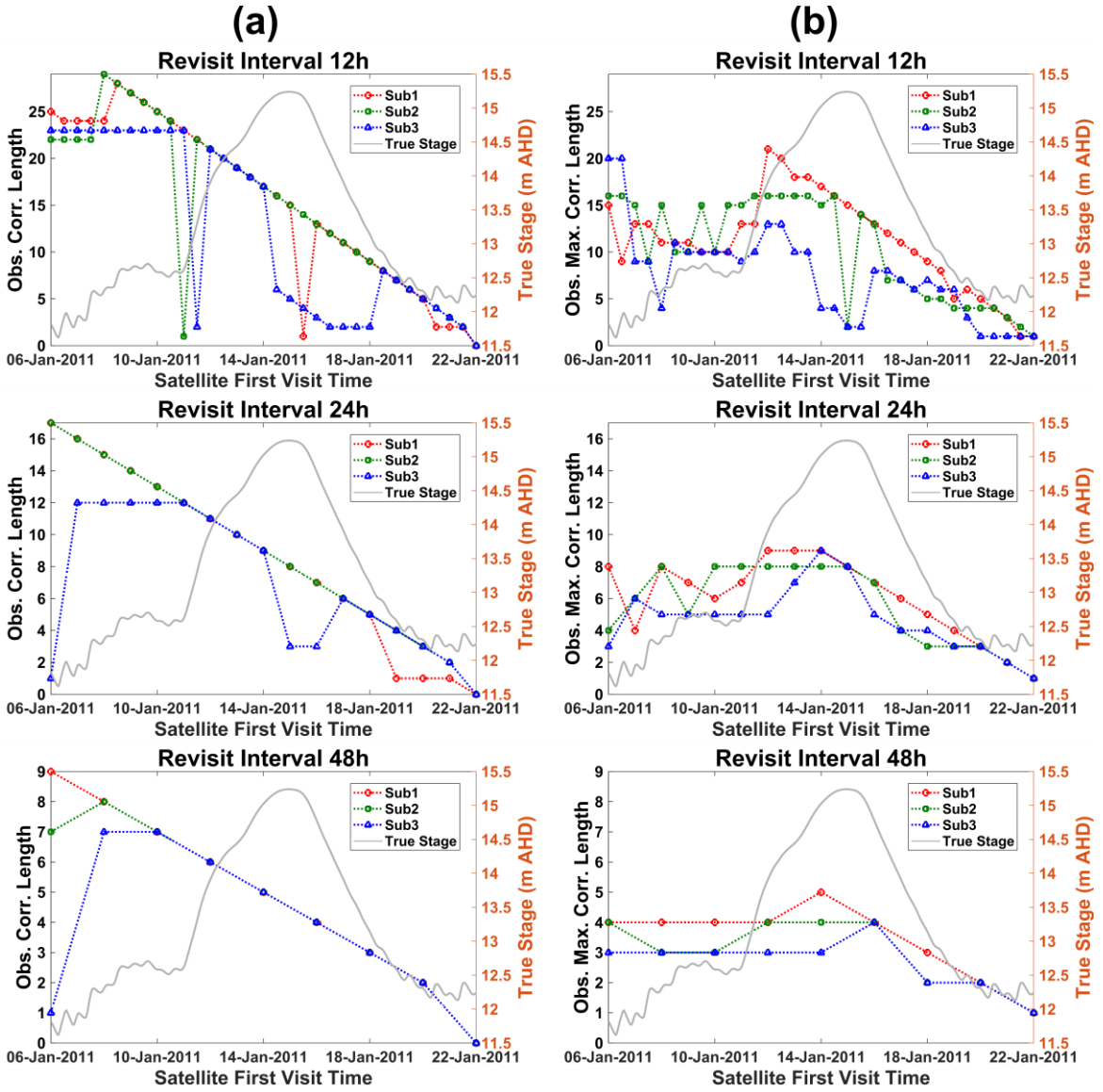

Fig. 8.10 Observation correlation lengths with respect to different first visit times and revisit intervals are shown in the left column (a) of this plot, with correlation length defined as the number of images for which sequential assimilation with weights carried forward through multiplication had a positive impact. Positive impact was defined in terms of the next 12h BSS as calculated in Fig. 8.5, with the main difference being that the weights for each image were multiplied forward in time. The right column (b) shows the number of images after the first visit at which the maximum improvement in the BSS was observed. The true stage at Lawrence is shown as a reference in all the sub-plots.

Correlation length also varied with respect to the revisit interval under consideration, as well as the spatial location and timing of the observation. For instance, when considering a revisit interval of 48h, correlation length extended to the maximum number of images considered in the experiment, for all first visit times evaluated in Sub-reach 1 and starting from

10th Jan for all the sub-reaches. However, when the revisit frequency increased such that multiple observations covered the highly variable transitioning flows during and after the channel peak, different particles were selected which did not perform equally well. Consider the 12h revisit case, in which the correlation length for Sub-reach 3 dropped quite noticeably for observations with first visit times between 14th and 18th Jan, i.e. during the channel peak. This effect reduced to just the 15th and 16th Jan for the 24h revisit and completely disappeared for the 48h case, as the number of observations during the channel peak decreased. In Sub-reach 2, the correlation length fell to just 12h for the first visit on 11th Jan, being when the stage in the channel exhibits a steep increase within the next assimilation interval, indicating that the selected particles did not continue to perform well. Interestingly, the maximum correlation length for the same first visit time in column (b) was significantly longer ~7.5 days, indicating that the maximum positive effects from the multiple image assimilation could follow even after the first negative impacts were observed.

For first visit times starting in the rising limb, maximum improvements were consistently observed in the falling limb, especially for Sub-reaches 2 and 3 exhibiting more hydrodynamic flows. In Sub-reach 1 where flows are predominantly kinematic, observation correlation length was higher as the extent contains effectively the same information after valley filling. The correlation length in this reach was also nearly equal to the maximum correlation length after the 12th of Jan, implying that the assimilation efficiency increases as more images are considered together with weight multiplication. This is also corroborated by the multiple image efficiency figures in the previous section, illustrating that the BSS in this reach was close to 1 for all first visit times starting from 12th Jan. A similar trend was not evident for the other two sub-reaches exhibiting more dynamic flows, as the transferability of observational information from one time step to the next was rather limited. In general, the maximum correlation length was longer in Sub-reach 2 than in Sub-reach 3 for all examined revisit intervals, specifically for first visit times starting in the early rising limb. The trend was reversed for the first visit times starting at the falling limb of the hydrograph, but this was only noticeable for the 12h revisit interval case. With the exception of first visit times starting on 6th Jan in Sub-reach 3 for the 12h imaging frequency where maximum improvements were evident after ~10 days, the average time for maximum correlation was ~5-6 days in Sub-reach 3 and ~7-8 days in Sub-reach 2. As previously noted, the variability in the values of observation correlation and maximum correlation length was maximum for the 12h case where the most number of time

slices were observed during the period of most variable flows, being during and around the peak.

8.3.4 Maximum Possible Improvements through Flood Extent Assimilation

The maximum improvements obtained through flood extent assimilation are quantified in Table 8.1. Specifically, for each sub-reach and each revisit interval, the average BSS across all the gauges in the catchment was computed. The maximum value obtained was identified along with the temporal scenario under which it was observed, for both the single and multi-image assimilation cases. Interestingly, the maximum BSS (~1 indicating 100% improvement in forecast accuracy over the open loop) that was observed across all cases, was obtained through assimilating multiple (11) images with a 12h revisit frequency in Sub-reach 1. Although, it should be noted that almost similar maximum BSS values were obtained for the 12h revisit case in Sub-reaches 2 and 3, after assimilating 11 and 8 images, respectively. For Sub-reach 1, the ~100% improvement in the assimilated forecast was only achieved during the 12h revisit case, while for the other revisit intervals the efficiency declined with increasing gap between observations. However, for Sub-reach 2 and 3 the maximum improvement achievable in the forecast after assimilating flood extents was only marginally different between the different revisit scenarios and was >99% in all cases.

In general, the best performing first visit time for all sub-reaches was several days before the channel peak, for which the maximum improvement in mean squared errors in channel water levels represented by the BSS was observed a few days after the peak. Here the timing of the peak refers to the time at which the channel peak was observed at the hydrometric gauges located within that particular sub-domain. This implies that it would be expedient to consider when and where the most improvements in the inundation forecast accuracy are desired, in order to identify the optimum targeted observation design strategy for flood extent assimilation. In Sub-reach 1 the maximum improvement achieved through single image assimilation was only ~65% through an image assimilated in the falling limb. In contrast, a single post-peak image considered in Sub-reach 2 produced >95% improvements, while in Sub-reach 3 the assimilation of a single image in the rising limb led to >88% improvement. The difference between the assimilation of 11 12-hourly images in Sub-reach 2 and assimilating a single image was ~5%, which does not justify the expensive acquisition of 10 more images. This implies that in a budget-limited scenario, a single image assimilated at the right place (e.g. reaches with

Table 8.1 Summary table of maximum possible improvements in BSS for gauged water level simulations within the channel through flood extent assimilation. BSS values were averaged across all gauges

Sub-reach	**Reach Hydraulic Behaviour**	**Dominant Flow Control**	**Revisit**	**Max. BSS**	**First Visit**	**Time of Max. Improvement**	**No. of Images Assimilated**
I	Kinematic	Topography	12h	0.9996	08-01-2011 12:00	14-01-2011 00:00	11
			24h	0.9825	10-01-2011 00:00	16-01-2011 00:00	6
			48h	0.8536	10-01-2011 00:00	18-01-2011 00:00	4
			Single	0.6516	15-01-2011 12:00:00	Full forecast	1
II	Hydrodynamic	Inflows during the rising limb and flood propagation during the falling limb	12h	0.9992	09-01-2011 00:00:00, 10-01-2011 00:00:00	14-01-2011 00:00:00, 15-01-2011 00:00:00	11
			24h	0.9988	12-01-2011 00:00	20-01-2011 00:00	8
			48h	0.9929	14-01-2011 00:00	22-01-2011 00:00	4
			Single	0.9529	15-01-2011 00:00:00	Full forecast	1
III	Hydrodynamic	Tidal backwater effects	12h	0.9994	10-01-2011 12:00	14-01-2011 12:00	8
			24h	0.9979	11-01-2011 00:00	17-01-2011 00:00	7
			48h	0.9955	06-01-2011 00:00	12-01-2011 00:00	7
			Single	0.8846	11-01-2011 00:00	Full forecast	1

hydraulic characteristics similar to Sub-reach 2) and at the right time (e.g. post-peak) can produce improvements comparable to assimilating multiple images. The importance of the observation spatiotemporal analysis shown here, is further highlighted for catchments with significant flood risk by the finding that assimilating a single image can produce improvements comparable to multiple images.

Currently, scientists have limited control on when the image can be acquired, but an increasing number of spaceborne SAR sensors are rapidly contributing to changing this scenario. Since satellites can be tasked with the acquisition of specific flood observations under the International Charter on "Space and Major Disasters", such spatiotemporal investigations are deemed necessary to optimise the benefits from flood extent assimilation in terms of improved forecast accuracy.

8.4 Chapter Summary

This chapter examined the sensitivity of the assimilation algorithm performance to the observation spatiotemporal characteristics. Images were assimilated at different points across the flood event in both single and multi-image assimilation scenarios, with spatial coverage being for one of the three sub-reach types identified from uniform flow behaviour through a DEM-based method. For the multi-image scenario, different first visit and revisit frequencies were considered, and impacts on forecast skill evaluated through the Brier Skill Score computed at nine gauges along the main stem of the channel. Furthermore, the spatial impact of the assimilation was quantified through spatiotemporal mean root mean squared error in the forecast floodplain water depth. The experimental results from this chapter demonstrate the sensitivity of assimilation performance to the spatiotemporal observation scenario. Results illustrated herewith also highlight that based on where the improvements are desired e.g. in the channel or in the floodplain, the optimum spatiotemporal observation scenario can differ. This chapter presented results concluding the final objective of this thesis. Following this Chapter 9 implements the methods developed in this thesis to a real-world flood event in the Clarence Catchment, to assess the operational potential of the proposed algorithm.

CHAPTER NINE

"If the flap of a butterfly's wings can be instrumental in generating a tornado, it can equally well be instrumental in preventing a tornado."

- Edward Lorenz. In a talk presented at the 139th *Annual Meeting of the American Association for the Advancement of Science*, Washington DC, USA, on 29th Dec 1972. Press release published in *Essence of Chaos* (1995), Appendix 1, 181.

9. Real-data Application

Chapter 7 demonstrated the potential of the proposed assimilation algorithm to reduce forecast inflow uncertainties, while Chapter 8 established that flood extent assimilation performance was sensitive to observation location, timing, and repeat interval. Here the ability to improve the modelled flood hydrodynamics when only hydrological forecast streamflow and SAR images were available was assessed through a real-world application, by applying the proposed flood extent assimilation framework to the 2011 event in the Clarence Catchment. The fuzzy flood maps were derived using the methods developed in Chapter 4 and assimilated into Lisflood-FP in this experiment using the PF-based flood extent assimilation framework of Chapter 6. Moreover, the real water level gauges along the main stem of the channel and the aerial photographs used in Chapter 4 were used to evaluate the assimilation performance. This chapter concludes the testing of the assimilation framework proposed in this thesis, with Chapters 10 and 11 presenting conclusions from this thesis and future perspectives.

9.1 Introduction

Given that the assimilation of flood extents into hydraulic models is a novel research direction with only a few studies that have tested its potential through proof-of-concept type studies, the application to real-world scenarios has often been neglected. With the notable exception of Hostache et al. (2018), who applied a particle filter based assimilation scheme to improve flood forecasts in the Severn Catchment, UK, the direct assimilation of flood extents has been restricted to the evaluation of synthetic data and cases. For example, Cooper et al. (2019) show that the RMSE between the forecast and the truth can be reduced by more than ~95% within an idealised setting with synthetic topography, inflows, and observations, using their backscatter operator which directly assimilates SAR intensity.

However, they also acknowledge that the approach would only work well where the wet-dry pixel distributions were well separated i.e. in regions of low SAR backscatter uncertainty. While the proposed approach is interesting and provides a new outlook on the assimilation of flood extents, the restricted applicability to areas of clear separation between flooded and non-flooded pixels, makes the operational potential of the algorithm quite limited. For global applications, acknowledging the uncertainty in observations is imperative to ensure that the methods can be applied universally. Moreover, studies have shown that observation errors may have large impacts on assimilation performance, and are also acutely sensitive to the choice of the observation operator chosen (Waller et al. 2018b).

Synthetic studies are undoubtedly important to assess the potential of newly developed assimilation strategies in a controlled environment. However, real-world applications reveal the capability of algorithms to deal with the myriad of uncertainties that are added when dealing with actual observations, inputs, and boundary data. Accordingly, this chapter applies the proposed assimilation algorithm to a real flood event in the Clarence Catchment, after verifying its potential and sensitivity to observation characteristics through synthetic experiments in Chapters 7 and 8, respectively. The open loop ensemble was identical to the one employed for the synthetic experiments, with the assumption that during a real flood event, observed inflows will be unavailable. Therefore, the positively biased streamflow forecast from a hydrological model simulated for the synthetic experiments, will be the only dataset available for the hydraulic flood inundation modelling. Consequently, the question investigated through the real-world application presented in this chapter was, how close to the observed flow dynamics is it possible to get if flood extents are assimilated as and when they become available. The main difference from the synthetic experiments lies primarily in the extent and source of the assimilated observations, which in this case were the actual Cosmo-Skymed SAR images with limited spatial coverage (See Fig. 9.1). The following sections summarize the main findings from this experiment.

9.2 Results and Discussion

Results from the real-world application of the proposed assimilation strategy are presented here. In this experiment, real SAR observations of flood extents having limited

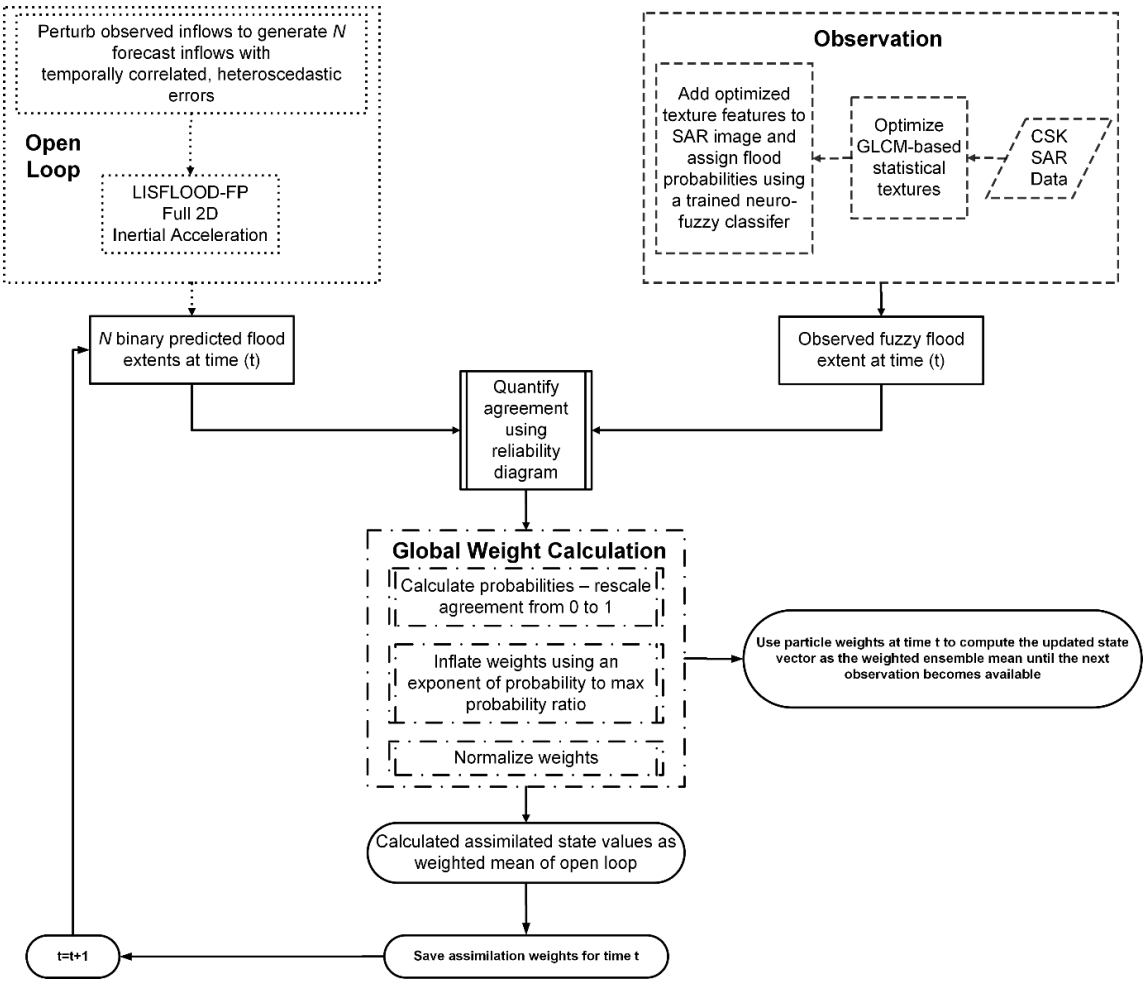

Fig. 9.1 Schematic of the real-world application of the proposed assimilation algorithm.

spatial coverage were used. The SAR images were converted to fuzzy flood maps using the approach outlined in Chapter 4, prior to their assimilation into the hydraulic model Lisflood-FP. This facilitated an evaluation of the proposed assimilation strategy for operational forecasting applications, by demonstrating its potential through an application to a real-world flood hindcasting problem. Forecast performance was first evaluated in terms of flood extent, using the optical aerial photographs described in Chapter 3. However, only the impact of assimilating Image I could be evaluated due to the timing of the aerial image acquisitions. This was followed by an assessment of the impacts on water level estimates within the channel using six hydrometric gauges along the main stem of the Clarence River available to this study (Locations shown in Fig.7.7). The satellite and aerial optical image coverage has been shown in Chapter 3, in Fig. 3.1 and Fig. 3.3, respectively.

9.2.1 Impact on channel water levels

The impact on channel water levels was first investigated, as several water level (WL) gauges were available along the main stem of the channel, with WL hydrographs presented in Fig. 9.2. The first observation was that the model calibration was inadequate, as the flood peak always arrived later than the observations for the open loop and the assimilated observations. This is due primarily to the calibration strategy used in this thesis, which relied solely on crowdsourced (CS) water levels. Although there is no direct relationship between the source of WL observations and the poor calibration, one possible reason for the wrongly simulated flood peak arrival time, is the lack of timing information on the CS data points. Recall that the water levels were derived from field photographs of wrack marks and high water marks, which provided maximum water depths but not the time at which they were observed. Consequently, when the maximum pixel-wise water depths simulated by the model were extracted and compared at the corresponding locations, information on when each of the model realizations should reach that depth value was not available. This explains why the shape of the hydrograph and the magnitude of the peak flow depth of the open loop and assimilated forecasts was similar to the observed values, yet differences of more than a day occurred in the flood peak arrival time.

Another thing that becomes evident from the hydrograph comparisons is that even though a positive bias was enforced on the inflows, the model does not overestimate the peak channel water level at all gauges. For instance, at Ulmarra the peak flow depth for the open loop as well as the increased value obtained after the assimilation, were lower than the observed peak flow. This indicates an underestimated Manning's, as a higher skin friction value would push the water height up. However, the lag in the arrival time indicates that the flood wave is travelling slower in the model than in the truth, which means that the friction value would most likely need to be calibrated and specified in a distributed fashion in this catchment. In fact, it appears as though the similarity observed in the channel water level peaks at Grafton, Brushgrove, and Lawrence was a function of the positively biased inflows, as otherwise the flows would have been further underestimated. Some of the observed differences in the non-uniform over- and under-estimation of flow values at the gauges, could also be explained by how the channel cross sections were specified in the model and in the real world. The temporal distance between the observed and simulated flood peaks increased from upstream to downstream as expected due to flood propagation

principles. For example, as the flood wave travelled from cell to cell in the domain and encountered the same incorrectly specified Manning's roughness, the water was further decelerated and the momentum dropped, reducing wave propagation speeds with the effects compounded from upstream to downstream.

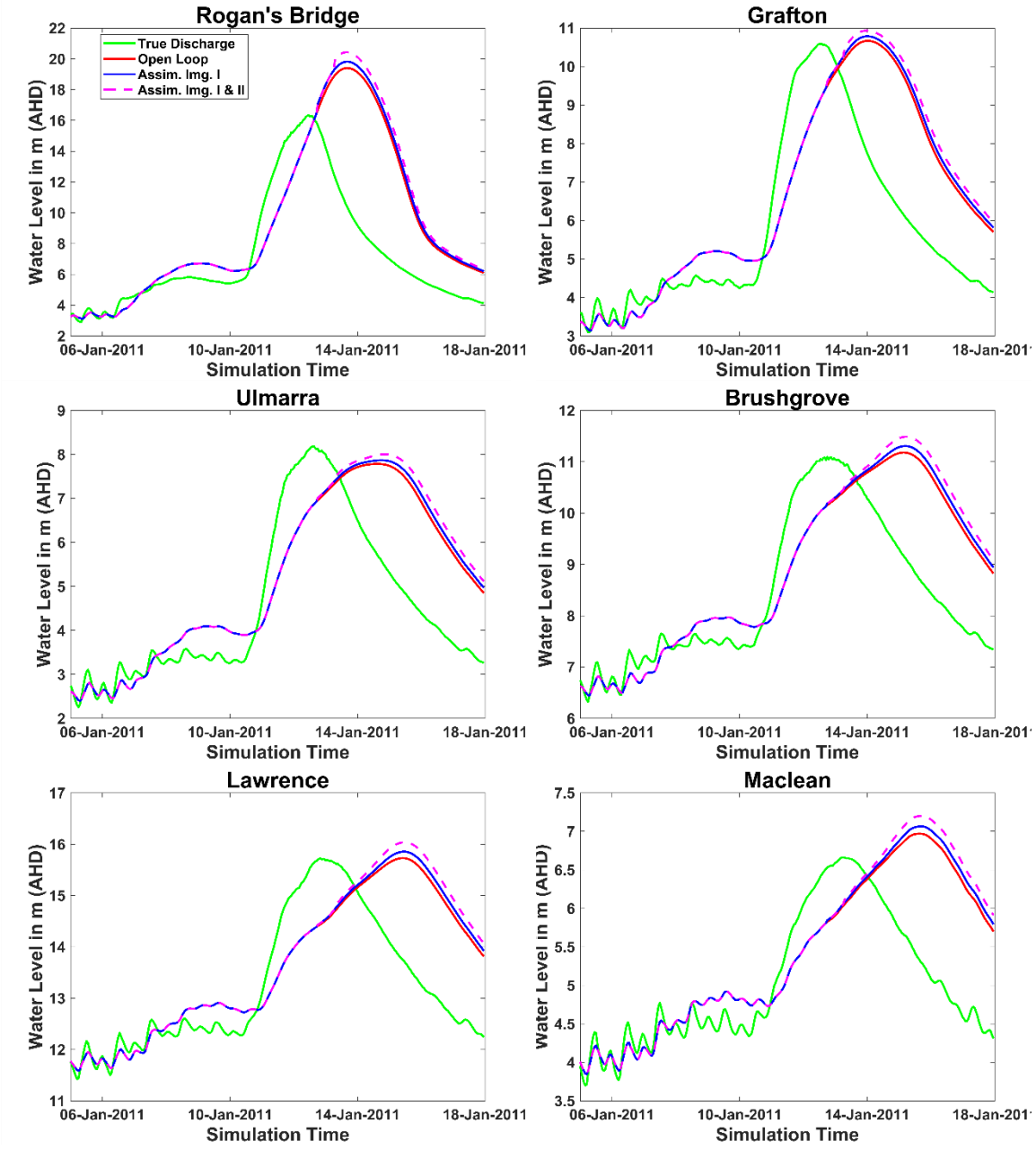

Fig. 9.2 Channel water level time series at the gauge locations along the main stem of the Clarence River.

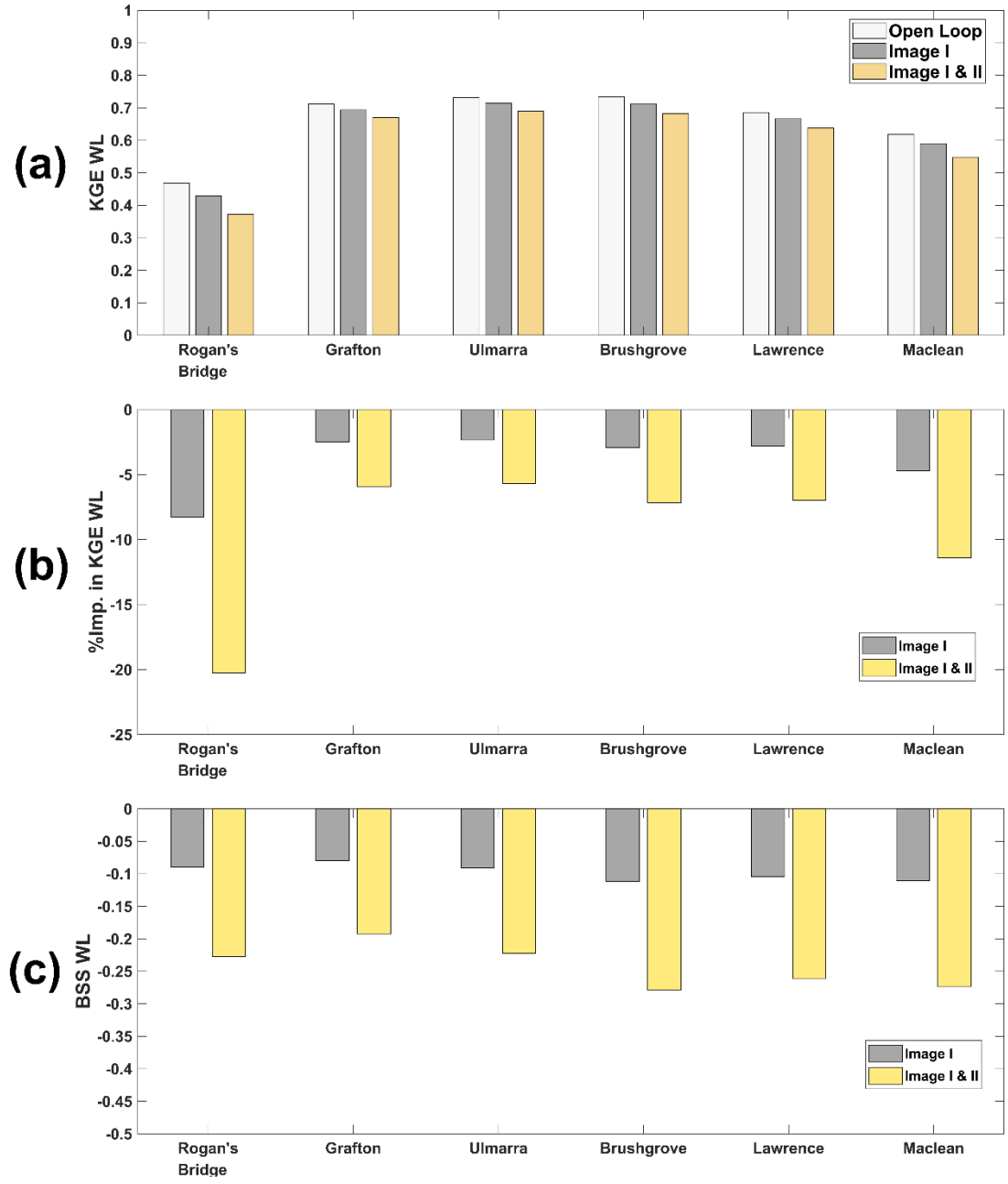

Fig. 9.3 Plots showing the (a) Kling Gupta Efficiency or KGE, (b) the percentage improvement in the KGE, and (c) the Brier Skill Scores for the six gauges along the main stem of the Clarence River.

For every gauge location examined, the assimilation increased the positive bias on the water levels. The main reason for this was that the inundation in the post-peak SAR images used for the assimilation was larger than what any of the ensemble members or particles exhibited at the corresponding time steps, as the flood peak had not arrived by that time for any of the particles. Typically, this can be assumed to be a function of an

underrepresented ensemble spread, but in this case it was a function of the time lag in the model flood wave arrival time, due to which the ensemble could not encapsulate the observation at the corresponding time steps. Accordingly, the highest weights were assigned to those particles that significantly overestimated the extent based on the observed extent. This implied that particles with higher input inflows were selected through the flood extent assimilation, pulling the model trajectory up, as evident from the channel water level hydrographs. The magnitude of assimilation impact was consistent from upstream to downstream, and the observed forecast error degradation was found to be a function of the assimilation timing with respect to the gauge specific flood wave arrival time. For example, at Rogan's Bridge where the assimilation time steps were just a few hours before the flood peak arrived, the relative degradation visible at each assimilation time step was greater than at Maclean where the arrival time was more than two days after the assimilation. There was a large difference in the magnitude of assimilation impacts between the first image and when both images were considered together (i.e. the distance between the red and the blue line versus the blue and dashed line in Fig. 9.2), which was nearly double across all gauges.

The performance metrics used for the statistical evaluation of gauged water level hydrographs – the Kling Gupta Efficiency, the percentage improvement in KGE, and the Brier Skill Scores which quantify the improvement in Mean Squared Errors – are shown in Fig. 9.3 (a), (b), and (c), respectively . The KGE values for the open loop and assimilated forecasts increased from Rogan's Bridge to Ulmarra and then started decreasing again for the gauges located further downstream. The magnitude of the degradation in the KGE resulting from the assimilation mirrored the pattern observed in the actual KGE values, and decreased until Ulmarra before starting to increase again. In contrast, the BSS plot did not exhibit clear trends with respect to the channel flow direction, implying that relative increase in assimilated forecast mean squared error with respect to the open loop was fairly constant across the domain. Average degradation in KGE values after the assimilation of the first image was ~4%, increasing to ~10% after the assimilation of both images on an average. Likewise, BSS values which can be interpreted as percentage change in MSE values, showed a degradation of ~7.5% at the first assimilation time step, which nearly tripled to ~21% after the assimilation of both images at the second time step.

In order to test the hypothesis that the main reason for the degradation was the wrongly simulated flood wave arrival time, the images were assimilated into the open loop

after accounting for the lag in the peak and the relative image acquisition timings with respect to the hydrograph. The lag was estimated by calculating the time step at which the flood extents simulated by the truth model used in the synthetic experiments best matched the observed inundation. As the open loop was based on observed inflows, this simple method helped to identify the lag in the flood peak arrival time, which was identical across all the particles due to the specification of a single common calibrated channel roughness value. Based on this analysis the lag time in inundation extent was quantified as ~50h across the catchment on average, and accordingly, the images were assimilated lagged by 50h from their actual acquisition times for this test. It is important to note that the lag was considered in only the time steps where the images were assimilated, meaning that only the timing of the particles was shifted to accommodate the observation. It is also possible for the assimilation strategy to search across a time window as well as across the particles as another way to overcome this problem, and future work should investigate the development of an automated framework to accomplish this. Plots identical to the previous figures for the gauge assessment were generated to facilitate this assessment. Specifically, the channel water level hydrograph comparisons are shown in Fig. 9.4 and the performance evaluation metrics are illustrated in Fig. 9.5.

It is immediately apparent that the flood extent assimilation produces positive impacts, confirming the hypothesis that the degradation previously observed is a function of the incorrectly simulated flood wave arrival time and not the assimilation algorithm. Model state trajectories were pulled down this time, reducing the positive bias in the forecast mean as expected. This experiment highlighted the importance of a well calibrated model and the need for updating parameters in addition to the states through flood extent assimilation. However, many studies have investigated the issue of hydraulic model calibration (Mason et al. 2003; Tarpanelli et al. 2013; Domeneghetti et al. 2014b; Wood et al. 2016; Gobeyn et al. 2017), and in this thesis the objective was to assess the potential of quantitative model calibration using crowdsourced data which had never been used before (Assumpção et al. 2018). Moreover, it was also desirable to avoid the use of the same data for the model calibration and then the evaluation of the assimilation, which could then bias the conclusions from the experiments. Furthermore, having the model perfectly calibrated would also have prohibited the assimilation from making any notable improvements, and thus restricting a more realistic evaluation of the framework.

The other option was to consider parametric uncertainty in the model ensembles, through a dual state-parameter assimilation framework. An initial attempt to include parametric and topographic uncertainties in the ensemble was made, but the ensemble size of 128 particles was inadequate for sampling across the entire state parameter space. To ensure that the spread captures the observations, the ensemble size needed to be increased substantially, which was not practically possible due to the computational expense of running hydraulic models operationally. Accordingly, the assumption that some of the particles or their weighted mean are close to the unknown "true" state failed in the case of dual state-parameter assimilation. Consequently, even the assimilation of observations was unable to nudge the forecast mean trajectory towards the truth because the ensembles did not encompass the true value. An investigation of ensemble generation methods which can optimize the sampling to ensure that the truth is included, was beyond the scope of the present research. Future research should therefore further examine this question and strive to develop optimized sampling methods, which can aid the representation of multiple state, parameter, and boundary data uncertainties within smaller ensembles. Further details on the choice of the ensemble size were provided in Chapter 7 in Section 7.2.2 Ensemble Generation.

For the lagged hydrograph comparison, assimilation of the first image produced a large positive impact but adding the second image did not further improve the state trajectory by much, as evident from the difference between the solid blue and dashed magenta lines. This is also illustrated in the gauge statistics, where the assimilation of image one produced an average improvement of ~10% in the KGE and ~20% in the MSE as shown by the Brier Skill Scores but adding image two resulted in only a marginal increase in improvement. The primary reason for this is expected to be the close temporal spacing between the images at a time when the floodplain storage capacity was nearly full. This implies that either most of the particles behave similarly in this region of the hydrograph or that the same particles were selected by the assimilation of both images. As both the images are just after the peak with only a 12h gap between them, the latter is more plausible as only a few particles can be expected to perform well during similar flow conditions. Therefore, the likelihood of selecting the same particles between assimilation time steps is quite high when flows are not changing significantly. The trends observed in the magnitude of forecast improvements measured through the KGE from upstream to downstream were identical to the patterns demonstrated for the degradation, in the assimilation without the

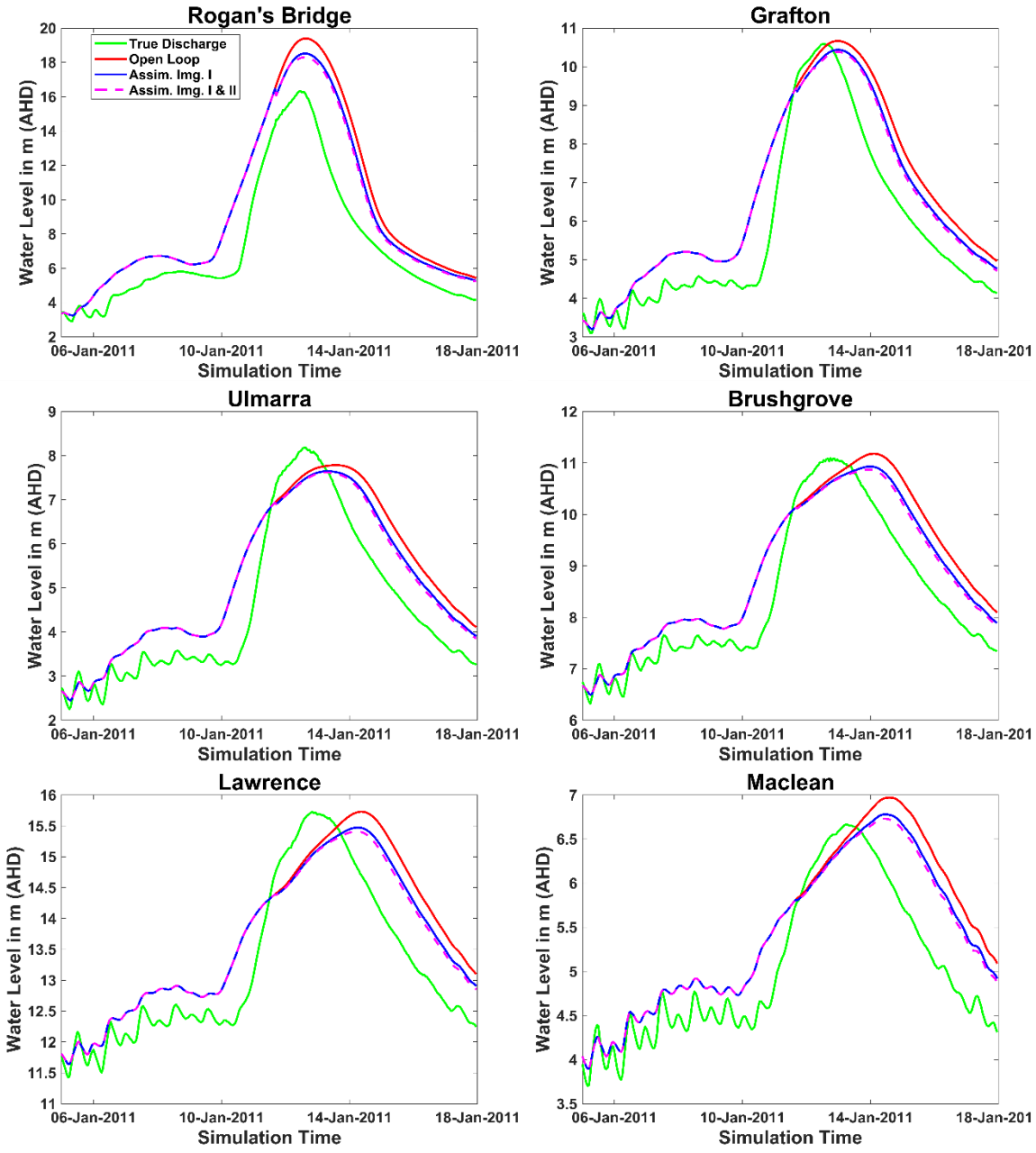

Fig. 9.4 As for Fig. 9.2 but for the flood extent assimilation performed with consideration of the flood peak arrival time lag.

consideration of the peak time lag. As in the previous case, the BSS exhibited no noticeable trends with respect to flow distance. Note that even in the lag adjusted case, the temporal distance between the observed and simulated flood peaks increased from upstream to downstream.

9.2.2 Impact on simulated inundation extent

The impact on the forecast inundation extent was assessed using aerial photographs acquired close to the image acquisition time after assimilating the first CSK SAR image. The contingency maps comparing the open loop and assimilated forecasts to the aerial photographs are given in Fig. 9.6 and Fig. 9.7. The binary flood maps derived in Chapter 4 for the reliability diagram assessment were used for the comparison to maintain consistency with the synthetic experiments, and because the model forecast extents were binary. The observed flood maps were aggregated from ~10cm to 90m to allow for a one-to-one comparison with the model cells. Note that of the 543,222 total modelled cells, the maximum wet cells for all model realizations averaged ~63,500, with the coverage of the aerial photographs being only 3,929 pixels or 0.7% of the domain.

Table 9.1 Contingency matrix evaluating the forecast flood extent maps against the corresponding extents simulated by the truth model, at the first assimilation time step.

No Lag	**Junction Hill**				**Ulmarra**			
	Open Loop		**Assimilation**		**Open Loop**		**Assimilation**	
	MF	MNF	MF	MNF	MF	MNF	MF	MNF
OF	615	2	615	2	1802	3	1804	1
ONF	465	568	543	490	502	29	507	24
CSI	**0.568**		**0.530**		**0.781**		**0.780**	
%age imp. CSI	**-6.724**				**-0.106**			
Kappa	**0.475**		**0.400**		**0.079**		**0.067**	
%age imp. Kappa	**-15.678**				**-15.099**			

OF=Observed Flooded; ONF=Observed Non-Flooded; MF=Modelled Flooded; MNF=Modelled Non-Flooded; CSI= Critical Success Index

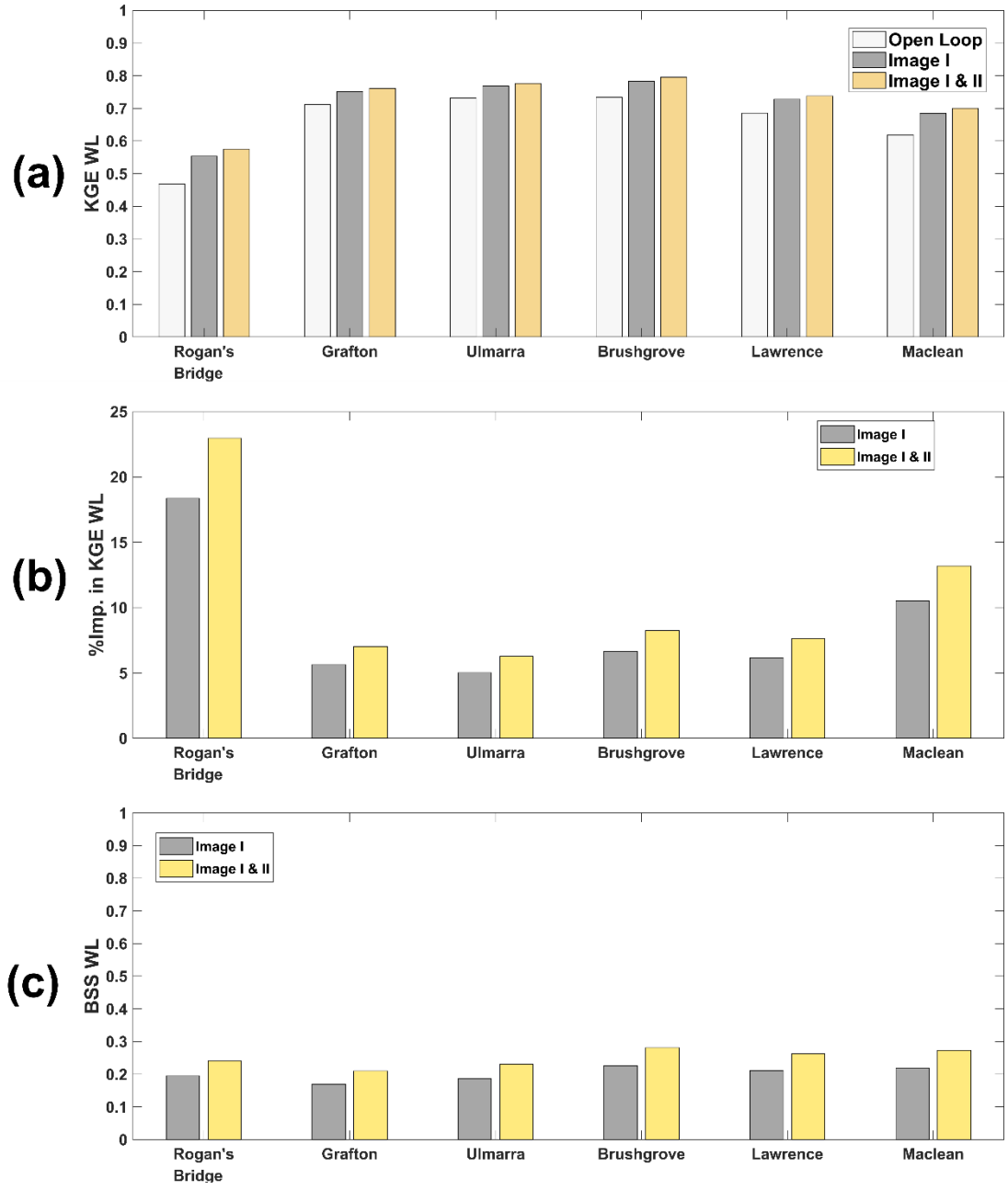

Fig. 9.5 As for Fig. 9.3 but for the lagged assimilation accounting for the delay in the flood peak timing.

The contingency maps revealed almost no visible differences between the assimilated and open loop forecasts at both Junction Hill and Ulmarra, although closer examination revealed some increase in False Alarms at Junction Hill after the assimilation. This degradation can be explained by the wrongly simulated flood peak timing as well. Due to the selection of particles with higher flows when trying to match the observed post peak

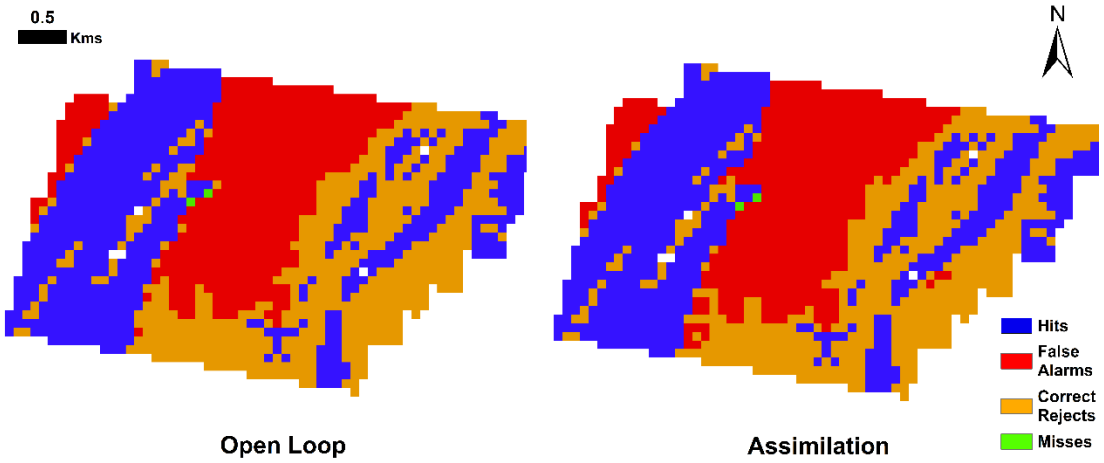

Fig. 9.6 Contingency maps comparing the forecast versus observed flood extents derived from the aerial photographs of Junction Hill, for the open loop and assimilation runs at the first assimilation time step.

images at a time when the peak had not yet arrived for most particles, the corresponding inundation extents were also larger.

Note that this discussion pertains to the models before the peak lag adjustment. This is also evident from the significant over prediction observed at both the evaluation sites, and the relative increase in overestimation of extents after the assimilation. The inundation in the open loop forecast mean even before the peak was much higher than the observed flooded area. The artificially enforced positive bias in the inflows, used to replicate the

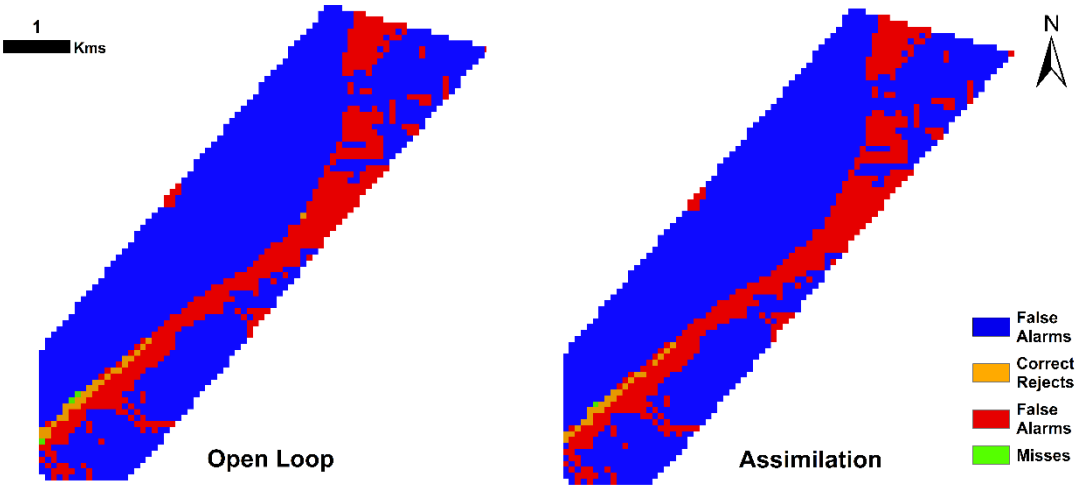

Fig. 9.7 As for Fig. 9.6 but for the Ulmarra region.

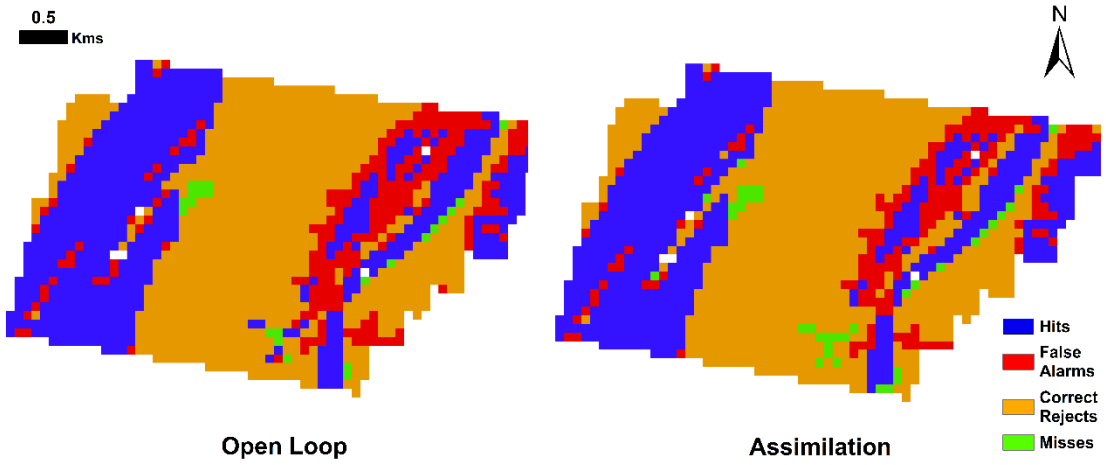

Fig. 9.8 Contingency maps comparing the forecast versus observed flood extents derived from the aerial photographs of Junction Hill, for the open loop and assimilation runs at the first assimilation time step after considering peak time lag.

typical nature of errors observed in forecast streamflow obtained from coupled numerical weather prediction and hydrological models, led to a large number of false alarms in the open loop.

Examining the pixel statistics from the contingency matrix presented in Table 9.1 it can be observed that the assimilation had a negative impact for both evaluation sites. At Junction Hill, the values of the CSI and Kappa statistics were low and quite close to each other. The assimilation of image one introduced degradations of ~6% and ~15%,

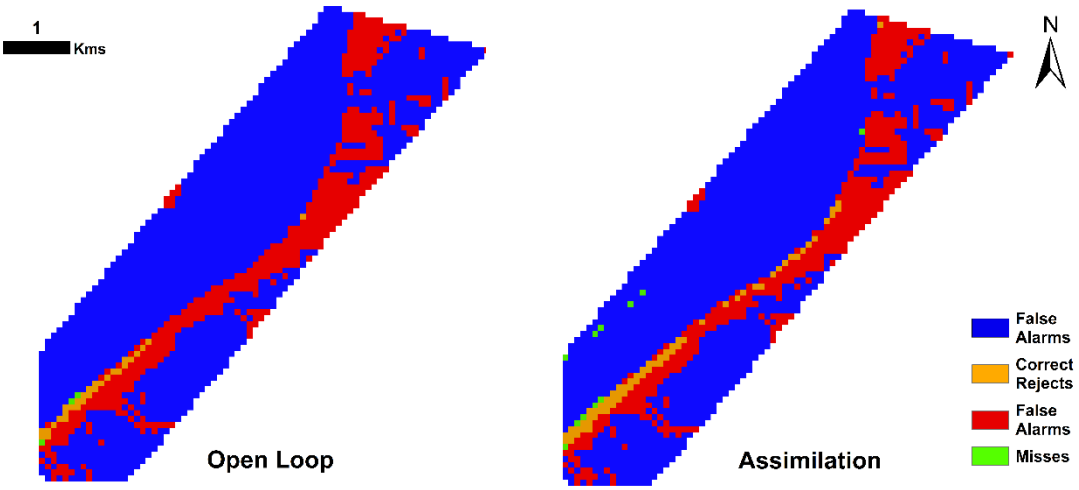

Fig. 9.9 As for Fig. 9.8 but for the Ulmarra region.

Table 9.2 Contingency matrix evaluating the forecast flood extent maps against the corresponding extents simulated by the truth model, at the first assimilation time step after considering peak time lag.

Lagged	**Junction Hill**				**Ulmarra**			
	Open Loop		**Assimilation**		**Open Loop**		**Assimilation**	
	MF	MNF	MF	MNF	MF	MNF	MF	MNF
OF	594	23	580	37	1802	3	1794	11
ONF	232	801	188	845	502	29	465	66
CSI	**0.700**		**0.720**		**0.781**		**0.790**	
%age imp. CSI	**2.980**				**1.179**			
Kappa	**0.691**		**0.722**		**0.079**		**0.169**	
%age imp. Kappa	**4.549**				**113.660**			

OF=Observed Flooded; ONF=Observed Non-Flooded; MF=Modelled Flooded; MNF=Modelled Non-Flooded; CSI= Critical Success Index

respectively. At Ulmarra the CSI values were spuriously high as the extent was significantly over predicted and almost the entire area covered by the aerial photograph was simulated as wet. This is in line with previous studies on binary pattern matching measures and their bias towards models which over estimate extents (Wealands et al. 2005; Stephens et al. 2012, 2014; Stephens and Bates 2015; Landuyt et al. 2018). Moreover, the low values of Kappa highlight that a large percentage of the match apparent in the simulated and observed extents can be explained by random chance because of the skewed class distributions in the binary maps. The problems with the statistics notwithstanding, the assimilation degraded the values for both the CSI and the Kappa, although due to the CSI being positive biased towards models which over predict, the percentage degradation observed at Ulmarra was nearly negligible.

An analysis similar to the water level hydrographs was carried out in this case as well. The particle weights obtained by assimilating the images after considering the peak time lag were used to compute the weighted mean flood extents, which were then evaluated against the inundation observed in the aerial photographs. Fig. 9.8 and Fig. 9.9 show the

contingency maps for Junction Hill and Ulmarra, respectively, while the confusion matrix and the corresponding statistics have been summarized in Table 9.2. Note that the time step at which the simulated extents were extracted for the evaluation in this case have also been lagged from the actual aerial image acquisition time to account for the flood wave arrival lag. The process used to identify the lag was the same as before, the relative position with respect to the hydrograph was quantified through comparisons with the synthetic truth or calibrated model used in Chapter 7.

Adjusting the arrival time lag in the model ensembles allowed for a better comparison with the evaluation data, as evident from the contingency maps. The previously observed large overestimations nearly disappeared at Junction Hill, while at Ulmarra the difference was not as obvious but became apparent upon closer examination. The over prediction visible next to the channel in the open loop as well as the assimilated forecast at junction Hill in the north east quarter of the maps can also be thought to be a function of the binarization of the evaluation maps. Recall from the aerial photo figures shown in Chapter 4 that this region was most likely covered with floating vegetation. The observed values of inundation had therefore been interpolated in this region for the fuzzy image matching, as the underlying flooding was uncertain. When the fuzzy aerial flood maps were binarized with a 0.5 threshold that embodies the maximum entropy, some of the uncertain wet pixels were hard classified as non-flooded whereas in reality their wet-dry status is uncertain. Nevertheless, the assimilation reduced false alarms along the eastern edge of the maps at Junction Hill and near the river bank at Ulmarra.

Summary statistics from the confusion matrices corroborate that the false alarms were indeed reduced as a consequence of the assimilation, leading to an increased agreement with the observed aerial photo based flood maps. Specifically, the assimilation of the first image improved the values of the CSI and Kappa statistics at Junction Hill by ~3% and ~4.5% respectively, while at Ulmara the values improved by ~1% and ~114%, respectively. Absolute values of the Kappa were very low at Ulmarra, thus explaining the more than hundred percent improvement after the assimilation. The relative increase in the Kappa was also greater at Junction Hill, as even though the Kappa corrects for expected agreement, the simulation of correct rejects is also given some weightage unlike in the CSI. As the assimilation primarily reduced false alarms by correctly rejecting them and even introduced some missed pixels, the relative improvements in CSI were limited at both

evaluation sites. Nevertheless, the magnitude of the reduction in over prediction was greater than the increase in under prediction, resulting in improved values for both statistics at both the evaluation sites. Overall the assimilation impacts on the simulated inundation extent were verified to be positive by this analysis, establishing the potential of the algorithm proposed in this thesis for operational inundation forecasting applications.

9.3 Chapter Summary

This chapter applied the flood extent assimilation algorithm proposed in this thesis to a real-world situation and evaluated the performance against actual observations. The experimental design and the performance metrics used were identical to those used in the synthetic experiments for the same variables. Assimilation impacts were evaluated in terms of simulated flood extent through aerial photographs and channel water levels through hydrometric gauge observations. Due to an inadequate hydraulic model calibration, the forecasts could not capture the flood peak timing correctly, however the overall effects of the assimilation were positive when the lag in the simulated flood peak arrival time was considered. Assimilating the first image which was just after the peak had passed at the gauges located within the observational coverage, had a notable and consistent positive effect on the channel water level forecast, corroborating the findings from both Chapters 7 and 8. The relative forecast improvement on adding the second image was not significant, in contrast to what was observed in Chapter 7. In terms of the flood extent, only the assimilation of the first image could be evaluated due to the acquisition timing of the aerial imagery. Estimation of inundation extents was improved at both evaluation sites, after considering the peak timing lag. These results demonstrate that the flood extent assimilation algorithm proposed in this thesis successfully reduced flood forecast uncertainties for a real-world application.

PART IV
SUMMARY
AND
PERSPECTIVES

CHAPTER TEN

“Our imagination is stretched to the utmost, not, as in fiction, to imagine things which are not really there, but just to comprehend those things which 'are' there.”

- Richard P. Feynman, In *“Surely You're Joking, Mr. Feynman!” - Adventures of a Curious Character* (1985).

10. Conclusions

This thesis investigated the potential of SAR-based flood observations to improve hydraulic flood inundation forecasts using a novel particle filter based flood extent assimilation scheme. The hydraulic model was initialised using an ensemble of uncertain inflows, simulated to imitate those from a hydrological streamflow forecast model, and the ensemble of inundation maps generated evaluated against SAR-derived flood extents at assimilation time steps when available, using a novel reliability diagram based cost function. The assimilation of SAR-derived flood extents was used to update the model trajectory to account for forecast inflow and hydraulic flood inundation model errors. The principal aim of this thesis was to develop a methodology for improved near real-time estimation and forecasting of flood inundation by using a combination of SAR-based flood extent and a hydraulic model. Specifically, the following objectives were addressed:

- Improved operational single image SAR-based flood mapping.
- Use of crowd-sourced water levels for quantitative hydraulic model calibration.
- A new cost function for flood extent assimilation.
- The sensitivity of assimilation performance to observation spatiotemporal characteristics.

This thesis is the first to optimize the assimilation of SAR-derived flood extents into hydraulic models, to maximize the possible improvements in inundation forecast accuracy. Therefore, an effort was made to introduce operational process and/or improvements at every major step. This chapter outlines the progress made towards achieving this goal and is divided into four parts aligned with the specific objectives above.

10.1 SAR-based Flood Extent Mapping

While texture is often utilized in SAR based flood mapping approaches it has seldom been optimized within this context, to derive the maximum possible pattern information in a minimum number of bands and reduce subjectivity in window size, direction, and texture feature selection. An approach to optimize omnidirectional GLCM-based image features, derived using the range of the spatial autocorrelation as the window size, was proposed through an Independent Component Transform. The optimized texture bands were added to the SAR image pre-classification using a Gaussian neuro-fuzzy classifier, which resulted in reduced uncertainties. The classification performance was evaluated against the addition of arbitrarily selected texture features, and without any texture addition to SAR. Results indicate that the texture optimization approach was able to extract the most useful texture information, showing drastic error reductions over the other approaches tested. The proposed approach was evaluated at two validation sites situated in the Clarence catchment, Australia - Junction Hill and Ulmarra - for which aerial photographs were available close to the time of acquisition of one of the SAR images. Accuracy was evaluated using a fuzzy set approach and through reliability diagrams based on pixel count ratios. Pixels for individual land cover classes were subsequently extracted and the classification performance within each class quantified. This led to the following findings:

- Texture based image enhancements represent a viable approach to improve flood mapping from single SAR images, leading to a reduction in uncertainties particularly in water look-a-like regions e.g. smooth urban surfaces like tarmac. The addition of optimized texture therefore ensured that linear urban features like roads were not wrongly detected as flooded in the Junction Hill region.
- The addition of optimized texture also led to decreased classifier sensitivity due to speckle noise, corroborated by a notable shift in pixels from uncertain to more certain bin values in the reliability diagrams. The study also demonstrated that the use of randomly selected texture features could actually degrade flood mapping accuracy, highlighting the importance of the proposed texture optimization.

- The neuro-fuzzy classifier was able to characterize backscatter uncertainties through the use of fuzzy membership functions, while the model selection and testing phase also explicitly defined the classification uncertainty.
- A fuzzy set validation approach was introduced, for the first time, to assess SAR-based flood maps. While these methods cannot replace more traditional approaches like reliability diagrams, which provide a detailed assessment of over- and under-prediction, they can complement the evaluation strategy to make the validation more robust. Moreover, the fuzzy set approach presents a unique opportunity to represent the uncertainties in the validation data and measure the performance while remaining cognizant of these errors.

10.2 Hydraulic Model Calibration using Crowd-sourced Data

The use of crowd-sourced water levels for a quantitative calibration of a 2D-hydraulic flood inundation model has never been attempted before. The channel roughness parameter for the hydraulic model Lisflood-FP was calibrated using a collection of 32 distributed floodplain water levels, derived from crowd-sourced field photographs of high water marks whose timing of acquisition was unknown. Assuming that these were representative of the maximum water depth throughout the associated model grid cell, quantitative performance measures were used to estimate absolute and relative model errors, to demonstrate the potential for application to data sparse regions. As a first step of model verification, the calibrated parameter value was inter-compared with similar information derived from hydrometric gauges, demonstrating that crowd-sourcing could be a viable data collection option. Further, plots of the maximum water depths simulated by the calibrated model against those obtained through crowd-sourcing and gauges revealed only minimal deviations from the observations. Finally, the inundation extent simulated by the calibrated model was evaluated against an optical remote sensing image, demonstrating acceptable agreement. The experiments presented in this thesis towards this particular objective, led to the following conclusions:

- This study showed that it is possible to use a limited number of accurate crowd-sourced water levels to constrain a 2D-hydraulic model. Crowd-sourcing can thus be used to support remote sensing data in ungauged or flashy catchments.

- The methods developed in this thesis can easily be extended to large volumes of crowd-sourced data, albeit the availability of an associated time stamp and geolocation is deemed necessary.

10.3 Flood Extent Assimilation Algorithm

A novel cost function based on the concept of mutual information was introduced in this thesis. This cost function quantified the reduction in uncertainty about the observations given the model prediction. Each particle representing a particular model realization with a unique erroneous inflow input was evaluated against a probabilistic observation and assigned a global weight using the mutual information at each assimilation time step. The erroneous inflows were generated based on the 2011 flood event in the Clarence Catchment, with the assumption of being generated from a hydrological model. Temporally correlated variance changing errors with a positive multiplicative bias were thus emulated, to imitate hydrological streamflow forecasts generated from forecast precipitation datasets. The efficiency of the proposed assimilation algorithm was demonstrated through synthetic and real world experiments. Experiments using synthetic observations generated at the same time and with the same error characteristics of the actual CSK SAR images, evaluated assimilation impacts on flood extent, floodplain water depths, floodplain flow velocities, channel depth, and flow. Results were assessed against the "truth" model set up with observed inflows and the best available topography, downstream boundary, and parameters. Real experiments using the same open loop ensemble and actual SAR-based fuzzy flood maps, evaluated assimilation performance against independent aerial photographs and hydrometric water level gauges along the main stem of the Clarence River. The following key conclusions could be drawn from this study:

- Improvements observed in flood extents after the assimilation, assessed through binary pattern matching measures, were limited in magnitude. This might be a function of the 1cm threshold used for the flood extent delineation as not many cells change their wet-dry status at every time step. Moreover, as the open loop is positively biased the binary pattern matching measures also lose sensitivity as all the models over predict.
- Simulated water depths were evaluated both in the channel and in the floodplain. For each case there was an improvement of more than 40% when the cumulative impact of

both images was considered. The synthetic assessment of forecast flow velocities in the floodplain show that the assimilation efficiency increased with lead time, as the momentum in the model domain builds over time. Correcting at a point after the peak where the images considered here were located, is thus able to effectively constrain the model flood propagation which dominates inundation patterns at this stage. Impacts on channel discharge and water depth were consistent from upstream to downstream and only differed in absolute magnitudes not in relative performance.

- Improvements on spatial simulation of water levels and velocities were only evaluated for lead times of up to 72h, with improvements remaining consistent over the duration of the analysis. For the hydrograph assessment of channel flow and depth, minor improvements could be observed until up to 7 days after the assimilation time steps, although the magnitude decayed with lead time. This was perhaps a consequence of the post-peak timing of the observations available, as the inflow errors continuously being added to the domain reduce from this point and so the forecast quality is dominated by the influence of flood propagation.
- The assessment of channel discharge revealed that the assimilation was able to correct for unsystematic bias, as it pulled the forecast trajectory towards the truth irrespective of the direction of the bias. This implies the observation is able to facilitate the selection of particles that perform well across the entire domain for post-peak images. The confidence in the ensemble size with respect to the uncertainty in the open loop was also substantiated, as the assimilation performance revealed adequate sampling within the state-space.
- The assimilation of SAR-derived flood extents is able to improve hydraulic flood inundation forecast skill in spite of it being a prognostic state variable of hydraulic models. For example, both for the synthetic and real experiments the assimilation was able to improve forecast skill consistently and persistently in the channel and in the floodplain. This implies that the assimilation of flood extents into hydraulic inundation forecasts could be significant from an emergency management perspective.
- For the real experiment the assimilation degraded the forecast from that of the open loop, as the flood wave arrival time was not effectively captured. This was a direct consequence of the lack of timing information on the crowd-sourced water depth observations used for the channel friction calibration. However, when the lag in the

ensemble was considered the assimilation algorithm improved the forecast for both the inundation extent as well as the point gauge locations.

The assimilation algorithm proposed here not only introduced the principles of information theory to hydraulic data assimilation, but through a synthetic experiment evaluated for the first time the impacts of the assimilation on all of the state variables. Moreover, the distributed spatial impact on simulated floodplain water depths and velocities, along with the evolution of these errors over different lead times, was examined for the first time.

10.4 Observation Spatiotemporal Sensitivity Analysis

This section presents the major conclusions from the sensitivity analysis of the assimilation performance to the observation spatiotemporal characteristics. Three sub-reaches of the Clarence River exhibiting uniform flow behaviour, were experimentally identified through DEM-based methods. Images were assimilated at each sub-reach at different points across the hydrograph. Impacts on the entire model domain and within the observed sub-domain were subsequently evaluated. Brier Skill Scores and spatiotemporal mean RMSE in floodplain water depth from single and multiple image assimilation experiments were presented, implicitly comparing the performance of the assimilated forecast to the open loop. Results revealed the following major findings:

- The assimilation efficiency was extremely sensitive to the spatial location and timing of the image. Specifically, the coverage of the images with respect to reach hydraulic characteristics and flood wave arrival time were found to be important governing factors for the assimilation efficiency. For example, the assimilation of images in Sub-reach 2 and 3, which exhibit more diffusive flow behaviour, demonstrated an almost similar response while the assimilation effects observed in Sub-reach 1 differed significantly.
- Improvements at gauge locations within the assimilation sub-domain were maximum when the images were acquired at and after the peak arrival time in the particular sub-reach. However, for spatial water depth simulation improvements, the best performing images were those acquired just before and after the inflow peak timing.

- Assimilation efficiency increased from upstream to downstream for the gauges, which is in line with expectations. In numerical models, as one erroneous value is computed based on another for each subsequent grid cell encountered in the flow direction, errors are typically amplified from upstream to downstream. The main reason being that the errors are added to the system through the inflow boundary, and are then propagated through the numerical model towards the downstream end. This is consistent with the synthetic experiments of Cooper et al. (2018) who showed that the assimilation of satellite-derived water levels has more persistent impacts in longer domains, especially downstream of the observation location.
- When multiple images were assimilated the timing of the first visit as well as the observation correlation length were critical to ensure that the forecast was not degraded by the assimilation. In general, the observation correlation length was a function of the first visit time and reach hydraulic behaviour. For example, in Sub-reach 1 the assimilation resulted in greater improvements when a larger number of observations were considered together through weight multiplication However, in the other more hydrodynamic sub-reaches, as the inundation response of the floodplain with respect to the channel was highly non-linear, this correlation typically dwindled after a few time steps. Different particles performed well at different times and were therefore selected by the assimilation. The sensibility of multiplying the previous observation weights thus decays rapidly during the peak flow as the extent variability in the domain is maximized. Moreover, the experiments highlighted that the observation correlation length, or the time for which the weight multiplication continued to produce positive results, could be starkly different from the time at which the maximum improvements were observed. This implies that the temporal observation correlation, increases to a local maxima and then starts to drop again.
- An analysis of the maximum improvements possible through the flood extent assimilation revealed that the difference between assimilating a single image and a set of eleven images at Sub-reach 2 resulted in only slight differences in the forecast skill. This implies that the assimilation performance was mainly limited by the timing and the spatial location of the image used for the assimilation. This is an important conclusion and implies that optimizing the spatiotemporal observation characteristics could potentially result in more cost effective SAR acquisitions to support flood forecasts.

Overall, this experiment conclusively showed that the assimilation algorithm is sensitive to the location, timing, and frequency of the observations. Moreover, the study also demonstrated that the assimilation of a single SAR image can produce impacts similar to those from multiple acquisitions.

CHAPTER ELEVEN

“Imagination is more important than knowledge.”

Albert Einstein (2015). “Bite-Size Einstein: Quotations on Just About Everything from the Greatest Mind of the Twentieth Century”, p.32, St. Martin's Press

11. Perspectives

This thesis has developed innovative methods to improve the accuracy of flood extent delineation from single SAR images and to subsequently use these to enhance hydraulic flood inundation forecasting skill. Although the methods presented here are novel, the experimental results revealed that many open research questions still remain. This chapter briefly outlines some of these scientific challenges, which future studies in this direction should endeavour to investigate. The chapter is divided into four sections in line with the stated objectives and structure in Chapter 10.

11.1 SAR-based Flood Extent Mapping

The texture optimization and neuro-fuzzy classification approach proposed in this thesis successfully improved flood detection capabilities when only a single SAR image of the event is available. This assumption of data availability is typically true in most operational cases. Although the use of a supervised classification technique requires the use of manually selected training data at the moment, this can be automated by using a combination of image segmentation, thresholding and region growing, by extending the methods proposed by Matgen et al. (2011). Once the semivariogram ranges have been precomputed for available SAR sensors, and the training process has been automated, the proposed approach could be operationally implemented. Moreover, if additional data are available, they can be utilized to further refine the resulting flood maps. However, it is important to choose appropriate ancillary data for post classification map refinement, to ensure that correctly identified flood pixels are not removed from the analysis. For instance, the use of Height Above Nearest Drainage models (Rennó et al. 2008; Nobre et al. 2016) to remove areas of “unlikely” flooding based on terrain

indicators, risks the removal of pixels inundated through pluvial processes, and therefore should be used with caution. Future work towards single SAR-image based flood extent mapping should aim towards answering some of the following challenges which still remain:

- The training of the supervised classifier requires manual intervention at this point which could be automated in the future using self-learning convolutional neural networks, by extracting sample training pixel locations from global surface water extents maps.
- In the case of some texture bands, the nature of fuzzy membership functions may deviate from the Gaussian distribution, and thus the sensitivity of this textured approach to the use of different fuzzy membership functions needs to be investigated.
- Topographic information from the DEM can be added along with the texture bands, within the neurofuzzy classifier, to further increase the information content.
- Contextual information about neighbouring pixel classes can be expected to further enhance the classification outputs, and thus misclassifications or fuzziness within homogeneous image segments can be further reduced.
- The potential of the fuzzy kappa statistic needs to be further demonstrated through testing different neighbourhood influence functions, different resolution SAR images, and different algorithms.

11.2 Crowdsourced Observations in Hydraulic Modelling

This study presented a first attempt towards quantitative flood model calibration using crowd-sourced water levels. The experiments detailed in this thesis successfully demonstrated the utility of crowdsourced data for channel roughness parameterization of hydraulic models. Pressing research questions include:

- The objective quantification of the often subjective uncertainties in crowdsourced data is a significant challenge. As the data are usually sourced from human sensor networks, a variety of socio-cultural or even mental health factors can alter the accuracy of the data eventually acquired. Hydrological data sources, and therefore scientists, have never dealt with uncertainties of this nature, which makes the characterization of these uncertainties

extremely difficult. In the future, it might help to borrow ideas from the emerging field of socio-hydrology or encourage collaborative interdisciplinary investigations to improve the value of these data.

- Another significant challenge is the data collection process itself, which can be classified into direct and indirect active citizenship, which as the names suggest respectively refer to whether the data was purposefully collected for an experiment or volunteered. As one would imagine, the data that is directly collected meets predefined quality standards and can immediately be utilised to support experiments. However, the vast multitude of volunteered data available on social media sites during an emergency, for instance, is still not utilised to its full potential. The primary reason for this is the lack of appropriate methods to automatically derive quantitative information, e.g. water levels which can directly support hydraulic modelling. Developing maturity in these methods require interdisciplinary efforts in collaboration with computer scientists, where there has been evolutionary progress in semantic analysis and computer vision, which can help to parse text and images respectively into useful information. Tapping into this repository of volunteered geographic information, can significantly alter flood preparedness and management practices in the dense populated urban centres around the world, where flood risk has increased manifold in recent decades (Jha et al. 2011).
- Finally, methods to deal with the large uncertainties in these crowd-sourced observations need to be explored. For example, in case of uncertain data in terms of the water level estimates, approaches suggested by Hostache et al. (2009) could be used, where the model is forced to lie within observation error limits rather than replicate the measurements. While in the presence of geolocation errors, the approach of Schumann et al. (2008) could be used to shift the pixel randomly in all directions within the limits of the horizontal accuracy. This helps to derive a range of possible uncertain values which can then be utilised using the aforementioned technique of Hostache et al. (2009).

11.3 Optimizing Flood Extent Assimilation

This thesis presented a first attempt towards optimizing flood extent assimilation to maximize forecast skill improvements in inundation patterns generated by hydraulic models.

First, a novel flood extent assimilation algorithm was designed and then its sensitivity to observation spatiotemporal characteristics was evaluated. Results demonstrated the potential of the proposed flood extent assimilation algorithm to reduce forecast uncertainty, with persistent improvements for up to a week. Research on hydraulic data assimilation of flood extents is very new and many scientific challenges remain:

- Ways to interpret modelled water depths as fuzzy flood extents need to be developed so that the fuzziness can be incorporated into the assimilation cost functions. Binarization of the water depth maps into extents reduces the available information content. For example, for areas in which high resolution DEMs are available over the entire domain, this could be achieved by fitting a linear function to the elevation values observed in each modelled grid cell, and used to assign fuzzy values of flooding to each pixel (Pappenberger et al. 2007b).
- Methods to adequately encompass larger uncertainties within smaller model ensembles need to be developed, due to the computational effort required to run hydraulic flood inundation simulations. This inhibits the inclusion of uncertainties from channel geometry, floodplain topography, and distributed friction parameters, as then the size of the state parameter space would become too large to sample from efficiently using just 128 particles. Any assimilation efforts can only be as good as the quality of the particles, as the assumption is that the truth is somewhere included in the ensemble representation. This may require the use of forward state sensitivity methods which allow the estimation of optimum state-parameter perturbations dynamically in an online fashion (Lakshmivarahan and Lewis 2010). Such investigations will be especially relevant for operationalizing these flood extent assimilation techniques on global flood forecasting platforms.
- The integration of hydraulic flood inundation forecasts with other alternative data sources such as, point observations of water levels and velocities from crowdsourcing, depth Doppler maps of inundation from GNSS reflectometry, and water level maps from SWOT also needs to be considered when designing future assimilation frameworks.
- Machine learning can offer several advantages when applied to the problem of model-data integration, particularly in terms of computational efficiency in real-time (once the integration model has “learned” the data characteristics) as compared to some advanced filtering techniques. However, since the algorithm has no knowledge of the model physics,

a key benefit of data assimilation which lies in the bridging of observations using system dynamics might be lost (Lahoz 2010; Keith Beven, personal communication). Indeed, this might lead to negligible impacts in practice, but extensive benchmarking against established assimilation methods is necessary to evaluate the universality and potential of such methods.

11.4 Observation Spatiotemporal Sensitivity

The critical impact of observation location, timing, and interval on flood extent assimilation performance was also demonstrated in this thesis. Synthetic experiments were designed where flood extent was assimilated at hydraulically uniform reaches at different times across the hydrograph. Impacts on the entire model domain were quantified, which revealed that observations acquired at the peak or earlier could be significantly detrimental to the flood forecast quality. Moreover, reaches with high variable hydrodynamic flows were identified as locations of interest for flood extent assimilation, i.e. the inundation forecast quality could benefit most from observations in these reaches. This investigation was the first of its kind for flood extent assimilation, and future research should also examine:

- The impacts of different input data and parametric uncertainties on the optimum satellite acquisition location, timing, and frequency need to be quantified. Intuitively, the reaches which exhibit hydrodynamic out of bank flows, should benefit most from the assimilation in any scenario. However, in the case of localized errors in topography and/or channel geometry, model forecasts can be degraded if the observations are centred over a highly erroneous sub-region, as Andreadis and Schumann (2014) found for water level assimilation. It will especially be interesting to understand how targeted observation designs should change to accommodate the dominant sources of uncertainty in a catchment. This could lead to more tailored and accurate flood inundation forecasts.
- Methods to pre-emptively select only the most informative observations for the assimilation should be explored, as the coverage of flood events from satellite data is only expected to increase in the future. When multiple images are available, incorporating them all within the forecasting chain, does not necessarily lead to forecast improvements as this thesis demonstrated. Accordingly, the utilization of a limited number of optimal

observations, is expected to have a larger overall impact on flood forecast accuracy, and therefore objective methods should be designed to identify them *a priori*.

Publications

Published

Peer-Reviewed Journal Papers

- **Dasgupta, A.,** Grimaldi, S., Ramsankaran, R., Pauwels, V. R. N., & Walker, J. P. (2018). Towards operational SAR-based flood mapping using neuro-fuzzy texture-based approaches. *Remote Sensing of Environment*, *215*(15 September 2018), 313–329. http://doi.org/10.1016/j.rse.2018.06.019 (Ten Citations so far, with four in Q1 journals).

Peer-Reviewed Conference Papers

- **Dasgupta, A.,** Grimaldi, S., Ramsankaran, R., and Walker, J.P. (2017) Optimized GLCM -based texture features for improved SAR-based flood mapping, *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Fort Worth, TX, 23-28 July 2017, pp. 3258-3261. DOI: 10.1109/IGARSS.2017.8127692 (Poster)

Invited Book Chapters

- **Dasgupta, A.,** Grimaldi, S., Ramsankaran, R., Pauwels, V. R. N., Walker, J. P., Chini, M., Hostache, R. and Matgen, P. (2018). Flood Mapping Using Synthetic Aperture Radar Sensors from Local to Global Scales. In *Global Flood Hazard (eds G. J. Schumann, P. D. Bates, H. Apel and G. T. Aronica).* AGU Books. John Wiley Publications. doi:10.1002/9781119217886.ch4 (Six Citations so far, with one in Q1 journals).

Conference Abstracts

- **Dasgupta, A.,** Hostache, R., Ramsankaran, R., Pauwels, V.R.N., Schumann, G.J.P., Grimaldi, S., and Walker, J.P. (2019). Evaluating the Impact of Flood Extent Assimilation on Hydraulic Model Forecast Skill. In proceedings of *MODSIM2019, 23rd International Congress on Modelling and Simulation*, Canberra, Australia, 1-6 December, 2019. (Poster)

- **Dasgupta, A.,** Hostache, R., Ramsankaran, R., Pauwels, V.R.N., Schumann, G.J.P., Grimaldi, S., and Walker, J.P. (2019). Optimizing Targeted SAR Acquisitions for Flood Extent Assimilation to Improve Inundation Forecasts. In *Proceedings American Geophysical Union Fall Meeting*, San Fransisco, USA, 9-13 December, 2019. (Poster)

- **Dasgupta, A.,** Hostache, R., Ramsankaran, R., Pauwels, V.R.N., Schumann, G.J.P., Grimaldi, S., and Walker, J.P. (2019). On the impacts of location, timing, and frequency of inundation extent assimilation on flood forecast skill. In *Proceedings of Hydrologic Ensemble Prediction EXperiment* (HEPEX), Reading, UK, 25-28 November, 2019. (Oral)

- **Dasgupta, A.,** Hostache, R., Ramsankaran, R., Pauwels, V.R.N., Schumann, G.J.P., Grimaldi, S., and Walker, J.P. (2019). Improving Flood Forecasts using SAR-based Flood Extent Assimilation. In *Proceedings of the European Geosciences Union General Assembly*, Vienna, Austria, 7-12 April, 2019. (Oral)

- **Dasgupta, A.,** Grimaldi, S., Ramsankaran, R., Hostache, R., Schumann, G., Pauwels, V.R.N., & Walker, J.P. On the Spatiotemporal Impacts of Flood Extent Assimilation. In *Proceedings American Geophysical Union Fall Meeting*, Washington DC, 10-14 December, 2018. (Oral)

- Jain, P., Ramsankaran, R., & **Dasgupta, A.** GIS based Integrated Multi-Criteria Modelling Framework for Watershed Prioritization in India. In *Proceedings American Geophysical Union Fall Meeting*, Washington DC, 10-14 December, 2018. (Poster)

- **Dasgupta, A.,** Hostache, R., Ramsankaran, R., Pauwels, V.R.N., Schumann, G.J.P., Chini, M., Grimaldi, S., Matgen, P., and Walker, J.P. (2018). Evolutionary assimilation of SAR-derived inundation extents into a flood model. In *Proceedings of the European Geosciences Union General Assembly*, Vienna, Austria, 8-14 April, 2018. (Poster+Oral)

- **Dasgupta, A.,** Schumann, G.J.P., Grimaldi, S., Ramsankaran, R., Pauwels, V.R.N., & Walker, J.P. (2018). Diagnostic Assessment of Localized Flood Flow Behaviour. In *Proceedings of Hydrologic Ensemble Prediction EXperiment* (HEPEX), Melbourne, VIC, 6-8 February, 2018. (Poster)

- **Dasgupta, A.,** Grimaldi, S., Ramsankaran, R., Pauwels, V.R.N., & Walker, J.P. (2017). Use of Crowd-sourced Data to Assess the Skill of Hydraulic Flood Forecast Models. In Syme, G., Hatton MacDonald, D., Fulton, B. and Piantadosi, J. (eds) *MODSIM2017, 22nd International Congress on Modelling and Simulation.* Modelling and Simulation Society of Australia and New Zealand, 3-8 December 2017, pp. 516. ISBN: 978-0-9872143-7-9. (Oral)

In Process

Peer-Reviewed Journal Papers

- **Dasgupta, A.,** Schumann, G.J.P., Ramsankaran, R., Grimaldi, S., Pauwels, V.R.N., Hostache, R., and Walker, J.P. From "Data-poor" to "Data-rich": Advances in Flood Observations and Implications for Hydraulic Modelling. *Reviews of Geophysics* (Manuscript in preparation).

- **Dasgupta, A.,** Hostache, R., Ramsankaran, R., Pauwels, V.R.N., Schumann, G.J.P., Grimaldi, S., and Walker, J.P. On the Sensitivity of flood extent assimilation algorithms to observation location, frequency, and timing. *Water Resources Research* (Manuscript in preparation – Chapter 8).

- **Dasgupta, A.,** Hostache, R., Ramsankaran, R., Pauwels, V.R.N., Schumann, G.J.P., Grimaldi, S., and Walker, J.P. Improving the performance of SAR-derived flood extent assimilation. *Water Resources Research* (Manuscript in preparation – Chapter 7).

- **Dasgupta, A.,** Hostache, R., Ramsankaran, R., Pauwels, V.R.N., Schumann, G.J.P., Grimaldi, S., and Walker, J.P. Flood forecast improvements using remote sensing data assimilation with coupled hydrological-hydraulic models. *Remote Sensing of Environment* (Manuscript in preparation – Chapter 9).

- **Dasgupta, A.,** Grimaldi, S., Ramsankaran, R., Pauwels, V.R.N., and Walker, J.P. Hydraulic model calibration using crowdsourced water levels *Remote Sensing of Environment: Short Communications* (Manuscript in preparation – Chapter 5).

Invited Book Chapters

- **Dasgupta, A.,** Grimaldi, S., Ramsankaran, R., Hostache, R., Matgen, P., Chini, M., Pauwels, V. R. N., and Walker, J. P. (2019). Earth Observation for Improved Hydraulic Flood Forecasts. In *Earth Observation for Flood Applications: Progress and Perspectives (ed. G. J.-P. Schumann).* Elsevier. (Accepted).

References

Abaza M, Garneau C, Anctil F (2014) Comparison of Sequential and Variational Streamflow Assimilation Techniques for Short-Term Hydrological Forecasting. J Hydrol Eng 4014042. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001013

Adam S, Wiebe J, Collins M, Pietroniro A (1998) Radarsat flood mapping in the Peace-Athabasca Delta, Canada. Can J Remote Sens 24:69–79. https://doi.org/10.1080/07038992.1998.10874693

Aggarwal A, Rafique F, Rajesh E, Ahmed S (2014) Urban flood hazard mapping using change detection on wetness transformed images. Hydrol Sci J 0:1–10. https://doi.org/10.1080/02626667.2014.952638

Ahamed A, Bolten J, Doyle C, Fayne J (2017) Near Real-Time Flood Monitoring and Impact Assessment Systems Case Study: 2011 Flooding in Southeast Asia. In: C RE, Restrepo JD, Brakenridge GR, Kettner AJ (eds) Remote Sensing of Hydrological Extremes. pp 105–118

Alfieri L, Burek P, Dutra E, et al (2013) GloFAS-global ensemble streamflow forecasting and flood early warning. Hydrol Earth Syst Sci 17:. https://doi.org/10.5194/hess-17-1161-2013

Alfieri L, Cohen S, Galantowicz J, et al (2018) A global network for operational flood risk reduction. Environ Sci Policy 84:149–158. https://doi.org/10.1016/j.envsci.2018.03.014

Alfonso L, Mukolwe MM, Di Baldassarre G (2016) Probabilistic Flood Maps to support decision-making: Mapping the Value of Information. Water Resour Res 52:1026–1043. https://doi.org/10.1002/2015WR017378

Alsdorf DE, Rodriguez E, Lettenmaier DP (2007) Measuring Surface Water from Space. Rev Geophys 45:1–24. https://doi.org/10.1029/2006RG000197.1.INTRODUCTION

Amitrano D, Belfiore V, Cecinati F, et al (2016) Urban Areas Enhancement in Multitemporal SAR RGB Images Using Adaptive Coherence Window and Texture Information. IEEE J Sel

Top Appl Earth Obs Remote Sens 9:3740–3752. https://doi.org/10.1109/JSTARS.2016.2555340

Amitrano D, Di Martino G, Iodice A, et al (2018) Unsupervised Rapid Flood Mapping Using Sentinel-1 GRD SAR Images. IEEE Trans Geosci Remote Sens 1–10. https://doi.org/10.1109/TGRS.2018.2797536

Andreadis KM (2018) Data Assimilation and River Hydrodynamic Modeling Over Large Scales. In: Global Flood Hazard: Applications in Modeling, Mapping, and Forecasting, Volume 233 of Geophysical Monograph Series. pp 229–237

Andreadis KM, Clark EA, Lettenmaier DP, Alsdorf DE (2007) Prospects for river discharge and depth estimation through assimilation of swath-altimetry into a raster-based hydrodynamics model. Geophys Res Lett 34:1–5. https://doi.org/10.1029/2007GL029721

Andreadis KM, Schumann GJP (2014) Estimating the impact of satellite observations on the predictability of large-scale hydraulic models. Adv Water Resour 73:44–54. https://doi.org/10.1016/j.advwatres.2014.06.006

Andreadis KM, Schumann GJP, Pavelsky T (2013) A simple global river bankfull width and depth database. Water Resour Res 49:7164–7168. https://doi.org/10.1002/wrcr.20440

Antonova S, Kääb A, Heim B, et al (2016) Spatio-temporal variability of X-band radar backscatter and coherence over the Lena River Delta, Siberia. Remote Sens Environ 182:169–191. https://doi.org/10.1016/j.rse.2016.05.003

Arcement GJ, Schneider VR (1989) Guide for Selecting Manning ' s Roughness Coefficients for Natural Channels and Flood Plains United States Geological Survey Water-supply Paper 2339

Archer L, Neal JC, Bates PD, House JI (2018) Comparing TanDEM-X Data with Frequently-Used DEMs for Flood Inundation Modelling. Water Resour Res 1–18. https://doi.org/10.1029/2018WR023688

Aronica G, Bates PD, Horritt MS (2002) Assessing the uncertainty in distributed model predictions using observed binary pattern information within GLUE. Hydrol Process 16:2001–2016. https://doi.org/10.1002/hyp.398

Arulampalam MS, Maskell S, Gordon N, Clapp T (2002) A tutorial on particle filters for online nonlinear/non-GaussianBayesian tracking. Signal Process IEEE Trans 50:174–188. https://doi.org/10.1109/78.978374

Assumpção TH, Popescu I, Jonoski A, Solomatine DP (2018) Citizen observations contributing to flood modelling : opportunities and challenges. Hydrol Earth Syst Sci 22:1473–1489

Astrium Services (2013) SPOT 6 & SPOT 7 imagery user guide. Astrium Serv vi + 77pp. https://doi.org/SI/DC/13034-v1.0

Bai Y, Li X (2011) Evolutionary Algorithm-Based Error Parameterization Methods for Data Assimilation. Mon Weather Rev 139:2668–2685. https://doi.org/10.1175/2011MWR3641.1

Balaguer A, Ruiz LA, Hermosilla T, Recio JA (2010) Definition of a comprehensive set of texture semivariogram features and their evaluation for object-oriented image classification. Comput Geosci 36:231–240. https://doi.org/10.1016/j.cageo.2009.05.003

Banister AR, Nichols N (2012) Ensemble methods in data assimilation. 1–38

Baratelli F, Flipo N, Rivière A, Biancamaria S (2018) Retrieving river baseflow from SWOT spaceborne mission. Remote Sens Environ 218:44–54. https://doi.org/10.1016/j.rse.2018.09.013

Barreto TLM, Almeida J, Cappabianco FAMFAM (2016) Estimating Accurate Water Levels for Rivers and Reservoirs by using SAR Products: A Multitemporal Analysis. Pattern Recognit Lett 0:1–10. https://doi.org/10.1016/j.patrec.2016.05.015

Bates P, Trigg M, Neal J, Dabrowa A (2013) LISFLOOD-FP User manual. Univ Bristol 1–49

Bates PD (2012) Integrating remote sensing data with flood inundation models: how far have we got? Hydrol Process 26:2515–2521. https://doi.org/10.1002/hyp.9374

Bates PD, De Roo a PJ (2000a) A simple raster based model for flood inundation simulation. J Hydrol 236:54–77. https://doi.org/10.1016/S0022-1694(00)00278-X

Bates PD, De Roo APJ (2000b) A simple raster-based model for flood inundation simulation. J Hydrol 236:54–77

Bates PD, Horritt MS, Fewtrell TJ (2010) A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. J Hydrol 387:33–45. https://doi.org/10.1016/j.jhydrol.2010.03.027

Bates PD, Neal J, Sampson C, et al (2017) Progress Toward Hyperresolution Models of Global Flood Hazard. Elsevier Inc.

Bates PD, Neal JC, Alsdorf D, Schumann GJP (2014a) Observing Global Surface Water Flood Dynamics. Surv Geophys 35:839–852. https://doi.org/10.1007/s10712-013-9269-4

Bates PD, Pappenberger F, Romanowicz RJ (2014b) Uncertainty in Flood Inundation Modelling. In: Beven KJ, Hall J (eds) Applied Uncertainty Analysis for Flood Risk Management. London: Imperial College Press., pp 232–269

Bates PD, Wilson MD, Horritt MS, et al (2006) Reach scale floodplain inundation dynamics observed using airborne synthetic aperture radar imagery: Data analysis and modelling. J Hydrol 328:306–318. https://doi.org/10.1016/j.jhydrol.2005.12.028

Bauer-Gottwein P, Jensen IH, Guzinski R, et al (2015) Operational river discharge forecasting in poorly gauged basins: the Kavango River basin case study. Hydrol Earth Syst Sci 19:1469–1485. https://doi.org/10.5194/hess-19-1469-2015

Bazi Y, Bruzzone L, Melgani F (2007) Image thresholding based on the EM algorithm and the generalized Gaussian distribution. Pattern Recognit 40:619–634. https://doi.org/10.1016/j.patcog.2006.05.006

Berberoglu S, Curran PJ, Lloyd CD, Atkinson PM (2007) Texture classification of Mediterranean land cover. Int J Appl Earth Obs Geoinf 9:322–334. https://doi.org/10.1016/j.jag.2006.11.004

Berberoglu S, Lloyd CD, Atkinson PM, Curran PJ (2000) The integration of spectral and textural information using neural networks for land cover mapping in the Mediterranean. Comput Geosci 26:385–396. https://doi.org/10.1016/S0098-3004(99)00119-3

Beven K (2006) A manifesto for the equifinality thesis. J Hydrol 320:18–36. https://doi.org/10.1016/j.jhydrol.2005.07.007

Beven K (2016) Facets of uncertainty: Epistemic uncertainty, non-stationarity, likelihood, hypothesis testing, and communication. Hydrol Sci J 61:1652–1665. https://doi.org/10.1080/02626667.2015.1031761

Beven K (2014) A Framework for Uncertainty Analysis. In: Applied Uncertainty Analysis for Flood Risk Management. pp 39–59

Beven K (1993) Prophecy , reality and uncertainty in distributed hydrological modelling. Adv Water Resour 16:41–51

Beven K, Binley A (1992) The Future of Distributed Models: Model Calibration and Uncertainty Prediction. Hydrol Process. https://doi.org/10.1002/hyp.3360060305

Beven K, Hall J (2014) Use of Models in Flood Risk Management. In: Applied Uncertainty Analysis for Flood Risk Management. pp 25–35

Beven KJ (2012) RAINFALL-RUNOFF MODELING - The Primer, 2nd edn. Wiley-Blackwell

Beven KJ, Almeida S, Aspinall WP, et al (2018a) Epistemic uncertainties and natural hazard risk assessment – Part 1: A review of different natural hazard areas. Nat Hazards Earth Syst Sci 18:2741–2768. https://doi.org/https://doi.org/10.5194/nhess-18-2741-2018

Beven KJ, Aspinall WP, Bates PD, et al (2018b) Epistemic uncertainties and natural hazard risk assessment - Part 2: What should constitute good practice? Nat Hazards Earth Syst Sci 18:2769–2783. https://doi.org/10.5194/nhess-18-2769-2018

Beven KJ, Young PC (2003) Comment on "Bayesian recursive parameter estimation for hydrologic models" by M. Thiemann, M. Trosset, H. Gupta, and S. Sorooshian. Water Resour Res 39:1–4. https://doi.org/doi:10.1029/2001WR001183

Biancamaria S, Lettenmaier DP, Pavelsky TM (2016) The SWOT Mission and Its Capabilities for Land Hydrology. Surv Geophys 37:307–337. https://doi.org/10.1007/s10712-015-9346-y

Birch CE, Dottori F, Trigg MA, et al (2016) The credibility challenge for global fluvial flood risk analysis. Environ Res Lett 11:094014. https://doi.org/10.1088/1748-9326/11/9/094014

Blasco F, Bellan MF, Chaudhury MUU (1992) Estimating the extent of floods in Bangladesh using SPOT data. Remote Sens Environ 39:167–178. https://doi.org/10.1016/0034-4257(92)90083-V

Blöschl G, Bierkens MFP, Chambel A, et al (2019) Twenty-three unsolved problems in hydrology (UPH) – a community perspective. Hydrol Sci J 0:1–18. https://doi.org/10.1080/02626667.2019.1620507

Boni G, Ferraris L, Pulvirenti L, et al (2016) A Prototype System for Flood Monitoring Based on Flood Forecast Combined with COSMO-SkyMed and Sentinel-1 Data. IEEE J Sel Top Appl Earth Obs Remote Sens 9:2794–2805. https://doi.org/10.1109/JSTARS.2016.2514402

Bovolo F, Bruzzone L (2007) A Split-Based Approach to Unsupervised Change Detection in Large-Size Multitemporal Images: Application to Tsunami-Damage Assessment. IEEE Trans Geosci Remote Sens 45:1658–1669. https://doi.org/10.1109/TGRS.2007.895835

Briggs J, Dowd M, Meyer R (2013) Data assimilation for large-scale spatio-temporal systems using a location particle smoother. Environmetrics 24:81–97. https://doi.org/10.1002/env.2184

Brisset P, Monnier J, Garambois PA, Roux H (2018) On the assimilation of altimetric data in 1D Saint–Venant river flow models. Adv Water Resour 119:41–59. https://doi.org/10.1016/j.advwatres.2018.06.004

Brocca L, Tarpanelli A, Filippucci P, et al (2018) How much water is used for irrigation? A new approach exploiting coarse resolution satellite soil moisture products. Int J Appl Earth Obs Geoinf 73:752–766. https://doi.org/10.1016/j.jag.2018.08.023

Brown KM, Brownett JM (2016) Progress in operational flood mapping using satellite synthetic aperture radar (SAR) and airborne light detection and ranging (LiDAR) data. Prog Phys Geogr 40:196–214. https://doi.org/10.1177/0309133316633570

Browne PA (2016) A comparison of the equivalent weights particle filter and the local ensemble transform Kalman filter in application to the barotropic vorticity equation. Tellus, Ser A Dyn Meteorol Oceanogr 68:. https://doi.org/10.3402/tellusa.v68.30466

Butts MB, Payne JT, Kristensen M, Madsen H (2004) An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation. J Hydrol 298:242–266. https://doi.org/10.1016/j.jhydrol.2004.03.042

Camacho RA, Martin JL, Mcanally W, et al (2015) A Comparison of Bayesian Methods for Uncertainty Analysis in Hydraulic and Hydrodynamic Modeling. J Am Water Resour Assoc 51:1372–1393. https://doi.org/10.1111/1752-1688.12319

Canisius F, Brisco B, Murnaghan K, et al (2019) SAR Backscatter and InSAR Coherence for Monitoring Wetland Extent, Flood Pulse and Vegetation: A Study of the Amazon Lowland. Remote Sens 11:720. https://doi.org/10.3390/rs11060720

Carr JR (1996) Spectral and textural classification of single and multiple band digital images. Comput Geosci 22:849–865. https://doi.org/10.1016/S0098-3004(96)00025-8

Carr JR, De Miranda FP (1998) The semivariogram in comparison to the co-occurrence matrix for classification of image texture. IEEE Trans Geosci Remote Sens 36:1945–1952. https://doi.org/10.1109/36.729366

Carrivick JL (2006) Application of 2D hydrodynamic modelling to high-magnitude outburst floods: An example from Kverkfj{ö}ll, Iceland. J Hydrol 321:187–199. https://doi.org/10.1016/j.jhydrol.2005.07.042

Chaabani C, Chini M, Abdelfattah R, et al (2018) Flood Mapping in a Complex Environment Using Bistatic TanDEM-X/TerraSAR-X InSAR Coherence. Remote Sens 10:1873. https://doi.org/10.3390/rs10121873

Chen H, Liang Q, Liu Y, Xie S (2018) Hydraulic correction method (HCM) to enhance the efficiency of SRTM DEM in flood modeling. J Hydrol 559:56–70. https://doi.org/10.1016/j.jhydrol.2018.01.056

Chen H, Yang D, Hong Y, et al (2013) Hydrological data assimilation with the Ensemble Square-Root-Filter: Use of streamflow observations to update model states for real-time flash flood forecasting. Adv Water Resour 59:209–220. https://doi.org/10.1016/j.advwatres.2013.06.010

Chen HM, Varshney PK, Arora MK (2003) Performance of Mutual Information Similarity Measure for Registration of Multitemporal Remote Sensing Images. IEEE Trans Geosci Remote Sens 41:2445–2454. https://doi.org/10.1109/TGRS.2003.817664

Chertok DL, Lardner RW (1996) Variational data assimilation for a nonlinear hydraulic model. Appl Math Model 20:675–682. https://doi.org/10.1016/0307-904X(96)00048-0

Chica-Olmo M, Abarca-Hernández F (2000) Computing geostatistical image texture for remotely sensed data classification. Comput Geosci 26:373–383. https://doi.org/10.1016/S0098-3004(99)00118-1

Chini M, Albano M, Saroli M, et al (2015) Coseismic liquefaction phenomenon analysis by COSMO-SkyMed: 2012 Emilia (Italy) earthquake. Int J Appl Earth Obs Geoinf 39:65–78. https://doi.org/10.1016/j.jag.2015.02.008

Chini M, Bignami C, Stramondo S, Pierdicca N (2008) Uplift and subsidence due to the 26 December 2004 Indonesian earthquake detected by SAR data. Int J Remote Sens 29:3891–3910. https://doi.org/10.1080/01431160701871112

Chini M, Member S, Hostache R, et al (2017) A Hierarchical Split-Based Approach for Parametric Thresholding of SAR Images : Flood Inundation as a Test Case. IEEE Trans Geosci Remote Sens 1–14

Chini M, Pelich R, Pulvirenti L, et al (2019) Sentinel-1 InSAR Coherence to Detect Floodwater in Urban Areas: Houston and Hurricane Harvey as A Test Case. Remote Sens 11:107. https://doi.org/10.3390/RS11020107

Chini M, Piscini A, Cinti FR, et al (2013) The 2011 Tohoku (Japan) tsunami inundation and liquefaction investigated through optical, thermal, and SAR data. IEEE Geosci Remote Sens Lett 10:347–351. https://doi.org/10.1109/LGRS.2012.2205661

Chini M, Pulvirenti L, Pierdicca N (2012) Analysis and interpretation of the COSMO-SkyMed observations of the 2011 Japan tsunami. IEEE Geosci Remote Sens Lett 9:467–471. https://doi.org/10.1109/LGRS.2011.2182495

Christensen NS, Lettenmaier DP (2007) A multimodel ensemble approach to assessment of climate change impacts on the hydrology and water resources of the Colorado River Basin. Hydrol Earth Syst Sci 11:1417–1434. https://doi.org/10.5194/hess-11-1417-2007

Chu H, Chang L (2009) Applying Particle Swarm Optimization to Parameter Estimation of the Nonlinear Muskingum Model. J Hydrol Eng 14:1024–1027. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000070

Chu T, Lindenschmidt KE (2017) Comparison and Validation of Digital Elevation Models Derived from InSAR for a Flat Inland Delta in the High Latitudes of Northern Canada. Can J Remote Sens 43:109–123. https://doi.org/10.1080/07038992.2017.1286936

Chu W, Gao X, Sorooshian S (2010) Improving the shuffled complex evolution scheme for optimization of complex nonlinear hydrological systems: Application to the calibration of the Sacramento soil-moisture accounting model. Water Resour Res 46:1–12. https://doi.org/10.1029/2010WR009224

Clement MA, Kilsby CG, Moore P (2017) Multi-temporal synthetic aperture radar flood mapping using change detection. J Flood Risk Manag 1–17. https://doi.org/10.1111/jfr3.12303

Cloke H, Pappenberger F, Thielen J, Thiemig V (2013) Operational European Flood Forecasting. Environ Model Find Simplicity Complex Second Ed 415–434. https://doi.org/10.1002/9781118351475.ch25

Cohen J, Riihimaki H, Pulliainen J, et al (2016) Implications of boreal forest stand characteristics for X-band SAR flood mapping accuracy. Remote Sens Environ 186:47–63. https://doi.org/10.1016/j.rse.2016.08.016

Cooper ES, Dance SL, Garcia-Pintado J, et al (2018) Observation impact, domain length and parameter estimation in data assimilation for flood forecasting. Environ Model Softw 104:199–214. https://doi.org/10.1016/j.envsoft.2018.03.013

Cooper ES, Dance SL, García-Pintado J, et al (2019) Observation operators for assimilation of satellite observations in fluvial inundation forecasting. Hydrol Earth Syst Sci 23:2541–2559. https://doi.org/10.5194/hess-2018-589

Coulthard TJ, Neal JC, Bates PD, et al (2013) Integrating the LISFLOOD-FP 2D hydrodynamic model with the CAESAR model: implications for modelling landscape evolution. Earth Surf Process Landforms 38:1897–1906. https://doi.org/10.1002/esp.3478

Coulthard TJJ, Hicks DMM, Van De Wiel MJJ (2007) Cellular modelling of river catchments and reaches: Advantages, limitations and prospects. Geomorphology 90:192–207. https://doi.org/10.1016/j.geomorph.2006.10.030

Cover TM, Thomas JA (1991) Elements of Information Theory

Cox RJ, Shand, T.D.Blacka MJ (2010) Australian Rainfall & Runoff revision project 10: Appropriate safety criteria for people

CRED, UNISDR (2016) Poverty & Death: Disaster Mortality 1996-2015

D'Addabbo A, Refice A, Pasquariello G, et al (2016) A Bayesian Network for Flood Detection Combining SAR Imagery and Ancillary Data. IEEE Trans Geosci Remote Sens 54:3612–3625. https://doi.org/10.1109/TGRS.2016.2520487

Dasgupta A (2015) Reduction of Uncertainties in a 2D Hydrodynamic Model using Remote Sensing Data. University of Twente

Dasgupta A, Grimaldi S, Ramsankaran RAAJ, et al (2018) Towards operational SAR-based flood mapping using neuro-fuzzy texture-based approaches. Remote Sens Environ 215:313–329. https://doi.org/10.1016/j.rse.2018.06.019

De-yong H, Xiao-juan L, Wen-ji Z, Hui-li G (2008) Texture Analysis and its Application for Single-Band SAR Thematic Information Extraction. IGARSS 2008 - 2008 IEEE Int Geosci Remote Sens Symp II-935-II–938. https://doi.org/10.1109/IGARSS.2008.4779149

de Almeida G a. M, Bates P (2013) Applicability of the local inertial approximation of the shallow water equations to flood modeling. Water Resour Res 49:4833–4844. https://doi.org/10.1002/wrcr.20366

de Almeida G a. M, Bates P, Freer JE, Souvignet M (2012) Improving the stability of a simple formulation of the shallow water equations for 2-D flood modeling. Water Resour Res 48:n/a-n/a. https://doi.org/10.1029/2011WR011570

De Grandi GF, Mayaux P, Malingreau JP, et al (2000) New perspectives on global ecosystems from wide-area radar mosaics: Flooded forest mapping in the tropics. Int J Remote Sens 21:1235–1249. https://doi.org/10.1080/014311600210155

De Lannoy GJM, Houser PR, Pauwels VRN, Verhoest NEC (2006) Assessment of model uncertainty for soil moisture through ensemble verification. J Geophys Res 111:D10101. https://doi.org/10.1029/2005JD006367

Devegowda D, Arroyo-Negrete E, Datta-Gupta A (2010) Flow relevant covariance localization during dynamic data assimilation using EnKF. Adv Water Resour 33:129–145. https://doi.org/10.1016/j.advwatres.2009.10.001

Di Baldassarre G, Montanari A (2009) Uncertainty in river discharge observations: A quantitative analysis. Hydrol Earth Syst Sci 13:913–921. https://doi.org/10.5194/hess-13-913-2009

Di Baldassarre G, Schumann G, Bates PD (2009) A technique for the calibration of hydraulic models using uncertain satellite observations of flood extent. J Hydrol 367:276–282. https://doi.org/10.1016/j.jhydrol.2009.01.020

Di Baldassarre G, Schumann G, Bates PD, et al (2010) Flood-plain mapping: a critical discussion of deterministic and probabilistic approaches. Hydrol Sci J 55:364–376. https://doi.org/10.1080/02626661003683389

Di Baldassarre G, Schumann G, Brandimarte L, Bates P (2011) Timely Low Resolution SAR Imagery To Support Floodplain Modelling: A Case Study Review. Surv Geophys 32:255–269. https://doi.org/10.1007/s10712-011-9111-9

Do HX, Westra S, Leonard M, Gudmundsson L (2010) Global-Scale Prediction of Flood Timing Using Atmospheric Reanalysis. 1–27

Domeneghetti A, Schumann GJP, Frasson RPM, et al (2018a) Characterizing water surface elevation under different flow conditions for the upcoming SWOT mission. J Hydrol 561:848–861. https://doi.org/10.1016/j.jhydrol.2018.04.046

Domeneghetti A, Schumann GJP, Tarpanelli A (2019) Preface: Remote sensing for flood mapping and monitoring of flood dynamics. Remote Sens. 11:11–14

Domeneghetti A, Tarpanelli A, Brocca L, et al (2014a) The use of remote sensing-derived water surface data for hydraulic model calibration. Remote Sens Environ 149:130–141. https://doi.org/10.1016/j.rse.2014.04.007

Domeneghetti A, Tarpanelli A, Brocca L, et al (2014b) The use of remote sensing-derived water surface data for hydraulic model calibration. Remote Sens Environ 149:130–141. https://doi.org/10.1016/j.rse.2014.04.007

Domeneghetti A, Tarpanelli A, Grimaldi L, Brath A (2018b) Flow Duration Curve from Satellite : Potential of a Lifetime SWOT Mission. Remote Sens 10:1–23. https://doi.org/10.3390/rs10071107

Dottori F, Salamon P, Bianchi A, et al (2016) Development and evaluation of a framework for global flood hazard mapping. Adv Water Resour 94:87–102. https://doi.org/10.1016/j.advwatres.2016.05.002

Dowling TI, Brooks M, Read AM (2011) Continental hydrologic assessment using the 1 second (30m) resolution Shuttle Radar Topographic Mission DEM of Australia. MODSIM 2011 - 19th Int Congr Model Simul - Sustain Our Futur Underst Living with Uncertain 2395–2401

Duan Q, Ajami NK, Gao X, Sorooshian S (2007) Multi-model ensemble hydrologic prediction using Bayesian model averaging. Adv Water Resour 30:1371–1386. https://doi.org/10.1016/j.advwatres.2006.11.014

Dumedah G (2012) Formulation of the Evolutionary-Based Data Assimilation, and its Implementation in Hydrological Forecasting. Water Resour Manag 26:3853–3870. https://doi.org/10.1007/s11269-012-0107-0

Dumedah G (2015) Toward essential union between evolutionary strategy and data assimilation for model diagnostics: An application for reducing the search space of optimization problems using hydrologic genome map. Environ Model Softw 69:342–352. https://doi.org/10.1016/j.envsoft.2014.09.025

Dumedah G, Coulibaly P (2014) integration of evolutionary algorithm in particle filter and enkf. J Hydroinformatics 16:74–94

Dumedah G, Coulibaly P (2013a) Evaluating forecasting performance for data assimilation methods: The ensemble Kalman filter, the particle filter, and the evolutionary-based assimilation. Adv Water Resour 60:47–63. https://doi.org/10.1016/j.advwatres.2013.07.007

Dumedah G, Coulibaly P (2013b) Evolutionary assimilation of streamflow in distributed hydrologic modeling using in-situ soil moisture data. Adv Water Resour 53:231–241. https://doi.org/10.1016/j.advwatres.2012.07.012

Dumedah G, Walker JP (2013) Joint model state-parameter retrieval through the evolutionary data assimilation approach. UsersMonashEduAu 1–6

Durand M, Andreadis KM, Alsdorf DE, et al (2008) Estimation of bathymetric depth and slope from data assimilation of swath altimetry into a hydrodynamic model. Geophys Res Lett 35:1–5. https://doi.org/10.1029/2008GL034150

Elvira V, Miguez J, Djurie PM (2017) Adapting the Number of Particles in Sequential Monte Carlo Methods Through an Online Scheme for Convergence Assessment. IEEE Trans Signal Process 65:1781–1794. https://doi.org/10.1109/TSP.2016.2637324

Emerton RE, Stephens EM, Pappenberger F, et al (2016) Continental and global scale flood forecasting systems. Wiley Interdiscip Rev Water 3:391–418. https://doi.org/10.1002/wat2.1137

Evans TL, Costa M, Tomas WM, Camilo AR (2014) Large-scale habitat mapping of the Brazilian Pantanal wetland: A synthetic aperture radar approach. Remote Sens Environ 155:8–108. https://doi.org/10.1016/j.rse.2013.08.051

Evensen G (2003) The Ensemble Kalman Filter: Theoretical formulation and practical implementation. Ocean Dyn 53:343–367. https://doi.org/10.1007/s10236-003-0036-9

Evensen G (2004) Sampling strategies and square root analysis schemes for the EnKF. Ocean Dyn 54:539–560. https://doi.org/10.1007/s10236-004-0099-2

Farr A, Huxley C (2013) Lower Clarence Flood Model Update 2013

Fearnhead P, Künsch H (2017) Particle Filters and Data Assimilation. 1–31. https://doi.org/10.1146/annurev-statistics-031017-100232

Fewtrell T, Bates P, Wit a De, et al (2009) Comparison of varying complexity numerical models for the prediction of flood inundation in Greenwich, UK

Fewtrell TJ, Duncan A, Sampson CC, et al (2011) Benchmarking urban flood models of varying complexity and scale using high resolution terrestrial LiDAR data. Phys Chem Earth 36:281–291. https://doi.org/10.1016/j.pce.2010.12.011

Fleischmann A, Paiva R, Collischonn W (2019) Can regional to continental river hydrodynamic models be locally relevant? A cross-scale comparison. J Hydrol X 3:100027. https://doi.org/10.1016/j.hydroa.2019.100027

Fletcher SJ (2018a) Observation Space Variational Data Assimilation Methods. Data Assim Geosci 1:753–763. https://doi.org/http://dx.doi.org/10.1016/B978-0-12-804444-5.00018-0

Fletcher SJ (2018b) Variational Data Assimilation. In: Data Assimilation for the Geosciences. pp 673–703

Fohringer J, Dransch D, Kreibich H, Schroter K (2015) Social media as an information source for rapid flood inundation mapping. Nat Hazards Earth Syst Sci 15:2725–2738. https://doi.org/10.5194/nhess-15-2725-2015

Franklin SE, Wulder MA, Lavigne MB (1996) Automated derivation of geographic window sizes for use in remote sensing digital image texture analysis. Comput Geosci 22:665–673. https://doi.org/10.1016/0098-3004(96)00009-X

Frappart F, Calmant S, Cauhope M, et al (2006) Preliminary results of ENVISAT RA-2-derived water levels validation over the Amazon basin. Remote Sens Environ 100:252–264. https://doi.org/10.1016/j.rse.2005.10.027

Frasson RP de M, Pavelsky TM, Fonstad MA, et al (2019) Global Relationships Between River Width, Slope, Catchment Area, Meander Wavelength, Sinuosity, and Discharge. Geophys Res Lett 46:3252–3262. https://doi.org/10.1029/2019GL082027

Fu L, Rodriguez E (2004) High-Resolution Measurement of Ocean Surface Topography by Radar Interferometry for Oceanographic and Geophysical Applications. In: Sparks RSJ, Hawkesworth CJ (eds) The State of the Planet: Frontiers and Challenges in Geophysics, Geophysical Monograph Series. American Geophysical Union, pp 209–224

Fujita I, Watanabe H, Tsubaki R (2007) Development of a non-intrusive and efficient flow monitoring technique: The space-time image velocimetry (STIV). Int J River Basin Manag 5:105–114. https://doi.org/10.1080/15715124.2007.9635310

Gallant JC, Dowling TI, Read a. M, et al (2011) 1 second SRTM Derived Products User Guide. 106

Garambois P, Larnier K, Monnier J, et al (2019) Effective channel and ungauged braided river discharge estimation by assimilation of multi-satellite water heights of different spatial sparsity To cite this version : HAL Id : hal-02308560 E ective channel and ungauged braided river discharge estimation

García-pintado J, Mason DC, Dance SL (2014) Moderation of ensemble covariances for data assimilation of satellite-based water level observations into flood modelling. 16:11618

García-Pintado J, Mason DC, Dance SL, et al (2015) Satellite-supported flood forecasting in river networks: A real case study. J Hydrol 523:706–724. https://doi.org/10.1016/j.jhydrol.2015.01.084

García-Pintado J, Neal JC, Mason DC, et al (2013) Scheduling satellite-based SAR acquisition for sequential assimilation of water level observations into flood modelling. J Hydrol 495:252–266. https://doi.org/10.1016/j.jhydrol.2013.03.050

Georgakakos KP, Seo DJ, Gupta H, et al (2004) Towards the characterization of streamflow simulation uncertainty through multimodel ensembles. J Hydrol 298:222–241. https://doi.org/10.1016/j.jhydrol.2004.03.037

Gill MK, Kaheil YH, Khalil A, et al (2006) Multiobjective particle swarm optimization for parameter estimation in hydrology. Water Resour Res 42:n/a-n/a. https://doi.org/10.1029/2005WR004528

Gilles D, Moore M (2010) Review of Hydraulic Flood Modeling Software used in Belgium , The Netherlands , and The United Kingdom. Int Perspect Water Resour Manag

Giustarini L, Chini M, Hostache R, et al (2015a) Flood hazard mapping combining hydrodynamic modeling and multi annual remote sensing data. Remote Sens 7:14200–14226. https://doi.org/10.3390/rs71014200

Giustarini L, Hostache R, Kavetski D, et al (2016) Probabilistic Flood Mapping Using Synthetic Aperture Radar Data. IEEE Trans Geosci Remote Sens 54:6958–6969

Giustarini L, Hostache R, Matgen P, et al (2013) A Change Detection Approach to Flood Mapping in Urban Areas Using TerraSAR-X. IEEE Trans Geosci Remote Sens 51:2417–2430. https://doi.org/10.1109/TGRS.2012.2210901

Giustarini L, Matgen P, Hostache R, et al (2011) Assimilating SAR-derived water level data into a hydraulic model: a case study. Hydrol Earth Syst Sci 15:2349–2365. https://doi.org/10.5194/hess-15-2349-2011

Giustarini L, Matgen P, Hostache R, Dostert J (2012a) From SAR-derived flood mapping to water level data assimilation into hydraulic models. In: Remote Sensing for Agriculture, Ecosystems, and Hydrology XIV. SPIE Remote Sensing, Edinburgh, United Kingdom, p 85310U

Giustarini L, Matgen P, Hostache R, Dostert J (2012b) From SAR-derived flood mapping to water level data assimilation into hydraulic models. Remote Sens Agric Ecosyst Hydrol XIV 8531:85310U. https://doi.org/10.1117/12.974655

Giustarini L, Vernieuwe H, Verwaeren J, et al (2015b) Accounting for image uncertainty in SAR-based flood mapping. Int J Appl Earth Obs Geoinf 34:70–77. https://doi.org/10.1016/j.jag.2014.06.017

Gobeyn S, Neill J, Lievens H, et al (2015) Impact of the SAR acquisition timing on the calibration of a flood inundation model. 17:3822

Gobeyn S, Van Wesemael A, Neal J, et al (2017) Impact of the timing of a SAR image acquisition on the calibration of a flood inundation model. Adv Water Resour 100:126–138. https://doi.org/10.1016/j.advwatres.2016.12.005

Godsill S, Clapp T (2001) Improvement Strategies for Monte Carlo Particle Filters. In: Sequential Monte Carlo Methods in Practice. pp 139–158

Godsill SJ, Doucet A, West M (2004) Monte carlo smoothing for nonlinear time series. J Am Stat Assoc 99:156–168. https://doi.org/10.1198/016214504000000151

Gong W, Gupta H V., Yang D, et al (2013) Estimating epistemic and aleatory uncertainties during hydrologic modeling: An information theoretic approach. Water Resour Res 49:2253–2273. https://doi.org/10.1002/wrcr.20161

Grimaldi S, Li Y, Pauwels VRN, Walker JP (2016) Remote Sensing-Derived Water Extent and Level to Constrain Hydraulic Flood Forecasting Models: Opportunities and Challenges. Surv Geophys 37:977–1034. https://doi.org/10.1007/s10712-016-9378-y

Grimaldi S, Li Y, Walker JP, Pauwels VRN (2017). Bathymetric survey of the Upper Clarence. Monash University. DOI: 10.4225/03/5a20708405ecd

Grimaldi S, Li Y, Walker JP, Pauwels VRN (2018) Effective Representation of River Geometry in Hydraulic Flood Forecast Models. Water Resour Res. https://doi.org/10.1002/2017WR021765

Grimaldi S, Schumann GJ -P., Shokri A, et al (2019) Challenges, opportunities and pitfalls for global coupled hydrologic-hydraulic modeling of floods. Water Resour Res 55:1–24. https://doi.org/10.1029/2018WR024289

Grimaldi S, Xu J, Li Y, et al (2020) Flood mapping under vegetation using single SAR acquisitions. Remote Sens Environ 237:111582. https://doi.org/10.1016/j.rse.2019.111582

Gupta H (2003) Reply to comment by K. Beven and P. Young on "Bayesian recursive parameter estimation for hydrologic models." Water Resour Res 39:1–5. https://doi.org/10.1029/2002WR001405

Gupta HV, Kling H (2011) On typical range , sensitivity , and normalization of Mean Squared Error and Nash-Sutcliffe Efficiency type metrics. 47:2–4. https://doi.org/10.1029/2011WR010962

Gupta HV, Sorooshian S, Yapo PO (1998) Toward improved calibration of hydrologic models : Multiple and noncommensurable measures of information. Water Resour Res 34:751–763

Gupta H V, Kling H, Yilmaz KK, Martinez GF (2009) Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J Hydrol 377:80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Haack B, Bechdol M (1999) Multisensor remote sensing data for land use/cover mapping. Comput Environ Urban Syst 23:53–69. https://doi.org/10.1016/S0198-9715(99)00003-4

Haack B, Bechdol M (2000) Integrating multisensor data and RADAR texture measures for land cover mapping. Comput Geosci 26:411–421. https://doi.org/10.1016/S0098-3004(99)00121-1

Habert J, Ricci S, Le Pape E, et al (2016) Reduction of the uncertainties in the water level-discharge relation of a 1D hydraulic model in the context of operational flood forecasting. J Hydrol 532:52–64. https://doi.org/10.1016/j.jhydrol.2015.11.023

Hagen-Zanker A (2006) Map comparison methods that simultaneously address overlap and structure. J Geogr Syst 8:165–185. https://doi.org/10.1007/s10109-006-0024-y

Hagen-Zanker A, Straatman B, Uljee I (2005) Further developments of a fuzzy set map comparison approach. Int J Geogr Inf Sci 19:769–785. https://doi.org/10.1080/13658810500072137

Hagen A (2003) Fuzzy set approach to assessing similarity of categorical maps. Int J Geogr Inf Sci 17:235–249. https://doi.org/10.1080/13658810210157822

Haile AT (2005) Integrating Hydrodynamic Models and High Resolution DEM(LIDAR) For Flood Modelling. University of Twente

Haile AT, Rientjes THM (2005) Effects Of Lidar Dem Resolution In Flood Modelling : A Model Sentitivity Study For The City Of Tegucigalpa , Honduras. In: ISPRS WG III/3, III/4, V/3 Workshop "Laser scanning 2005", Enschede, The Netherlands. pp 168–173

Haile AT, Rientjes THM (2007) Uncertainty Issues in Hydrodynamic Flood Modeling. In: Proceedings of the 5th International symposium on Spatial Data Quality SDQ 2007, Modelling qualities in space and time, ITC, Enschede, The Netherlands, 13–15 June 2007.

Haralick R, Shanmugan K, Dinstein I (1973) Textural features for image classification. IEEE Trans. Syst. Man Cybern. 3:610–621

Hartanto IM, van der Kwast J, Alexandridis TK, et al (2017) Data assimilation of satellite-based actual evapotranspiration in a distributed hydrological model of a controlled water system. Int J Appl Earth Obs Geoinf 57:123–135. https://doi.org/10.1016/j.jag.2016.12.015

Hastie T, Tibshirani R, Friedman J (2009) The Elements of Statistical Learning. Elements 1:337–387. https://doi.org/10.1007/b94608

Hawker L, Neal JC, Bates P (2019) Accuracy assessment of the TanDEM-X 90 Digital Elevation Model for selected floodplain sites. Remote Sens Environ 232:. https://doi.org/10.1016/j.rse.2019.111319

Haworth RJ, Ollier CD (1992) Continental rifting and drainage reversal: The clarence river of Eastern Australia. Earth Surf Process Landforms 17:387–397. https://doi.org/10.1002/esp.3290170408

He D-C, Wang L (1991) Texture features based on texture spectrum. Pattern Recognit 24:391–399. https://doi.org/10.1016/0031-3203(91)90052-7

Henry J -B., Chastanet P, Fellah K, Desnos Y -L. (2006) Envisat multi-polarized ASAR data for flood mapping. Int J Remote Sens 27:1921–1929. https://doi.org/10.1080/01431160500486724

Herbst E, Schorfheide F (2019) Tempered particle filtering. J Econom 210:26–44. https://doi.org/10.1016/j.jeconom.2018.11.003

Hess LL, Melack JM, Simonett DS (1990) Radar detection of flooding beneath the forest canopy: a review. Int J Remote Sens 11:1313–1325. https://doi.org/10.1080/01431169008955095

Hirpa FA, Salamon P, Beck HE, et al (2018a) Calibration of the Global Flood Awareness System (GloFAS) using daily streamflow data. J Hydrol 566:595–606. https://doi.org/10.1016/j.jhydrol.2018.09.052

Hirpa FA, Salamon P, Beck HE, et al (2018b) Calibration of the Global Flood Awareness System (GloFAS) using daily streamflow data. J Hydrol. https://doi.org/10.1016/j.jhydrol.2018.09.052

Hirschmüller H (2008) Stereo processing by semiglobal matching and mutual information. IEEE Trans Pattern Anal Mach Intell 30:328–341. https://doi.org/10.1109/TPAMI.2007.1166

Hirt C (2018) Artefact detection in global digital elevation models (DEMs): The Maximum Slope Approach and its application for complete screening of the SRTM v4.1 and MERIT DEMs. Remote Sens Environ 207:27–41. https://doi.org/10.1016/j.rse.2017.12.037

Hoeting J a, Madigan D, Raftery a E, Volinsky CT (1999) Bayesian model averaging: A tutorial. Stat Sci 14:382–401. https://doi.org/10.2307/2676803

Horkaew P, Puttinaovarat S (2017) Entropy-based fusion of water indices and DSM derivatives for automatic water surfaces extraction and flood monitoring. ISPRS Int J Geo-Information 6:. https://doi.org/10.3390/ijgi6100301

Horritt MS (2000) Calibration of a two-dimensional finite element flood flow model using satellite radar imagery. Water Resour Res 36:3279–3291

Horritt MS (2006) A methodology for the validation of uncertain flood inundation models. J Hydrol 326:153–165. https://doi.org/10.1016/j.jhydrol.2005.10.027

Horritt MS (1999) A statistical active contour model for SAR image segmentation. Image Vis Comput 17:213–224. https://doi.org/10.1016/S0262-8856(98)00101-2

Horritt MS, Bates PD (2001) Effects of spatial resolution on a raster based model of flood flow. J Hydrol 253:

Horritt MS, Mason DC, Cobby DM, et al (2003) Waterline mapping in flooded vegetation from airborne SAR imagery. Remote Sens Environ 85:271–281. https://doi.org/10.1016/S0034-4257(03)00006-3

Horritt MS, Mason DC, Luckman AJ (2001) Flood boundary delineation from Synthetic Aperture Radar imagery using a statistical active contour model. Int J Remote Sens 22:2489–2507. https://doi.org/10.1080/01431160116902

Horritt MSS, Bates PD (2002) Evaluation of 1D and 2D numerical models for predicting river flood inundation. J Hydrol 268:87–99

Hossain F, Maswood M, Siddique-E-Akbor AH, et al (2014) A promising radar altimetry satellite system for operational flood forecasting in flood-prone bangladesh. IEEE Geosci Remote Sens Mag 2:27–36. https://doi.org/10.1109/MGRS.2014.2345414

Hostache R, Chini M, Giustarini L, et al (2018a) Near real-time assimilation of SAR derived flood maps for improving flood forecasts . Water Resour Res 1–46. https://doi.org/10.1029/2017WR022124

Hostache R, Chini M, Giustarini L, et al (2018b) Near-Real-Time Assimilation of SAR-Derived Flood Maps for Improving Flood Forecasts. Water Resour Res 54:5516–5535. https://doi.org/10.1029/2017WR022205

Hostache R, Chini M, Giustarini L, et al (2018c) Near real-time assimilation of SAR derived flood maps for improving flood forecasts . Water Resour Res 1–46

Hostache R, Lai X, Monnier J, Puech C (2010) Assimilation of spatially distributed water levels into a shallow-water flood model. Part II: Use of a remote sensing image of Mosel River. J Hydrol 390:257–268. https://doi.org/10.1016/j.jhydrol.2010.07.003

Hostache R, Matgen P, Schumann G, et al (2009) Water Level Estimation and Reduction of Hydraulic Model Calibration Uncertainties Using Satellite SAR Images of Floods. IEEE Trans Geosci Remote Sens 1–10

Hostache R, Matgen P, Wagner W (2012) Change detection approaches for flood extent mapping: How to select the most adequate reference image from online archives? Int J Appl Earth Obs Geoinf 19:205–213. https://doi.org/10.1016/j.jag.2012.05.003

Hostache R, Schumann G, Matgen P, et al (2006) 3D Flood Information from SAR as a mean for Reducing Uncertainties in Flood Inundation Modelling. In: XXX ISPRS Congress, Commission VII. pp 217–222

Hou Y hang, Li Y jia, Liang X (2019) Mixed aleatory/epistemic uncertainty analysis and optimization for minimum EEDI hull form design. Ocean Eng 172:308–315. https://doi.org/10.1016/j.oceaneng.2018.12.003

Huang C, Chen Y, Zhang S, Wu J (2018a) Detecting, Extracting, and Monitoring Surface Water From Space Using Optical Sensors: A Review. Rev Geophys 56:333–360. https://doi.org/10.1029/2018RG000598

Huang Q, Long D, Du M, et al (2018b) An improved approach to monitoring Brahmaputra River water levels using retracked altimetry data. Remote Sens Environ 211:112–128. https://doi.org/10.1016/j.rse.2018.04.018

Hunt BR, Kostelich EJ, Szunyogh I (2007) Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. Phys D Nonlinear Phenom 230:112–126. https://doi.org/10.1016/j.physd.2006.11.008

Hunter NM, Bates PD, Horritt MS, Wilson MD (2007) Simple spatially-distributed models for predicting flood inundation: A review. Geomorphology 90:208–225. https://doi.org/10.1016/j.geomorph.2006.10.021

Hunter NM, Bates PD, Neelz S, et al (2008) Benchmarking 2D hydraulic models for urban flooding. Proc ICE - Water Manag 161:13–30. https://doi.org/10.1680/wama.2008.161.1.13

Hutchinson MF (2011) ANUDEM VERSION 5.3 USER GUIDE

Huxley C, Beaman F (2014) Additional crossing of the Clarence River at Grafton: flood impact, levee upgrade, and structural considerations. In: Hydraulic structures and society - Engineering challenges and extremes. Brisbane, Australia, pp 1–8

Intergovernmental Committee On Surveying & Mapping (2008) ICSM Guidelines for Digital Elevation Data. Geosci Aust 1–49

International Charter “Space & Major Disasters” U (2017) Annual Report

Italian Space Agency (2009) COSMO-SkyMed SAR Products Handbook. 105

Jain SK, Singh RD, Jain MK, Lohani a. K (2005) Delineation of flood-prone areas using remote sensing techniques. Water Resour Manag 19:333–347. https://doi.org/10.1007/s11269-005-3281-5

James G, Witten D, Hastie T, Tibshirani R (2000) An introduction to Statistical Learning

Jang JSR (1993) ANFIS: Adaptive-Network-Based Fuzzy Inference System. IEEE Trans Syst Man Cybern 23:665–685. https://doi.org/10.1109/21.256541

Jarihani AA, Callow JN, Mcvicar TR, et al (2015) Satellite-derived Digital Elevation Model (DEM) selection , preparation and correction for hydrodynamic modelling in large , low-gradient and data-sparse catchments. J Hydrol 524:489–506. https://doi.org/10.1016/j.jhydrol.2015.02.049

Jha A, Lamond J, Bloch R, et al (2011) Five feet high and rising: cities and flooding in the 21st century

Johansen AM (2015) On Blocks, Tempering and Particle MCMC for Systems Identification. IFAC-PapersOnLine 48:969–974. https://doi.org/10.1016/j.ifacol.2015.12.256

Jouin M, Gouriveau R, Hissel D, et al (2016) Particle filter-based prognostics: Review, discussion and perspectives. Mech Syst Signal Process 72–73:2–31. https://doi.org/10.1016/j.ymssp.2015.11.008

Jung HC, Jasinski M, Kim J-W, et al (2012) Calibration of two-dimensional floodplain modeling in the central Atchafalaya Basin Floodway System using SAR interferometry. Water Resour Res 48:n/a-n/a. https://doi.org/10.1029/2012WR011951

Kantas N, Doucet A, Singh SS, et al (2015) On Particle Methods for Parameter Estimation in State-Space Models. Stat Sci 30:328–351. https://doi.org/10.1214/14-STS511

Kim D, Yu H, Lee H, et al (2019) Ensemble learning regression for estimating river discharges using satellite altimetry data: Central Congo River as a Test-bed. Remote Sens Environ 221:741–755. https://doi.org/10.1016/j.rse.2018.12.010

Kim S, Sharma A (2019) The role of floodplain topography in deriving basin discharge using passive microwave remote sensing. Water Resour Res. https://doi.org/10.1029/2018WR023627

Kitagawa G (1996) Monte Carlo Filter and Smoother for Non-Gaussian Nonlinear State Space Models. J Comput Graph Stat 5:1–25. https://doi.org/https://doi.org/10.1080/10618600.1996.10474692

Koriche SA (2012) Remote Sensing Based Hydrological Modelling for Flood Early Warning in the Upper and Middle Awash River Basin. University of Twente

Krause P, Boyle DP (2005) Comparison of different efficiency criteria for hydrological model assessment. Adv Geosci 5:89–97

Kropatsch WG, Strobl D (1990) The Generation of SAR Layover and Shadow Maps From Digital Elevation Models. IEEE Trans Geosci Remote Sens 28:98–107. https://doi.org/10.1109/36.45752

Kumar A, Dasgupta A, Lokhande S, Ramsankaran RAAJ (2019) Benchmarking the Indian National CartoDEM against SRTM for 1D hydraulic modelling. Int J River Basin Manag 0:1–10. https://doi.org/10.1080/15715124.2019.1606816

Kuplich TM, Curran PJ, Atkinson PM (2005) Relating SAR image texture to the biomass of regenerating tropical forests. Int J Remote Sens 26:4829–4854. https://doi.org/10.1080/01431160500239107

Kutija V, Bertsch R, Glenis V, et al (2014) Model Validation Using Crowd-Sourced Data From a Large Pluvial Flood. 11th Int Conf Hydroinformatics 9

Lacava T, Ciancia E, Faruolo M, et al (2019) On the Potential of RST-FLOOD on Visible Infrared Imaging Radiometer Suite Data for Flooded Areas Detection. Remote Sens 11:598. https://doi.org/10.3390/rs11050598

Lahoz W, Khattatov B, Menard R, et al (2010) Data Assimilation: Making Sense of Observations, First. Springer-Verlag Berlin Heidelberg, London

Lai X, Liang Q, Yesou H (2013) Variational assimilation of remotely sensed flood extents using a two-dimensional flood model. Hydrol Earth Syst Sci Discuss 10:11185–11220. https://doi.org/10.5194/hessd-10-11185-2013

Lai X, Liang Q, Yesou H, Daillet S (2014) Variational assimilation of remotely sensed flood extents using a 2-D flood model. Hydrol Earth Syst Sci 18:4325–4339. https://doi.org/10.5194/hess-18-4325-2014

Lai X, Monnier J (2009) Assimilation of spatially distributed water levels into a shallow-water flood model. Part I: Mathematical method and test case. J Hydrol 377:1–11. https://doi.org/10.1016/j.jhydrol.2009.07.058

Lakshmivarahan S, Lewis JM (2010) Forward Sensitivity Approach to Dynamic Data Assimilation. 2010:. https://doi.org/10.1155/2010/375615

Landuyt L, Van Wesemael A, Schumann GJ-P, et al (2018) Flood Mapping Based on Synthetic Aperture Radar: An Assessment of Established Approaches. IEEE Trans Geosci Remote Sens PP:1–18. https://doi.org/10.1109/TGRS.2018.2860054

Lang MW, Kasischke ES (2008) Using C-band synthetic aperture radar data to monitor forested wetland hydrology in Maryland's coastal plain, USA. IEEE Trans Geosci Remote Sens 46:535–546. https://doi.org/10.1109/TGRS.2007.909950

Langland RH (2006) Issues in targeted observing. Q J R Meteorol Soc 131:3409–3425. https://doi.org/10.1256/qj.05.130

Le Boursicaud R, Pénard L, Hauet A, et al (2016) Gauging extreme floods on YouTube: Application of LSPIV to home movies for the post-event determination of stream discharges. Hydrol Process 30:90–105. https://doi.org/10.1002/hyp.10532

Le Coz J, Patalano A, Collins D, et al (2016) Crowdsourced data for flood hydrology: Feedback from recent citizen science projects in Argentina, France and New Zealand. J Hydrol 541:766–777. https://doi.org/10.1016/j.jhydrol.2016.07.036

Li J, Wong DWSS (2010) Effects of DEM sources on hydrologic applications. Comput Environ Urban Syst 34:251–261. https://doi.org/10.1016/j.compenvurbsys.2009.11.002

Li L, Chen Y, Yu X, et al (2015a) Sub-pixel flood inundation mapping from multispectral remotely sensed images based on discrete particle swarm optimization. ISPRS J Photogramm Remote Sens 101:10–21. https://doi.org/10.1016/j.isprsjprs.2014.11.006

Li M, Pang B, He Y, Nian F (2013a) Particle filter improved by genetic algorithm and particle swarm optimization algorithm. J Softw 8:666–672. https://doi.org/10.4304/jsw.8.3.666-672

Li M, Zhang P, Wu Y, et al (2012) Unsupervised Change Detection on SAR Images Using Triplet Markov Field Model. IEEE Geosci Remote Sens Lett 10:1–5. https://doi.org/10.1109/LGRS.2012.2219494

Li T, Sun S, Sattar TP, Corchado JM (2014a) Fight sample degeneracy and impoverishment in particle filters: A review of intelligent approaches. Expert Syst Appl 41:3944–3954. https://doi.org/10.1016/j.eswa.2013.12.031

Li Y, Martinis S, Wieland M (2019) Urban flood mapping with an active self-learning convolutional neural network based on TerraSAR-X intensity and interferometric coherence. ISPRS J Photogramm Remote Sens 152:178–191. https://doi.org/10.1016/j.isprsjprs.2019.04.014

Li Y, Ryu D, Western AW, Wang QJ (2013b) Assimilation of stream discharge for flood forecasting: The benefits of accounting for routing time lags. Water Resour Res 49:1887–1900. https://doi.org/10.1002/wrcr.20169

Li Y, Ryu D, Western AW, Wang QJ (2014b) Towards improved operational flood forecasting by assimilation of real- time streamflow measurements Research outline Objective

Li Y, Ryu D, Western AW, Wang QJ (2015b) Assimilation of stream discharge for flood forecasting: Updating a semidistributed model with an integrated data assimilation scheme. Water Resour Res n/a-n/a. https://doi.org/10.1002/2014WR016667

Lievens H, De Lannoy GJM, Al Bitar A, et al (2016) Assimilation of SMOS soil moisture and brightness temperature products into a land surface model. Remote Sens Environ 180:292–304. https://doi.org/10.1016/j.rse.2015.10.033

Liu J, Kalnay E (2008) Estimating observation impact without adjoint model in an ensemble Kalman filter. Q J R Meteorol Soc 134:1327–1335. https://doi.org/10.1002/qj.280

Liu X, Chen S, Zhuo L, et al (2018) Multi-sensor image registration by combining local self-similarity matching and mutual information. Front Earth Sci 12:779–790. https://doi.org/10.1007/s11707-018-0717-9

Liu Y, Gupta H V. (2007) Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. Water Resour Res 43:n/a-n/a. https://doi.org/10.1029/2006WR005756

Long Y, Hu X (2017) Spatial Partition-Based Particle Filtering for Data Assimilation in Wildfire Spread Simulation. ACM Trans Spat Algorithms Syst 3:5:1--5:33. https://doi.org/10.1145/3099471

Lu S, Wu B, Yan N, Wang H (2011) Water body mapping method with HJ-1A/B satellite imagery. Int J Appl Earth Obs Geoinf 13:428–434. https://doi.org/10.1016/j.jag.2010.09.006

Lymburner L, Tan P, Mueller N, et al (2011) The National Dynamic Land Cover Dataset

Madsen H, Canizares R (1999) Comparison of Extended and Ensemble Kalman Filters . . . Int J Numer Methods Fluids 31:961–981

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

Martinis S, Kersten J, Twele A (2015a) A fully automated TerraSAR-X based flood service. ISPRS J Photogramm Remote Sens 104:203–212. https://doi.org/10.1016/j.isprsjprs.2014.07.014

Martinis S, Kuenzer C, Twele A (2015b) Flood Studies Using Synthetic Aperture Radar Data. Remote Sens Water Resour Disasters, Urban Stud 145–173. https://doi.org/doi:10.1201/b19321-10

Martinis S, Kuenzer C, Wendleder A, et al (2015c) Comparing four operational SAR-based water and flood detection approaches. Int J Remote Sens 36:3519–3543. https://doi.org/10.1080/01431161.2015.1060647

Martinis S, Rieke C (2015) Backscatter analysis using multi-temporal and multi-frequency SAR data in the context of flood mapping at River Saale, Germany. Remote Sens 7:7732–7752. https://doi.org/10.3390/rs70607732

Martinis S, Twele a., Voigt S (2009) Towards operational near real-time flood detection using a split-based automatic thresholding procedure on high resolution TerraSAR-X data. Nat Hazards Earth Syst Sci 9:303–314. https://doi.org/10.5194/nhess-9-303-2009

Martinis S, Twele A (2010) A hierarchical spatio-temporal Markov model for improved flood mapping using multi-temporal X-band SAR data. Remote Sens 2:2240–2258. https://doi.org/10.3390/rs2092240

Martinis S, Twele A, Voigt S (2011) Unsupervised extraction of flood-induced backscatter changes in SAR data using markov image modeling on irregular graphs. IEEE Trans Geosci Remote Sens 49:251–263. https://doi.org/10.1109/TGRS.2010.2052816

Mason DC, Cobby DM, Horritt MS, Bates PD (2003) Floodplain friction parameterization in two-dimensional river flood models using vegetation heights derived from airborne scanning laser altimetry. Hydrol Process 17:1711–1732. https://doi.org/10.1002/hyp.1270

Mason DC, Davenport IJ, Neal JC, et al (2012a) Near Real-Time Flood Detection in Urban and Rural Areas Using High-Resolution Synthetic Aperture Radar Images. IEEE Trans Geosci Remote Sens 50:3041–3052. https://doi.org/10.1109/TGRS.2011.2178030

Mason DC, Davenport LJ (1996) Accurate and efficient determination of the shoreline in ERS-1 SAR images. IEEE Trans Geosci Remote Sens 34:1243–1253. https://doi.org/10.1109/36.536540

Mason DC, Garcia-Pintado J, Cloke HL, Dance SL (2015) The potential of flood forecasting using a variable-resolution global Digital Terrain Model and flood extents from Synthetic Aperture Radar images. Front Earth Sci 3:1–14. https://doi.org/10.3389/feart.2015.00043

Mason DC, Giustarini L, Garcia-Pintado J, Cloke HL (2014) Detection of flooded urban areas in high resolution Synthetic Aperture Radar images using double scattering. Int J Appl Earth Obs Geoinf 28:150–159. https://doi.org/10.1016/j.jag.2013.12.002

Mason DC, Horritt MS, Dall'Amico JT, et al (2007) Improving river flood extent delineation from synthetic aperture radar using airborne laser altimetry. IEEE Trans Geosci Remote Sens 45:3932–3943. https://doi.org/10.1109/TGRS.2007.901032

Mason DC, Schumann GJ-P, Neal JC, et al (2012b) Automatic near real-time selection of flood water levels from high resolution Synthetic Aperture Radar images for assimilation into hydraulic models: A case study. Remote Sens Environ 124:705–716. https://doi.org/10.1016/j.rse.2012.06.017

Mason DC, Schumann GJP, Bates PD (2010a) Data Utilization in Flood Inundation Modelling. In: Flood Risk Science and Management. pp 209–233

Mason DC, Speck R, Devereux B, et al (2010b) Flood Detection in Urban Areas Using TerraSAR-X. Ieee Trans Geosci Remote Sens 48:882–894. https://doi.org/10.1109/tgrs.2009.2029236

Mason DC, Trigg M, Garcia-Pintado J, et al (2016) Improving the TanDEM-X Digital Elevation Model for flood modelling using flood extents from Synthetic Aperture Radar images. Remote Sens Environ 173:15–28. https://doi.org/10.1016/j.rse.2015.11.018

Mason DCC, Bates PDD, Dall' Amico JT, Dall' Amico JT (2009) Calibration of uncertain flood inundation models using remotely sensed water levels. J Hydrol 368:224–236. https://doi.org/10.1016/j.jhydrol.2009.02.034

Matgen P, Giustarini L, Chini M, et al (2016) Creating a water depth map from SAR flood extent and topography data. Int Geosci Remote Sens Symp 2016-Novem:7635–7638. https://doi.org/10.1109/IGARSS.2016.7730991

Matgen P, Henry J-B, Pappenberger F, et al (2004) Uncertainty in Calibrating Flood Propagation Models with Flood Boundaries Derived from SAR Imagery. In: XX ISPRS Congress, Commission VII

Matgen P, Hostache R, Schumann G, et al (2011) Towards an automated SAR-based flood monitoring system: Lessons learned from two case studies. Phys Chem Earth 36:241–252. https://doi.org/10.1016/j.pce.2010.12.009

Matgen P, Montanari M, Hostache R, et al (2010) Towards the sequential assimilation of SAR-derived water stages into hydraulic models using the Particle Filter: proof of concept. Hydrol Earth Syst Sci 14:1773–1785. https://doi.org/10.5194/hess-14-1773-2010

Matgen P, Schumann G, Henry J-B, et al (2007a) Integration of SAR-derived river inundation areas, high-precision topographic data and a river flow model toward near real-time flood management. Int J Appl Earth Obs Geoinf 9:247–263. https://doi.org/10.1016/j.jag.2006.03.003

Matgen P, Schumann GUY, Pappenberger F, Pfister L (2007b) Sequential assimilation of remotely sensed water stages in flood inundation models. In: Proceedings of Symposium HS3007 at IUGG2007. Perugia, Italy, pp 78–88

Mattern JP, Dowd M, Fennel K (2013) Particle filter-based data assimilation for a three-dimensional biological ocean model and satellite observations. J Geophys Res Ocean 118:2746–2760. https://doi.org/10.1002/jgrc.20213

Mazzoleni M (2016) Improving Flood Prediction Assimilating Uncertain Crowdsourced Data into Hydrologic and Hydraulic Models. Delft University of Technology and UNESCO-IHE Institute for Water Education

Mazzoleni M, Alfonso L, Chacon-Hurtado J, Solomatine D (2015) Assimilating uncertain, dynamic and intermittent streamflow observations in hydrological models. Adv Water Resour 83:323–339. https://doi.org/10.1016/j.advwatres.2015.07.004

Mazzoleni M, Chacon-Hurtado J, Noh SJ, et al (2018) Data Assimilation in Hydrologic Routing: Impact of Model Error and Sensor Placement on Flood Forecasting. J Hydrol Eng 23:04018018. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001656

Mazzoleni M, Verlaan M, Alfonso L, et al (2017) Can assimilation of crowdsourced streamflow observations in hydrological modelling improve flood prediction? Hydrol Earth Syst Sci 21:839–861. https://doi.org/10.5194/hessd-12-11371-2015

McFeeters SK (1996) The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. Int J Remote Sens 17:1425–1432. https://doi.org/10.1080/01431169608948714

McFeeters SK (2013) Using the normalized difference water index (ndwi) within a geographic information system to detect swimming pools for mosquito abatement: A practical approach. Remote Sens 5:3544–3561. https://doi.org/10.3390/rs5073544

McMillan HK, Hreinsson EO, Clark MP, et al (2013) Operational hydrological data assimilation with the recursive ensemble Kalman filter. Hydrol Earth Syst Sci 17:21–38. https://doi.org/10.5194/hess-17-21-2013

Mcqueen KG (2016) Landscape evolution of the Clarence River catchment : Weird rivers and wild ideas. In: Fourth Australian Regolith Geoscientists Association Conference. Thredbo, NSW

Michailovsky CI, McEnnis S, Berry PAM, et al (2012) River monitoring from satellite radar altimetry in the Zambezi River basin. Hydrol Earth Syst Sci 16:2181–2192. https://doi.org/10.5194/hess-16-2181-2012

Moore MR (2011) Development of a high-resolution 1D/2D coupled flood simulation of Charles City, Iowa. University of Iowa

Moradkhani H (2008) Hydrologic Remote Sensing and Land Surface Data Assimilation. Sensors 8:2986–3004. https://doi.org/10.3390/s8052986

Moradkhani H, Dechant CM, Sorooshian S (2012) Evolution of ensemble data assimilation for uncertainty quantification using the particle filter-Markov chain Monte Carlo method. Water Resour Res 48:. https://doi.org/10.1029/2012WR012144

Moradkhani H, Hsu K-L, Gupta H, Sorooshian S (2005a) Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter. Water Resour Res 41:1–17. https://doi.org/10.1029/2004WR003604

Moradkhani H, Sorooshian S, Gupta H V., Houser PR (2005b) Dual state-parameter estimation of hydrological models using ensemble Kalman filter. Adv Water Resour 28:135–147. https://doi.org/10.1016/j.advwatres.2004.09.002

Moulatlet GM, Rennó CD, Costa FRC, et al (2015) Mapping hydrological environments in central Amazonia: Ground validation and surface model based on SRTM DEM data corrected for deforestation. Earth Syst Sci Data 7:29–34. https://doi.org/10.5194/essd-7-29-2015

Mueller N, Lewis A, Roberts D, et al (2016) Water observations from space: Mapping surface water from 25years of Landsat imagery across Australia. Remote Sens Environ 174:341–352. https://doi.org/10.1016/j.rse.2015.11.003

Mujumdar PP (2001) Flood wave propagation. Resonance 6:66–73. https://doi.org/10.1007/bf02839085

Mukherjee NR, Samuel C (2016) Assessment of the temporal variations of surface water bodies in and around Chennai using landsat imagery. Indian J Sci Technol 9:. https://doi.org/10.17485/ijst/2016/v9i18/92089

Mukherjee SS, Joshi PK, Mukherjee SS, et al (2013) Evaluation of vertical accuracy of open source Digital Elevation Model (DEM). Int J Appl Earth Obs Geoinf 21:205–217. https://doi.org/10.1016/j.jag.2012.09.004

Musa ZN, Popescu I, Mynett a. (2015) A review of applications of satellite SAR, optical, altimetry and DEM data for surface water modelling, mapping and parameter estimation. Hydrol Earth Syst Sci Discuss 12:4857–4878. https://doi.org/10.5194/hessd-12-4857-2015

Muste M, Ho HC, Kim D (2011) Considerations on direct stream flow measurements using video imagery: Outlook and research needs. J Hydro-Environment Res 5:289–300. https://doi.org/10.1016/j.jher.2010.11.002

Nakamura K, Hirose N, Choi BH, Higuchi T (2009) Particle Filtering in Data Assimilation and Its Application to Estimation of Boundary. https://doi.org/10.1007/978-3-540-71056-1

National Flood Risk Advisory Group, Australia-New Zealand Emergeny Management Committee (2014) Technical flood risk management guideline : Flood hazard

Neal J, Schumann G, Bates P, et al (2009) A data assimilation approach to discharge estimation from space. Hydrol Process 23:3641–3649. https://doi.org/10.1002/hyp.7518

Neal J, Schumann G, Fewtrell T, et al (2011) Evaluating a new LISFLOOD-FP formulation with data from the summer 2007 floods in Tewkesbury, UK. J Flood Risk Manag 4:88–95. https://doi.org/10.1111/j.1753-318X.2011.01093.x

Neal J, Villanueva I, Wright N, et al (2012) How much physical complexity is needed to model flood inundation? Hydrol Process 26:2264–2282. https://doi.org/10.1002/hyp.8339

Neal JC, Atkinson PM, Hutton CW (2007) Flood inundation model updating using an ensemble Kalman filter and spatially distributed measurements. J Hydrol 336:401–415. https://doi.org/10.1016/j.jhydrol.2007.01.012

Neal JC, Odoni N a., Trigg MA, et al (2015) Efficient incorporation of channel cross-section geometry uncertainty into regional and global scale flood inundation models. J Hydrol 529:169–183. https://doi.org/10.1016/j.jhydrol.2015.07.026

Nester T, Komma J, Viglione A, Blöschl G (2012) Flood forecast errors and ensemble spread-A case study. Water Resour Res 48:1–19. https://doi.org/10.1029/2011WR011649

Neuman SP (2003) Maximum likelihood Bayesian averaging of uncertain model predictions. Stoch Environ Res Risk Assess 17:291–305. https://doi.org/10.1007/s00477-003-0151-7

Nichols NK (2010) Mathematical Concepts of Data Assimilation. In: Lahoz W (ed) Data Assimilation. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 13–39

Nico G, Pappalepore M, Pasquariello G, et al (2000) Comparison of SAR amplitude vs. coherence flood detection methods -a GIS application. Int J Remote Sens 21:1619–1631. https://doi.org/10.1080/014311600209931

NLWRA NL and WRA (2000) Australian Water Resource Assessment

Nobre AD, Cuartas LA, Momo MR, et al (2016) HAND contour: A new proxy predictor of inundation extent. Hydrol Process 30:320–333. https://doi.org/10.1002/hyp.10581

Noh SJIN, Tachikawa Y, Shiiba M, Kim S (2013) Sequential data assimilation for streamflow forecasting using a distributed hydrologic model : particle filtering and ensemble Kalman filtering. Floods From Risk to Oppor 357:341–349

O'Grady D, Leblanc M, Bass A (2014) The use of radar satellite data from multiple incidence angles improves surface water mapping. Remote Sens Environ 140:652–664. https://doi.org/10.1016/j.rse.2013.10.006

O'Grady D, Leblanc M, Gillieson D (2013) Relationship of local incidence angle with satellite radar backscatter for different surface conditions. Int J Appl Earth Obs Geoinf 24:42–53. https://doi.org/10.1016/j.jag.2013.02.005

O'Hagan A (2006) Bayesian Analysis of Computer Code Outputs. Reliab Eng Syst Saf 91:1290–1300. https://doi.org/10.1007/978-1-4471-0657-9_11

O'Loughlin FE, Paiva RCD, Durand M, et al (2016) A multi-sensor approach towards a global vegetation corrected SRTM DEM product. Remote Sens Environ 182:49–59. https://doi.org/10.1016/j.rse.2016.04.018

Oakley JE, O'Hagan A (2004) Probabilistic sensitivity analysis of complex models: a Bayesian approach. J R Stat Soc 66:751–769. https://doi.org/10.1111/j.1467-9868.2004.05304.x

Ogilvie A, Belaud G, Delenne C, et al (2015) Decadal monitoring of the Niger Inner Delta flood dynamics using MODIS optical data. J Hydrol 523:368–383. https://doi.org/10.1016/j.jhydrol.2015.01.036

Ohki M, Tadono T, Itoh T, et al (2019) Flood Area Detection Using PALSAR-2 Amplitude and Coherence Data: The Case of the 2015 Heavy Rainfall in Japan. IEEE J Sel Top Appl Earth Obs Remote Sens PP:1–11. https://doi.org/10.1109/jstars.2019.2911596

Oliveira ER, Disperati L, Cenci L, et al (2019) Multi-Index Image Differencing Method (MINDED) for Flood Extent Estimations. Remote Sens 11:1–29

Ordoyne C, Friedl M a. (2008) Using MODIS data to characterize seasonal inundation patterns in the Florida Everglades. Remote Sens Environ 112:4107–4119. https://doi.org/10.1016/j.rse.2007.08.027

Otsu N (1979) A Threshold Selection Method from Gray-Level Histograms. IEEE Trans Syst Man Cybern 9:62–66. https://doi.org/10.1109/TSMC.1979.4310076

Oubanas H, Gejadze I, Malaterre PO, et al (2018a) Discharge Estimation in Ungauged Basins Through Variational Data Assimilation: The Potential of the SWOT Mission. Water Resour Res 54:2405–2423. https://doi.org/10.1002/2017WR021735

Oubanas H, Gejadze I, Malaterre PO, Mercier F (2018b) River discharge estimation from synthetic SWOT-type observations using variational data assimilation and the full Saint-Venant hydraulic model. J Hydrol 559:638–647. https://doi.org/10.1016/j.jhydrol.2018.02.004

Ouled Sghaier M, Hammami I, Foucher S, Lepage R (2018) Flood Extent Mapping from Time-Series SAR Images Based on Texture Analysis and Data Fusion. Remote Sens 10:237. https://doi.org/10.3390/rs10020237

Paiva RCD, Collischonn W, Bonnet MP, et al (2013) Assimilating in situ and radar altimetry data into a large-scale hydrologic-hydrodynamic model for streamflow forecast in the Amazon. Hydrol Earth Syst Sci 17:2929–2946. https://doi.org/10.5194/hess-17-2929-2013

Panegrossi G, Ferretti R, Pulvirenti L, Pierdicca N (2011) Impact of ASAR soil moisture data on the MM5 precipitation forecast for the Tanaro flood event of April 2009. Nat Hazards Earth Syst Sci 11:3135–3149. https://doi.org/10.5194/nhess-11-3135-2011

Pappenberger F, Beven K, Frodsham K, et al (2007a) Grasping the unavoidable subjectivity in calibration of flood inundation models: A vulnerability weighted approach. J Hydrol 333:275–287. https://doi.org/10.1016/j.jhydrol.2006.08.017

Pappenberger F, Beven K, Horritt M, Blazkova S (2005) Uncertainty in the calibration of effective roughness parameters in HEC-RAS using inundation and downstream level observations. J Hydrol 302:46–69. https://doi.org/10.1016/j.jhydrol.2004.06.036

Pappenberger F, Frodsham K, Beven KJ, et al (2007b) Fuzzy set approach to calibrating distributed flood inundation models using remote sensing observations. Hydrol Earth Syst Sci 11:739–752. https://doi.org/10.5194/hess-11-739-2007

Pappenberger F, Ramos MH, Cloke HL, et al (2015) How do I know if my forecasts are better? Using benchmarks in hydrological ensemble prediction. J Hydrol 522:697–713. https://doi.org/10.1016/j.jhydrol.2015.01.024

Patil A, Ramsankaran R (2018) Improved streamflow simulations by coupling Soil Moisture Analytical Relationship in EnKF based hydrological data assimilation framework. Adv Water Resour 121:173–188. https://doi.org/10.1016/J.ADVWATRES.2018.08.010

Patro S, Chatterjee C, Singh R, Raghuwanshi NS (2009) Hydrodynamic modelling of a large flood-prone river system in India with limited data. Hydrol Process 23:2774–2791. https://doi.org/10.1002/hyp.7375

Pekel JF, Cottam A, Gorelick N, Belward AS (2016) High-resolution mapping of global surface water and its long-term changes. Nature 540:418–422. https://doi.org/10.1038/nature20584

Penny SG, Miyoshi T (2016) A local particle filter for high-dimensional geophysical systems. Nonlinear Process Geophys 23:391–405. https://doi.org/10.5194/npg-23-391-2016

Perdigão RAP, Blöschl G (2014) Spatiotemporal flood sensitivity to annual precipitation: Evidence for landscape-climate coevolution. Water Resour Res 50:5492–5509. https://doi.org/10.1002/2014WR015365

Pham HT, Marshall L, Johnson F, Sharma A (2018) A method for combining SRTM DEM and ASTER GDEM2 to improve topography estimation in regions without reference data. Remote Sens Environ 210:229–241. https://doi.org/10.1016/j.rse.2018.03.026

Piazzi G, Thirel G, Campo L, Gabellani S (2018) A Particle Filter scheme for multivariate data assimilation into a point-scale snowpack model in Alpine environment. Cryosph Discuss 1–37. https://doi.org/10.5194/tc-2017-286

Pierdicca N, Chini M, Pulvirenti L, Macina F (2008) Integrating Physical and Topographic Information Into a Fuzzy Scheme to Map Flooded Area by SAR. Sensors 8:4151–4164. https://doi.org/10.3390/s8074151

Pierdicca N, Pulvirenti L, Boni G, et al (2017) Mapping Flooded Vegetation Using COSMO-SkyMed: Comparison With Polarimetric and Optical Data Over Rice Fields. IEEE J Sel Top Appl Earth Obs Remote Sens 10:1–13. https://doi.org/10.1109/JSTARS.2017.2711960

Pierdicca N, Pulvirenti L, Chini M, et al (2013) Observing floods from space: Experience gained from COSMO-SkyMed observations. Acta Astronaut 84:122–133. https://doi.org/10.1016/j.actaastro.2012.10.034

Pierdicca N, Pulvirenti L, Chini M, et al (2014) Flood mapping by SAR: Possible approaches to mitigate errors due to ambiguous radar signatures. Int Geosci Remote Sens Symp 3850–3853. https://doi.org/10.1109/IGARSS.2014.6947324

Pitcher LH, Pavelsky TM, Smith LC, et al (2018) AirSWOT InSAR mapping of surface water elevations and hydraulic gradients across the Yukon Flats Basin, Alaska. Water Resour Res. https://doi.org/10.1029/2018WR023274

Plank S, Jüssi M, Martinis S, Twele A (2017) Mapping of flooded vegetation by means of polarimetric Sentinel-1 and ALOS-2/PALSAR-2 imagery. Int J Remote Sens 38:3831–3850. https://doi.org/10.1080/01431161.2017.1306143

Plaza DA, De Keyser R, De Lannoy GJM, et al (2012) The importance of parameter resampling for soil moisture data assimilation into hydrologic models using the particle filter. Hydrol Earth Syst Sci 16:375–390. https://doi.org/10.5194/hess-16-375-2012

Plaza Guingla DA, De Keyser R, De Lannoy GJM, et al (2013) Improving particle filters in rainfall-runoff models: Application of the resample-move step and the ensemble Gaussian particle filter. Water Resour Res 49:4005–4021. https://doi.org/10.1002/wrcr.20291

Pradhan B, Hagemann U, Shafapour Tehrany M, Prechtel N (2014) An easy to use ArcMap based texture analysis program for extraction of flooded areas from TerraSAR-X satellite image. Comput Geosci 63:34–43. https://doi.org/10.1016/j.cageo.2013.10.011

Pradhan B, Tehrany MS, Jebur MN (2016) A New Semiautomated Detection Mapping of Flood Extent From TerraSAR-X Satellite Image Using Rule-Based Classification and Taguchi Optimization Techniques. IEEE Trans. Geosci. Remote Sens. 54:4331–4342

Pramanik N, Panda RK, Sen D (2010) One dimensional hydrodynamic modeling of river flow using DEM extracted river cross-sections. Water Resour Manag 24:835–852. https://doi.org/10.1007/s11269-009-9474-6

Prigent C, Lettenmaier DP, Aires F, Papa F (2016) Toward a High-Resolution Monitoring of Continental Surface Water Extent and Dynamics, at Global Scale: from GIEMS (Global Inundation Extent from Multi-Satellites) to SWOT (Surface Water Ocean Topography). Surv Geophys 37:339–355. https://doi.org/10.1007/s10712-015-9339-x

Puech C, Hostache R, Raclot D, Matgen P (2007) Estimation of flood water levels by merging DEM and satellite imagery using hydraulics laws through AI to enhance the estimates. In: Proceedings of Second Space for Hydrology Workshop. ESA. Geneva, pp 1–7

Pulvirenti L, Chini M, Marzano FS, et al (2012) Detection of floods and heavy rain using Cosmo-SkyMed data: The event in Northwestern Italy of November 2011. Int Geosci Remote Sens Symp 3026–3029. https://doi.org/10.1109/IGARSS.2012.6350788

Pulvirenti L, Chini M, Pierdicca N, et al (2011a) Flood monitoring using multi-temporal COSMO-SkyMed data: Image segmentation and signature interpretation. Remote Sens Environ 115:990–1002. https://doi.org/10.1016/j.rse.2010.12.002

Pulvirenti L, Chini M, Pierdicca N, Boni G (2015) Use of SAR Data for Detecting Floodwater in Urban and Agricultural Areas: The Role of the Interferometric Coherence. IEEE Trans Geosci Remote Sens 54:1532–1544. https://doi.org/10.1109/TGRS.2015.2482001

Pulvirenti L, Marzano FS, Pierdicca N, et al (2014a) Discrimination of water surfaces, heavy rainfall, and wet snow using COSMO-SkyMed observations of severe weather events. IEEE Trans Geosci Remote Sens 52:858–869. https://doi.org/10.1109/TGRS.2013.2244606

Pulvirenti L, Pierdicca N, Boni G, et al (2014b) Flood damage assessment through multitemporal COSMO-SkyMed data and hydrodynamic models: The Albania 2010 case study. IEEE J Sel Top Appl Earth Obs Remote Sens 7:2848–2855. https://doi.org/10.1109/JSTARS.2014.2328012

Pulvirenti L, Pierdicca N, Chini M, Guerriero L (2011b) An algorithm for operational flood mapping from Synthetic Aperture Radar (SAR) data using fuzzy logic. Nat Hazards Earth Syst Sci 11:529–540. https://doi.org/10.5194/nhess-11-529-2011

Pulvirenti L, Pierdicca N, Chini M, Guerriero L (2013) Monitoring flood evolution in vegetated areas using cosmo-skymed data: The tuscany 2009 case study. IEEE J Sel Top Appl Earth Obs Remote Sens 6:1807–1816. https://doi.org/10.1109/JSTARS.2012.2219509

Quinn N, Bates PD, Neal J, et al (2019) The Spatial Dependence of Flood Hazard and Risk in the United States. Water Resour Res 55:1890–1911. https://doi.org/10.1029/2018WR024205

Rakovec O (2014) Improving operational flood forecasting using data assimilation. Wageningen University

Refice A, Capolongo D, Pasquariello G, et al (2014) SAR and InSAR for flood monitoring: Examples with COSMO-SkyMed data. IEEE J Sel Top Appl Earth Obs Remote Sens 7:2711–2722. https://doi.org/10.1109/JSTARS.2014.2305165

Rémy S, Pannekoucke O, Bergot T, Baehr C (2009) Adaptation of a particle filtering method for data assimilation in a 1D numerical model used for fog forecasting. 19:1–19. https://doi.org/10.1002/qj

Rennó CD, Nobre AD, Cuartas LA, et al (2008) HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia. Remote Sens Environ 112:3469–3481. https://doi.org/10.1016/j.rse.2008.03.018

Revel M, Ikeshima D, Yamazaki D, Kanae S (2019) A physically based empirical localization method for assimilating synthetic SWOT observations of a continental-scale river: A case study in the Congo basin. Water (Switzerland) 11:. https://doi.org/10.3390/w11040829

Revilla-Romero B, Hirpa FA, del Pozo JT, et al (2015) On the use of global flood forecasts and satellite-derived inundation maps for flood monitoring in data-sparse regions. Remote Sens 7:15702–15728. https://doi.org/10.3390/rs71115702

Revilla-Romero B, Wanders N, Burek P, et al (2016) Integrating remotely sensed surface water extent into continental scale hydrology. J Hydrol 543:659–670. https://doi.org/10.1016/j.jhydrol.2016.10.041

Richards JA, Woodgate PW, Skidmore AK (1987) An explanation of enhanced radar backscattering from flooded forests. Int J Remote Sens 8:1093–1100. https://doi.org/10.1080/01431168708954756

Ridolfi E, Alfonso L, Di Baldassarre G, et al (2014) An entropy approach for the optimization of cross-section spacing for river modelling. Hydrol Sci J 59:126–137. https://doi.org/10.1080/02626667.2013.822640

Robinson N, Regetz J, Guralnick RP (2014) EarthEnv-DEM90: A nearly-global, void-free, multi-scale smoothed, 90m digital elevation model from fused ASTER and SRTM data. ISPRS J Photogramm Remote Sens 87:57–67. https://doi.org/10.1016/j.isprsjprs.2013.11.002

Rodell M, Houser PR, Jambor U, et al (2004) The Global Land Data Assimilation System. Bull Am Meteorol Soc 85:381–394. https://doi.org/10.1175/BAMS-85-3-381

Rogencamp G (2004) Lower Clarence River Flood Study Review – Final Report: March 2004: Volume 1 of 2 Main Text

Sala J, Lopez A, Romero L, Koudogbo F (2016) A Sentinel-1 Flood map generation QGIS plugin. 18:14693

Samuel J, Coulibaly P, Dumedah G, Moradkhani H (2014) Assessing Model State and Forecasts Variation in Hydrologic Data Assimilation. J Hydrol 513:127–141. https://doi.org/10.1016/j.jhydrol.2014.03.048

Sanders BF (2007) Evaluation of on-line DEMs for flood inundation modeling. Adv Water Resour 30:1831–1843. https://doi.org/10.1016/j.advwatres.2007.02.005

Sanders BF, Schubert JE (2019) PRIMo: Parallel raster inundation model. Adv Water Resour 126:79–95. https://doi.org/10.1016/J.ADVWATRES.2019.02.007

Sanyal J, Lu XX (2004) Application of Remote Sensing in Flood Management with Special Reference to Monsoon Asia: A Review. Nat Hazards 33:283–301. https://doi.org/10.1023/B:NHAZ.0000037035.65105.95

Schlaffer S, Chini M, Dettmering D, Wagner W (2016) Mapping wetlands in Zambia using seasonal backscatter signatures derived from ENVISAT ASAR time series. Remote Sens 8:1–24. https://doi.org/10.3390/rs8050402

Schlaffer S, Chini M, Giustarini L, Matgen P (2017) Probabilistic mapping of flood-induced backscatter changes in SAR time series. Int J Appl Earth Obs Geoinf 56:77–87. https://doi.org/10.1016/j.jag.2016.12.003

Schlaffer S, Matgen P, Hollaus M, Wagner W (2015) Flood detection from multi-temporal SAR data using harmonic analysis and change detection. Int J Appl Earth Obs Geoinf 38:15–24. https://doi.org/10.1016/j.jag.2014.12.001

Schnebele E, Cervone G, Waters N (2014) Road assessment after flood events using non-authoritative data. Nat Hazards Earth Syst Sci 14:1007–1015. https://doi.org/10.5194/nhess-14-1007-2014

Schneider R, Tarpanelli A, Nielsen K, et al (2018) Evaluation of multi-mode CryoSat-2 altimetry data over the Po River against in situ data and a hydrodynamic model. Adv Water Resour 112:17–26. https://doi.org/10.1016/j.advwatres.2017.11.027

Schumann G, Bates PD, Horritt MS, et al (2009a) Progress in integration of remote sensing-derived flood extent and stage data and hydraulic models. Rev Geophys 47:1–20. https://doi.org/10.1029/2008RG000274

Schumann G, Brakenridge G, Kettner A, et al (2018) Assisting Flood Disaster Response with Earth Observation Data and Products: A Critical Assessment. Remote Sens 10:1230. https://doi.org/10.3390/rs10081230

Schumann G, Cutler M, Black A, et al (2008a) Evaluating uncertain flood inundation predictions with uncertain remotely sensed water stages. Int J River Basin Manag 5124:37–41. https://doi.org/10.1080/15715124.2008.9635347

Schumann G, Di Baldassarre G, Alsdorf D, Bates PD (2010) Near real-time flood wave approximation on large rivers from space: Application to the River Po, Italy. Water Resour Res 46:1–8. https://doi.org/10.1029/2008WR007672

Schumann G, Di Baldassarre G, Bates PD (2009b) The utility of spaceborne radar to render flood inundation maps based on multialgorithm ensembles. IEEE Trans Geosci Remote Sens 47:2801–2807. https://doi.org/10.1109/TGRS.2009.2017937

Schumann G, Hostache R, Puech C, et al (2007a) High-resolution 3-D flood information from radar imagery for flood hazard management. IEEE Trans Geosci Remote Sens 45:1715–1725. https://doi.org/10.1109/TGRS.2006.888103

Schumann G, Matgen P, Cutler MEJEJ, et al (2008b) Comparison of remotely sensed water stages from LiDAR, topographic contours and SRTM. ISPRS J Photogramm Remote Sens 63:283–296. https://doi.org/10.1016/j.isprsjprs.2007.09.004

Schumann G, Matgen P, Hoffmann L, et al (2007b) Deriving distributed roughness values from satellite radar data for flood inundation modelling. J Hydrol 344:96–111. https://doi.org/10.1016/j.jhydrol.2007.06.024

Schumann G, Matgen P, Pappenberger F (2008c) Conditioning Water Stages From Satellite Imagery on Uncertain Data Points. IEEE Geosci Remote Sens Lett 5:810–813

Schumann G, Pappenberger F, Matgen P (2008d) Estimating uncertainty associated with water stages from a single SAR image. Adv Water Resour 31:1038–1047. https://doi.org/10.1016/j.advwatres.2008.04.008

Schumann GJ-P, Bates PD (2018) The Need for a High-Accuracy, Open-Access Global DEM. Front Earth Sci 6:225. https://doi.org/10.3389/FEART.2018.00225

Schumann GJ-P, Bates PD, Neal JC, Andreadis KM (2014a) Technology: Fight floods on a global scale. Nature 507:169–169. https://doi.org/10.1038/507169e

Schumann GJ-P, Frye S, Wells G, et al (2016) Unlocking the full potential of Earth observation during the 2015 Texas flood disaster. Water Resour Res 52:3288–3293. https://doi.org/10.1002/2015WR017126.Received

Schumann GJ-PP, Di Baldassarre G (2010a) The direct use of radar satellites for event-specific flood risk mapping. Remote Sens Lett 1:75–84. https://doi.org/10.1080/01431160903486685

Schumann GJ (2008) Water stages from remotely sensed imagery for improved flood inundation modelling. University of Dundee

Schumann GJ -P. (2019) The need for scientific rigour and accountability in flood mapping to better support disaster response. Hydrol Process hyp.13547. https://doi.org/10.1002/hyp.13547

Schumann GJ, Moller DK (2015) Microwave remote sensing of flood inundation. Phys Chem Earth 83–84:84–95. https://doi.org/10.1016/j.pce.2015.05.002

Schumann GJ, Neal JC, Voisin N, et al (2013) A first large-scale flood inundation forecasting model. Water Resour Res 49:6248–6257. https://doi.org/10.1002/wrcr.20521

Schumann GJP, Andreadis KM (2016) A method to assess localized impact of better floodplain topography on flood risk prediction. Adv Meteorol 2016:. https://doi.org/10.1155/2016/6408319

Schumann GJP, Andreadis KM, Bates PD (2014b) Downscaling coarse grid hydrodynamic model simulations over large domains. J Hydrol 508:289–298. https://doi.org/10.1016/j.jhydrol.2013.08.051

Schumann GJP, Bates PD, Di Baldassarre G, Mason DC (2012) The Use of Radar Imagery in Riverine Flood Inundation Studies. In: Fluvial Remote Sensing for Science and Management. pp 115–140

Schumann GJP, Bates PD, Neal JC, Andreadis KM (2014c) Measuring and Mapping Flood Processes. Elsevier Inc.

Schumann GJP, Di Baldassarre G (2010b) The direct use of radar satellites for event-specific flood risk mapping. Remote Sens Lett 1:75–84. https://doi.org/10.1080/01431160903486685

Schumann GJP, Domeneghetti A (2016) Exploiting the proliferation of current and future satellite observations of rivers. Hydrol Process 30:2891–2896. https://doi.org/10.1002/hyp.10825

Schumann GJP, Neal JC, Mason DC, Bates PD (2011) The accuracy of sequential aerial photography and SAR data for observing urban flood dynamics, a case study of the UK summer 2007 floods. Remote Sens Environ 115:2536–2546. https://doi.org/10.1016/j.rse.2011.04.039

See L, Fritz S, Dias E, et al (2016) Supporting earth-observation calibration and validation: A new generation of tools for crowdsourcing and citizen science. IEEE Geosci Remote Sens Mag 4:38–50. https://doi.org/10.1109/MGRS.2015.2498840

Sene K, Weerts AH, Beven K, et al (2014) Uncertainty Estimation in Fluvial Flood Forecasting Applications. In: Applied Uncertainty Analysis for Flood Risk Management. pp 462–498

Senthilnath J, Shenoy HV, Rajendra R, et al (2013) Integration of speckle de-noising and image segmentation using Synthetic Aperture Radar image for flood extent extraction. J Earth Syst Sci 122:559–572. https://doi.org/10.1007/s12040-013-0305-z

Shaad K, Ninsalam Y, Padawangi R, Burlando P (2016) Towards high resolution and cost-effective terrain mapping for urban hydrodynamic modelling in densely settled river-corridors. Sustain Cities Soc 20:168–179. https://doi.org/10.1016/j.scs.2015.09.005

Shannon CE, Weaver W (1964) The Mathematical Theory of Communication

Shastry A, Durand M (2019) Utilizing Flood Inundation Observations to Obtain Floodplain Topography in Data-Scarce Regions. Front Earth Sci 6:1–10. https://doi.org/10.3389/feart.2018.00243

Shen X, Anagnostou EN, Allen GH, et al (2019) Near-real-time non-obstructed flood inundation mapping using synthetic aperture radar. Remote Sens Environ 221:302–315. https://doi.org/10.1016/j.rse.2018.11.008

Sinclair Knight Merz F, Roads and Traffic Authority of NSW TPS (2011) Wells Crossing to Iluka Road: upgrading the Pacific Highway: Tyndale to Maclean alternative alignment: decision report

Slinski KM, Hogue TS, McCray JE (2019) Active-Passive Surface Water Classification: A New Method for High-Resolution Monitoring of Surface Water Dynamics. Geophys Res Lett 4694–4704. https://doi.org/10.1029/2019GL082562

Slivinski L, Spiller E, Apte A (2015) A Hybrid Particle-Ensemble Kalman Filter for Lagrangian Data Assimilation. Mon Weather Rev 143:195–211. https://doi.org/http://dx.doi.org/10.1175/MWR-D-14-00051.1

Smith LC (1997) Satellite remote sensing of river inundation area, stage, and discharge: A review. Hydrol Process 11:1427–1439. https://doi.org/10.1002/(sici)1099-1085(199708)11:10<1427::aid-hyp473>3.0.co;2-s

Smith PJ, Dance SL, Nichols NK (2011) A hybrid data assimilation scheme for model parameter estimation: Application to morphodynamic modelling. Comput Fluids 46:436–441. https://doi.org/10.1016/j.compfluid.2011.01.010

Smith PJ, Thornhill GD, Dance SL, et al (2013) Data assimilation for state and parameter estimation: Application to morphodynamic modelling. Q J R Meteorol Soc 139:314–327. https://doi.org/10.1002/qj.1944

Stedinger JR, Vogel RM, Lee SU, Batchelder R (2008) Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. 44:1–17. https://doi.org/10.1029/2008WR006822

Stephens E, Bates P (2015) Assessing the reliability of probabilistic flood inundation model predictions. Hydrol Process n/a-n/a. https://doi.org/10.1002/hyp.10451

Stephens E, Schumann G, Bates P (2014) Problems with binary pattern measures for flood model evaluation. Hydrol Process 28:4928–4937. https://doi.org/10.1002/hyp.9979

Stephens EMM, Bates PDD, Freer JEE, Mason DCC (2012) The impact of uncertainty in satellite data on the assessment of flood inundation models. J Hydrol 414–415:162–173. https://doi.org/10.1016/j.jhydrol.2011.10.040

Suri S, Reinartz P (2010) Mutual-information-based registration of TerraSAR-X and ikonos imagery in Urban areas. IEEE Trans Geosci Remote Sens 48:939–949. https://doi.org/10.1109/TGRS.2009.2034842

Takagi T, Sugeno M (1985) Fuzzy identification of systems and its applications to modeling and control. Syst Man Cybern IEEE Trans SMC-15:116–132. https://doi.org/10.1109/TSMC.1985.6313399

Tarekegn TH, Haile AT, Rientjes T, et al (2010) Assessment of an ASTER-generated DEM for 2D hydrodynamic flood modeling. Int J Appl Earth Obs Geoinf 12:457–465. https://doi.org/10.1016/j.jag.2010.05.007

Tarpanelli A, Brocca L, Melone F, Moramarco T (2013) Hydraulic modelling calibration in small rivers by using coarse resolution synthetic aperture radar imagery. Hydrol Process 27:1321–1330. https://doi.org/10.1002/hyp.9550

Thiele A, Cadario E, Schulz K, Thoennessen U (2007) Model Based Building Recognition from Multi-Aspect InSAR Data in Urban Areas. IEEE Trans Geosci Remote Sens 45:3583–3593

Thiemann M, Trosset M, Gupta H, Sorooshian S (2001) Bayesian recursive parameter estimation for hydrologic models. Water Resour Res 37:2521–2535. https://doi.org/10.1029/2000WR900405

Thirel G, Salamon P, Burek P, Kalas M (2013) Assimilation of MODIS snow cover area data in a distributed hydrological model using the particle filter. Remote Sens 5:5825–5850. https://doi.org/10.3390/rs5115825

Tong X, Luo X, Liu S, et al (2018) An approach for flood monitoring by the combined use of Landsat 8 optical imagery and COSMO-SkyMed radar imagery. ISPRS J Photogramm Remote Sens 136:144–153. https://doi.org/10.1016/j.isprsjprs.2017.11.006

Tourian MJ, Schwatke C, Sneeuw N (2017) River discharge estimation at daily resolution from satellite altimetry over an entire river basin. J Hydrol 546:230–247. https://doi.org/10.1016/j.jhydrol.2017.01.009

Townsend P (2001) Mapping seasonal ooding in forested wetlands using multi-temporal Radarsat SAR. Photogramm Eng Remote Sensing 67:

Trudel M, Leconte R, Paniconi C (2014) Analysis of the hydrological response of a distributed physically-based model using post-assimilation (EnKF) diagnostics of streamflow and in situ soil moisture observations. J Hydrol 514:192–201. https://doi.org/10.1016/j.jhydrol.2014.03.072

Tuozzolo S, Lind G, Overstreet B, et al (2019) Estimating river discharge with swath altimetry: A proof of concept using AirSWOT observations. Geophys Res Lett. https://doi.org/10.1029/2018GL080771

Twele A, Cao W, Plank S, Martinis S (2016) Sentinel-1-based flood mapping: a fully automated processing chain. Int J Remote Sens 37:2990–3004. https://doi.org/10.1080/01431161.2016.1192304

Uhe P, Mitchell D, Bates P, et al (2019) Enhanced flood risk with 1.5°C global warming in the Ganges-Brahmaputra-Meghna basin. Environ Res Lett. https://doi.org/10.1088/1748-9326/ab10ee

Ulaby F, Kouyate F, Brisco B, Williams TH (1986) Textural Information in SAR Images. IEEE Trans Geosci Remote Sens GE-24:235–245. https://doi.org/10.1109/TGRS.1986.289643

van Leeuwen P (2017) Particle Filters for nonlinear data assimilation in high-dimensional systems. Ann la Fac des Sci Toulouse Mathématiques 26:1051–1085. https://doi.org/10.5802/afst.1560

Van Wesemael A, Landuyt L, Lievens H, Verhoest NEC (2019) Improving flood inundation forecasts through the assimilation of in situ floodplain water level measurements based on alternative observation network configurations. Adv Water Resour 130:229–243. https://doi.org/10.1016/j.advwatres.2019.05.025

Villarini G, Krajewski WF, Ntelekos A a., et al (2010) Towards probabilistic forecasting of flash floods: The combined effects of uncertainty in radar-rainfall and flash flood guidance. J Hydrol 394:275–284. https://doi.org/10.1016/j.jhydrol.2010.02.014

Voormansik K, Praks J, Antropov O, et al (2014) Flood Mapping With TerraSAR-X in Forested Regions in Estonia. IEEE J Sel Top Appl Earth Obs Remote Sens 7:562–577. https://doi.org/10.1109/JSTARS.2013.2283340

Vrugt J a. (2003) A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. Water Resour Res 39:. https://doi.org/10.1029/2002WR001642

Vrugt J a., Diks CGH, Gupta H V., et al (2005) Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation. Water Resour Res 41:1–17. https://doi.org/10.1029/2004WR003059

Wagener T, McIntyre N, Lees MJ, et al (2003) Towards reduced uncertainty in conceptual rainfall-runoff modelling: dynamic identifiability analysis. Hydrol Process 17:455–476. https://doi.org/10.1002/hyp.1135

Walker JP, Houser PR (2005) Hydrologic Data Assimilation. Adv water Sci Methodol 233. https://doi.org/10.5772/1112

Waller JA, Dance SL, Nichols NK (2016) Theoretical insight into diagnosing observation error correlations using observation-minus-background and observation-minus-analysis statistics. Q J R Meteorol Soc 142:418–431. https://doi.org/10.1002/qj.2661

Waller JA, García-Pintado J, Mason DC, et al (2018a) Technical note: Analysis of observation uncertainty for flood assimilation and forecasting. Hydrol Earth Syst Sci Discuss 1–13. https://doi.org/10.5194/hess-2018-43

Waller JA, García-Pintado J, Mason DC, et al (2018b) Technical note: Assessment of observation quality for data assimilation in flood models. Hydrol Earth Syst Sci 22:3983–3992. https://doi.org/10.5194/hess-22-3983-2018

Wang D, Cai X (2008) Robust data assimilation in hydrological modeling - A comparison of Kalman and H-infinity filters. Adv Water Resour 31:455–472. https://doi.org/10.1016/j.advwatres.2007.10.001

Wang X, Holland DM, Gudmundsson GH (2018) Accurate coastal DEM generation by merging ASTER GDEM and ICESat/GLAS data over Mertz Glacier, Antarctica. Remote Sens Environ 206:218–230. https://doi.org/10.1016/j.rse.2017.12.041

Wang Y, Ruan R, She Y, Yan M (2011) Extraction of water information based on RADARSAT SAR and Landsat ETM+. Procedia Environ Sci 10:2301–2306. https://doi.org/10.1016/j.proenv.2011.09.359

Ward PJ, Jongman B, Salamon P, et al (2015) Usefulness and limitations of global flood risk models. Nat Clim Chang 5:712–715. https://doi.org/10.1038/nclimate2742

Wdowinski S, Kim SW, Amelung F, et al (2008) Space-based detection of wetlands' surface water level changes from L-band SAR interferometry. Remote Sens Environ 112:681–696. https://doi.org/10.1016/j.rse.2007.06.008

Wealands SR (2006) Quantitative methods for hydrological spatial field comparison. The University of Melbourne

Wealands SR, Grayson RB, Walker JP (2005) Quantitative comparison of spatial fields for hydrological model assessment - Some promising approaches. Adv Water Resour 28:15–32. https://doi.org/10.1016/j.advwatres.2004.10.001

Wellmann J (2013) Information Theory for Correlation Analysis and Estimation of Uncertainty Reduction in Maps and Models. Entropy 15:1464–1485. https://doi.org/10.3390/e15041464

Werner MGF (2010) A comparison of flood extent modelling approaches through constraining uncertainties on gauge data. Hydrol Earth Syst Sci 8:1141–1152. https://doi.org/10.5194/hess-8-1141-2004

Werner MGF, Hunter NM, Bates PD (2005) Identifiability of distributed floodplain roughness values in flood extent estimation. J Hydrol 314:139–157. https://doi.org/10.1016/j.jhydrol.2005.03.012

Westerhoff RS, Kleuskens MPHH, Winsemius HC, et al (2013) Automated global water mapping based on wide-swath orbital synthetic-aperture radar. Hydrol Earth Syst Sci 17:651–663. https://doi.org/10.5194/hess-17-651-2013

Wetterhall F, Pappenberger F, Alfieri L, et al (2013) HESS Opinions "forecaster priorities for improving probabilistic flood forecasts." Hydrol Earth Syst Sci 17:4389–4399. https://doi.org/10.5194/hess-17-4389-2013

Weydahl DJ (1996) Flood monitoring in Norway using ERS-1 SAR images. IGARSS '96 1996 Int Geosci Remote Sens Symp 1:151–153. https://doi.org/10.1109/IGARSS.1996.516274

Winsemius HC, Van Beek LPH, Jongman B, et al (2013) A framework for global river flood risk assessments. Hydrol Earth Syst Sci 17:1871–1892. https://doi.org/10.5194/hess-17-1871-2013

Wöhling T, Vrugt J a. (2008) Combining multiobjective optimization and Bayesian model averaging to calibrate forecast ensembles of soil hydraulic models. Water Resour Res 44:1–18. https://doi.org/10.1029/2008WR007154

Wood M (2016) Improving hydraulic model parameterization using SAR data. University of Bristol

Wood M, Hostache R, Neal J, et al (2016) Calibration of channel depth and friction parameters in the LISFLOOD-FP hydraulic model using medium-resolution SAR data and identifiability techniques. Hydrol Earth Syst Sci 20:4983–4997. https://doi.org/10.5194/hess-20-4983-2016

Wood M, Jong SM De, Straatsma MW (2018) Locating flood embankments using SAR time series: A proof of concept. Int J Appl Earth Obs Geoinf 70:72–83. https://doi.org/10.1016/j.jag.2018.04.003

Woodhead SPB (2007) Bayesian calibration of flood inundation simulators using an observation of flood extent. University of Bristol

Wu Q, Tinka A, Weekly K, et al (2015) Variational Lagrangian Data Assimilation in open channel networks. Water Resour Res 51:1916–1938. https://doi.org/10.1002/2014WR015270.Received

Xie X, Zhang D (2010) Data assimilation for distributed hydrological catchment modeling via ensemble Kalman filter. Adv Water Resour 33:678–690. https://doi.org/10.1016/j.advwatres.2010.03.012

Yamazaki D, Baugh C a., Bates PD, et al (2012) Adjustment of a spaceborne DEM for use in floodplain hydrodynamic modeling. J Hydrol 436–437:81–91. https://doi.org/10.1016/j.jhydrol.2012.02.045

Yamazaki D, Ikeshima D, Sosa J, et al (2019) MERIT Hydro: A high-resolution global hydrography map based on latest topography datasets. Water Resour Res 55:5053–5073. https://doi.org/10.1029/2019WR024873

Yamazaki D, Ikeshima D, Tawatari R, et al (2017) A high-accuracy map of global terrain elevations. Geophys Res Lett 44:5844–5853. https://doi.org/10.1002/2017GL072874

Yamazaki D, Revel M, Kanae S (2018) Model Based Observation Localization Weighting Function for Amazon Mainstream

Yan H, Moradkhani H (2016) Combined assimilation of streamflow and satellite soil moisture with the particle filter and geostatistical modeling. Adv Water Resour 94:364–378. https://doi.org/10.1016/j.advwatres.2016.06.002

Yan H, Zarekarizi M, Moradkhani H (2018) Toward improving drought monitoring using the remotely sensed soil moisture assimilation: A parallel particle filtering framework. Remote Sens Environ 216:456–471. https://doi.org/10.1016/j.rse.2018.07.017

Yan K, Di Baldassarre G, Solomatine DP, Schumann GJ-P (2015) A review of low-cost space-borne data for flood modelling: topography, flood extent and water level. Hydrol Process 3387:n/a-n/a. https://doi.org/10.1002/hyp.10449

Ye W, Hansen DP, Jakeman AJ, et al (1997) Assessing the natural variability of runoff: Clarence Basin catchments, NSW, Australia. Math Comput Simul 43:251–260. https://doi.org/10.1016/S0378-4754(97)00007-4

Yoon Y, Durand M, Merry CJ, et al (2012) Estimating river bathymetry from data assimilation of synthetic SWOT measurements. J Hydrol 464–465:363–375. https://doi.org/10.1016/j.jhydrol.2012.07.028

Yu D, Yin J, Liu M (2016) Validating city-scale surface water flood modelling using crowd-sourced data. Environ Res Lett 11:1748–9326

Yue L, Shen H, Zhang L, et al (2017) High-quality seamless DEM generation blending SRTM-1, ASTER GDEM v2 and ICESat/GLAS observations. ISPRS J Photogramm Remote Sens 123:20–34. https://doi.org/10.1016/j.isprsjprs.2016.11.002

Zadeh L a. (1965) Fuzzy sets. Inf Control 8:338–353. https://doi.org/10.1016/S0019-9958(65)90241-X

Zaji AH, Bonakdari H, Gharabaghi B (2018) Remote Sensing Satellite Data Preparation for Simulating and Forecasting River Discharge. IEEE Trans Geosci Remote Sens 56:3432–3441. https://doi.org/10.1109/TGRS.2018.2799901

Zalite K, Voormansik K, Olesk A, et al (2014) Effects of inundated vegetation on X-band HH-VV backscatter and phase difference. IEEE J Sel Top Appl Earth Obs Remote Sens 7:1402–1406. https://doi.org/10.1109/JSTARS.2013.2279552

Zaussinger F, Dorigo W, Gruber A, et al (2018) Estimating irrigation water use over the contiguous United States by combining satellite and reanalysis soil moisture data. Hydrol Earth Syst Sci 1–42

Zebker HA, Villasenor J (1992) Decorrelation in interferometric radar echoes. IEEE Trans Geosci Remote Sens 30:950–959. https://doi.org/10.1109/36.175330

Zheng F, Tao R, Maier HR, et al (2018) Crowdsourcing Methods for Data Collection in Geophysics: State of the Art, Issues, and Future Directions. Rev Geophys 56:698–740. https://doi.org/10.1029/2018RG000616

Zheng X, Xiong H, Yue L, Gong J (2016) An improved ANUDEM method combining topographic correction and DEM interpolation. Geocarto Int 31:492–505. https://doi.org/10.1080/10106049.2015.1059899

Ziliani MG, Ghostine R, Ait-El-Fquih B, et al (2019) Enhanced flood forecasting through ensemble data assimilation and joint state-parameter estimation. J Hydrol 577:123924. https://doi.org/10.1016/j.jhydrol.2019.123924

Appendix A Impact of Different Observation Operators

The impact of using different observation operators on the model ranking is quite significant. Several cost functions for extent comparison were tested before selecting the mutual information based agreement metric. This appendix summarizes the impacts of using different observation operators on the assimilation efficiency through the experimental setup used in Chapter 8. Synthetic observations were assimilated in the three hydraulically uniform sub-reaches using different observation operators, and the Brier Skill Scores (BSS) were computed for the assimilated forecast for a lead time of 12h until the next assimilation time step. Recall that BSS values represent the change in Mean Squared Errors after the assimilation, implying that negative values indicate an increase in errors while positive values signify an error reduction. First, deterministic cost functions were tested to assess whether the extent-based ranking of models (particles) results in forecast improvements. Then probabilistic cost functions for the flood extent assimilation were developed and tested.

The ideal cost function was expected to result in forecast improvements for a 12h window for all cases. In a synthetic experiment, since observations are simulated based on the "truth" model, for a 12h lead time the assimilation should have no effect in the worst-case scenario. This is because the particle weights will not change as a consequence of the assimilation if the observation is not informative. Negative impacts of the assimilation then imply that the cost function is not weighting the particles correctly and cannot discriminate between models effectively. Cost functions that result in strongly negative impacts are actually ranking the models based on errors, i.e. models with higher errors get higher weights and thus the overall forecast quality degrades. The following sections show the impacts of using both deterministic and probabilistic observation operators and then discuss the choice of the cost function in this thesis.

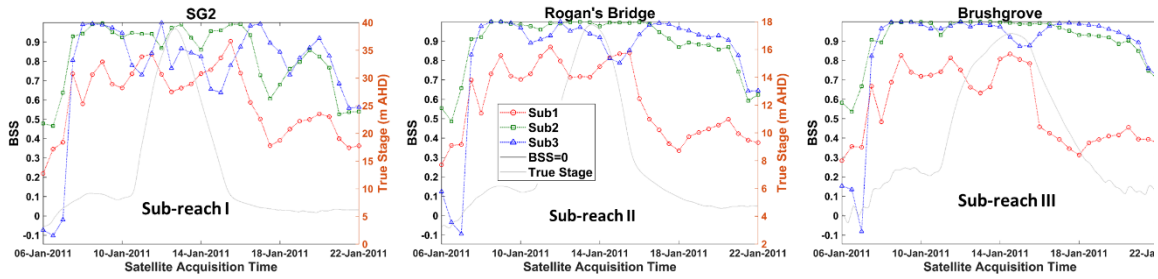

Fig. A.1 Brier Skill Scores (BSS) obtained for single image assimilation in each sub-reach, for a lead time of 12h from the time of the satellite overpass. Observations were independently considered each 12h starting from the 6th of Jan with BSS calculated at three water level gauges along the channel (one in each sub-domain); the true stage at the location is shown in all subplots as a reference. Positive values of BSS imply forecast improvements, while negative values imply degradation and 0 implies no change from the open loop. Each point on each curve is representative of the satellite acquisition time and the corresponding BSS. The cost function used here is the Critical Success Index or CSI.

A.1 Deterministic Observation Operators

The operators tested in this section were primarily tested to ensure that updating inundation forecasts using flood extents has positive impacts, with the understanding that they cannot be used for assimilation due to their inability to account for observational uncertainty. As a first sanity check, the Critical Success Index was used for comparing the models and the synthetic observations at the assimilation time steps. Due to its ubiquity in flood modelling and mapping literature, this evaluation metric was chosen to serve as a benchmark in spite of its limitations highlighted in literature (Stephens et al. 2014). The objective was to identify a cost function which performed at least as well as the CSI driven "assimilation".

The CSI and different powers of the CSI (4 and 8), were tried to assess the sensitivity of the rankings to the metric used. As evident from the figures Fig. A.1, Fig. A.2, and Fig.

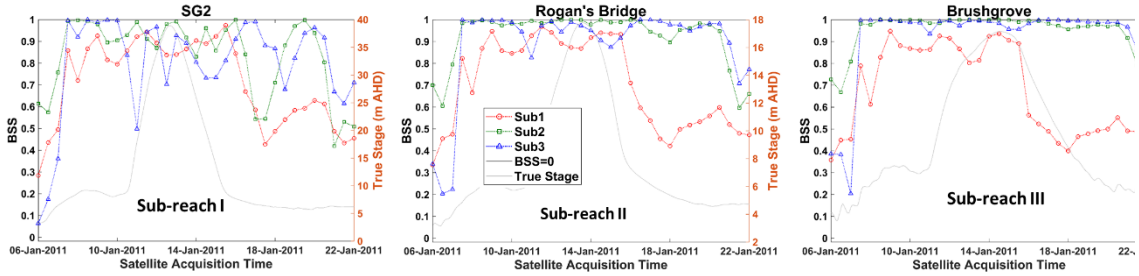

Fig. A.2 As for Fig. A.1 but using CSI^4 as a cost function for the assimilation.

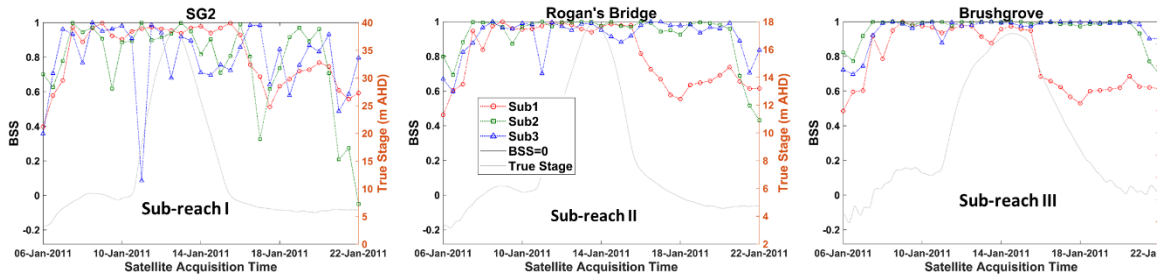

Fig. A.3 As for Fig. A.1 but using CSI^8 as a cost function for the assimilation.

A.3 flood extent assimilation results in forecast improvements. Increasing the power of the function, does increase sensitivity and the resulting improvements at the location of the assimilation, but degrades at gauges further upstream. This confirmed that an extent based ranking does improve the forecast but increasing the sensitivity of the cost function artificially using a power function, has several limitations and results in inconsistent performance. Next probabilistic functions were tested.

A.2 Probabilistic Observation Operators

In this section the metrics based on reliability diagrams and probability binning are illustrated. Reliability diagrams allow the unique advantage of comparing the binary model outputs with a probabilistic observation or vice versa, by plotting the proportion of wet cells in each probability bin. The deviation from the 1:1 line is then the error in the model or the classification algorithm. In this regard, root mean squared errors (RMSE) have been recommended (Horritt 2006) to quantify this deviation, and more recently the RMSE weighted by bin sizes has gained popularity (Giustarini et al. 2016; Schlaffer et al. 2017). The weighted RMSE was found to be insensitive to subtle changes in extent as due to the

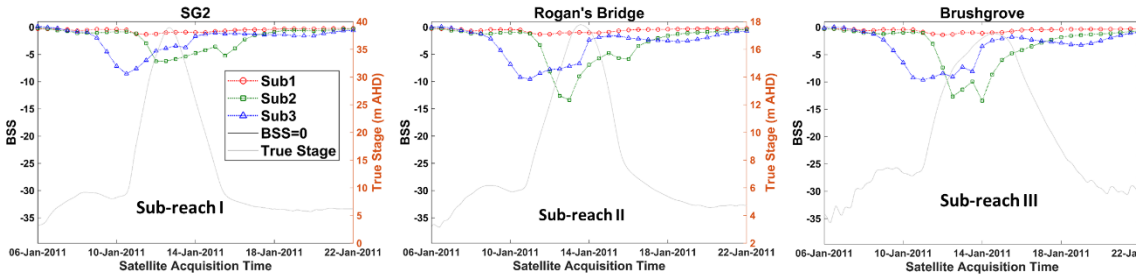

Fig. A.4 As for Fig. A.1 but using the fuzzy CSI as a cost function. The fuzzy CSI is calculated as the ratio of True Positives and the sum of True Positives, False Positives, and False Negatives in each probability class in the reliability diagram.

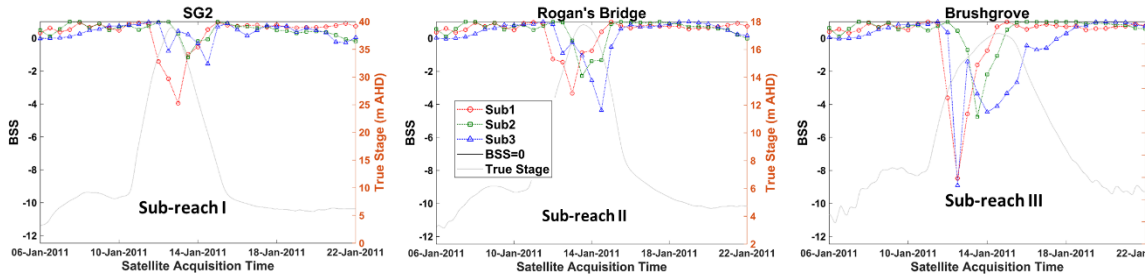

Fig. A.5 As for Fig. A.1 but using the product of RMSE and Mean Bias calculated from the Reliability Diagrams. Note that the product is a measure of errors and therefore, the cost function used here ranks the models with higher errors higher. This is by design, to assess whether the metrics work well for the quantification of errors.

weighting using bin sizes, it results in asymmetric penalties for the usually larger non-flooded area. Some new metrics were calculated and tested for the reliability diagrams. Mean bias, mean percent bias, and the CSI, which are all commonly used statistics but have not been explored for reliability diagrams were tested here. The main reason for testing these was to increase the sensitivity of the cost function.

As evident from the plots of the fuzzy CSI shown in Fig. A.4 and the plots of error based rankings based on the RMSE and Bias shown in Fig. A.5, it seems that the lumped metrics calculated based on the reliability diagrams are inadequate for extent comparison. In fact, the forecast seems to improve for most images when ranked based on errors, with the exception of a few images at and around the peak, which means that the errors are not being correctly characterized. The fuzzy CSI based ranking fails because the assumption of proportion of simulated wet cells in each observed probability class being equal to the mean bin value does not consistently hold. In other words, the CSI exclusively works for

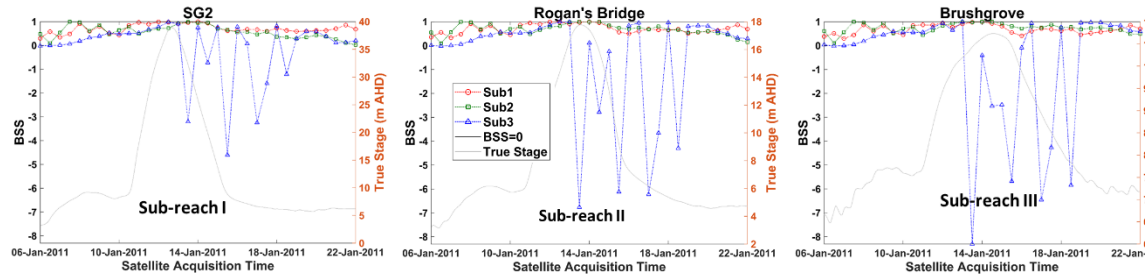

Fig. A.6 As for Fig. A.1 but using the absolute value of the log of the product of the RMSE and Mean Bias to rank the models. Here the errors are compressed by using the log function, but not inverted into accuracies as the value of the error product exceeds unity. Absolute values are used to convert the negative values from the log function into positive values of probability.

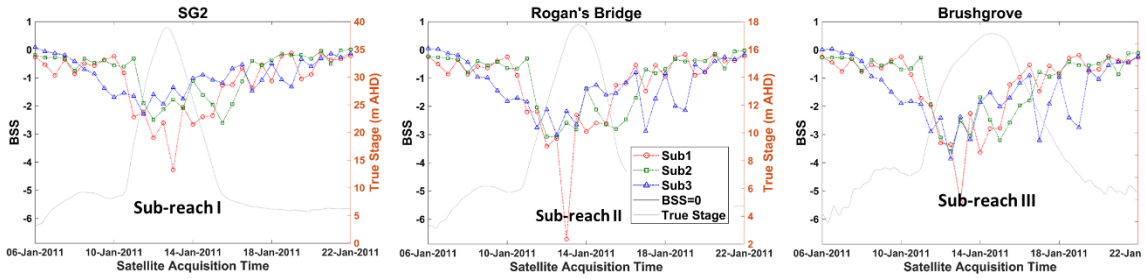

Fig. A.7 As for Fig. A.6 but with the values of the error product restricted to less than unity, resulting in an inversion into accuracies using the absolute value of the logarithm.

binary maps and the threshold of 0.5 representing the maximum uncertainty or entropy of the observation is the only one that results in a useful assessment. Different thresholds were also tried (0.4-0.6) but the CSI fails to accurately quantify model performance in all other cases.

Fig. A.6 used the absolute values of the log of the error product as the cost function but did not invert the error scale into accuracies as the values were larger than unity. The error ranges were compressed and hence the rankings are different from Fig. A.5, thus a different pattern is evident in the graphs. Here too the error based ranking results in improvements, implying that the lumped metrics from the reliability diagrams do not actually correct quantify the errors of classification. Fig. A.7 shows the plots for when the error metric is effectively translated into accuracies, and as evident from the plots, the forecast performance is consistently degraded. This implies that the metrics are not able to characterize errors at all as the ranking based on the "errors" results mostly in improvements and when the ranking is inverted based on accuracies, the forecast degrades every time. There are several reasons which might be responsible for this. The first, is that the assumption that the ideal model should predict a proportion identical to the predicted probability class is wrong. The second possibility is that the lumping of the metrics, or the

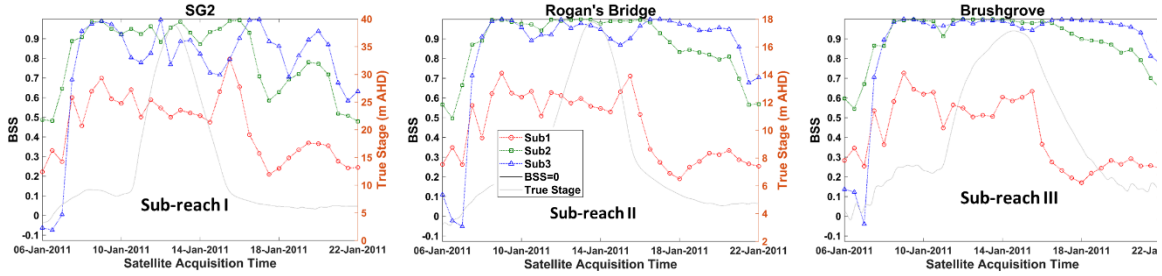

Fig. A.8 As for Fig. A.1 but using Mutual Information (MI) as a cost function for the assimilation.

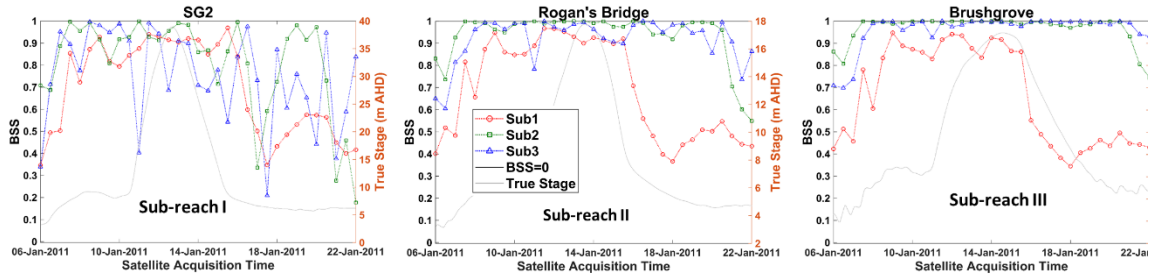

Fig. A.9 As for Fig. A.1 but using MI^4 as a cost function for the assimilation.

mean error values like mean bias and RMSE are not representative due to the aggregation. Here it appears as though lumped metrics based on the reliability diagrams are actually quite dangerous to use when assessing models or flood classification algorithms as they work only in select cases. This does not mean that reliability diagrams are not useful, given that they are the only evaluation method which allow the assessment of over and under estimation for probabilistic versus binary mapping and modelling. However, the lumped metrics do not provide useful information, as was also observed in Chapter 4 where the weighted RMSE reliability could not capture the improvement in the fuzzy maps, which was visually apparent through the figures.

The next test was for the mutual information based metric, which is an almost similar approach to the reliability diagrams when comparing binary and probabilistic outputs. The key difference being that the number of bins for the analysis can be increased and are evaluated for non-linear relationships as well. Take the reliability diagram for example, it uses a scatter plot of the simulated proportion and observed probabilities. This works well when the relationship between the two variables is linear. Mutual information on the other hand, plots a 2D histogram by dividing the scatterplot into squares, and counting the number of points inside each square. Here the number of "squares" or bins for the analysis, can be increased indefinitely but for this case increasing the bin size beyond 256 ceased to have notable impacts. As evident from Fig. A.8, the mutual information based cost function results in consistent improvements. In order to test the sensitivity enhancements through the use of a power function, the MI based metric raised to the power of 4 was tested, illustrated here in Fig. A.9. However, on increasing the power, the global consistency of forecast improvements in the domain declined, and images which performed well downstream resulted in model degradation in the upstream. Based on all of these tests, a mutual information based cost function was chosen for this thesis.