UNIVERSITY OF READING

Department of Meteorology

Improving flood prediction using data assimilation

Elizabeth S. Cooper

A thesis submitted for the degree of Doctor of Philosophy

March 2019

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Elizabeth Cooper

Publications

Chapters 4 and 5 of this thesis are reproduced from the following publications respectively:

- Cooper, E. S., Dance, S. L., Garcia-Pintado, J., Nichols, N. K. and Smith, P. J. Observation impact, domain length and parameter estimation in data assimilation for flood forecasting. Environmental Modelling and Software, 104. pp. 199-214. ISSN 1364-8152, https://doi.org/10.1016/j.envsoft.2018.03.013 (2018)
- Cooper, E. S., Dance, S. L., Garca-Pintado, J., Nichols, N. K., and Smith, P.: Observation operators for assimilation of satellite observations in fluvial inundation forecasting, Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-589, in review, 2018.

All work undertaken in these publications was carried out by Elizabeth Cooper with coauthors providing guidance and review.

Abstract

River flooding is a costly problem worldwide. Timely, accurate prediction of the behaviour of flood water is vital in helping people make preparations. Mathematical hydrodynamic models can predict the behaviour of flood water given information about inputs such as river bathymetry, local topography, inflows, and values for model parameters. Uncertainty in these inputs leads to inaccuracies in model predictions; data assimilation can be used to improve forecasts by combining model predictions with observational information, taking into account uncertainties in both.

In this thesis we investigate ways to maximize the impact of observational information from satellite-based synthetic aperture (SAR) instruments in data assimilation for inundation forecasting. We show in synthetic twin experiments using an ensemble transform Kalman filter that using joint state-parameter data assimilation techniques to correct the model channel friction parameter as well as water levels provides a significant, long lasting benefit to the model forecast. We show that errors in the channel friction parameter and inflow are interdependent.

We propose a novel observation operator that allows direct use of measured SAR backscatter values, potentially allowing the use of many more observations per SAR image. We test our new observation operator in synthetic experiments, showing that we can successfully update inundation forecasts and the value of the model channel friction parameter using our new approach. We show that different observation operator approaches can generate significantly different updates to model forecasts and illustrate the physical mechanisms responsible. Lastly, we use our new observation operator to assimilate backscatter values from real SAR images, showing that our new approach can be used to improve inundation forecasts in a real case study. Improved understanding of the physical mechanisms by which updates are generated by different observation operators provides insights into improving the observation impact of SAR data.

Dedication

for Richard, James, Matthew, David and Daniel

Acknowledgements

I would like to thank my team of excellent supervisors, Prof. Sarah Dance, Dr. Polly Smith, Prof. Nancy Nichols and Dr. Javier Garcia-Pintado for huge amounts of support and encouragement. I would also like to thank David Mason for some really helpful input at various stages of the project. I gratefully acknowledge financial support from the National Environmental Research Council (NERC) through the SCENARIO DTP, and from my CASE sponsors the Satellite Applications Catapult.

The Meteorology Department at Reading has been a great place to study and I'm grateful to everyone I've had useful discussions and tea breaks with over the last 4.5 years. Special thanks to members of DARC, to current office-mates Aga, Michael and Kaja, and to Hannah Bloomfield who has been a marvel.

Lastly I want to thank my wonderful family, especially my Mum and Dad, for their endless support. Massive thanks to Mum and Fran for looking after the boys while I worked; without your help I couldn't possibly have done this. And thanks to Richard for so many things, including encouragement, laughter, sympathy and wine. Finally, thank you to my sons James, Matthew, David and Daniel, who make all things worthwhile.

Contents

Chapter 1:

Introduction

1.1 Motivation

River flooding is a costly problem in the UK and worldwide (e.g. Ward et al. [2018]) and the number of people exposed to flood risk is expected to increase with future climate change (see e.g. Yamazaki et al. [2018]). Real-time, accurate inundation forecasts from numerical models can help to mitigate the damage caused by river inundation events by alerting people to when and where they are likely to be affected. However, predictions from such numerical inundation models are inevitably subject to uncertainty due to approximations in the governing equations and uncertainties in model parameters and driving data. The work in this thesis is part of an effort to improve inundation forecasts using data assimilation, a powerful mathematical tool which can be used to combine predictions from a numerical inundation model with observations in order to give better predictions of flood extent and depth. In particular, the work in this thesis explores questions about how to make the best use of observational information. We wish to maximise the observation impact of SAR-derived information, i.e. we seek to improve the accuracy of inundation forecasts but also to increase the time over which the observations influence the forecast.

A number of authors have used data assimilation to improve inundation forecasting in a number of different ways, as we discuss in chapter 2. Some studies (e.g. Ricci et al. [2011], Lai and Monnier [2009] and Hostache et al. [2010]) have focussed on using observational information from river gauges, either alone or in combination with information from satellite images; in this thesis we consider situations in which synthetic aperture radar (SAR) observations are our only source of information. SAR has the advantage over optical satellite monitoring that it works day and night and is not affected by cloud cover or rain, which could otherwise cause problems in flood situations. The reason for taking a SAR-only approach is that the number of gauges worldwide is in decline (Vrsmarty et al. [2001]) and we wish to have a method that will be applicable in ungauged catchments. Although SAR observations typically show a clear delineation between wet and dry areas, and can cover a large spatial area, the information they contain is limited in time. Using SAR-based information to correct only water levels can therefore lead to short-lived improvement in the forecast. Some authors (e.g. Andreadis et al. [2007], Matgen et al. [2010], Giustarini et al. [2011], Garcia-Pintado et al. [2015], Garcia-Pintado et al. [2013] and Mason et al. [2015]) have demonstrated increased observation impact by correcting errors in inflow (amount of water flowing into the domain of interest) as well as water levels, but less attention has been paid the role of model parameters. It is well known that channel friction processes are important in the evolution of flood events (e.g. Hostache et al. [2010], James et al. [2016]) and in this thesis we therefore investigate whether using data assimilation to correct the value of the parameter controlling channel friction can increase observation impact.

There are a number of ways in which information from a SAR image could be used in data assimilation. Authors such as Neal et al. [2009], Giustarini et al. [2011], Giustarini et al. [2012] and Garcia-Pintado et al. [2015] have used SAR-derived water levels as observations. This involves intersecting SAR images with digital elevation models in order to extract information about water level. Such observations are typically sparse in real case studies (e.g.Mason et al. [2012]). We have designed and implemented a new approach to extracting information from SAR images, in which we directly use measured backscatter values as observations. This approach allows for the use of many more observations than the derived water level approach and requires less processing of the SAR image, potentially speeding up the time taken between acquisition and an updated forecast. A related approach is used in Hostache et al. [2018], in which the authors use SAR images to produce flood probability maps and use these as observations in a particle filtering method.

1.2 Thesis aims

The aims of this thesis are to answer the following research questions:

1. **How does estimation of the channel friction parameter affect observation impact in data assimilation for inundation forecasting?**

Can we use joint state-parameter data assimilation techniques to retrieve the correct channel friction parameter in synthetic experiments using SAR-like observations, and does this significantly improve the forecast? How is error in the channel friction parameter related to error in inflows?

2. **Can we design and implement a new method of extracting observational information from a SAR image for use in data assimilation?**
   How do existing approaches generate an update to the forecast via the observation operator? Can we design and implement a new observation operator in order to use raw backscatter observations in data assimilation? Does our new observation operator work to improve the forecast in synthetic experiments, and can we use it to recover the true value of the channel friction coefficient?

3. **Can we apply our new observation approach to a case study using real topography and SAR data?**
   Does our new backscatter observation operator improve the forecast in a real case study, and do we observe the same physical mechanisms as in the synthetic experiments? How can we measure improvement to the forecast when we have no synthetic 'truth' to compare to?

1.3 Principal new results

The work in this thesis provides the following answers to the research questions posed:

1. In our synthetic experiments using a simplified domain we found that it was possible to retrieve the correct value for the channel friction parameter using synthetic SAR-derived water level observations in a joint state-parameter data assimilation approach. We found that correcting the channel friction parameter has a large, positive impact on subsequent water depth forecasts, significantly improving observation impact by extending the time over which observations influence the forecast. We found that inflow error and channel friction parameter error are interdependent, making the two sources of error difficult to separate out.

2. We have designed and implemented a new observation operator that directly uses SAR backscatter values as observations. We have shown that the method can be used to successfully update a model forecast and concurrently correct the value of the channel friction parameter in synthetic experiments. We have shown that

different approaches to observations from SAR images (water level observations vs backscatter observations) and different observation operators can produce substantially different updates to modelled inundation predictions. We explain these differences by illustrating the physical mechanisms responsible for the updates in the different approaches.

3. We successfully applied our novel backscatter observation operator to a real case study, using a series of SAR images capturing a flood event near Tewkesbury in the UK in the winter of 2012. We used a series of binary flood match measures to demonstrate the benefit to the forecast of updating water levels in this way.

1.4 Thesis outline

This thesis is structured as follows:

- Chapter 2 outlines the principles of data assimilation for inundation prediction, and describes the work of other authors in approaching this problem. We briefly describe SAR measurements and approaches to inundation modelling.
- Chapter 3 introduces and describes the Clawpack software that was used to build our inundation model and run all of the inundation simulations in this thesis.
- Chapter 4 addresses the first research question posed in section 1.2. We describe a series of synthetic experiments in which we show that we can use joint state-parameter estimation techniques to successfully retrieve a good value for the channel friction parameter, and that this has a significant impact on the forecast. We also show that assuming zero momentum immediately following an assimilation causes a shock in the solution and demonstrate a simple method for avoiding this problem. This chapter has been published as Cooper et al. [2018b].
- Chapter 5 addresses the second research question. We outline two existing methods that can be used to assimilate SAR-derived flood edge water elevation measures. We then introduce a novel method which directly assimilates backscatter values from a SAR image. We describe the three different observation operators and show that the choice of observation operator can have large impact on the updated inundation forecast. We compare the results of using the three different operators in synthetic experiments and discuss the physical mechanisms associated with generating an update in each case. This chapter has been published as Cooper et al. [2018a].

- Chapter 6 addresses the final research question and shows the results of applying our new observation operator (described in chapter 5) to a real case study. We show that assimilating SAR-derived backscatter values generates an update to model predicted water levels, and that this update improves the forecast at the assimilation time.
- Chapter 7 summarizes the main conclusions of the thesis and outlines possible areas for future research.

Chapter 2:

Data assimilation for inundation forecasting

In this chapter we introduce some of the key topics used in this thesis. In section 2.1 we briefly outline the theory of data assimilation. In section 2.2 we discuss approaches to fluvial inundation modelling; in section 2.3 we introduce the Synthetic Aperture Radar (SAR) images used to provide observations for data assimilation in this thesis. We summarise a number of current approaches to data assimilation for inundation forecasting in section 2.4. In section 2.5 we introduce the forecast verification measures that are used in chapter 6 to assess the agreement between inundation model predictions and observed data.

2.1 Introduction to data assimilation

Data assimilation is a powerful mathematical technique in which predictions from a numerical model of a dynamical physical system are combined with observational information; this allows us to predict the current and future states of the physical system more accurately than using either a numerical model or observations in isolation. In this thesis we consider river flooding, but data assimilation methods are very well established in operational numerical weather prediction (e.g. Rawlins et al. [2007]) and are employed in a wide variety of applications including modelling of land surface processes, (e.g. Pinnington et al. [2018]), ocean modelling (e.g. Thomas and Haines [2017]) and neuroscience (e.g. Moye and Diekman [2018]).

In data assimilation we describe the evolution of a dynamical physical process, such as a flood event, using a forecast model,

$$\mathbf{x}(t_{k+1}) = \mathfrak{M}(\mathbf{x}(t_k)), \tag{2.1}$$

where $\mathbf{x} \in \mathbb{R}^N$ is the model state vector, t represents time and $\mathfrak{M} : \mathbb{R}^N \rightarrow \mathbb{R}^N$ is the forecast model; in this case an inundation model.

We assume that at particular times we have observations of the system, $\mathbf{y}_{obs} \in \mathbb{R}^p$, related to the true state, $\mathbf{x}^t$, by

$$\mathbf{y}_{obs} = \mathbf{h}(\mathbf{x}^t) + \epsilon. \tag{2.2}$$

Observations may be indirect measurements of state variables and not located at model cell centres, so an observation operator, $\mathbf{h} : \mathbb{R}^N \rightarrow \mathbb{R}^p$ is required to map the state vector into observation space. Observation noise, ϵ, is assumed to be Gaussian, stochastic and unbiased, with covariance $\mathbf{R} \in \mathbb{R}^{p \times p}$.

There are a number of approaches to the data assimilation problem, in which we seek the state (called the 'analysis') that best matches both the model prediction and the observations, subject to the uncertainties in each. One approach is to minimize a cost function of the form

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^f)\mathbf{P}^{-1}(\mathbf{x} - \mathbf{x}^f) + \frac{1}{2}(\mathbf{y}_{obs} - \mathbf{h}(\mathbf{x}^f))\mathbf{R}^{-1}(\mathbf{y}_{obs} - \mathbf{h}(\mathbf{x}^f)), \tag{2.3}$$

where $\mathbf{x}^f$ is the model prediction at the observation time (often called the 'background' state and referred to in this thesis as the 'forecast' state). The matrix $\mathbf{P} \in \mathbb{R}^{N \times N}$ is the forecast error covariance matrix, and $\mathbf{R} \in \mathbb{R}^{p \times p}$ is the observation error covariance matrix. Minimisation of equation 2.3 corresponds to the maximum a posteriori (MAP) solution. The MAP is equivalent to the minimum variance solution when h and $\mathfrak{M}$ are linear, under the assumption that the errors in the forecast state and observations are Gaussian (Nichols [2010]). The first term on the right hand side of equation 2.3 describes the difference between the model state and the forecast state, the second term describes the difference between the model state and the observations, and these terms are weighted according to the relative uncertainties in the forecast and the observations through $\mathbf{P}$ and $\mathbf{R}$.

2.1.1 Variational approaches

In variational data assimilation, the cost function is minimised using numerical iterative methods. Variational techniques generally use the same forecast error covariance matrix at all assimilation times, i.e. uncertainty in the forecast state is assumed to be fixed and not updated based on observational information. The sequential variational approach is known as 3D-Var (e.g. Lorenc et al. [2000]). Sequential data assimilation schemes comprise two steps: a forecast step in which the system is evolved according to equation 2.1 and an update step, in which we update the model forecast according to the observations. The updated (analysis) state is then used as the starting point for the next forecast step and the process can be repeated many times. An alternative to the sequential approach is the 4D-Var approach (e.g. Rawlins et al. [2007]), in which an assimilation time window is defined. In 4D-Var techniques, the aim is to find the initial state which best fits the forecast state at the initial time (the background) and all of the observations available in the defined window, subject to the model equation 2.1. The advantage of the 4D-Var approach is that observations can be distributed across an assimilation window.

2.1.2 Kalman filter approaches

The Kalman filter (KF) (Kalman [1960]) provides a solution that minimises equation 2.3 in systems with linear forecast models and observation operators. An advantage of the Kalman filter approach over variational methods is that the forecast state error covariance matrix is updated at each assimilation time along with the model state. The Kalman filter is only applicable to systems with linear h and $\mathfrak{M}$; the extended Kalman filter (EKF) is an approximation to the KF which can be used in the case of non-linearity. The EKF requires linearisation of the observation operator and forward model; this can lead to poor results in highly non-linear cases (Gelb [1974]).

2.1.3 Ensemble Kalman filter approaches

In this thesis we use an ensemble transform Kalman filter (ETKF) to find the analysis state. The ETKF is a sequential technique, based on the Kalman filter, which uses an ensemble of state vectors to represent a number of possible realisations of the dynamic system. There are a number of different ensemble Kalman filter (EnKF) methods but in all of them the ensemble mean is used to represent the best estimate of the true state of the system, and the spread of the ensemble members around the mean gives an estimate of the uncertainty in the mean prediction ([Evensen, 1994]). One formulation of the EnKF uses perturbed observations to ensure that the analysis ensemble retains the correct statistics (Evensen [2003]); we have used an ETKF which is sometimes known as a 'deterministic' EnKF method as it does not require observation perturbations. Variants of the deterministic approach are reviewed in Tippett et al. [2003]. The ETKF is one of a number of square root filters, which are reviewed in Livings et al. [2008]. We use an ETKF here following the approach of Garcia-Pintado et al. [2013] and Garcia-Pintado et al. [2015] for a similar application. The ETKF has the advantage of being numerically efficient in implementation.

The ETKF is valid for non-linear forecast models and observation operators, and retains the advantage of flow-dependent forecast state covariances. However, since the ensemble is a statistical sample of state space, there are a number of potential problems with practical implementation of the technique. Undersampling can lead to underestimation in ensemble spread (ensemble collapse), causing the forecast to give insufficient weight to observations and diverge from the truth (filter divergence)(Petrie and Dance [2010]). Inflation techniques can be used to counter this problem (see e.g. Anderson [2007]). Undersampling can also lead to noise in the state forecast covariance matrix, producing spurious updates (Petrie and Dance [2010]). Techniques to deal with this problems include localization (e.g. Hamill et al. [2001]). The Kalman filter and ETKF update equations are given in section 4.3.2.

2.1.4 Particle filter approaches

Particle filter techniques represent another sequential approach to the data assimilation problem (e.g. van Leeuwen [2009], Doucet et al. [2013]). In particle filter methods the requirement for errors to be Gaussian is relaxed, and non-linear models and observation operators are permitted. Particle filter approaches use an ensemble of states, or particles, to represent forecast and analysis uncertainty in a similar way to EnKF systems. Standard particle filter techniques are known to suffer from problems such as particle degeneracy and impoverishment, particularly when the size of the state is large (e.g. Snyder et al. [2008]).

2.2 Inundation models

In this section we discuss some approaches to modelling river flood inundation processes. In chapter 3 we discuss in more detail the particular hydrodynamic model used in the remainder of this thesis, which uses Clawpack software (Clawpack Development Team [2014]).

There are a number of well developed numerical inundation models that can predict the behaviour of flood water given information about the topography of the domain and the amount of water flowing into the area, e.g. HEC-RAS [HEC-RAS Development Team], Telemac [Hervouet, 2000], LISFLOOD-FP [Neal et al., 2012a], Dassflow [Honnorat et al., 2007]. Most models use a variation of the shallow water equations to predict changes in water depth in river channels. The shallow water equations are an approximation to the Navier-Stokes equations, which can be used to model the behaviour of fluids. The shallow water equations are valid for situations in which

- the vertical scales are much smaller than the horizontal scales, allowing vertical velocity to be neglected;
- viscous forces can be neglected;
- the fluid being modelled is incompressible.

These approximations are generally met for inundation modelling applications. The shallow water equations are given and discussed in the context of Clawpack in section 3.1.

Hydrodynamic inundation models provide approximate solutions to the shallow water equations using various numerical schemes. Models can use implicit methods (using information from the current time as well as the past time to calculate a solution) or explicit methods (which use information from past timesteps only to calculate the solution). HEC-RAS and Telemac-2d use implicit methods, whereas Clawpack and LISFLOOD-FP use explicit solvers. Time steps in the model can be user-specified or can be adaptive to preserve numerical stability. Models can solve 1D or 2D forms of the shallow water equations, and use a variety of discretisation approaches. HEC-RAS uses finite difference methods to solve the full 1D shallow water equations in river channels and makes assumptions about the conveyance of water onto the floodplain. Telemac-2d uses finite volume methods to to solve the 2D depth-averaged shallow water equations. Dassflow has both 1D and 2D capabilities and uses finite volume methods on an unstructured mesh. LISFLOOD-FP uses an explicit finite difference technique to solve the 1D shallow water equations in river channels and a raster-based approach on the floodplain.

The 1D approach is considerably less expensive than using the full 2D shallow water equations, but has the disadvantage that momentum is not exchanged between water in the river and on the floodplain. Three different model approaches are compared in Horritt and Bates [2002], in which HEC-RAS (1D), LISFLOOD-FP (1D and raster approach) and Telemac-2d (2D) are all separately calibrated for two flood events of the river Severn, UK, two years apart. The predictions of models calibrated on the first flood are then compared with satellite-derived inundation extents from the second flood and vice versa. The study shows that the three models all had similar predictive power when optimally calibrated, but that the performance of LISFLOOD-FP was sensitive to the type of calibration data used, requiring calibration against inundation data rather than discharge data to give good results. In Neal et al. [2012b] the authors use a series of benchmarking tests to compare the performance of three different approaches to solving the 2D shallow water equations, and show that for gradually varying flows the simplest representation of the physical processes performed similarly well to more complex representations, and to industry standards for commercial inundation codes. However, simpler models are unable to deal with shocks (hydraulic jumps). In the hydrodynamic model we use for the simulations in this thesis we use a finite volume, explicit method to solve the 2D shallow water equations throughout the domain. Our method is able to simulate shocks in the solution.

Numerical models require appropriate boundary conditions in order to generate realistic solutions, and there are a number of different approaches to specifying these. In many models it is possible to set an inflow rate at the upstream boundary, which can be based on measured gauging station data. A downstream boundary condition can be set by specifying the water depth at the boundary, or an outflow rate; either of these conditions may be available from gauging stations. In our work we use an extrapolating outflow boundary, to allow water to freely leave the domain (see section 3.1.2 for more details of the boundary condition options in Clawpack); we treated inflow as a source term in the equations (see section 4.8.2).

In inundation modelling, the boundary between wet and dry areas on the floodplain moves as the flood event develops and this moving boundary can cause problems for some solvers. Approaches that use the 1D equations do not suffer from this problem as the shallow water equations are only valid in the channel with e.g. a 2D or raster-based volume filling approach on the floodplain. Clawpack deals with the problem of wet/dry interfaces by assuming that there is water everywhere in the domain, but that the water depth is zero in dry areas.

All inundation models require good topographical information to give accurate results. Digital Terrain Models (DTMs) are often available for flood plains, although are not always of a sufficiently high quality for hydrological modelling (see Schumann and Bates [2018]). DTMs do not typically contain information about river and stream bed elevations, and accurate, recent representation of river channels can be difficult to obtain. Further, combining information about flood plain and river topography is non-trivial. When high resolution topography information is available, the resolution at which the model is run will depend on the available compute resources; higher resolution models provide more accurate results but at a higher cost and a longer run time. To address this problem, LISFLOOD-FP uses a sophisticated sub-grid parameterization technique for river channels, allowing accurate representations of flow in river channels with dimensions smaller than the grid-scale (see Neal et al. [2012a]); this reduces the cost of the model by removing the need for a grid resolution that can resolve river channels.

2.3 Synthetic aperture radar (SAR) observations

Synthetic Aperture Radar instruments on satellites (such as CosmoSkymed and Sentinel 1) can provide a large amount of observational information about a river flood event (e.g. Brown et al. [2016], Schumann et al. [2009], Mason et al. [2012]). SAR instruments are side-looking active instruments that emit radiation (of wavelength cm to m) towards the surface of the Earth. SAR instruments use cloud penetrating radiation, giving the instruments all-weather capability, and can be operational day and night, unlike passive sensors that rely on solar radiation. The strength of the returned signal measured at the SAR sensor depends strongly on the roughness properties of the surface from which it has been reflected.

Figure 2.1 shows an example of a SAR image in a flood situation. Areas of open water are very smooth and therefore reflect a large proportion of the incoming radiation away from the sensor. Pixels in flooded or other wet areas therefore have low backscatter values and appear as dark areas on SAR images. Dry areas have rougher surface properties and reflect a larger proportion of incoming radiation back towards the sensor; dry areas have higher backscatter values and therefore appear paler than wet areas. During a flood event SAR images therefore generally show a clear difference between flooded and non-flooded areas.

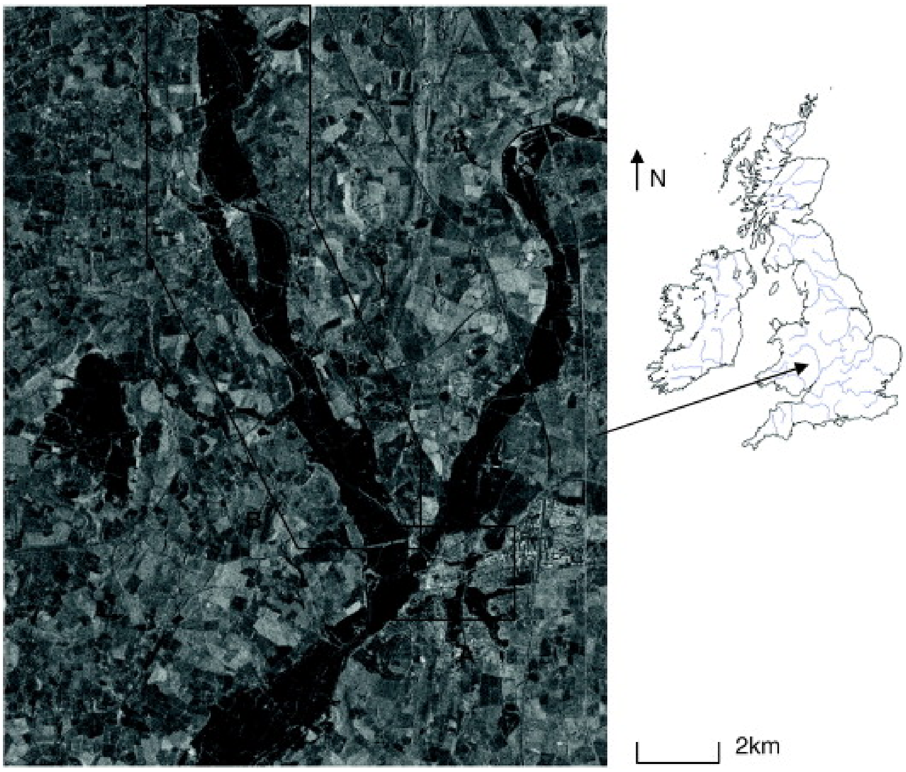


Figure 2.1: TerraSAR-X 2.5m resolution SAR image of the Severn catchment near Tewkesbury, UK, July 2007. Reproduced from Mason et al. [2012]. Dark areas are water.

SAR images can be processed to various different levels of calibration. In chapter 6 we use digital number (DN) values for our observations, whereas in Giustarini et al. [2011], backscatter values measured in dB are used for data assimilation. When purchasing SAR data, DN values are available at lower cost than backscatter values. This is because backscatter values are processed to a higher level than DN values; a radiometric calibration is required to produce backscatter values. Radiometric calibration is essential when comparing SAR images taken at different times, but it is possible to use DN values to identify wet and dry pixels contained in one particular image.

There are a number of approaches to identifying pixels as wet or dry based on SAR backscatter or DN measurements. Techniques include thresholding (e.g. Henry et al. [2006]) with varying levels of user-interpretation (as compared in Brown et al. [2016]) and region growing/clustering (e.g. Horritt et al. [2001]); it is also possible to use change detection methods if a number of co-located images are available (e.g. Hostache et al. [2012]). Identifying wet and dry areas using these methods can be used for flood mapping and monitoring (as in e.g. Brown et al. [2016], Matgen et al. [2011]) and for validation and calibration of inundation models (e.g. Mason et al. [2009], Baldassarre et al. [2009] Wood et al. [2016]).

Intersection of flood edges with a DTM provides information about the elevation of water, by using the fact that water elevation must be the same as the terrain elevation at the location of the flood edge. For model calibration, Mason et al. [2009] and Stephens et al. [2013] show that better calibration can be achieved by comparing modelled water levels with SAR derived water levels, rather than using binary wet-dry comparisons. However, it is not clear whether derived water levels provide better observation impact than wet/dry observations in data assimilation; this question is addressed in chapter 5.

A difficulty in determining flood edge water levels arises when a flood edge is found to be on a steeply sloping area of topography, since small errors in pinpointing the position of the flood edge cause large errors in the elevation assigned to the water. Flood edges in positions with slopes above a threshold are therefore often excluded from calculations e.g. Mason et al. [2012]. Complications in accurately determining the flood edge in a SAR image can include:

- Vegetation: If the flooded area contains plants that are taller than the flood depth, this can affect the SAR signal return strength, causing wet areas to appear dry.
- In windy conditions the surface of a body of water can be rough due to ripples and therefore have a higher backscatter values than usual. This potentially risks wet areas appearing to be dry.
- In urban areas, buildings can prevent substantial amounts of ground being observed by SAR instruments due to layover and shadowing (see Mason et al. [2014], Mason et al. [2018], Tanguy et al. [2017]).

2.4 Approaches to data assimilation for inundation forecasting

There have been a number of attempts to use data assimilation to improve hydrodynamic forecasting, many of which have used data from gauges to provide observational information. For example, in Madsen [2003] water level data from gauges was used to successfully update forecasts of water levels on the Sesia River basin in Italy. In Srensen et al. [2004] and Srensen and Madsen [2004] the authors test various ensemble Kalman filter techniques, assimilating tidal gauge data into hydrodynamic models. Synthetic tidal gauge data is assimilated in a twin experiment in Srensen and Madsen [2004], and in Srensen et al. [2004] real tidal gauge data is assimilated in a hydrodynamic model of the North Sea and Baltic Sea system. In Mure-Ravaud et al. [2016] river level and flow information is used to update real-time operational river forecasts for the city of Dijon in France.

A number of authors have highlighted the potential of using various different satellite-based observations in conjunction with flood forecasting models, and a comprehensive review of current research and future opportunities is presented in Grimaldi et al. [2016]. In this thesis we assimilate only information from SAR images, since a method which can work on an ungauged catchment is highly desirable. In section 2.4.1 we outline the work of other authors who have used only SAR-derived observations. In section 2.4.2 we briefly describe some studies in which gauge observations are assimilated alongside SAR-derived observations.

2.4.1 Assimilating SAR-derived observations

A number of studies have focussed on using SAR-derived observations to improve inundation forecasting. These include Matgen et al. [2007a] and Matgen et al. [2007b], in which water level observations (WLOs) are extracted from SAR images capturing a flood event in 2003 on the river Alzette in Luxembourg. These observations are used to successfully update water level forecasts from a pre-calibrated HEC-RAS model towards measured high water marks using a direct-insertion technique. In Andreadis et al. [2007] the authors use LISFLOOD-FP and a square root EnKF technique to perform an identical twin experiment. The method uses synthetic SAR-derived WLOs to update water level and discharge forecasts, and uses an auto-regressive error forecast model to simultaneously update inflow. The study found that water levels and discharge could be successfully updated towards the true values, and that the system showed more sensitivity to observation frequency than to observation error. The authors point out that the contribution from errors in channel characteristics such as the model friction parameter values are not investigated in their study, but are likely to play an important role in data assimilation for depth and discharge estimation.

A perturbed observation ensemble Kalman filter technique is used in Neal et al. [2009] to update water depth and discharge (flow rate) forecasts from HEC-RAS using water level observations derived from a SAR image showing flooding of the river Alzette. Updated water level and discharge estimates are shown to be closer to independent measurements from wrack marks and gauge data than open loop (no assimilation) predictions, demonstrating the potential of data assimilation techniques in a real case study of river inundation forecasting. The study also shows that forecasts produced by the hydrodynamic model are sensitive to the way the river channel geometry is represented.

In Matgen et al. [2010] a particle filter approach is used with HEC-RAS and synthetic water level observations; the authors show that only updating water levels in this way gave a very short lived improvement to forecast water levels. Simultaneously updating the inflow, assuming a constant error, is shown to give a longer benefit to the forecast. A similar approach is used in Giustarini et al. [2011] to assimilate SAR-derived water levels for a case study on the river Alzette. The authors show that a constant error approach to inflow estimation has limitations, leading to better results at short timescales, but overcorrection of inflow at mid to longer time scales. In Giustarini et al. [2012] the particle filter method with inflow error correction is applied to two different case studies; on the Po river in Italy and the Sure river in Luxembourg. The authors demonstrate a correction of modelled water levels towards observed water levels in both cases.

In Garcia-Pintado et al. [2013] an ETKF approach is used with LISFLOOD-FP to assimilate synthetic water level observations, showing that water levels can be successfully updated towards the truth. The study also shows that simultaneously updating the inflow assuming a constant error gives better results than only updating the state (water levels), and that SAR-like observations from early in a flood event have a larger influence on the subsequent forecast improvement. In Garcia-Pintado et al. [2015] the authors use similar techniques in a real case study; SAR-derived water level observations are assimilated into a LISFLOOD-FP model to produce better estimates of water depth and discharge than using the model alone.

Various authors have demonstrated that although SAR images cover a large spatial area, the information they contain is limited in time, e.g. Lai and Monnier [2009], Matgen et al. [2007b] and Schumann et al. [2009]. This is because the behaviour of the water in a flood situation is only weakly dependent on past water levels, and depends much more strongly on the inflow and model parameter values. This leads to a situation in which data assimilation can effectively correct water levels at the time of an observation, but with incomplete information about inflow and parameter values, the forecast quickly moves away from the truth.

In order to address the short-lived improvement in forecast, several authors - e.g. Andreadis et al. [2007], Matgen et al. [2010], Giustarini et al. [2011], Garcia-Pintado et al. [2015], Garcia-Pintado et al. [2013] and Mason et al. [2015], have devised data assimilation schemes that use observational information to correct inflow (or inflow error) as well as water levels. These have all been shown to produce longer-lived observation impact than state-only estimation. Much less attention has been paid to the role of model parameters controlling processes such as channel friction, despite the fact that the evolution of a flood is highly sensitive to this parameter in particular; Andreadis and Schumann [2014] and the comprehensive review paper Grimaldi et al. [2016], indicate that model parameters are likely to have an important influence on the behaviour of the flow and therefore on observation impact.

One study which did examine model parameter values is Garcia-Pintado et al. [2015], in which the channel friction was estimated along with water levels, inflow error and a number of other parameters. The approach was found to give a good forecast, but correction of the channel friction parameter along with the inflow did not result in significant extra improvement to the forecast. Further, the parameter values obtained, though physically reasonable, could not be validated. This motivates the identical twin experiments we describe in chapter 4, which have been carried out with a known, unbiased inflow, allowing us to isolate and examine the effect of estimating the channel friction parameter. This work has allowed us to conclude that channel friction parameter estimation is possible, and produced a large improvement in observation impact in our experiments. We showed that simultaneous estimation of inflow error may obscure this improvement.

Another aspect to observation impact is the way in which observational information is extracted from a SAR image in order to use it in a data assimilation context. All of the studies mentioned so far assimilate derived water-level observations. In Wood [2016] and Hostache et al. [2018] flood probability maps are produced from SAR images (see Giustarini et al. [2016] for details of probability maps) and these are used as observations in experiments with LISFLOOD-FP and a particle filter. An advantage of this method is that is removes the need for integration of the SAR image with a DTM in order to derive water levels. This is also true of the new observation operator approach that we describe and apply in chapters 4 and 5.

2.4.2 Assimilating SAR-derived and gauge observations

Other approaches to data assimilation for flood prediction use both satellite-derived and in situ gauge data as observations. These include Ricci et al. [2011] in which a 1D river model is used with an extended Kalman filter and two sets of observations in a two-step assimilation process. The authors first test the EKF in an idealised setting and then apply it to a real case study. In the first step of the data assimilation process, the upstream flow is corrected using downstream flow measurements. In the second step, the authors directly assimilate water levels and discharge (flow) measurements sequentially, once an hour. The effect of correcting the inflow with data assimilation is found to be much greater than the effect of correcting water levels and discharge, particularly in the longer term. This two step process is applied to some real data in a number of catchments in France and found to give good results in some cases. However, the authors point out that the success of the method depends on good calibration of parameters (e.g friction) in the shallow water equations before the assimilation. In common with studies using SAR observations only, the authors find that the sensitivity of the forecast to the initial water levels is short lived (hours); here it is shown that the sensitivity of the forecast to initial water levels is negligible compared to the sensitivity to the upstream flow at longer times.

A 4D variational approach is used to assimilate water levels with a 2D numerical flood model in the two linked papers Lai and Monnier [2009] and Hostache et al. [2010]. The authors use DassFlow software to numerically solve the 2D shallow water equations. In part I (Lai and Monnier [2009]) the mathematical method is described and used in a simple test domain to correct inflow estimation and water levels; in part II (Hostache et al. [2010]) the method is applied to a real flood event in order to extract Manning's friction coefficients for the domain. The authors show in Hostache et al. [2010] that calibrating a flood model (i.e. retrieving friction coefficients) with SAR derived data alongside in situ hydrograph data leads to a better result than when using the hydrograph data only. An identical twin experiment is carried out in order to show that the method can accurately identify the Manning's friction coefficients in the channel and on the floodplain. Water depths extracted from a real SAR image (RADARSAT) are then assimilated along with hydrograph data; the addition of the SAR-derived water depth information allows the friction coefficients to be determined. A sensitivity analysis indicates that the channel friction is much more important to the development of the flood than the friction coefficients on the floodplain in this case; this motivates our work to better understand the effect of the channel friction parameter in chapter 4.

The approaches of Ricci et al. [2011], Lai and Monnier [2009] and Hostache et al. [2010] all require the assimilation of some information about the flow in the river, i.e. from gauges. In this thesis we consider assimilation only of SAR-derived information. This means that our method can work in ungauged catchments, or in cases where gauge data are unreliable. Since gauge data are not trustworthy in the very high flow situations associated with flooding, a method that does not require such information is extremely valuable.

2.5 Measuring improvement in inundation forecasting

In the synthetic experiments described in chapters 4 and 5 measuring improvement to the forecast due to assimilation is straightforward; we have a ‘truth’ flood to which we can compare the model forecast and analysis predictions. In chapter 6 we apply our methods to a real case study in which we do not have a true solution for comparison. In this case we measure the improvement to the forecast using a binary wet/dry comparison with observations; despite some problems, such measures may be used to assess the performance of flood models (see e.g. Stephens et al. [2013]).

In order to use binary measures we first create a contingency table. The form of the contingency table is as shown in table 2.1.

		Modelled	
		Wet	Dry
Observed	Wet	A (Correctly predicted wet)	C (Wrongly predicted dry)
	Dry	B (Wrongly predicted wet)	D (Correctly predicted dry)

Table 2.1: Contingency table, after Mason [2003], Stephens et al. [2013] and Schumann et al. [2009].

Each model cell in which there is also an observation is assigned to a category as defined in table 2.1. The number of cells in each category is then counted to give one value of A, B, C and D per model prediction. For each case, A can be thought of as the number of ‘hits’, B can be thought of as the number of ‘false alarms’, C can be thought of as ‘misses’ and D as ‘correct rejections’. In flooding applications, much of the domain is likely to be in the D category as there will likely be large areas of the model domain where the river is never likely to flood.

In the application of this thesis, the model cells and the observation cells are always assumed to be co-located. In the case where this is not true, some interpolation of either the observation or the model prediction would be necessary, and the manner of this interpolation would likely affect the results. We do not consider this question here.

There are various ways to combine the values in the contingency table into measures that give an idea of how well the model predictions match the observations. Each of the measures tells us something different about the match between a model prediction and observations. We outline eight such measures here; for further details see Stephens et al. [2013], Schumann et al. [2009] or Hunter [2005]. The measures are referred to slightly differently by different authors; here we follow the naming conventions of Stephens et al. [2013].

- **Bias**, defined as
$$\frac{A+B}{A+C}. \tag{2.4}$$
The bias score expresses the balance between model overprediction and model underprediction. A forecast that is unbiased in terms of under and overprediction would score 1; a model that overpredicts flooding scores > 1, and a model that underpredicts flooding scores < 1.

- **Proportion Correct (PC) or $F^{<1>}$**, defined as
$$\frac{A+D}{A+B+C+D}. \tag{2.5}$$
This score measures the proportion of the total cells that are correctly identified by the model, according to the observations. This score varies between 0 for a forecast that wrongly predicts in every cell, and 1 for a perfect forecast. This score is dominated by D, as in any domain there is likely to be a large area in which there is never likely to be any flooding. Comparison of this score between domains is very difficult due to the different numbers of cells in category D in different domains. This weakness is not important in our application, as we only use these scores for one domain.

- **Hit rate (HR)**, defined as
$$\frac{A}{A+C}. \tag{2.6}$$
The hit rate is the fraction of the observed flood that is correctly predicted by the model to be wet. Like the PC score, the value of HR varies between 0 and 1, with scores closer to 1 for a good forecast.

- **False alarm rate (F)**, defined as
$$\frac{B}{B+D}. \tag{2.7}$$
The false alarm rate is the proportion of observed dry area that is predicted by the model to be wet. This score value varies between 0 for a good forecast and 1 for a

poorer forecast.

- **Pierce Skill Score (PSS)**, defined as

$$\frac{AD - BC}{(A + C)(B + D)}. \tag{2.8}$$

  This is equal to $HR - F$, that is the hit rate minus the false alarm rate. This score can vary between -1 and $+1$. A score of more than zero indicates that the forecast has more hits than misses and is therefore likely to be useful. A perfect forecast, (where $HR = 1$ and $F = 0$) will have a PSS score equal to 1.

- **Critical Success Index (CSI) or $\mathbf{F}^{<2>}$**, defined as

$$\frac{A}{A + B + C}. \tag{2.9}$$

  The CSI is similar to the PC measure, but excludes category D from the calculations since this can dominate the PC measure. CSI varies between 0 and 1, with a value of 1 indicating a perfect forecast.

- $\mathbf{F}^{<3>}$, defined as

$$\frac{A - C}{A + B + C}. \tag{2.10}$$

  The $F^{<3>}$ measure is similar to the CSI but penalises underprediction for use in cases when overprediction is preferable to underprediction. A perfect forecast would give a score of 1.

- $\mathbf{F}^{<4>}$, defined as

$$\frac{A - B}{A + B + C}. \tag{2.11}$$

  Like $F^{<4>}$, the $F^{<3>}$ measure is similar to the CSI but penalises overprediction.

All of the measures detailed here give us different information about model inundation predictions compared to observations. In Stephens et al. [2013] the authors show that there are problems when using the CSI in particular to compare model performance in different catchments and for different sized floods. In the work presented in this thesis we are only comparing how well forecasts (with and without assimilation) represent the same floods, so this limitation is not relevant here. We note also that binary measures generally give better results for overpredictions compared to underpredictions for floods covering a large area since once the flood water is constrained by valley-like topography an increase in predicted flood volume leads to an increase in predicted depth of floodwater rather than prediction of a greater number of flooded cells.

2.6 Chapter summary

In this chapter we have reviewed and discussed a number of topics relevant to the thesis. In section 2.1 we presented an overview of data assimilation techniques. The ETFK method, which we use in chapters 4, 5 and 6 is discussed in more detail in section 4.3.2. In section 2.2 we outlined various approaches to fluvial inundation modelling; the inundation model used in this thesis is discussed further in chapter 4. In section 2.3 we discussed Synthetic Aperture Radar (SAR) images, which we use as observational information in this thesis. In section 2.4 we outlined a number of studies in which SAR data has been used in data assimilation for inundation forecasting. In section 2.5 we presented the forecast verification measures that are used in chapter 6 to test the agreement between our inundation model predictions and observed data.

Chapter 3:

Hydrodynamic model

In order to apply data assimilation to river flooding, we require an inundation model. There are a number of approaches to numerical inundation modelling, as discussed in section 2.2. In this thesis we used Clawpack [Clawpack Development Team, 2014, Mandli et al., 2016, LeVeque, 2002] to run our inundation simulations. Clawpack (‘Conservation LAW’) is an open source collection of FORTRAN and python code that can be used to solve a wide variety of conservation laws. The software can be downloaded by the user and adapted to fit the problem of interest. Clawpack uses finite volume methods and sophisticated Riemann solvers to treat systems of partial differential equations; in this work the equations of interest are the 2D shallow water equations that describe how river and flood water will move in space and time. The model splits the user-defined domain of interest into a number of cells and calculates the water depth in each cell. The code is capable of dealing with shocks in the solution, such as bores that may occur following a sudden increase of inflow into a particular river stretch. Clawpack deals effectively with the wet-dry interfaces that are present in an inundation event, and preserves depth non-negativity [George, 2008].

We chose Clawpack for the work presented in this thesis rather than a model specifically designed to simulate flooding. The advantages of Clawpack are

- the code is open source and transparent;
- code is very flexible in terms of boundary conditions and topography specification;
- the code uses fast, robust, well developed solvers;
- it is possible to specify the computational mesh resolution separately from the topography resolution.

The remainder of this chapter is adapted from a University of Reading Mathematics Report available online as Cooper et al. [2016].

3.1 Modelling the shallow water equations using Clawpack

Systems of partial differential equations (PDEs) can be used to model a wide variety of physical situations. The Navier-Stokes equations are a set of such equations that describe how pressure, velocity, temperature and density change with time and in space in a moving fluid. The Navier-Stokes equations are derived from laws of conservation of momentum and mass. The equation can be simplified in certain cases to give the shallow water equations. The shallow water equations hold in situations for which the viscosity of the fluid can be neglected, and the horizontal scale of the system is much larger than the vertical scale, i.e. the horizontal domain is large compared to the depth of the water. A further assumption is that the fluid is incompressible, so that its density is constant.

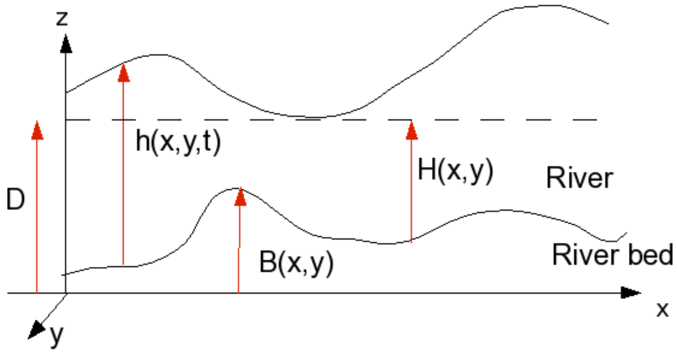


Figure 3.1: Shallow water equations schematic

The 2D shallow water equations can be written as (e.g. LeVeque [2002])

$$\frac{\partial \mathbf{q}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{q})}{\partial x} + \frac{\partial \mathbf{G}(\mathbf{q})}{\partial y} = \mathbf{R}(\mathbf{q}), \tag{3.1}$$

where $\mathbf{q}$ is a vector of conserved quantities

$$\mathbf{q} = \begin{bmatrix} h \\ hu \\ hv \end{bmatrix}, \tag{3.2}$$

h represents the depth of the fluid in the z direction (see figure 3.1), g is acceleration due to gravity, and u and v represent velocity in the x and y directions respectively. The bathymetry of the river bed depicted in figure 3.1 is represented by $B(x, y)$, assumed here to be constant in time. For applications of the shallow water equations to e.g. ocean modelling it is sometimes useful to refer to a reference or steady state depth; this is shown in figure 3.1 with a dashed line.

In equation 3.1, $\mathbf{F(q)}$ and $\mathbf{G(q)}$ represent fluxes of the conserved quantities in the x and y directions respectively. For the shallow water equations these are

$$\mathbf{F(q)} = \begin{bmatrix} hu \\ hu^2 + \frac{1}{2}gh^2 \\ huv \end{bmatrix} \qquad \textit{and} \qquad \mathbf{G(q)} = \begin{bmatrix} hv \\ huv \\ hv^2 + \frac{1}{2}gh^2 \end{bmatrix}. \tag{3.3}$$

The term on the right hand side of equation 3.1, $\mathbf{R(q)}$, is a source term; when $\mathbf{R(q)} = 0$, the system is said to be homogeneous. The homogeneous equations describe a system in which, within a given volume, any change in the conserved quantities with time is equal to the value of the flux of the quantities at the boundary of the volume. When $\mathbf{R(q)} \neq 0$ in a volume it means that there is a source of one or more of the conserved quantities within that volume. (When $\mathbf{R(q)}$ has negative values, this is sometimes referred to as a sink, but source will be used here for both positive and negative $\mathbf{R(q)}$).

Various methods can be used for the numerical solution of PDEs such as the shallow water equations. These include finite element methods (FEM), finite difference methods, finite volume methods (FVM) and boundary element methods. Finite volume methods are particularly suited to situations in which the behaviour of the system is not smooth; by using the integral (also called the ‘weak’) form of conservation laws it is possible for FVM solutions to ‘capture’ or ‘track’ shocks in the solutions of the equations (LeVeque [2002]). Clawpack (Clawpack Development Team [2014]) uses finite volume methods and can solve homogeneous and non-homogeneous equations.

3.1.1 Source terms in Clawpack

Clawpack can solve systems of partial differential equations with source terms. We have developed and implemented a new source term to add water to the domain in order to model river-like flow. We describe the pre-existing approach to treating friction as a source term in Clawpack in section 4.8.1 and our new inflow source term is described in section 4.8.2. This material appears in the paper published as Cooper et al. [2018b], reproduced in this thesis as chapter 4.

3.1.2 Boundary conditions in Clawpack

Correct specification of the solution at the boundaries of the computational domain is vital for the stability of any numerical scheme. To achieve this, Clawpack adds a user-specified number (2 by default) of ‘ghost cells’ next to each cell at a domain boundary. The domain is effectively extended in all directions by the addition of ghost cells. The behaviour of the solution at the boundaries then depends strongly on the values of calculated model quantities in the ghost cells.

Clawpack allows for different treatment of domain boundaries, which work by setting the values of $\mathbf{q}$ in ghost cells beyond the edge of the computational domain. There are four options in the code for specifying the boundary conditions. For the shallow water equations the available boundary conditions are:

- Non-reflecting outflow (extrapolating). This type of boundary acts as a free boundary, so that water can flow across the it without any (minimal) spurious reflections. This is most useful when the boundary of the computational domain does not have any physical significance but is an arbitrary edge of the domain of interest. In this case, the values of $\mathbf{q}$ are extrapolated from the cell next to the boundary into the ghost cells at each time step. This acts as a no-flow boundary in the case that the water is not flowing, but for non-zero velocity water flows over the boundary and leaves the domain. This is called a ‘zero order’ extrapolation in LeVeque [2002], and can be thought of as letting the numerical scheme collapse to a purely upwind scheme at the boundary. A first order extrapolation scheme can also be used in the code, but this has been shown to cause instabilities LeVeque [2002].

- Solid wall. The momentum of the water in the ghost cells is effectively reflected about the boundary, while the water depth is extrapolated as in the non-reflecting outflow case.

- Periodic. This is for situations where the behaviour of the system is periodic so that all water leaving the system at one edge re-enters at the opposite edge. The values for $\mathbf{q}$ in the inflow ghost cells are therefore copied from the cells at the same distance from the outflow boundary.

- User specified boundary conditions. These could be e.g. outflow conditions, if the flow across a boundary is known, or a boundary water depth condition if that is a known parameter for the system.

3.1.2.1 Topography at the boundary

The behaviour of water at the boundaries of the domain is highly dependent on the boundary conditions as described above. Another important factor for inundation modelling is the representation of the domain topography at and across the domain boundaries. In Clawpack, the value of the domain elevation is by default copied from the cells next to the boundary into the ghost cells. This represents a situation where there is no slope in bathymetry or topography across any boundaries. This is an adequate description of the left and right boundaries of the domains used here, since the domain is designed to be large enough in x that no water is expected to flow across these boundaries. The default no-slope condition is also used at the upstream boundary here, since very little water is likely to flow across it. However, at the downstream boundary the situation is very different; lots of water flows across the downstream boundary as it leaves the domain. A more physically realistic situation for the downstream boundary is therefore to extrapolate the slope of the domain into the ghost cells at the boundary. Changes have been made to the code to accommodate this, and it is possible to treat the downstream topography slope as a model parameter, as in Garcia-Pintado et al. [2015]. The effect of varying the parameter controlling the bathymetry slope at the boundary for a simple simulation can be seen in section 3.2.2.

3.2 Sensitivity studies

The following sensitivity studies were carried out in a domain with topography as shown in figure 3.2 with grid-spacing of approximately 10 m in both the x and y directions.

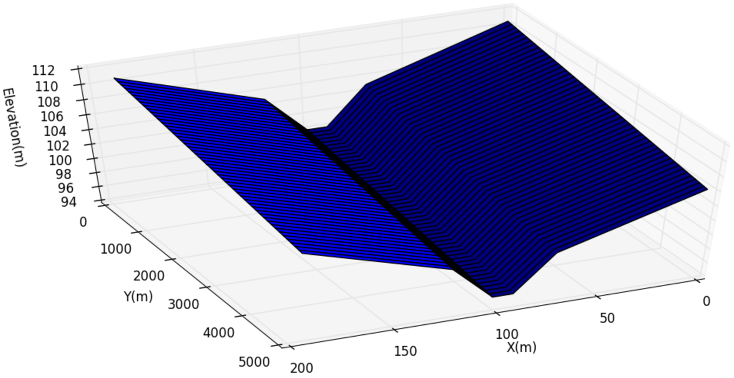


Figure 3.2: Elevation of the test domain in metres.

The upstream-downstream slope for the river and flood plain is 0.09% and the slope from the outside of the domain down to the river is 0.8%. The maximum depth of the channel below the bank is 8.5m and its width is 50m; these are values based on measured cross sections from the river Severn.

For the results shown in this section, the upstream boundary of the domain was set to be a solid wall to avoid water being able to leave the domain in that direction, and also to avoid water effectively being generated in the ghost cells. The other three boundaries are free (extrapolating) boundaries with an extrapolated slope for the downstream boundary.

3.2.1 Channel friction

In the following simulations, the domain shown in figure 3.2 was initially empty. Water was added into the domain close to the upstream boundary at a rate of $160m^3s^{-1}$ for a total simulation time of 3000 seconds. A gauge measuring depth (relative to topography) was simulated in a central position as shown in figure 3.3.

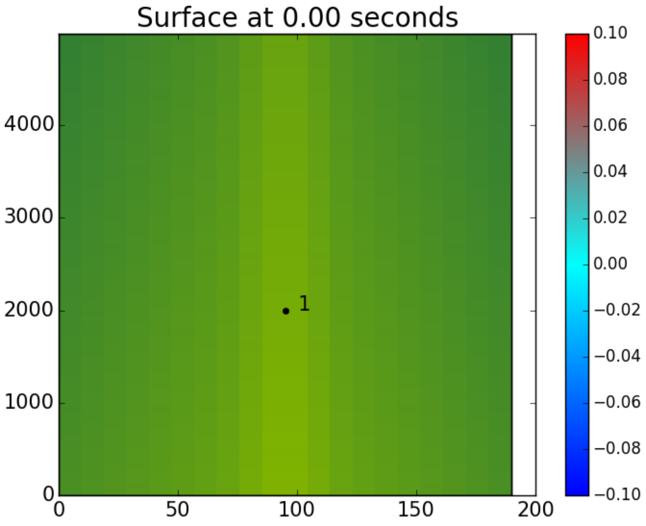


Figure 3.3: Location of gauge in domain marked with black dot and ‘1’. There is no water in the domain at the time shown; the green colours correspond to ‘dry land’ and are not shown on the colourbar.

The gauge recorded depth throughout the simulations, which were carried out for three different values of Manning’s friction coefficient,n in the domain. The friction values used were $n = 0.02$, $n = 0.002$ and $n = 0.07$. The first of these values is a reasonable estimate for a river channel; the other values are chosen as extremes of physically realistic possible values. For each case the friction coefficient was the same in the whole of the domain, i.e. the value of n in the rest of the domain was the same as in the channel. The results are shown in figure 3.4.

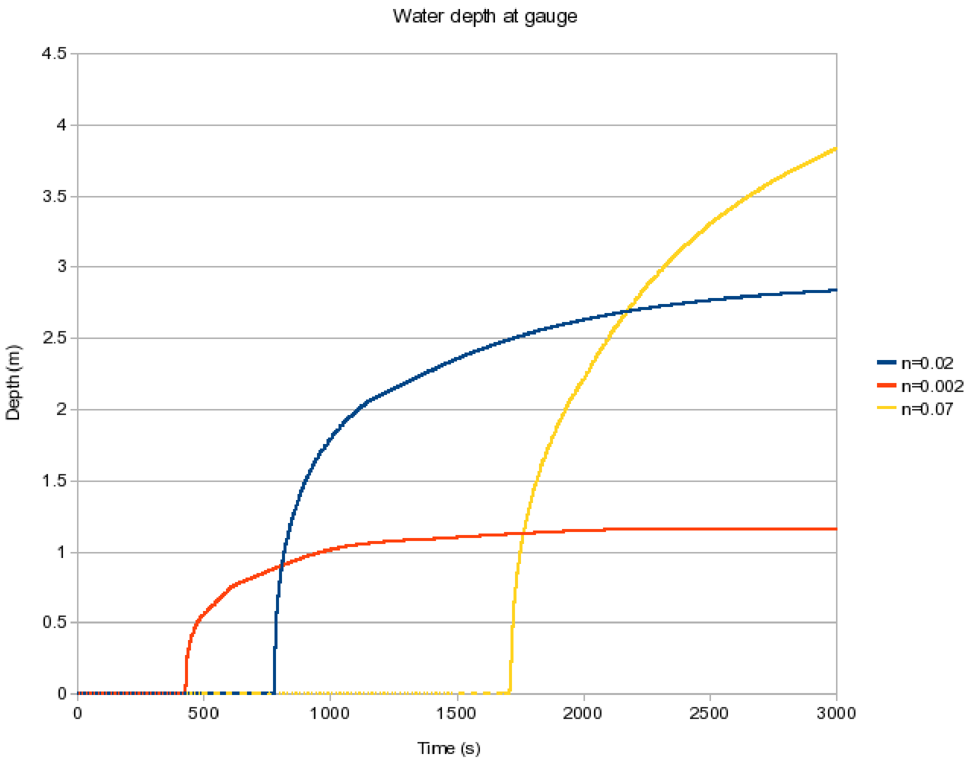
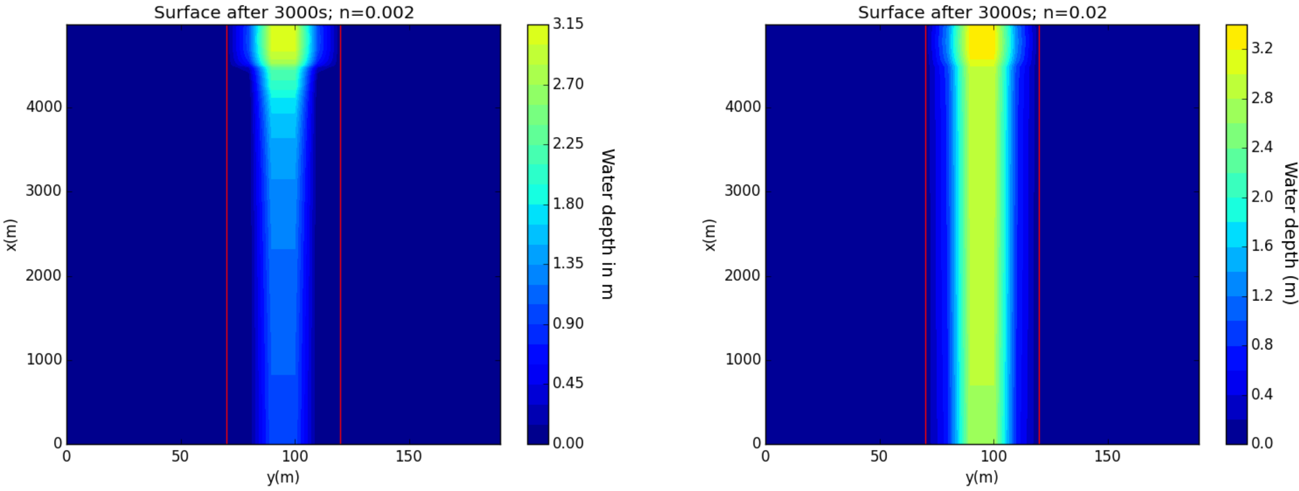
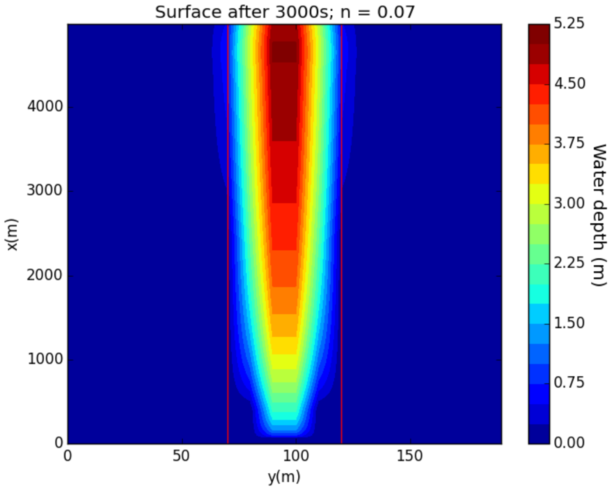


Figure 3.4: Water depth at gauge with time for different channel friction parameters

Figure 3.4 shows that the channel friction parameter influences both the time taken to reach equilibrium, and the equilibrium depth in the channel, where the channel is defined as the central 5 grid cells in the x direction for all y. The final water levels (after 3000 seconds) are shown in plan view for the three different values of n in figure 3.5. The edges of the channel are shown by the thin red lines and we can see in figure 3.5c that in the case in which $n = 0.07$, some flooding of the domain took place (where $x >$3000m). This is because under these conditions water travels so slowly in the channel in the x direction that it is forced out onto the banks. In all of the cases shown here the Manning's friction coefficient on the flood plain was the same as in the channel; making these values different will clearly affect the evolution of any inundation event.



(a) $n = 0.002$ (b) $n = 0.02$



(c) $n = 0.07$

Figure 3.5: Water depth in the domain at the end of the simulation for three different values of n in the channel. The domain is shown in plan view. Note the different scales for the colourbars.

3.2.2 Slope at downstream boundary

The effect of changing the parameter controlling the bathymetry slope at the downstream boundary can be seen in the following simulations. In each case, a symmetrical domain with a central channel was initially filled with water to a depth of approximately 4 m as shown in cross section in figure 3.6. The domain has a slope of 0.09% which caused the water to flow downhill under the influence of gravity and leave the domain at the downstream end. No extra water was added to the domain during the simulations, which ran for 1500 seconds. Some water remained in the domain at the end of the simulations. For one case the downstream slope was extrapolated into the ghost cells, and for the other case the 'no slope' condition was used, where the elevation in each ghost cells is set to be the same as the domain cell next to it on the inside of the boundary.

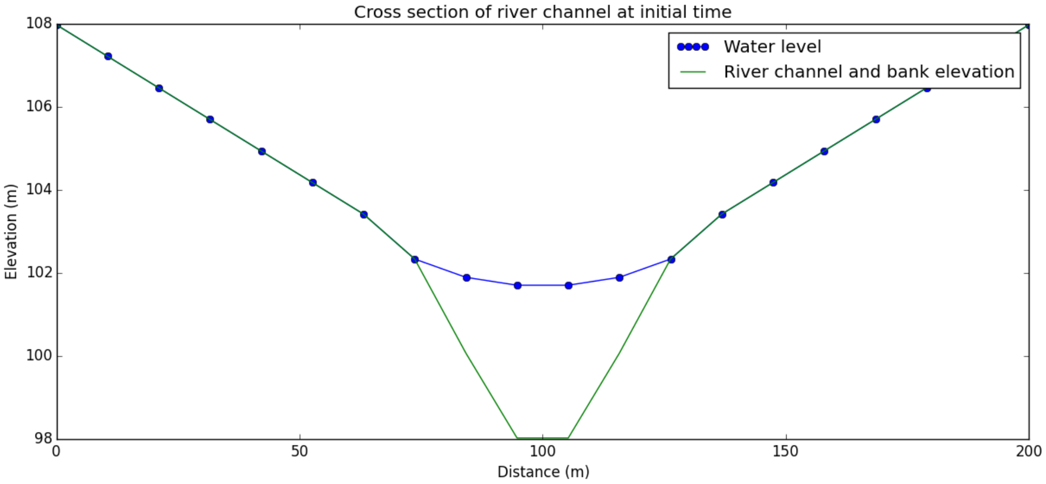
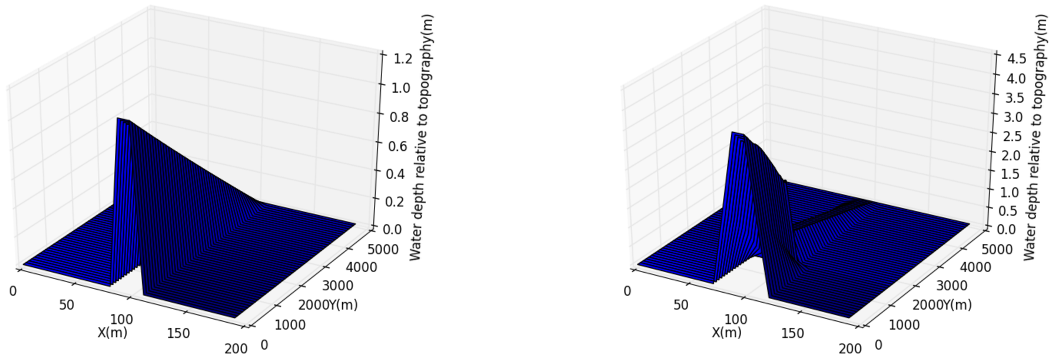
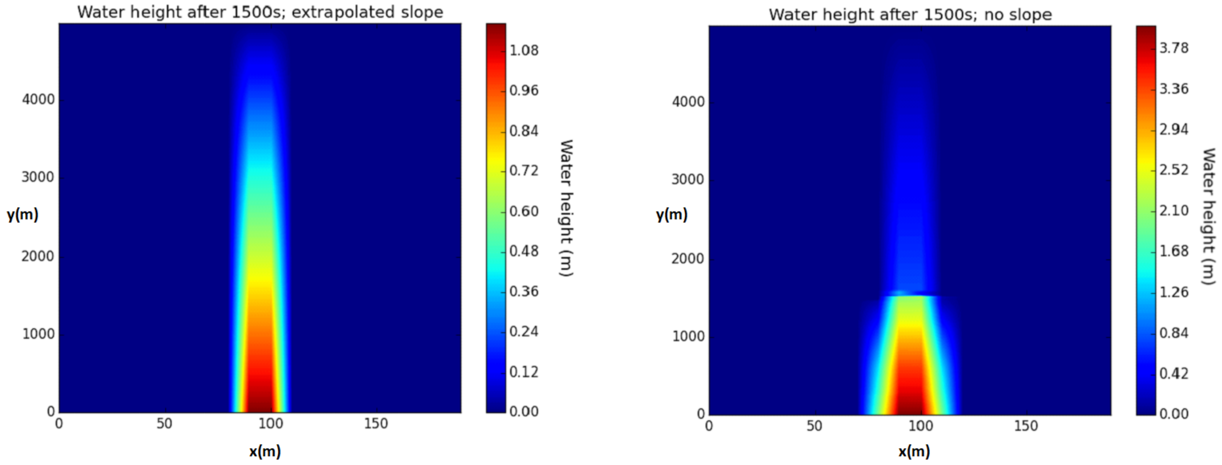


Figure 3.6: Cross section of the domain, showing the channel filled with water at the start of the simulations. The green line shows the elevation of the domain and the blue points show the water depth.

At the end of the simulations, the water profile is different for the two cases as shown in figure 3.7. For the extrapolated boundary slope condition, the water is able to leave the domain cleanly; the flow over the boundary is the same as elsewhere in the domain. For the no slope condition, water cannot leave the domain as fast and therefore builds up at the boundary.



(a) Extrapolated slope; water levels relative to topography in 3D (b) No slope; water levels relative to topography in 3D



(c) Extrapolated slope; water levels relative to topography in plan view (d) No slope; water levels relative to topography in plan view

Figure 3.7: Water profiles for the slope and no slope boundary conditions after 1500s.

3.3 Chapter summary

In this chapter we have described the inundation model used for the simulations in chapters 4, 5 and 6 of this thesis. We have described the boundary condition options in Clawpack; our description of the treatment of source terms is presented as part of our paper Cooper et al. [2018b], reproduced in this thesis as chapter 4 (see section 4.8 for treatment of source terms). In this chapter we have also presented results of studies investigating the sensitivity of modelled water levels to the domain friction parameter and the downstream topography slope.

Chapter 4:

Effect of channel friction estimation on observation impact

In this chapter we address the first research question outlined in chapter 1; How does estimation of the channel friction parameter affect observation impact in data assimilation for inundation forecasting? In particular we wish to find out:

- Can the ETKF retrieve the correct channel friction parameter in synthetic experiments using SAR-like observations, and does that improve the forecast?
- Is error in the channel friction parameter distinguishable from error in inflows?
- Does it matter if we assume zero momentum when restarting after assimilation?

We describe a series of synthetic experiments. The remainder of this chapter, except for the chapter summary (section 4.9), has been published and is reproduced from Cooper et al. [2018b].

4.1 Abstract

Accurate inundation forecasting provides vital information about the behaviour of fluvial flood water. Using data assimilation with an Ensemble Transform Kalman Filter we combine forecasts from a numerical hydrodynamic model with synthetic observations of water levels. We show that reinitialising the model with corrected water levels can cause an initialization shock and demonstrate a simple novel solution. In agreement with others, we find that although assimilation can accurately correct water levels at observation times, the corrected forecast quickly relaxes to the open loop forecast. Our new work shows that the time taken for the forecast to relax to the open loop case depends on domain length; observation impact is longer-lived in a longer domain. We demonstrate that jointly correcting the channel friction parameter as well as water levels greatly improves the forecast. We also show that updating the value of the channel friction parameter can compensate for bias in inflow.

Keywords Data assimilation, inundation forecasting, fluvial flooding, observation impact, joint state-parameter estimation, ensemble Kalman filter.

Highlights

- Data assimilation is applied to simulated flood forecasts and SAR-like observations
- Reinitialisation shock due to water level correction is removed using a novel method
- Observation impact is linked to domain length when updating only water levels
- Updating the channel friction parameter leads to marked improvement in forecast skill
- Updating the channel friction parameter can compensate for biased inflow

Software Availability

The inundation simulations in this work were generated using Clawpack 5.2.2, a collection of FORTRAN and python code available from http://www.clawpack.org/. Details of the amended Clawpack source code as used in this work are freely available on request from the corresponding author, as is the python code used to perform data assimilation on the inundation simulation output. Please contact e.s.cooper@pgr.reading.ac.uk for details.

4.2 Introduction

Data assimilation can improve the accuracy of predictions from flood inundation models by combining forecasts from the model with observations of the system, taking into account uncertainty in both the model predictions and the observations. In this study we use a sequential data assimilation method comprising a forecast-update dynamic feedback loop. During each forecast step, the numerical model runs an inundation simulation.

When an observation (or set of observations) is available the simulation is interrupted and the update step is performed; updating combines observational data and model predictions to give a better estimate of the state. The next forecast step then starts, with the adjusted water levels as the initial condition. An update is carried out each time a new observation or set of observations is available.

There are a number of numerical inundation models that can predict the behaviour of flood water given information about the topography of the domain and the amount of water flowing into the area, e.g. HEC-RAS, Telemac, LISFLOOD-FP [HEC-RAS Development Team, Hervouet, 2000, Neal et al., 2012a]. In a real flood situation, topographical information is often available in the form of a digital terrain model (DTM) and inflow estimates may come from an upstream gauge, or as output from a hydrological model. Observations of the flood may be available from a variety of different sources. These include river depth and flow rate measurements from gauges, and authors have used these data in assimilation schemes , e.g. Mure-Ravaud et al. [2016]. However, many catchments are ungauged and the number of gauges worldwide is in decline [Vrsmarty et al., 2001]. Observations of flood extent can be obtained from aerial photos, although the cloudy conditions associated with heavy rainfall often limit the usefulness of this information source. Recently, much attention has been paid to the use of synthetic aperture radar (SAR) satellite images in delineating flood extent, since such observing systems have all weather and day and night capability. Water depth information can then be retrieved from SAR satellite images using a high quality digital terrain model (DTM) as described in Mason et al. [2012] and Brown et al. [2016]. Such techniques for extracting information from SAR images are well established, e.g.Thornhill et al. [2012], Mason et al. [2010], Scott et al. [2008] and Scott and Mason [2007].

Various authors e.g. Lai and Monnier [2009], Matgen et al. [2007b] and Schumann et al. [2009] have used data assimilation techniques to highlight the fact that although observations from SAR can cover a large spatial area, the usefulness of the information they contain is limited in time. Assimilating data from one or more river gauges can help to mitigate this, as shown by Lai and Monnier [2009] and Hostache et al. [2010], but we consider here the situation in which only time-sparse satellite derived water level data is available for assimilation. This leads to a situation in which data assimilation can provide a good analysis - i.e. can correct water levels very well at the time of observations, but the model forecast then moves quickly away from the true water levels during the subsequent forecast step. This short lived improvement in the water levels has been shown in studies such as Andreadis et al. [2007], Neal et al. [2009] and Garcia-Pintado et al. [2013], which use ensemble Kalman Filter data assimilation algorithms, as well as in Matgen et al. [2010], where a particle filter approach is taken. This result indicates that water levels in a river flood situation are not strongly sensitive to initial conditions. In fact, water levels are likely to be more dependent on inflow and model parameter values, and updating one or a combination of these is therefore necessary.

In order to address the short-lived nature of the forecast improvement, authors such as Andreadis et al. [2007], Matgen et al. [2010], Giustarini et al. [2011], Garcia-Pintado et al. [2015], Garcia-Pintado et al. [2013] and Mason et al. [2015] have carried out data assimilation including on-line correction of inflow along with water levels. Inflow correction is shown in all of these studies to give much better forecast accuracy over time than correcting water levels alone. Less attention has been paid to the effect of errors in model parameters in sequential data assimilation, despite the fact that several studies, including Andreadis and Schumann [2014] and the comprehensive review paper by Grimaldi et al. [2016], indicate that model parameters are likely to have an important influence on the behaviour of the flow.

One study in which parameter effects are investigated is Garcia-Pintado et al. [2015], in which water levels, inflows and several model parameter values were updated simultaneously using an ensemble Kalman filter technique. The study used LISFLOOD-FP to model the flooding of the river Severn and tributaries near Tewkesbury, UK, in 2014, assimilating real SAR-derived water level observations. A large improvement in forecast skill was seen when inflow was corrected along with water levels, leading to good agreement between the forecast and independently measured gauge data. In this case, channel friction parameter estimation alongside estimation of water levels, inflows and other parameters was not found to improve the forecast significantly, despite the fact that water behaviour is strongly influenced by this parameter. The question of whether the retrieved friction parameter value was correct was left open as the true value for the system was not known.

In this study we address open questions about the role of the channel friction parameter in data assimilation for inundation modelling. We use a similar data assimilation technique to that in Garcia-Pintado et al. [2015] in twin experiments with an idealised topography and an unbiased inflow. This allows us to separate out and further investigate the effect of channel friction retrieval on the forecast. We find that, in contrast with Garcia-Pintado et al. [2015], online estimation of the channel friction parameter along with water levels leads to a large improvement in the forecast skill in our experiments. The twin experiments also show that our data assimilation method is capable of finding an accurate value for the channel friction parameter, even when water depth observations are only available on the flood plain during a flood.

We also investigate the effect of domain length on forecast skill, showing that because the assimilation is able to correct water levels in areas where there are no observations, the time taken for corrected water levels to decay back to the open loop (no assimilation) case is longer for a physically longer domain. Further, we demonstrate that when reinitialising the numerical model after an assimilation, an initialisation shock can occur. We demonstrate an efficient and effective technique for removing this shock, leading to more accurate forecasts in the hours immediately following an assimilation.

This paper is organised as follows: In section 4.3 the numerical inundation model is described, the data assimilation method is outlined and our novel re-initialisation method is demonstrated. In section 4.4 the experimental configurations for various simulations are described. Section 4.5 shows the effect of including online channel friction parameter estimation along with water level estimation, and compares results from different length domains. Section 4.6 draws conclusions about the effects of domain length and channel friction parameter estimation.

4.3 Methodology

In this section we describe the methods used in this study. In section 4.3.1 the numerical inundation model is outlined. Section 4.3.2 contains information about the data assimilation method used. In section 4.3.3 we discuss the impact of assuming the water has only hydrostatic momentum at the start of a forecast step and describe our approach to dealing with problems caused by this assumption.

4.3.1 Numerical inundation model

In this study we use a numerical flood model we have developed using Clawpack [Clawpack Development Team, 2014, Mandli et al., 2016, LeVeque, 2002], an open source collection of FORTRAN and python code that can be used to solve a wide variety of conservation laws. Clawpack uses finite volume methods and sophisticated Riemann solvers to treat systems of partial differential equations; in this work the equations of interest are the 2D shallow water equations that describe how river and flood water will move in space and time. The model splits the domain of interest into N cells and calculates the water depth in each cell. The code is capable of dealing with shocks in the solution, such as bores that may occur following a sudden increase of inflow into a particular river stretch. Clawpack deals effectively with the wet-dry interfaces which are present in an inundation event, and preserves depth non-negativity [George, 2008].

The shallow water equations for two spatial dimensions, x and y, can be written as (e.g.LeVeque [2002])

$$\frac{\partial \mathbf{q}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{q})}{\partial x} + \frac{\partial \mathbf{G}(\mathbf{q})}{\partial y} = \mathbf{R}(\mathbf{q}), \tag{4.1}$$

where $\mathbf{R(q)}$ is a source term and $\mathbf{q}$ is a vector of conserved quantities

$$\mathbf{q} = \begin{bmatrix} h \\ hu \\ hv \end{bmatrix}, \tag{4.2}$$

h represents depth of the fluid, and u and v represent velocity in the x and y directions respectively.

In equation (4.1), $\mathbf{F(q)}$ and $\mathbf{G(q)}$ represent fluxes of the conserved quantities in the x and y directions respectively. For the shallow water equations these are

$$\mathbf{F}(\mathbf{q}) = \begin{bmatrix} hu \\ hu^2 + \frac{1}{2}gh^2 \\ huv \end{bmatrix} \quad \text{and} \quad \mathbf{G}(\mathbf{q}) = \begin{bmatrix} hv \\ huv \\ hv^2 + \frac{1}{2}gh^2 \end{bmatrix}, \tag{4.3}$$

where g is acceleration due to gravity.

The effect of friction is modelled as a source term in Clawpack, since the friction force acts to reduce the momentum of the water. The magnitude of the momentum reduction is strongly dependent on a Manning's friction coefficient, n, and the flow of water is very sensitive to the value of this parameter. We have also added an inflow source term to the Clawpack code to model water arriving in the river stretch of interest, as further described in Appendix section 4.8.2 and Cooper et al. [2016]. The time step for the hydrodynamic model is variable, and automatically adjusted in the code to preserve numerical stability.

Correct specification of the solution at the boundaries of the computational domain is vital for the stability of any numerical scheme. To achieve this, Clawpack adds a user-specified number of 'ghost' cells (2 by default) next to each cell at a domain boundary.

The domain is effectively extended in all directions by the addition of these ghost cells and the behaviour of the solution at the boundaries then depends strongly on the values of calculated model quantities in the ghost cells. We use non-reflecting outflow (extrapolating) boundary conditions in which values of $\mathbf{q}$ are extrapolated from the cell next to the boundary into the ghost cells at each time step. This is called a zero order extrapolation in LeVeque [2002].

Another important factor is the representation of the domain topography at and across the domain boundaries. The default in the code is to copy the value of the domain elevation at the boundary into the ghost cells. This represents a situation where there is no slope in bathymetry or topography across any boundaries. This is not suitable for the downstream boundary in our experiments, at which the majority of the water leaves the domain. A more physically realistic situation for the downstream boundary is to extrapolate the slope of the domain into the ghost cells at the boundary and changes have been made to the code to accommodate this.

4.3.2 Data assimilation

4.3.2.1 State estimation

In data assimilation, a state vector is used to represent the state of a physical system. In this work the state vector, $\mathbf{x} \in \mathbb{R}^N$, comprises water depths in each of N computational cells. Sequential data assimilation algorithms comprise two steps: a forecast (or prediction) step and an update (or analysis) step. In the prediction step, an estimate of the state, $\mathbf{x}$ is evolved forward in time using the forecast model

$$\mathbf{x}(t_{k+1}) = \mathfrak{M}(\mathbf{x}(t_k)), \tag{4.4}$$

where $\mathfrak{M}$ is the forecast model, in this case the non-linear numerical shallow water equation model described in section 4.3.1. In the update step the forecast is updated to take account of observations of the state. We assume that the observations can be described by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \epsilon, \tag{4.5}$$

where $\mathbf{y} \in \mathbb{R}^p$ is a vector of observations and $\mathbf{x}$ is the true state of the system. Since the observations may be indirect and not located at model cell centres, an observation operator, $\mathbf{H} : \mathbb{R}^N \rightarrow \mathbb{R}^p$ is required, which maps the state vector into observation space. For this work, $\mathbf{H}$ is assumed to be a linear operator. The observation noise, ϵ is assumed to be unbiased stochastic noise with covariance $\mathbf{R} \in \mathbb{R}^{p\times p}$. The ensemble Kalman filter used here is based on the Kalman filter. In the Kalman filter, whenever observations are available the state and error covariance matrix are updated Kalman [1960] according to

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K}(\mathbf{y}_{obs} - \mathbf{H}\mathbf{x}^f) \tag{4.6}$$

and

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^f, \tag{4.7}$$

where forecast and analysis quantities are denoted by f and a superscripts respectively, $\mathbf{I} \in \mathbb{R}^{N\times N}$ is the identity matrix and $\mathbf{P} \in \mathbb{R}^{N\times N}$ is the state error covariance matrix. The matrix $\mathbf{K} \in \mathbb{R}^{N\times p}$ is the Kalman gain, given by

$$\mathbf{K} = \mathbf{P}^f\mathbf{H}^T(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R})^{-1}, \tag{4.8}$$

and $\mathbf{R} \in \mathbb{R}^{p\times p}$ is the observation error covariance matrix.

In the ensemble Kalman filter [Evensen, 1994], an ensemble of state vectors is used to represent a statistical sample of the forecast or analysis uncertainty. Each ensemble member represents one possible realisation of the true state of the system, given uncertainties in initial conditions and/or model parameters. For an ensemble comprising M state vectors, $\mathbf{x}_i$, $(i = 1, 2...M)$, a mean state at any time can be defined as

$$\overline{\mathbf{x}} = \frac{1}{M}\sum_{i=1}^{M}\mathbf{x}_i. \tag{4.9}$$

The mean of the ensemble, $\overline{\mathbf{x}}$, represents an estimate of the true state of the system.

For any ensemble, an ensemble perturbation matrix $\mathbf{X} \in \mathbb{R}^{N\times M}$ can be defined as

$$\mathbf{X} = \frac{1}{\sqrt{M-1}}(\mathbf{x}_1 - \overline{\mathbf{x}} \;\; \mathbf{x}_2 - \overline{\mathbf{x}} \;\; ...... \;\; \mathbf{x}_M - \overline{\mathbf{x}}). \tag{4.10}$$

The ensemble error covariance matrix, $\mathbf{P} \in \mathbb{R}^{N\times N}$ can then be calculated from

$$\mathbf{P} = \mathbf{X}(\mathbf{X})^T. \tag{4.11}$$

The forecast step for an ensemble system requires each state vector in the ensemble to be evolved by the forecast model according to equation (4.4). In the update step the forecast ensemble is combined with observations of the state to produce a ‘corrected’ ensemble of state vectors called the analysis ensemble. The analysis ensemble is then used as a set of initial conditions for the next forecast step. This forecast-update cycle can be repeated many times and an analysis ensemble calculated whenever observations of the system are available. The ensemble update equations separately update the ensemble mean and the ensemble perturbations according to

$$\overline{\mathbf{x}^a} = \overline{\mathbf{x}^f} + \mathbf{K}(\mathbf{y} - \overline{\mathbf{H}\mathbf{x}^f}) \tag{4.12}$$

and

$$\mathbf{X}^a = \mathbf{X}^f\mathbf{T}. \tag{4.13}$$

The vector $\overline{\mathbf{x}^a}$ is the analysis state (the mean of the analysis ensemble), $\overline{\mathbf{x}^f}$ is the mean of the forecast ensemble and $\mathbf{K} \in \mathbb{R}^{N\times p}$ is an ensemble version of the Kalman gain (as shown in equation (4.16)). The matrix $\mathbf{T} \in \mathbb{R}^{M\times M}$ updates the perturbations such that the state error covariance calculated by using $\mathbf{X}^a$ in equation (4.11) matches that given by the Kalman covariance update, equation (4.7) [Kalman, 1960]. There is not a unique solution for $\mathbf{T}$; here we use an unbiased, symmetric square root formulation known as an ensemble transform Kalman Filter (ETKF), following the approach of Livings et al. [2008], Livings [2005] and Garcia-Pintado et al. [2013]. In this approach we define a forecast observation ensemble comprising M forecast observation vectors, $\mathbf{y}_i^f$, $(i = 1, 2 \ldots M)$ such that

$$\mathbf{y}_i^f = \mathbf{H}(\mathbf{x}_i^f). \tag{4.14}$$

The forecast observation ensemble has a mean, and a perturbation matrix $\mathbf{Y}^f$, defined in the same way as for the state ensemble matrix.

We define a matrix $\mathbf{D}$ as

$$\mathbf{D} = \mathbf{Y}^f(\mathbf{Y}^f)^T + \mathbf{R}; \tag{4.15}$$

the Kalman gain $\mathbf{K}$ can then be written in terms of the forecast perturbation matrices $\mathbf{X}^f$ and $\mathbf{Y}^f$,

$$\mathbf{K} = \mathbf{X}^f(\mathbf{Y}^f)^T\mathbf{D}^{-1}. \tag{4.16}$$

Substituting equation (4.16) for $\mathbf{K}$ on the right hand side of equation (4.7), and using equations (4.13) and (4.11) on the left hand side shows that the matrix $\mathbf{T}$ in equation (4.13) then needs to satisfy

$$\mathbf{T}(\mathbf{T})^T = \mathbf{I} - (\mathbf{Y}^f)^T\mathbf{D}^{-1}(\mathbf{Y}^f). \tag{4.17}$$

Using the Sherman-Woodbury-Morisson identity for the right hand side of equation (4.17), as in equation (15) of Tippett et al. [2003], this becomes

$$\mathbf{T}(\mathbf{T})^T = (\mathbf{I} + (\mathbf{Y}^f)^T\mathbf{R}^{-1}\mathbf{Y}^f)^{-1}. \tag{4.18}$$

A scaled forecast observation ensemble perturbation matrix can then be introduced,

$$\widehat{\mathbf{Y}^f} = \mathbf{R}^{-\frac{1}{2}}\mathbf{Y}^f. \tag{4.19}$$

Performing a singular value decomposition [Golub and Van Loan, 1996] of $(\widehat{\mathbf{Y}^f})^T$ gives a factorisation such that

$$(\widehat{\mathbf{Y}^f})^T = \mathbf{U}\Sigma\mathbf{V}^T, \tag{4.20}$$

where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices with dimensions (M by M) and (p by p) respectively. The columns of $\mathbf{U}$ and $\mathbf{V}$ are the left and right singular vectors of $(\widehat{\mathbf{Y}}^f)^T$ respectively, and the diagonal elements of the (M by p) matrix Σ are the singular values of $(\widehat{\mathbf{Y}}^f)^T$. Combining equations (4.18), (4.19) and (4.20) gives

$$\mathbf{T}(\mathbf{T})^T = \mathbf{U}(\mathbf{I} + \Sigma\Sigma^T)^{-1}\mathbf{U}^T, \tag{4.21}$$

and a solution for $\mathbf{T}$ is therefore

$$\mathbf{T} = \mathbf{U}(\mathbf{I} + \Sigma\Sigma^T)^{-\frac{1}{2}}\mathbf{U}^T. \tag{4.22}$$

This is the solution used in this work.

4.3.2.2 Joint state-parameter estimation

Section 4.3.2.1 describes how the ETKF can be used to update the water levels in a computational domain, given observational data. It is also possible to update values of uncertain forecast model parameters as part of the same process. This is achieved by state augmentation, in which parameters are appended to the state vector [Smith et al., 2013, Navon, 1998, Evensen et al., 1998, Smith et al., 2009, 2011]. The augmented state vector, $\mathbf{x}_{aug}$, is then given by

$$\mathbf{x}_{aug} = \begin{bmatrix} \mathbf{x} \\ \mathbf{b} \end{bmatrix}, \tag{4.23}$$

where $\mathbf{b} \in \mathbb{R}^m$ is a vector of m parameters and $\mathbf{x}_{aug} \in \mathbb{R}^{N+m}$. In this work, we are interested in updating just one parameter, the Manning's friction coefficient in the river channel, n_{ch}. This means that $\mathbf{b}$ is scalar in this case.

We assume that the value of n_{ch} does not change with time over the course of a particular flood. This means that the value of n_{ch} is constant during the forecast step and only updates at assimilation times. The forecast equation for the augmented state vector is then given by

$$\mathbf{x}_{aug}(t_{k+1}) = \begin{bmatrix} \mathfrak{M}(\mathbf{x}(t_k)) \\ \mathbf{b}(t_k) \end{bmatrix}, \tag{4.24}$$

where $\mathfrak{M}$ is the forecast model as in equation (4.4).

The ETKF update equations (4.12) and (4.13) can be applied to the augmented state vector in the same way as described in section 4.3.2.1. The ensemble assimilation scheme then takes into account covariances between errors in the state vectors and the parameter(s). These covariances act to correct the parameter value according to information from observations as part of the same process that corrects water levels in the domain.

Estimating parameter values in this way has a number of advantages over a more traditional offline calibration approach. Firstly, the updating of the parameter values is performed with information from current observations. Calibrating parameters with data from previous events risks using out of date information which does not take into account changes to the river bed due to, for example, erosion or sediment transport. Calculating parameters using data assimilation also allows the value to change on shorter timescales during a flood event as the value is assumed to be constant during forecast steps, but updates each time new observational information is available. Additionally, off-line calibration and tuning of parameters can be computationally costly and needs to be performed ahead of a flooding event which may occur with little warning.

In our synthetic experiments we have assumed that the channel friction parameter, n_{ch} is the same value for the whole channel. In a real setting, it would be necessary to take an approach like that in Garcia-Pintado et al. [2015], where different values of n_{ch} were assumed for different types of channel such as large rivers and small streams.

4.3.3 Hydrostatic initialisation shock

The ETKF is a sequential method and this means that each time observations are available an assimilation can be carried out. In order to perform an assimilation, the forecast model is interrupted. The ensemble states (water levels) are updated by the

ETKF and each ensemble member then restarts running in Clawpack, reinitialised with the new water levels. In some approaches to data assimilation for inundation modelling, e.g. Lai and Monnier [2009], Hostache et al. [2010], Ricci et al. [2011], the state vector contains information about water flow or discharge rates as well as water levels. The flow rates are therefore updated along with water levels as part of the assimilation process. In contrast, in this study, as in Garcia-Pintado et al. [2015], the state vector contains only water depth information (plus parameters when considering the joint state-parameter estimation problem).

In Garcia-Pintado et al. [2015], a hydrostatic assumption was made for reinitialisation, i.e. the water was assumed to have zero momentum at each forecast restart time. The effect of this assumption in our domain can be seen by interrupting a simulation and restarting without performing any data assimilation, i.e. reinitialising with identical water levels as before, but with zero momentum everywhere. Comparing the root mean square error (RMSE) between the water depths predicted by a reinitialised version and continuous version of the simulation then gives a measure of how the hydrostatic assumption at restart affects the forecast. The RMSEs are measured over the whole domain and defined as

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(h_i^r - h_i^c)^2}, \tag{4.25}$$

where h_i^r is the forecast water depth after reinitialisation in the ith cell and h_i^c is forecast water depth in the same cell without reinitialisation. The number of cells in the domain is N, as before.

The open circles in figure 4.1 show RMSEs between the reinitialised forecast and the continuous forecast. Figure 4.1 shows that the consequence of using a hydrostatic assumption is that the error between the continuous and restarted cases is large at times less than approximately four hours in this system. In figure 4.1 the maximum RMSE caused by reinitialisation is 0.3m compared to depths in the channel of 5 to 6m and up to 0.4m on the floodplain. This means that forecasting the behaviour of flood water at these times is problematic. The error becomes negligible by approximately four hours after the assimilation time.

In order to correct for this without adding flow information to the state vector in the update step, we assume that the water in each cell has the same velocity (u and v in the x and y directions respectively) after the assimilation as it did before the assimilation. This gives a state vector of lower dimension (approximately one third as many entries) than the approach of Lai and Monnier [2009], Hostache et al. [2010] and Ricci et al. [2011], thus reducing computational expense, while avoiding problems caused by assuming zero flow rates at the start of each forecast step as in Garcia-Pintado et al. [2015]. Note that the state vector for the forecast step still contains water depth and flow rate information. Reinitialising with the same flow rate (hu and hv) values in each cell would also be possible, but since the water levels will have changed in some cells due to the assimilation, this is likely to lead to unrealistic behaviour. Reinitialising the water with the pre-analysis velocity values removes the large shock shown in figure 4.1; the resulting RMSE values are shown in figure 4.1 with filled circles. The very small RMSE values shown by the filled circles show that the method is effective in removing the initialisation shock.

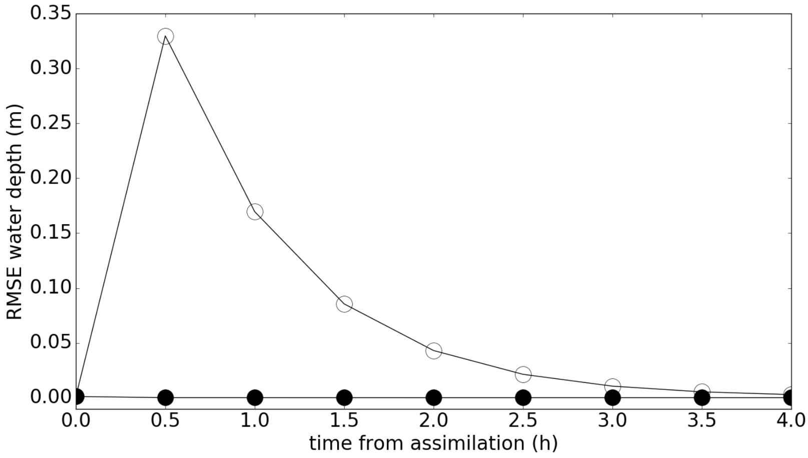
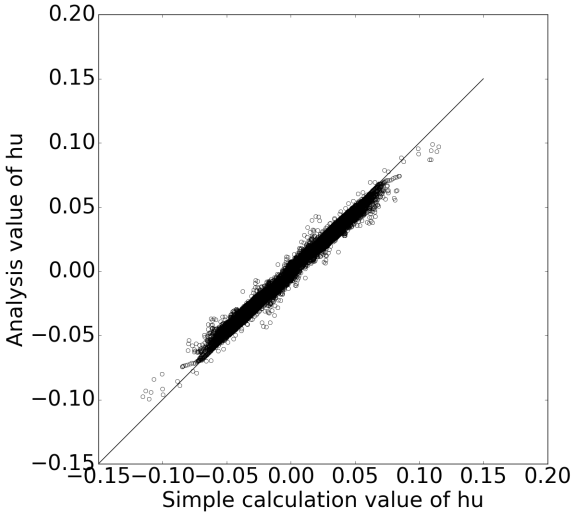
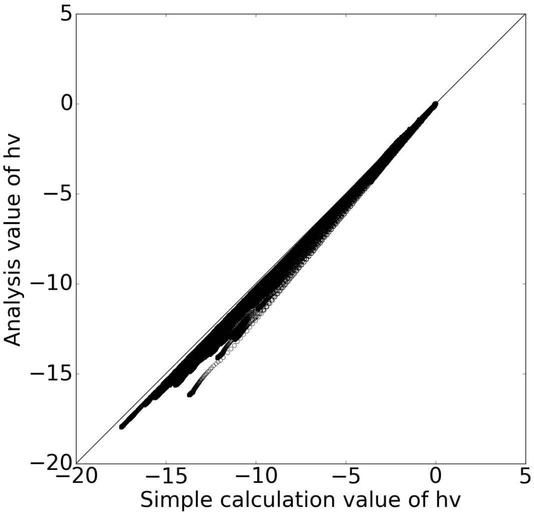


Figure 4.1: RMSE in water depth over the domain. Open circles show the RMSE between the continuous truth and restarted truth for a hydrostatic assumption. Filled circles show the RMSE when reinitialising with forecast velocities.

We performed a simple comparison of the values of hu and hv obtained using our approach ('simple calculation value') with those calculated during an assimilation in which hu and hv were included in the state vector ('analysis value'). We compared values at each assimilation time in an identical twin experiment in which we update both the water levels and the channel friction parameter. Figure 4.2 shows some typical results (from the assimilation at 28h in the SPL experiment as described in section 4.4.2).



(a)



(b)

Figure 4.2: Comparison of (a)hu and (b)hv values at one assimilation time.

Figure 4.2a shows the values of hu obtained using the simple method described here plotted against the values of hu calculated by including hu and hv in the assimilation state vector; figure 4.2b shows the same results for values of hv. The values are close to the identity line and therefore in broad agreement. We found that the agreement between the methods was better at later assimilation times, when the forecast error is low. This is because both methods provide values close to the pre-assimilation (forecast) values when the adjustment by the assimilation is relatively small. The values of n_{ch} obtained when including hu and hv in the state vector were almost identical (less than 0.001% difference) to the values obtained when the state vector comprises only water levels. We observed no instabilities in the solution at initialisation times using our technique.

4.4 Experimental design

4.4.1 Model domain

All of the experiments referred to in this paper are carried out in domains with a simplified river valley-like topography. We use two domains in this work, the 'long' domain describes an area of 20 km by 250m and is shown in figure 4.3. Note that the axes are not to the same scale. The 'regular' domain is 5km by 250m, and is identical to the long domain for $15 \leq y \leq 20km$ and $0 \leq x \leq 250m$. The domains are gently sloping symmetrical valleys with a 50m wide central river channel as shown in figure 4.3. The grid cell size for the computation is 10m by 10m in all cases. The river channel is defined to be the central 5 grid cells in the x direction for all values of y; the rest of the domain is defined as the flood plain. The cross section for both domains is the same. The domain has an upstream-downstream slope of 0.08% and the slope of the floodplain towards the river is 0.8%.

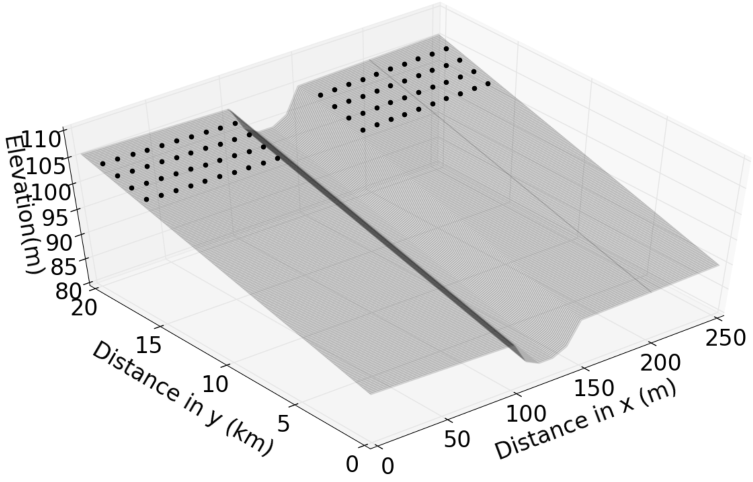


Figure 4.3: Elevation in metres of the long test domain used for the assimilation experiments. The regular domain is defined by $15 \leq y \leq 20$km and $0 \leq x \leq 250$m . The black circles indicate the positions of water elevation observations used in the data assimilation.

4.4.1.1 Identical Twin Experiments

In this study, we use an ETKF in identical twin experiments. Identical twin experiments are commonly carried out in order to test a data assimilation system as well as to generate information about the model to which data assimilation is applied, e.g. Evensen [1994]. In such experiments, a numerical model is used to generate a 'truth' output for a set of known initial conditions and model parameters. The truth run for these experiments is a continuous run of our inundation model for 112 hours using a time varying inflow shown by the solid black circles in figure 4.4. For the first four hours, the inflow is set to be constant in order to fill the river channel with water in a spin-up period. The inflow from t = 4 hours onwards is based on some real hydrograph data from a flood of the river Severn near Tewkesbury. The inflow comprises hourly values and linear interpolation is used to give flow rates between the hourly points. The truth run uses a value for the channel Manning's friction parameter of $n_{ch} = 0.04$, which is the value given for a natural stream by Maidment and Mays [1988]. The Manning's friction parameter on the flood plain, n_{fp}, is likely to be higher due to vegetation and here we use a value of $n_{fp} = 0.05$.

The inundation model is also used to generate a 100 member ensemble of flood realisations. This ensemble represents a forecast of the true flood given uncertainty in upstream inflow and the channel friction parameter values. The number of ensemble members is relatively small compared to the dimension of the state vector, which contains 12,500 water depths for the short domain and 50,000 for the long domain. However, none of the problems which indicate undersampling of a system (spurious correlations or ensemble collapse as outlined by Petrie and Dance [2010]) are seen, suggesting that this number of ensemble members is sufficient to represent the system in this case. It was therefore not necessary to apply any localisation or inflation in these experiments.

Each member of the ensemble is driven by a different inflow and has a different friction parameter. The ensemble inflows are generated by adding time correlated random errors to the 'true' inflow; the ensemble inflows and true inflow are shown in figure 4.4. The variance for the inflow distribution is a proportion of the inflow, since the error in measured or predicted flow is likely to be flow-dependent as in Garcia-Pintado et al. [2015]. The standard deviation for the generated inflow distribution is $0.15\times$ inflow, which is the same as the value used by Garcia-Pintado et al. [2015] and fits within the range of errors in measured flow rates (4% to 43%) reported in Di Baldassarre and Montanari [2009]. No bias was applied to the inflow ensemble so that the mean inflow is very similar to the true inflow. An unbiased inflow removes the effect of an incorrectly specified inflow, since this has already been studied elsewhere (e.g. Andreadis et al. [2007], Matgen et al. [2010], Giustarini et al. [2011], Garcia-Pintado et al. [2015], Garcia-Pintado et al. [2013] and Mason et al. [2015]). Choosing to use an unbiased inflow allows the effect of the incorrect channel friction parameter value to be studied in isolation; we briefly address the effect of biased inflow at the end of section 4.5.2.

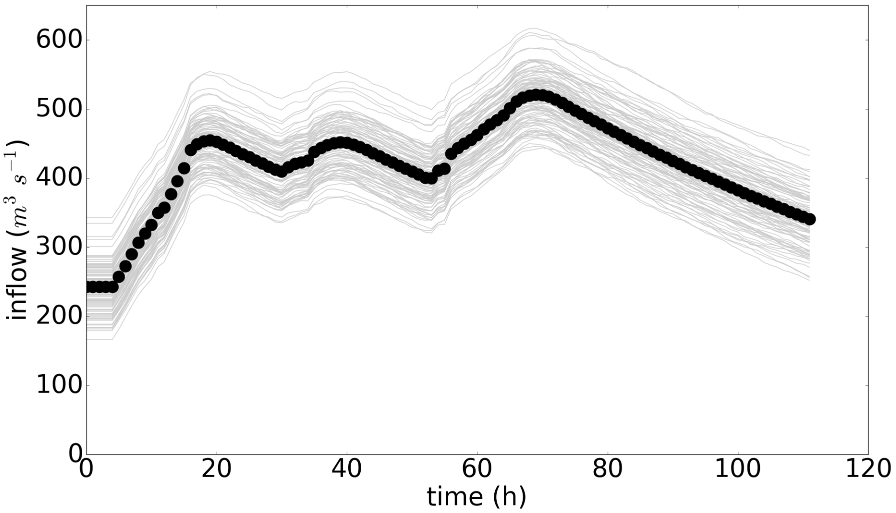


Figure 4.4: Inflow ensemble with time. Circles show the true inflow values and the grey lines show the ensemble inflows.

Ensemble channel friction parameters were generated by selecting from a Gaussian distribution centered on a ‘wrong’ initial value to reflect the fact that this parameter varies between catchments and will not generally be known before the start of a flood event. The channel friction parameter is also not likely to be directly measurable as it relates to the specific way in which Clawpack models friction. Different numerical models with different friction parameterisations have been shown to generate different optimal friction parameters for the same data for this reason [Horritt and Bates, 2002]. Centering the channel friction ensemble on an incorrect value also enables us to test whether the data assimilation scheme can retrieve the correct value from an incorrect starting point. The channel friction parameter for each ensemble member was selected from a Gaussian distribution, $\mathcal{N}(0.05, 0.01)$. The true value of $n_{ch} = 0.04$ falls within one standard deviation of the mean of this distribution. The value for the friction parameter on the flood plain was set to the true value for all ensemble members, i.e. $n_{fp} = 0.05$.

Data assimilation using an ETKF is carried out on the forecast ensemble, using synthetic observations generated from the truth as described in the next section. Since the ETKF is a sequential method the ensemble is evolved forward in time with Clawpack between observation times, and an assimilation is carried out every 12 hours. We also consider the 'open loop' case in which a forecast ensemble runs with no assimilation.

4.4.2 Experimental configurations

Various data assimilation experiments are carried out, each for a total period of 112 hours, including a 4 hour spin-up period with constant inflow to allow the river channel to fill up (as shown in figure 4.4). Synthetic observations of water depth are taken from the truth at 12 hourly intervals and assimilated with an ETKF. Twelve hourly assimilation intervals were chosen as this represents the smallest likely return time for SAR-equipped satellites at present.

Since we are running identical twin experiments, we know the true water levels everywhere in the domain, as well as the water levels forecast by the ensemble. It is theoretically possible to use SAR derived flood extent observations along with a high quality digital terrain model (DTM) to derive water levels in all parts of a real domain. One approach in a synthetic experiment is therefore to directly use water level observations in all parts of the domain as a proxy for SAR derived information. Such an approach is used by, for example, Lai and Monnier [2009]. In reality, SAR images can reliably provide information about water elevation only at a few points along the flood extent, as demonstrated in Mason et al. [2012].

In this paper, we do not use all the available water levels. Instead we use synthetic observations of water levels taken directly from the truth run in the positions shown by black circles in figure 4.3. This approach replicates a situation in which four reliable flood delineation positions are available from a SAR image at $y = 16, 17, 18$ and 19km. We assume that we have a SAR image covering the domain from $y = 15$ to 20km, and that water elevation at each of the four flood edge positions can be obtained from a DTM; this water level can then be extrapolated perpendicular to the channel to give water elevation in each floodplain cell in the cross sections where we have delineation observations (i.e. $y = 16, 17, 18, 19$km). Although extrapolating water elevation across a cross section in this way would also give information about the water elevation in the channel, we exclude observations in the channel in this experiment. This is because topography information in the channel is likely to be much less accurate than that for the floodplain, making water depths less certain. In a real case, water depth observations obtained in this way would have correlated errors, but since we use synthetic observations this is not the case here. Noise was added to the synthetic observations to represent uncertainty due to instrument error, flood extent determination and typical DTM errors. The noise added to the observations is Gaussian with a standard deviation of 0.25m; this is the same value as used in experiments with real data in Garcia-Pintado et al. [2015] and empirically determined in Mason et al. [2012].

We perform a series of experiments using both the regular and long domain. Further details of the individual experiments are as follows:

- **Case SOR**: State-only estimation in the regular domain, with the ensemble of channel friction parameters distributed about an incorrect 'first guess' as described in section 4.4.1.1.
- **Case SOL**: State-only estimation in the long domain; all other details as for case SOR.
- **Case SPR**: State and channel friction parameter estimation in the regular domain with the initial channel friction parameter distributed about an incorrect 'first guess' as described in section 4.4.1.1 .
- **Case SPL**: State and channel friction parameter estimation in the long domain; all other details as for case SPR.
- **Case OR**: Open loop ensemble forecast in the regular domain; this is a free running ensemble forecast with the same initial conditions as cases SOR and SPR but without assimilation of observations.
- **Case OL**: Open loop ensemble forecast for the long domain.

The positions of the observations are the same for both the long and regular domains. This corresponds to observations spread throughout the whole of the regular domain, and observations only in the upstream part of the long domain. The observation errors are the same for the four different configurations that use observations.

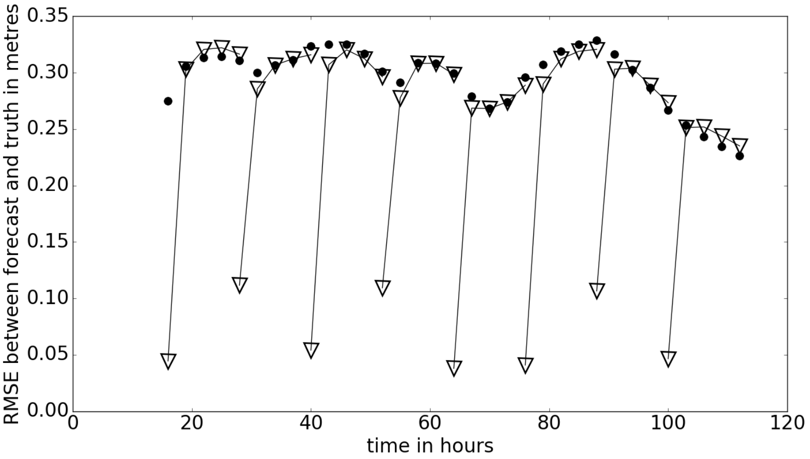
4.5 Results and discussion of assimilation

4.5.1 State only estimation (SOR and SOL)

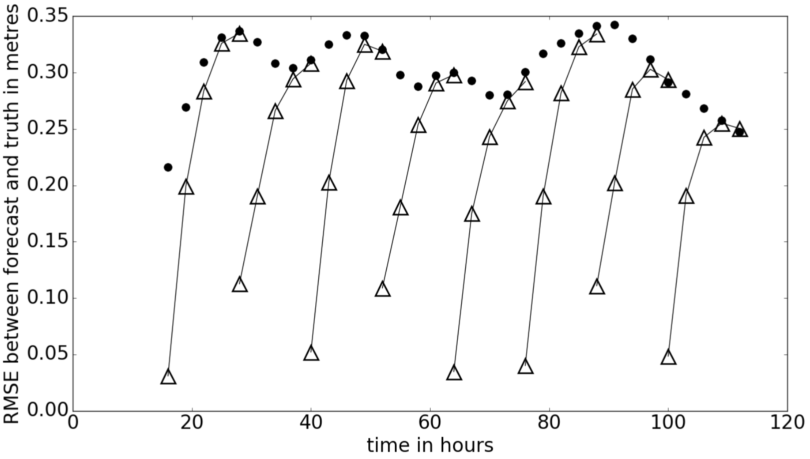
Results from state-only assimilation experiments are shown in figure 4.5. The graphs shows the RMSE over the whole domain between the forecast ensemble mean water depths and the true water depths at three hourly intervals from the time of the first assimilation at 16h. Here, RMSE is defined as

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\overline{(h_i^f} - h_i^t)^2}, \tag{4.26}$$

where h_i^f and h_i^t are the forecast ensemble mean and true water depth in cell i respectively. This definition means that the error is averaged over a larger area for the longer domain than for the regular domain. The values of RMSEs, though broadly similar, are therefore not directly comparable between domains.



(a) RMSE between the ensemble forecast mean and the true water depth ($\overline{\mathbf{x}^f} - \mathbf{x}^t$) in the regular domain. Triangles show results for state only estimation (SOR), with assimilation carried out at t=16h, 28h, 40h, 52h, 64h, 76h, 88h and 100h. Circles show the open loop (OR) case for the regular domain (no assimilation).

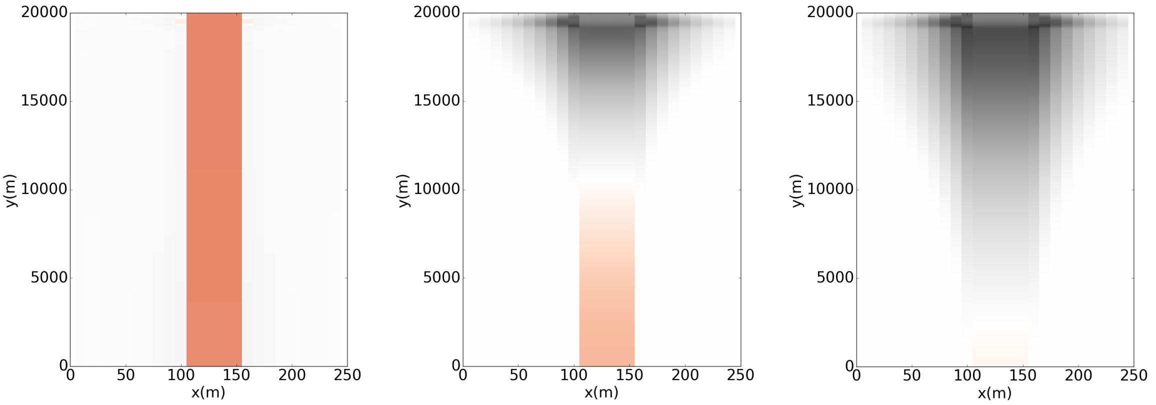


(b) RMSE between the ensemble forecast mean and the true water depth ($\overline{\mathbf{x}^f} - \mathbf{x}^t$) in the long domain. Triangles show results for state only estimation (SOL), with assimilation carried out at t=16h, 28h, 40h, 52h, 64h, 76h, 88h and 100h. Black circles show the open loop case (OL) for the long domain (no assimilation).

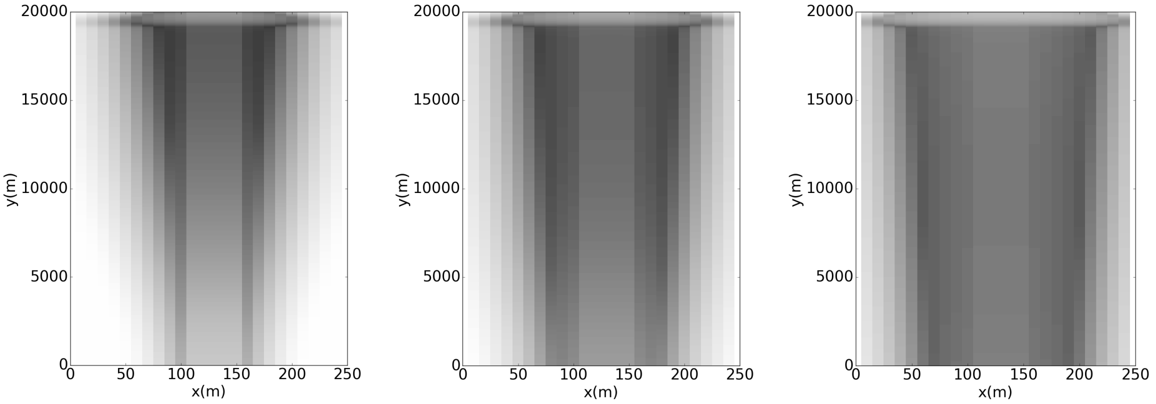
Figure 4.5: RMSEs for state-only estimation in the regular and long domains.

Figure 4.5a shows the RMSE between the forecast ensemble mean and true water depths with time for the regular domain (SOR). Figure 4.5b shows the same results for the longer domain (SOL). In both domains the ETKF produces a good analysis. The difference between the analysis ensemble mean water levels and the true water levels is very small at the time of each observation, and a large improvement is seen compared to the open loop forecast. However, for the regular domain in particular, the forecast skill is quickly lost during each subsequent forecast step and the RMSE quickly relaxes towards the open loop case. Comparing 4.5a and 4.5b shows that the RMSE at each analysis time is broadly similar for the two different domains. The results also show that the forecast skill persists for longer in the long domain; the forecast takes longer to relax to the open loop case in the longer domain than in the regular domain. This means that the same observations are having a longer-lived impact on the forecast when a longer stretch of river and floodplain is considered.

In order to further understand why the observation impact is longer-lived in the long domain, the evolution of the error during the forecast step can be investigated. Figure 4.6 shows the long domain in plan view with the error between the forecast ensemble mean and the true water levels in each cell. The errors are shown at several times during the forecast after assimilation at t = 52h and before assimilation at t = 64h. In this particular forecast period the inflow is increasing steadily, but similar patterns are seen for forecast periods in which the inflow is varying in other ways.



(a) $t = 0$ (b) $t = 1$ hours (c) $t = 2$ hours



(d) $t = 5$ hours (e) $t = 8$ hours (f) $t = 12$ hours

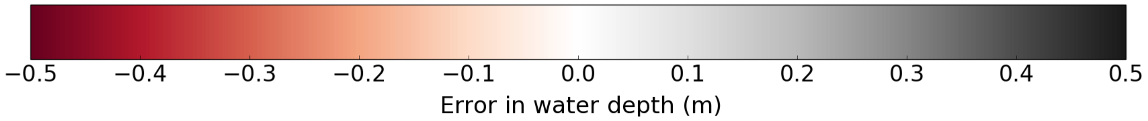


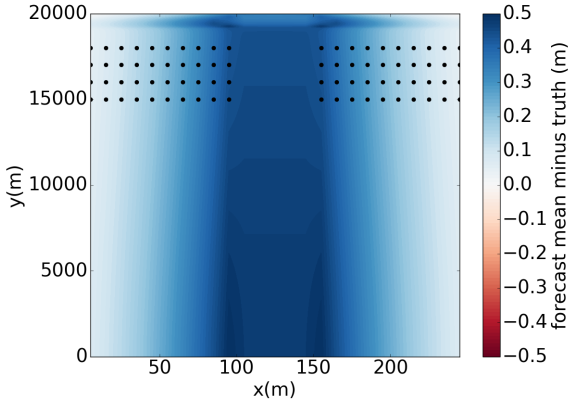
Figure 4.6: Forecast (ensemble mean) water depth minus true water depth in the long domain, shown in plan view for case SOL. Times are measured from assimilation at 52h, i.e. the forecast at 0h is the analysis calculated at 52 hours and figures 4.6c to 4.6f show results in the subsequent forecast step. For reference, the true water depth on the floodplain varied between 0 and 0.4m during this forecast period. Water depth in the centre of channel varied between 5.9m and 6.4m. Note that the pale areas close to $y = 20,000$m are due to water arriving in the domain with zero momentum in these areas.

Figure 4.6a shows the difference between the forecast ensemble mean water depths and the true water depths at the observation/assimilation time. The error between the forecast ensemble mean and true water depths is small at all points in the domain at this time. The difference between the forecast ensemble mean and the truth is highest in the channel; this reflects the fact that there are no observations of channel water depth used in the assimilation. Figure 4.6b shows the error between the forecast ensemble mean and the truth 1 hour after the assimilation. The error at the downstream end of the domain remains small, while a large, positive error can be seen in the upstream part. A positive error here means that the forecast is overestimating the water depth. In figure 4.6c, the errors at 2 hours after assimilation are shown. The physical area in which the forecast is overestimating water depths has grown in the hour between 4.6b and 4.6c, and penetrated further downstream into the domain. This pattern of error growth from upstream to downstream continues over time in figures 4.6d (5 hours after assimilation), 4.6e (8 hours after assimilation) and 4.6f (12 hours after assimilation).

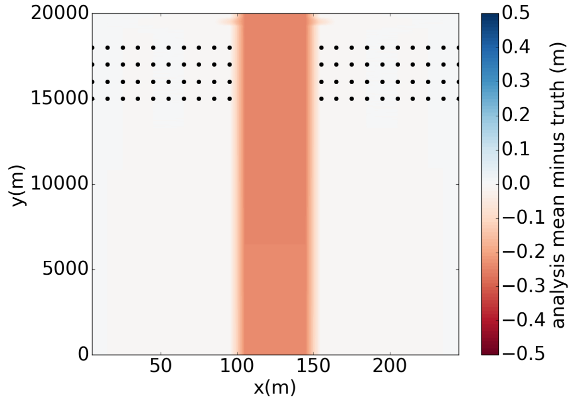
The observed pattern of error growth explains why the observations have impact in the longer domain for a longer time. The regular domain is defined by $15 \leq y \leq 20$km and $0 \leq x \leq 250$m (as shown in figure 4.3). The errors in water depth in this part of the long domain become very large within a few hours of the assimilation, as can be seen in figure 4.6c. This means that the error in the whole of the regular domain becomes large very quickly. In contrast, the RMSE for the long domain includes the downstream area where the water levels are still corrected from the assimilation; the RMSE is therefore lower for a longer time as the error takes a longer time to reach the downstream part of the longer domain.

During the forecast step, figure 4.6 shows that the error in the forecast resulting from incorrect specification of the n_{ch} parameter starts at the upstream end of the domain, and propagates downstream with time. This pattern of error growth is the same as that which would be expected from a bias in inflow, as noted in e.g. Andreadis et al. [2007]. A biased inflow acting on corrected water levels will clearly degrade the forecast water levels close to the upstream boundary first and this error will propagate downstream with the flow. The results shown in figure 4.6 therefore indicate that errors due to incorrect inflow specification and those due to incorrect specification of the channel friction parameter may be difficult to separate out in a real flood event.

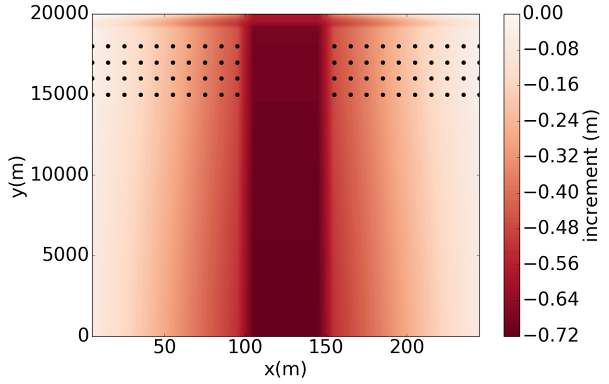
The low RMSEs between the analysis and the truth in the long domain highlight the fact that the ETKF is able to correct the water levels in areas for which there are no observations. The state error covariance matrix generated by the ensemble perturbations is such that information from the observations is spread throughout the domain, enabling corrections to be made to the state at the downstream part of the domain when observations are available only at the upstream end. This is further demonstrated in figure 4.7, which shows the difference between the forecast ensemble mean and the truth pre-assimilation (figure 4.7a) and post-assimilation (figure 4.7b), plotted in the long domain at t = 52h. Figure 4.7c shows the increments applied in the long domain as a result of observations in the upstream part of the domain only. The figure clearly shows that adjustments are made to water levels in the whole domain.



(a) Difference between the forecast ensemble mean and the truth ($\overline{\mathbf{x}^f} - \mathbf{x}^t$) plotted in the long domain (pre assimilation) at t = 52 hours.



(b) Difference between the analysis ensemble mean and the truth ($\overline{\mathbf{x}^a} - \mathbf{x}^t$) plotted in the domain (after assimilation) at t = 52 hours.



(c) Increments applied to the forecast ensemble mean for the assimilation at t = 52 hours

Figure 4.7: Difference between the forecast ensemble mean and the truth at t = 52 hours, (a) pre assimilation and (b) post assimilation; (c) shows the increment applied to the forecast to compute the analysis ($\overline{\mathbf{x}^a} - \overline{\mathbf{x}^f}$). In each plot the black circles show the position of the observations.

For a reliable ensemble, the RMSE should be close to the spread of the ensemble, where the spread is defined as the square root of the average ensemble variance (see e.g. Fortin et al. [2014]). Reliability plots (of RMSE vs. spread) should therefore produce points which lie close to the identity line. Such plots can be used to diagnose ensemble collapse, where the spread of an ensemble becomes unrealistically small.

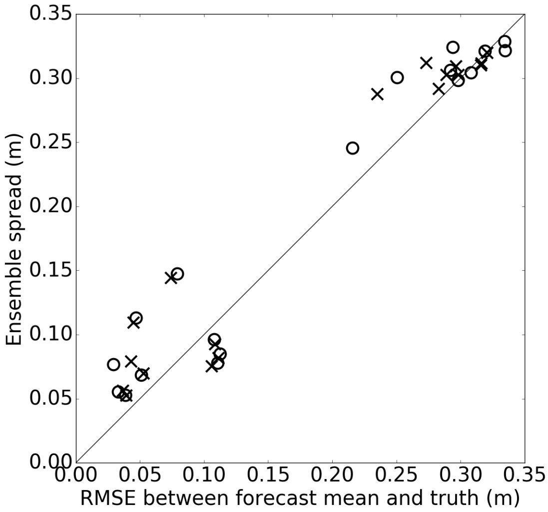


Figure 4.8: Reliability plot for the SOR and SOL experiments, showing forecast and analysis RMSE vs ensemble spread at each assimilation time. Circles are for the long domain, crosses for the regular domain. For an ideal ensemble, all points would lie on the identity line.

Figure 4.8 is a reliability plot for the SOR and SOL experiments. The points all lie close to the identity line, indicating that the ensemble spread is adequate to capture the uncertainty in the forecast. There is no indication of ensemble collapse. We note that in these experiments the time between observations is much larger than the time taken for the forecast to relax to the open loop case. If the observations were frequent enough to update the model before it relaxed back to the open loop, the risk of overfitting to the observations would be increased and in such cases inflation or localisation techniques may be required. The points in figure 4.8 form two clusters, with large error, large spread values before an assimilation and low error, small spread values for analysis ensembles.

4.5.2 State and parameter estimation (SPR and SPL)

Considering a longer stretch of river in our idealised domain, as in section 4.5.1, shows an improvement in the forecast skill of the ETKF, in that the RMSE increases more slowly. By design, much of the error between the truth and the forecast comes from the incorrect channel friction parameter in these experiments. In this section we show results from using data assimilation to jointly estimate the state and the channel friction parameter in both domains.

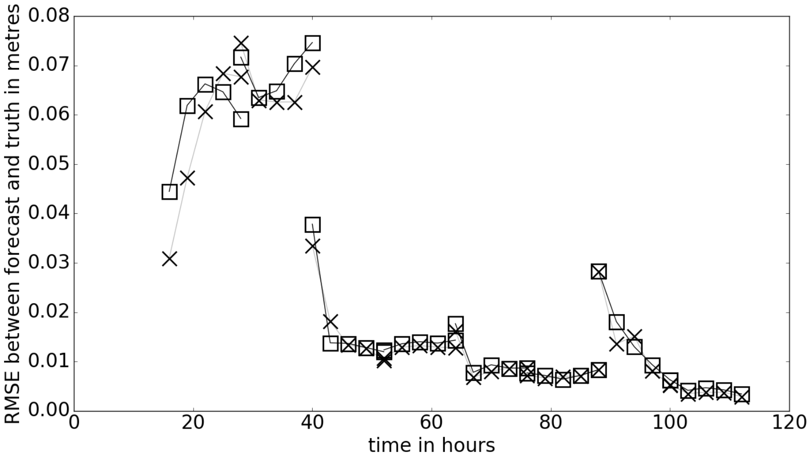
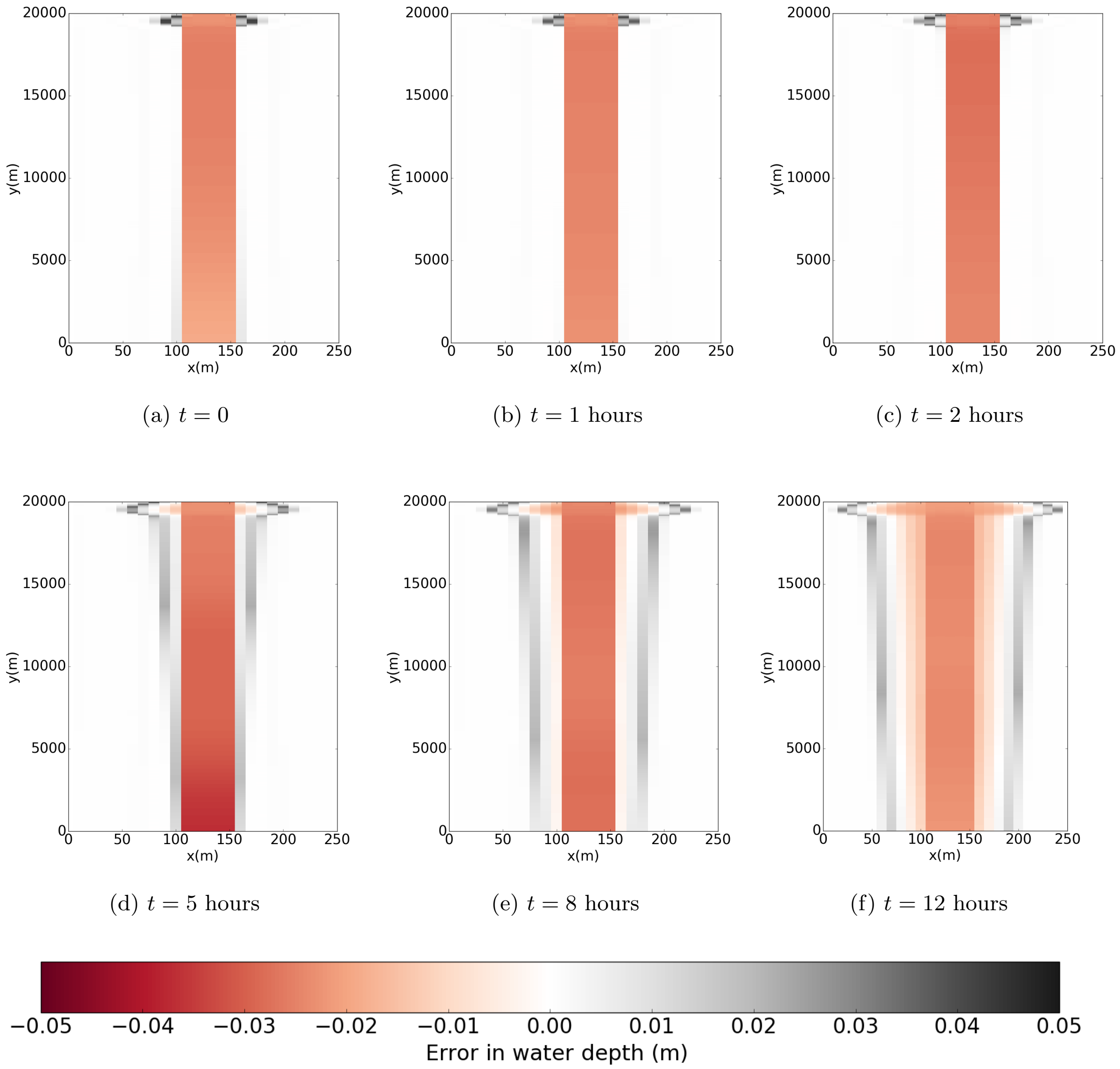


Figure 4.9: RMSE between the ensemble mean and the truth over the whole domain for joint state-parameter estimation; crosses denote RMSE for the long domain (SPL) and squares denote RMSE for the regular domain (SPR).

Figure 4.9 shows the RMSEs with time for the state-parameter estimation experiments in the regular and long domain. Comparison of figures 4.5 and 4.9 demonstrates a very clear improvement in the forecast for joint state-parameter estimation over the state only case (note the different scales on the y axes). Joint state-parameter estimation markedly increases the observation impact for exactly the same observations. The extra computational cost of estimating the friction parameter along with the state is extremely small, as it adds only a single extra component to the state vector. Joint state-parameter estimation is therefore a very efficient way of producing a much better forecast in this situation.

The small RMSE between the truth and the forecast for joint state-parameter estimation is smaller than the order of the observation error. There is no longer any significant difference between the results for the different length domains and this implies that the error growth must occur at similar rates along the length of the long domain when the friction parameter is estimated.



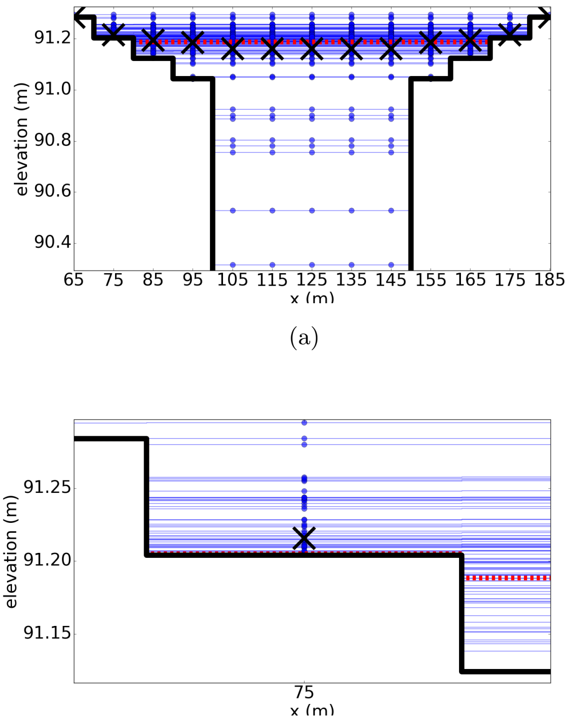
(a) $t = 0$ (b) $t = 1$ hours (c) $t = 2$ hours

(d) $t = 5$ hours (e) $t = 8$ hours (f) $t = 12$ hours

Figure 4.10: Forecast (ensemble mean) water depth minus true water depth in the long domain, shown in plan view for case SPL. Times are measured from assimilation at t=52h. For reference, the true water depth on the floodplain varied between 0 and 0.4m during this forecast period. Water depth in the centre of channel varied between 5.9m and 6.4m.

Figure 4.10 shows the the difference between the forecast ensemble mean and the true water levels, plotted in plan view in the long domain, at several times during the forecast step starting at t = 52 hours for the SPL experiment. Note that the scale in figure 4.10 is ten times smaller than in figure 4.6. Figure 4.10a shows the difference between the forecast ensemble mean water depths and the true water depths at the observation/assimilation time. As in the SOL experiment, the error between the forecast ensemble mean and true water depths is small at all points in the domain at this time. Figure 4.10b shows the error between the forecast and the truth 1 hour after the assimilation; figures 4.10c (2 hours after assimilation), 4.10d (5 hours after assimilation), 4.10e (8 hours after assimilation) and 4.10f (12 hours after assimilation) show how the error evolves. Unlike in the SOL experiment, the error does not propagate from upstream to downstream; instead, the error grows at a similar rate along the length of the domain. This further suggests that the upstream-downstream error growth seen in the SOL experiments is due to incorrect friction parameter specification.

Figure 4.10 shows that the mean forecast overestimates water depth on the floodplain but underestimates water depth in the channel. This reflects the fact that although each ensemble member predicts a physically realistic water level that is flat in cross section, the ensemble mean forecast is not necessarily flat and therefore not physically realistic. This is because the forecast mean water depth in each cell is the average value predicted by the ensemble members. An example of this situation is shown in figure 4.11.



(a)

(b)

Figure 4.11: Cross section of the domain showing bathymetry as a solid black line. The true water level is shown as a dashed red line, water levels predicted by each ensemble member are shown as blue circles. The mean forecast in each model cell is shown as a cross. Figure 4.11a shows the central part of the domain from $65 \leq x \geq 185$m. Figure 4.11b shows the forecast water levels and resulting forecast mean in the cell centred at 75m in greater detail.

Figure 4.11 shows the true flood level at one cross section of the domain and the water level predicted by each ensemble member. The crosses show the mean ensemble water level in each cell in the cross section. In the channel (e.g. at $x = 125$m), figure 4.11a shows the ensemble predictions are such that the ensemble mean is slightly lower than the true water level. Beyond the edge of the true flood on the flood plain, the true water depth relative to the topography is zero, and most ensemble members also predict zero water depth. However, as shown in figure 4.11b for the cell at $x = 75$m, there are cells beyond the flood edge in which some ensemble members predict non-zero water depth and the ensemble mean is therefore a very small positive water depth; this simply follows from equation 4.9. It is therefore possible for the ensemble mean to predict water levels deeper than the truth in cells beyond the true flood edge where there are a number of ensemble members which predict non-zero depth, even when the water level is under-predicted in the channel. It should be noted that the errors are very small.

We note that updated ensemble water levels used at restart times may also be non-physical (not flat). We did not encounter any problems related to this in our experiments but such water levels could potentially cause some initialisation shock behaviour as small amounts of water run off the flood plain into the river.

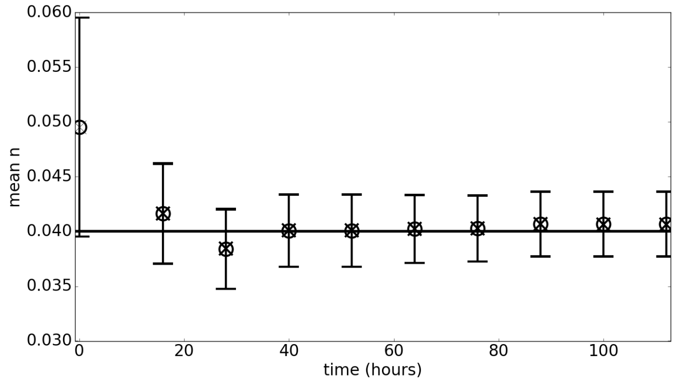
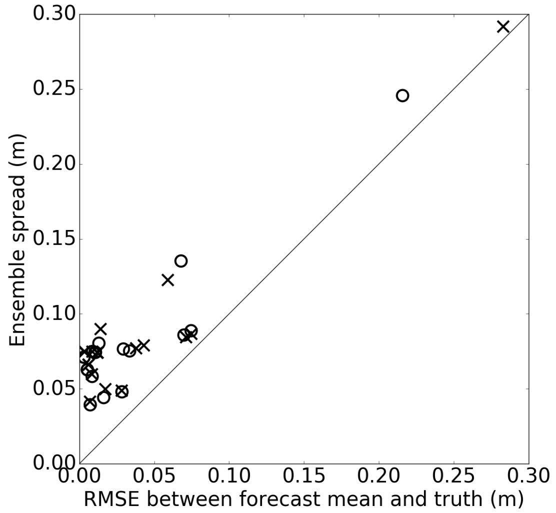
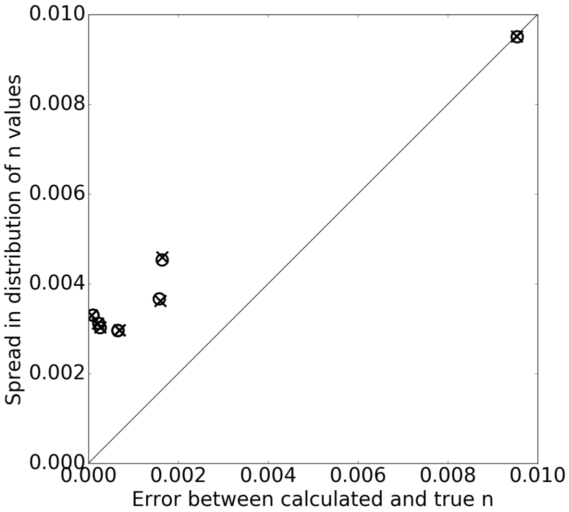


Figure 4.12: Calculated (analysis) mean channel friction parameter values at each assimilation time. The true value is shown by the horizontal line. Open circles show the values for the long domain, crosses for the regular domain. The error bars show one standard deviation of the analysis parameter distribution for the long domain; values for the regular domain are very similar.

Figure 4.12 shows the analysis ensemble mean value of the channel friction parameter at each assimilation time. The true value of the channel friction parameter, n_{ch}, indicated by the solid horizontal line, and the incorrect initial mean value is shown at time zero. The error bars show one standard deviation of the analysis n_{ch} distribution. The results show that the joint-state parameter data estimation produces a good estimate of the value of n_{ch}, and that the ensemble mean values are almost identical for both the regular and long domains. It is notable that the convergence of the estimated channel friction parameter value to the truth is achieved with water depth observations only taken on the floodplain.



(a)



(b)

Figure 4.13: Reliability plot for the SPR and SPL experiments. Circles are for the long domain, crosses for the regular domain. Plots show error vs spread for (a) the ensemble of water levels and (b) the n_{ch} distribution.

Reliability plots for the ensemble of state vectors are given in figure 4.13a and for the calculated parameter distributions in figure 4.13b. The plots show no evidence of ensemble collapse, which can be a problem in ensemble data assimilation schemes. In fact the spread of the state and parameter distributions remains broader than the size of the error in our experiments, which minimises the risk of overfitting the value of n_{ch} at a particular time in the simulation. If the spread of either the parameter or the state ensembles became too small it would be necessary to use inflation techniques (see e.g. Anderson [2007]).

We note that the joint state-parameter ensembles are over-dispersive; the state parameter spread relative to the error is higher than for the state-only experiments. This may be related to the fact that we see smaller errors between the forecast and the truth in the joint state parameter experiments.

These results show a clear advantage in jointly estimating the channel friction parameter alongside the model state; this contrasts with the findings of Garcia-Pintado et al. [2015], where no clear improvement in forecast water levels was seen. There are a number of possible reasons for this, one being that in Garcia-Pintado et al. [2015] convergence of the channel friction parameter value to a steady value was slow compared to the timescales of the flood event. The difference between our results and those in Garcia-Pintado et al. [2015] may also be related to the fact that there are more sources of uncertainty in Garcia-Pintado et al. [2015], which used real data and real topography, rather than the idealised situation in this study. One significant source of time-varying error not accounted for in the setup used by Garcia-Pintado et al. [2015] and not present in our experiments is lateral inflows (see Bermudez et al. [2017]). It may also be that the initial parameter value used in Garcia-Pintado et al. [2015] was already close to the true value, such that the error in the parameter was not a large source of error, whereas our initial guess was incorrect by design. A better initial guess would also explain the smaller changes to n_{ch} produced by the data assimilation in Garcia-Pintado et al. [2015]. Alternatively, it may be that the presence of the initialisation shock in Garcia-Pintado et al. (2015) prevented convergence to a more accurate channel friction parameter.

Another possible reason for the contrasting importance to the forecast of updating the channel friction parameter may be that in Garcia-Pintado et al. [2015] a bias correction was also made to the inflow as part of the data assimilation. It may be that inflow correction was also indirectly including some of the effects of the incorrect n_{ch} parameter. This seems likely, since the upstream-downstream error propagation pattern seen in figure 4.6 would also be expected for an incorrectly specified inflow. Thus it may be that no added benefit was gained from including n_{ch} estimation in addition to inflow bias estimation because correcting error in inflow bias also compensates for any error in n_{ch}. To test the interdependence of inflow error and error in n_{ch} we conducted some state-parameter estimation experiments exactly as for the SPL experiments but with biased inflows. No correction was made to the biased inflow, but the value of n_{ch} was updated at each assimilation time.

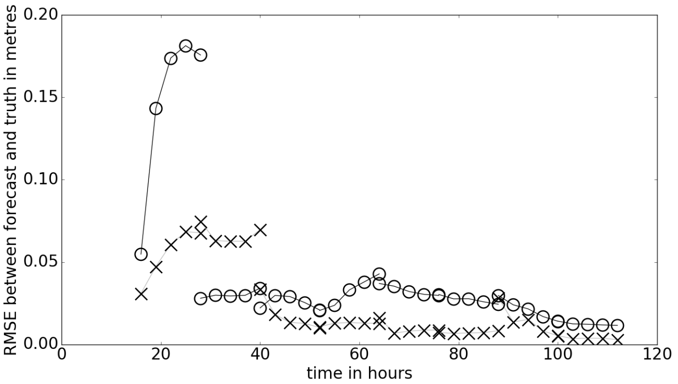


Figure 4.14: RMSEs between the mean forecast and the truth for positive inflow bias in the long domain shown as open circles. The corresponding RMSEs for the unbiased inflow are shown as crosses for comparison.

Figure 4.14 shows water level RMSEs with time for the long domain with biased and unbiased inflows, and demonstrates that correction of the channel friction parameter allows the forecast to predict accurate water levels, even with a biased inflow. The correction to the value of n_{ch} is therefore compensating for inflow bias.

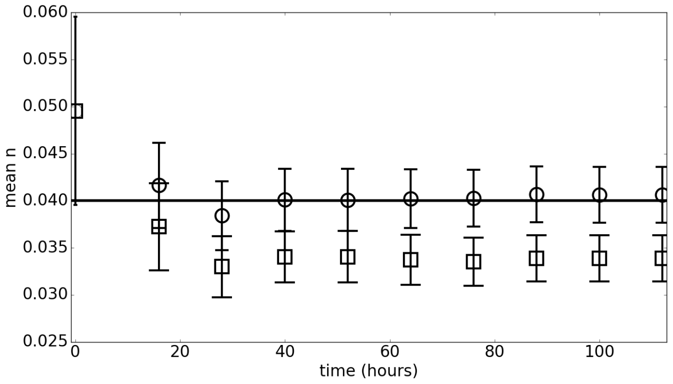


Figure 4.15: Calculated values of mean n_{ch} for biased inflow in the long domain shown as squares. Circles show the corresponding values for unbiased inflow. The horizontal line shows the true value and the error bars show one standard deviation of the calculated n_{ch} distribution.

The updated values of n_{ch} with time for biased inflow are shown in figure 4.15. When the inflow is biased, the value of n_{ch} does not converge to the true value, but instead to a value which allows it to compensate for inflow bias and minimise errors in water levels in the domain. This clearly indicates an interdependence between errors in channel friction parameter value and inflow.

Whatever the reason for the differences to the results in Garcia-Pintado et al. [2015], the work presented here shows that our inundation forecast model is sensitive to the channel friction parameter, and that the correct value of the parameter can be retrieved from observations of water depth by using joint state-parameter data assimilation during a flood event with unbiased inflow information.

4.6 Conclusions

In this study, we have investigated the effect of domain length and channel friction parameter estimation in data assimilation for flood inundation forecasting. We have also demonstrated that assuming water has zero momentum at the start of each forecast step can cause an initialisation shock. The period of time for which this shock then causes problems for the forecast is likely to be domain dependent; in this study we found that the impact of the shock disappeared within a period of approximately 4 hours. We developed a novel method to reinitialise water velocities in each model grid cell with pre-assimilation values, and showed that this approach eliminated the initialisation shock.

In agreement with Andreadis et al. [2007], Neal et al. [2009], Garcia-Pintado et al. [2013] and Matgen et al. [2010], we found that assimilating SAR-like water levels in a state-only data assimilation system gives a time limited improvement in the forecast skill, since such improvement can only persist for as long as the information is relevant in the domain. We have shown that considering a longer domain extends the time over which observations have an impact on the forecast, even when no extra observations are used. This is because the ETKF is able to correct water levels downstream of the observations due to strong covariances between the errors in water levels in different parts of the domain. In a domain with more realistic topography, the correlations between the errors in water depth in different parts of the domain may not be as strong, and are likely to be more complex. However, work by authors such as Garcia-Pintado et al. [2015] indicates that water level errors in large, real domains are correlated, as depths can be corrected at considerable distances from SAR-derived observations. Additionally, understanding the effects of domain length and friction parameter estimation without the extra complications of topographical features is important in understanding the fundamental sensitivities of such systems.

We have shown that in the forecast period following an assimilation, the difference between the forecast and the truth when the channel friction parameter is incorrectly specified and not updated grows faster in the upstream part of the domain, and then propagates downstream. This error propagation is the reason that the forecast ensemble retains skill for a longer time in the long domain; the errors grow more slowly in the downstream areas which are not considered in the regular domain.

Jointly estimating the channel friction parameter along with the water levels is shown to produce a significantly better forecast for the same observations at very little extra cost. This was not seen in Garcia-Pintado et al. [2015], in which inflow and parameters were estimated simultaneously. Correcting the channel friction parameter also eliminates the differences in forecast error growth for the two different domain lengths, as errors grow at similar rates in the upstream and downstream parts of the domain. We have shown that it is possible to estimate a good value for the channel friction parameter, even when water level observations are available only on the floodplain.

In summary, we have shown that in the case where there is no inflow bias but channel friction is incorrectly specified, assimilating SAR-like water levels from floodplains provides a time limited improvement in the forecast when only water levels are corrected. The time over which the forecast is improved depends on the length of the domain of interest. Authors such as Andreadis et al. [2007], Matgen et al. [2010], Giustarini et al. [2011], Garcia-Pintado et al. [2015, 2013] have shown that inflow correction can lead to a marked improvement in forecast skill. We have shown that jointly estimating the channel friction parameter along with the water levels also provides a clear improvement in the forecast at all times and can retrieve an accurate channel friction parameter value. Our results suggest that it may be difficult to separate out errors due to incorrect specification of inflow and incorrect specification of channel friction when carrying out ensemble data assimilation for inundation modelling. This is because the character of the errors in the forecast resulting from these two sources of uncertainty are similar; this interdependence explains our finding that updating the value of the channel friction parameter can compensate for a bias in inflow. Further study is required to see how well each of these conclusions are applicable to more complex and realistic topography, and for real satellite derived observations. In this way, the work here may serve to enhance operational flood forecasting potential.

4.7 Acknowledgements

The authors gratefully acknowledge the NERC studentship supporting Elizabeth Cooper, and CASE sponsorship from the Satellite Applications Catapult. This work was also supported in part by NERC grants NE/K00896X/1 and NE/K008900/1 as well as EPSRC grant EP/P002331/1 and the NERC National Centre for Earth Observation (NCEO). The software used in this article can be obtained from the corresponding author on request, subject to licensing conditions.

Appendices

4.8 Inflow and friction source terms

In order to model fluvial flooding in a stretch of river, the amount of water flowing into that stretch needs to be modelled. Here we describe the implementation of a new inflow source term in Clawpack. Clawpack is able to solve systems of partial differential equations with or without source terms and is designed such that the user can introduce new code to describe additional source terms. Our new inflow source term has been implemented in the code in a similar way to the pre-existing friction source term. The code for the friction source term is outlined in section 4.8.1 and the new inflow source term is described in section 4.8.2.

4.8.1 Pre-existing friction source term

Friction between the fluid and the channel in which it is flowing acts as a momentum source in the shallow water equations. This is represented in Clawpack with a source term of the form

$$\mathbf{R(q)}_{friction} = \begin{bmatrix} 0 \\ -\gamma(hu) \\ -\gamma(hv) \end{bmatrix}, \tag{4.8.1}$$

where γ is given by

$$\gamma = \frac{gn^2\sqrt{(hu)^2 + (hv)^2}}{h^{\frac{7}{3}}}, \tag{4.8.2}$$

h is water depth and u and v are velocities in the x and y directions. Acceleration due to gravity is denoted g and n is Manning's friction coefficient. This coefficient describes the roughness of the channel bed in which the water is flowing and in practical applications its value is usually determined empirically. The value of n is specified by the user in the Clawpack code, and can vary over the domain if specified in the simulation setup. The units of n are $sm^{-1/3}$.

In Clawpack, inhomogeneous sets of equations are solved using the method of fractional stepping described in LeVeque [2002] p.380-395. This method splits the equation into two simpler problems; one homogeneous conservation law and one inhomogeneous partial differential equation - which can be solved independently over the same time step. The solutions are then combined in an alternating fashion to give a solution to the whole problem. For a friction source term, the set of problems to be solved are the homogeneous system

$$\frac{\partial \mathbf{q}}{\partial t} + \frac{\partial \mathbf{F(q)}}{\partial x} + \frac{\partial \mathbf{G(q)}}{\partial y} = 0, \tag{4.8.3}$$

and the source term equation

$$\frac{\partial \mathbf{q}}{\partial t} = \mathbf{R(q)}_{friction}, \tag{4.8.4}$$

with $\mathbf{R(q)}_{friction}$ as in 4.8.1 and $\mathbf{q}$ a vector of conserved quantities (see equation (4.2)). For each time step in the code, equation (4.8.3) is advanced from t_i to t_{i+1} to give intermediate values of h_*, hu_* and hv_* in $\mathbf{q}_*$. The values in $\mathbf{q}_*$ are then used to solve equation (4.8.4) over the same time step. This introduces a 'splitting error' into the solution of order Δt $= t_{i+1}$-t_i, making the whole method only first order accurate. A more accurate splitting method ('Strang splitting') is available for implementation in the code, but the first order accuracy has been found to be more stable and sufficient in practice for a wide range of applications [LeVeque, 2002, 1997].

4.8.2 Inflow source term

For inundation simulations, water entering the domain of interest can be modelled as a source term. Many existing hydrodynamic models have such functionality; we have added an inflow source term to the Clawpack code in order to use it to model river-like flow. In operational situations, information regarding this source term may be available from an upstream gauge as a mass flow rate, Q, measured in m^3s^{-1}. In an ungauged catchment, the same information could be generated using a rainfall run-off model. Here we use hourly values of Q based on gauge data; linear interpolation is carried out between the hourly values. A water mass flow rate can be turned into a source term, S expressed in ms^{-1} (c.f. the term 'Sce' in Hervouet [2007] p.31, which has the same units and can include rainfall, infiltration etc) as long as the area of the domain or 'footprint' over which the water is added is known. For water added over an area A, $S = Q/A$. The equation for the inflow source term is then given by

$$\frac{\partial \mathbf{q}}{\partial t} = \mathbf{R}(\mathbf{q})_{inflow}, \tag{4.8.5}$$

where

$$\mathbf{R}(\mathbf{q})_{inflow} = \begin{bmatrix} S \\ 0 \\ 0 \end{bmatrix}. \tag{4.8.6}$$

Equations 4.8.5 and 4.8.6 show that for each time step Δt, the change in h due to the incoming water will depend on the value of the inflow source, S, over the same Δt. The extra water arriving in the domain creates extra water height, and is assumed here to arrive without any momentum; the water is subject only to hydrostatic momentum effects. This inflow source term has been implemented in the code in the following way

- determine in which grid cells the source term will be applied. This is reasonably arbitrary but must be such that the solution remains stable (we used trial and error in this experiment);
- calculate the total area that the inflow cells cover in the domain, A;
- for a given mass flow rate Q, calculate S for each value of Q by dividing by A;
- at the relevant grid points extract depth , h_*, as calculated from equation (4.8.3);
- calculate the change to h_* due to incoming water from a discretisation of equation (4.8.6) using a Crank-Nicholson scheme [Crank and Nicolson, 1996]

$$h = h_* + \Delta t \frac{S(t) + S(t + \Delta t)}{2}; \tag{4.8.7}$$

- use the new value of h from equation (4.8.7) to solve for the next time step.

4.8.3 Combining friction and inflow source terms

The source terms described in this section are applied in a sequential manner in the code. For each time step, the inflow source term calculates the new water depths in the relevant parts of the domain and then the friction source term is applied to the new water depths.

4.9 Chapter summary

In this chapter we have shown that in our simplified domain we were able to retrieve the correct channel friction parameter using SAR-like observations and an ETKF. We showed that correcting the value of the parameter at the same time as water levels led to a significant increase in the observation impact, i.e. an increase in the time over which the observations influenced the forecast. We showed that error in the channel friction parameter can compensate for error in inflows, making the two sources of error difficult to distinguish. We also found that assuming zero momentum when restarting our simulations after an assimilation caused an initialisation shock in the solution, and devised and implemented a novel method for dealing with this.

Chapter 5:

Observation operators for inundation forecasting - theory and idealised experiments.

This chapter is concerned with the second question from from chapter 1, which asks whether we can devise a new way to extract observational information from a SAR image. We would like to know the following:

- Can we find a new way to extract observational information from a SAR image, and can we design and implement a new observation operator in order to in order to use these observations in data assimilation?
- What are the physical mechanisms by which observation operators (including our new operator) update water levels?
- Does our new observation operator work to improve the forecast in synthetic experiments?

This work has been published as Cooper et al. [2018a]. The remainder of this chapter, except for the chapter summary (section 5.9), is reproduced from the revised manuscript.

5.1 Abstract

Images from satellite-based synthetic aperture radar (SAR) instruments contain large amounts of information about the position of flood water during a river flood event. This observational information typically covers a large spatial area, but is only relevant

for a short time if water levels are changing rapidly. Data assimilation allows us to combine valuable SAR-derived observed information with continuous predictions from a computational hydrodynamic model and thus to produce a better forecast than using the model alone. In order to use observations in this way a suitable observation operator is required. In this paper we show that different types of observation operator can produce very different corrections to predicted water levels; this impacts on the quality of the forecast produced. We discuss the physical mechanisms by which different observation operators update modelled water levels and introduce a novel observation operator for inundation forecasting. The performance of the new operator is compared in synthetic experiments with that of two more conventional approaches. The conventional approaches both use observations of water levels derived from SAR to correct model predictions. Our new operator is instead designed to use backscatter values from SAR instruments as observations; such an approach has not been used before in an ensemble Kalman filtering framework. Direct use of backscatter observations opens up the possibility of using more information from each SAR image and could potentially speed up the time taken to produce observations needed to update model predictions. We compare the strengths and weaknesses of the three different approaches with reference to the physical mechanisms by which each of the observation operators allow data assimilation to update water levels in synthetic twin experiments in an idealised domain.

5.2 Introduction

During a fluvial flood it is possible to use a numerical hydrodynamic model to predict future water levels and flood extents. Such predictions are subject to uncertainties and can be inaccurate; data assimilation can therefore be used to improve predictions by updating model forecasts based on various types of observations (e.g. Lai and Monnier [2009], Matgen et al. [2007a], Garcia-Pintado et al. [2013], Garcia-Pintado et al. [2015], Ricci et al. [2011], Barthlmy et al. [2016], Schumann et al. [2009], Oubanas [2018], Oubanas et al. [2018a], and Oubanas et al. [2018b]). For flooding, useful observations of river flow rate or water depth could come from river gauges. However the number of gauges is declining worldwide (Vrsmarty et al. [2001]) and a method that can work in ungauged catchments is therefore desirable. For this reason satellite images, and in particular synthetic aperture radar (SAR) images of flooded areas can be a good source of information (Grimaldi et al. [2016]).

SAR sensors are active, side-looking sensors included on several satellites, e.g. CosmoSkymed and Sentinel 1. Radiation (of wavelength *cm* to *m*) is emitted from the satellite and directed towards the surface of the Earth. The returning signal is recorded at a sensor and can be used to reconstruct information about the observed terrain. SAR radiation is cloud penetrating, giving the instruments all-weather capability. SAR instruments can also produce observations day and night, unlike passive sensors that rely on solar radiation.

The strength of the returned signal measured at the SAR sensor depends strongly on the roughness properties of the surface from which it has been reflected. During a flood event SAR images therefore generally show a clear difference between flooded and non-flooded areas. Pixels in flooded or other wet areas such as lakes and rivers have low backscatter values and appear as dark areas on SAR images; dry areas have higher backscatter values and dry pixels therefore appear paler. There are a number of techniques for separating pixels into wet and dry areas based on backscatter. Methods include thresholding (e.g. Henry et al. [2006]) with varying levels of user-interpretation (as compared in Brown et al. [2016]), region growing/clustering ('snakes') (e.g. Horritt et al. [2001]) and change detection (e.g. Hostache et al. [2012]). These techniques can be used to provide observational information for data assimilation frameworks, but are also used for flood mapping and monitoring (as in e.g. Brown et al. [2016], Matgen et al. [2011]) and for validation and calibration of inundation models (e.g. Mason et al. [2009], Baldassarre et al. [2009], Wood et al. [2016]). In the case of model calibration, Mason et al. [2009] and Stephens et al. [2013] suggest that comparing modelled and observed derived water level measures from SAR images results in better calibration than when using binary wet-dry comparisons. However, it is not clear whether derived water levels provide better observation impact than wet/dry observations in data assimilation.

In this work we consider different ways in which information from a SAR image can be used to correct inundation forecasts using data assimilation. The use of observations requires two steps. First, we must extract relevant, useable information from a SAR image. This involves processing the raw SAR data in some way to produce an observation, or set of observations, per image. In the second step we need to use an observation operator to map our model state vector into observation space - i.e. we extract the equivalent information from our model in order to compare it to the observations. The size of the difference between the observation and the equivalent information from the model forecast is then used to calculate an update or correction to the forecast. The observation operator depends on the type of observational information used and we show in this paper that the impact of observations on the forecast can be strongly dependent on the observation operator approach used. Despite this, the mechanisms through which different observation types and different observation operators update hydrodynamic forecasts have not received much attention in the literature.

In order to extract observational information from a SAR image, authors such as Mason et al. [2012], Giustarini et al. [2011], Neal et al. [2009] and Matgen et al. [2007a] have used an approach which relies on identifying the flood edge. Terrain information, e.g. from a digital terrain model, is then used to infer information about water levels on the floodplain. Water level observations (WLOs) can then be compared with model forecast water levels. Examples of two observation operators using flood edge WLOs are described further in section 5.4. A different type of observation is used for data assimilation in Wood [2016] and Hostache et al. [2018], in which flood probability maps are produced from SAR images using the method in Giustarini et al. [2016]. Particle filter data assimilation techniques are then used to update a hydrodynamic model using flood probability maps as observations.

We propose a new type of observation operator which directly uses pixel-by-pixel backscatter values as observations. As in Wood [2016] and Giustarini et al. [2016], we rely on the fact that SAR images yield distinct distributions of wet and dry backscatter values. However, our method employs an ensemble transform Kalman filter (ETKF) approach with a novel observation operator; we directly use measured SAR backscatter values as observations rather than derived flood probability measures.

In this paper we examine the performance of our new observation operator and that of two flood-edge observation operators in a series of synthetic experiments. We compare the physical mechanisms by which the different approaches update predicted water levels in the ETKF; to the authors' knowledge these physical mechanisms have not been discussed in the literature before. We outline the ETKF data assimilation algorithm in section 5.3 and in section 5.4 we describe the three observation operators which we have compared. Further details of our experiments are given in section 5.5, including an outline of the hydrodynamic model. In section 5.6 we demonstrate how well the assimilation can update model forecast water levels towards the truth with each observation operator and explore the different physical mechanisms by which updates are made. We also test the ability of the three operators to successfully update the model channel friction parameter through an augmented state vector approach. We find that our new backscatter operator generates better corrections to the state and parameter values than the simplest approach to assimilating flood edge observations, but does not always perform as well as the 'nearest wet pixel' approach. In section 5.7 we discuss issues relating to the application of our new observation operator to real data. In section 5.8 we conclude with a comparison of the relative strengths and weaknesses of the three different observation operators.

5.3 Data assimilation

In this paper we explore the use of observations from SAR images in updating forecasts from a hydrodynamic flood model. In section 5.3.1 we outline the ETKF data assimilation framework we use in our experiments (Bishop et al. [2001]), following the approach of Garcia-Pintado et al. [2013], Garcia-Pintado et al. [2015] and Cooper et al. [2018b]. In section 5.3.2 we describe the joint state-parameter estimation method we use to update the channel friction parameter value.

5.3.1 Ensemble transform Kalman filter (ETKF)

In data assimilation, forecasts from a numerical model are combined with observations of the same system. We use a state vector, $\mathbf{x}(t_k) \in \mathbb{R}^N$ to represent the state of the dynamical system at time t_k. Here, our model domain is split into N computational cells and the state vector contains N water depths at a given time. In this paper we use an ensemble of state vectors, where each state vector in the ensemble represents a possible state of the system. For an ensemble made up of M state vectors (members), $\mathbf{x}_i$, $(i = 1, 2, ..., M)$ the best estimate of the true state of the system is represented by the mean state, $\overline{\mathbf{x}}$, where

$$\overline{\mathbf{x}} = \frac{1}{M}\sum_{i=1}^{M}\mathbf{x}_i. \tag{5.3.1}$$

We can define a perturbation matrix, $\mathbf{X} \in \mathbb{R}^{N\times M}$, that can be used to derive a measure of uncertainty in the mean state. The perturbation matrix is

$$\mathbf{X} = \frac{1}{\sqrt{M-1}}(\mathbf{x}_1 - \overline{\mathbf{x}} \;\; \mathbf{x}_2 - \overline{\mathbf{x}} \;\; ...... \;\; \mathbf{x}_M - \overline{\mathbf{x}}). \tag{5.3.2}$$

We can then express the ensemble error covariance matrix, $\mathbf{P} \in \mathbb{R}^{N\times N}$ as

$$\mathbf{P} = \mathbf{X}(\mathbf{X})^T. \tag{5.3.3}$$

The ETKF is a two-step sequential data assimilation technique. In the forecast step, each vector $\mathbf{x}_i$, is evolved in time using the forecast equation,

$$\mathbf{x}_i(t_{k+1}) = \mathfrak{M}(\mathbf{x}_i(t_k)), \tag{5.3.4}$$

where $\mathfrak{M}$ is the forecast model. Here, $\mathfrak{M}$ is a hydrodynamic model built using Clawpack code (see section 5.5.1); the model evolves water levels in each ensemble member with time, generating an ensemble of flood forecast realisations.

In the update step the mean state vector and the error covariance matrix are both updated based on observational information. We use the ETKF in its standard application as a sequential filter. As such we perform an update step at the time of each available observation. We assume that the observations are related to the true state of the system, $\mathbf{x}^t$ according to

$$\mathbf{y}_{obs} = \mathbf{h}(\mathbf{x}^t) + \epsilon, \tag{5.3.5}$$

where the vector $\mathbf{y}_{obs} \in \mathbb{R}^p$ contains p observations. The vector ϵ represents observation error, which we assumed to be unbiased and stochastic with covariance $\mathbf{R} \in \mathbb{R}^{p \times p}$.The observation operator, $\mathbf{h} : \mathbb{R}^N \rightarrow \mathbb{R}^p$ maps the state vector into observation space. If the state vector and the observation vector contain the same quantity (e.g. water depth) then the observation operator is generally just required to pick out the values in the state vector corresponding to the spatial position of the observation(s); this may involve spatial interpolation if observations are not located at model grid points. However, it is commonly the case that observations are different quantities to those in the state vector and the observation operator therefore contains information about how the observed and state vector quantities are related as well as positional information. Different observation types (e.g. water elevation or wet/dry pixel information) therefore require different observation operators for the same model (i.e. for the same state vector).

In order to update the model forecast it is useful to create a forecast-observation ensemble, which contains M forecast-observation vectors, $\mathbf{y}_i^f$, $(i = 1, 2...M)$ such that

$$\mathbf{y}_i^f = \mathbf{h}(\mathbf{x}_i^f). \tag{5.3.6}$$

Equation (5.3.6) shows that the observation operator, $\mathbf{h}$, is applied to each state vector in order to extract observation equivalent information; each forecast-observation vector, $\mathbf{y}_i^f \in \mathbb{R}^p$ is what would be observed if the corresponding state vector, $\mathbf{x}_i^f$ represented the true state of the system. The model equivalent of the observation vector is given by the mean of the forecast-observation ensemble, $\overline{\mathbf{y}^f} \in \mathbb{R}^p$, where

$$\overline{\mathbf{y}^f} = \overline{\mathbf{h}(\mathbf{x})} = \frac{1}{M}\sum_{i=1}^{M}\mathbf{h}(\mathbf{x}_i). \tag{5.3.7}$$

Note that when the observation operator is nonlinear,

$$\overline{\mathbf{h}(\mathbf{x})} \neq \mathbf{h}(\overline{\mathbf{x}}). \tag{5.3.8}$$

We can also define a perturbation matrix $\mathbf{Y}^f \in \mathbb{R}^{p\times p}$ for the forecast-observation ensemble matrix:

$$\mathbf{Y} = \frac{1}{\sqrt{M-1}}(\mathbf{y}_1 - \overline{\mathbf{y}} \;\; \mathbf{y}_2 - \overline{\mathbf{x}} \;\; ...... \;\; \mathbf{y}_M - \overline{\mathbf{y}}). \tag{5.3.9}$$

The mean state vector and error perturbation matrix are updated separately in the ETKF. The mean state is updated according to

$$\overline{\mathbf{x}^a} = \overline{\mathbf{x}^f} + \mathbf{K}(\mathbf{y}_{obs} - \overline{\mathbf{y}^f}), \tag{5.3.10}$$

where $\overline{\mathbf{x}^a} \in \mathbb{R}^N$ and $\overline{\mathbf{x}^f} \in \mathbb{R}^N$ are the means of the analysis and forecast ensemble respectively. The ETKF uses an ensemble version of the Kalman gain, $\mathbf{K} \in \mathbb{R}^{N\times p}$ is, as defined in equation (5.3.13). The ensemble Kalman update (5.3.10) can be written in terms of the innovation, $\boldsymbol{\delta}_y$, where

$$\boldsymbol{\delta}_y = \mathbf{y}_{obs} - \overline{\mathbf{y}^f}. \tag{5.3.11}$$

The innovation is calculated in observation space. The term

$$\mathbf{K}(\boldsymbol{\delta}_y) \tag{5.3.12}$$

is known as the increment, and is the difference between $\overline{\mathbf{x}^a}$ and $\overline{\mathbf{x}^f}$. The increment is calculated in state space.

We use a square root formulation for the ETKF, following Ott et al. [2004], Livings et al. [2008] and Livings [2005]. This formulation is also used in Garcia-Pintado et al. [2013] and Cooper et al. [2018b]. In this approach the ensemble version of $\mathbf{K}$ is written as

$$\mathbf{K} = \mathbf{X}^f(\mathbf{Y}^f)^T(\mathbf{Y}^f(\mathbf{Y}^f)^T + \mathbf{R})^{-1}. \tag{5.3.13}$$

The state error perturbation matrix is updated in the ETKF according to

$$\mathbf{X}^a = \mathbf{X}^f \mathbf{T}. \tag{5.3.14}$$

The perturbation matrix is updated by the matrix $\mathbf{T} \in \mathbb{R}^{M \times M}$. We use an unbiased, symmetric square root formulation of the matrix $\mathbf{T}$, constructed in a way that ensures that the analysis state error covariance, $\mathbf{P}^a = \mathbf{X}^a(\mathbf{X}^a)^T$ is the same as the analysis error covariance calculated in the Kalman covariance update (in e.g. Kalman [1960]). The formulation makes use of a singular value decomposition (Golub and Van Loan [1996]),

$$(\mathbf{R}^{\frac{1}{2}}\mathbf{Y}^f)^T = \mathbf{U}\Sigma\mathbf{V}^T, \tag{5.3.15}$$

where $\mathbf{U} \in \mathbb{R}^{M \times M}$ and $\mathbf{V} \in \mathbb{R}^{p \times p}$ are orthogonal. The columns of $\mathbf{U}$ and $\mathbf{V}$ are the left and right singular vectors of $(\mathbf{R}^{\frac{1}{2}}\mathbf{Y}^f)^T$ respectively. The diagonal elements of the matrix $\Sigma \in \mathbb{R}^{M \times p}$ are the singular values of $(\mathbf{R}^{\frac{1}{2}}\mathbf{Y}^f)^T$. A solution for $\mathbf{T}$ is then

$$\mathbf{T} = \mathbf{U}(\mathbf{I} + \Sigma\Sigma^T)^{-\frac{1}{2}}\mathbf{U}^T, \tag{5.3.16}$$

where $\mathbf{I}$ is the identity matrix. See Livings et al. [2008], Cooper et al. [2018b] for further details of how $\mathbf{T}$ is computed.

5.3.2 Joint state-parameter estimation

State augmentation techniques can be used to correct values of uncertain forecast model parameters at the same time as the state is updated. In this approach, parameters are appended to the state vector (see Smith et al. [2013], Navon [1998], Evensen et al. [1998], Smith et al. [2009, 2011]), producing an augmented state vector, $\mathbf{x}_{aug}$:

$$\mathbf{x}_{aug} = \begin{bmatrix} \mathbf{x} \\ \mathbf{b} \end{bmatrix}, \tag{5.3.17}$$

where $\mathbf{x}_{aug} \in \mathbb{R}^{N+q}$. The vector $\mathbf{b} \in \mathbb{R}^q$ contains q parameters. In this paper only one parameter is being updated, so that $\mathbf{b}$ is scalar. The parameter we are updating in this paper is the Manning's friction coefficient in the river channel, n_{ch}, as the evolution of a flood is known to be very sensitive to this parameter.

The forecast equation for the case of an augmented state vector can be written as

$$\mathbf{x}_{aug}(t_{k+1}) = \begin{bmatrix} \mathfrak{M}(\mathbf{x}(t_k)) \\ \mathbf{b}(t_k) \end{bmatrix}. \tag{5.3.18}$$

Equation 5.3.18 shows that we assume the value of n_{ch} remains constant during the forecast step and changes only when the update equation is applied.

The augmented state vector is updated by the ETKF algorithm through equations (5.3.10) and (5.3.14). Parameter value(s) are updated according to the observations due to covariances between errors in the model state and errors in the parameter(s).

Model friction parameter values are more traditionally calculated using offline calibration techniques and data from previous flood events. Updating parameter values using a state augmentation approach has the advantage that it uses information from observations of the flood event of interest as it occurs. State augmentation can therefore take into account any recent changes to the river and its environment.

5.4 Observation operators for inundation forecasting

Much existing work on data assimilation for fluvial inundation forecasting has focussed on assimilating derived water level observations. Water level extraction is based on the fact that it is usually possible to differentiate between wet and dry areas in a SAR image; the contrast in backscatter between wet and dry pixels means that it is therefore possible to determine the position of the edge of a flooded area. Along this edge, the water elevation is the same as the elevation of the topography. This means that as long as a flood edge can be accurately identified and topographical information is available (e.g. a digital terrain model (DTM)), water levels at the flood edge can be derived from a SAR image. This approach has also been used for operational flood mapping, e.g. Brown et al. [2016]. In practise, it is not possible to accurately determine flood extents from SAR images over the whole ‘edge’ of a flooded area. This is clearly shown in Mason et al. [2012] and can lead to few, sparse observations of this type.

In the remainder of this section we describe the three different observation operators used in this study. In section 5.4.1 we describe the simplest way to use flood edge water level observations; the results in section 5.6.1.2 illustrate the problems with this approach. Section 5.4.2 gives an outline of the more sophisticated approach to using water level observations used in Garcia-Pintado et al. [2013] and Garcia-Pintado et al. [2015]. In section 5.4.3 we describe our new observation operator.

5.4.1 Observation operator $\mathbf{h}_s$: simple flood edge assimilation

In this approach, we assume $\mathbf{y}_{obs}$ comprises p water level observations at flood edge positions. The simplest way to use these observations to calculate an innovation is to extract water level information from each ensemble member at each observed flood edge location. The observation operator in this approach, $\mathbf{h}_s$, picks out water level predictions at the positions of the observed flood edges for each ensemble member. Some method of interpolation will generally be necessary in order to locate the closest cell to the measured flood edge location, but this was not needed in our identical twin experiments as the truth and forecast simulations use the same grid. The simple observation operator $\mathbf{h}_s$ in our case is therefore described by a sparse matrix, $\mathbf{H}_s$ dimension (p by N) containing one and zero values such that water elevation predictions corresponding to the positions of flood edge observations are mapped with weight one and all other values with weight equal to zero, i.e.

$$\mathbf{h}_s(\mathbf{x}_i^f) = \mathbf{H}_s \mathbf{x}_i^f. \tag{5.4.1}$$

The value of $\overline{\mathbf{y}^f}$ is then calculated according to equation (5.3.7).

This approach can lead to problems in application and is therefore not widely used, but we include it here to show the importance of how observations are used in data assimilation. The problem with this simple method is essentially that it does not use all of the available information. All ensemble members that predict shallower local water levels than the truth at the position of the observation will make the same contribution to $\overline{\mathbf{y}^f}$; they will all predict zero water depth at the flood edge position no matter how much shallower the ensemble prediction is than the truth.

5.4.2 Observation operator $\mathbf{h}_{np}$: nearest wet pixel approach

In this approach we assume again that $\mathbf{y}_{obs}$ comprises p water level observations at flood edge positions. In Garcia-Pintado et al. [2013] and Garcia-Pintado et al. [2015] the authors use flood edge water level observations with a more sophisticated observation operator, referred to here as the 'nearest wet pixel' method. The new observation operator, $\mathbf{h}_{np} \in \mathbb{R}^{p \times N}$ can be described as a sparse matrix containing values of one and zero, so that

$$\mathbf{h}_{np}(\mathbf{x}_i^f) = \mathbf{H}_{np}\mathbf{x}_i^f. \tag{5.4.2}$$

Now however, water elevation values are mapped differently. Each row of $\mathbf{h}_{np}$ contains a one at the positions corresponding to a flood edge locations observation only if the corresponding water elevation value in $\mathbf{x}_i^f$ is greater than the observed flood edge elevation. Where this is not the case, the entry in $\mathbf{x}_i^f$ corresponding to the 'nearest wet pixel' (i.e. the local flood edge position as predicted by the ith ensemble member) is instead given a weighting of one. Unlike the simple approach, this method allows information to be included from ensemble members that predict shallower water levels than the truth, since the contribution to $\overline{\mathbf{y}^f}$ will depend on the position of the flood edge predicted by each shallower ensemble member.

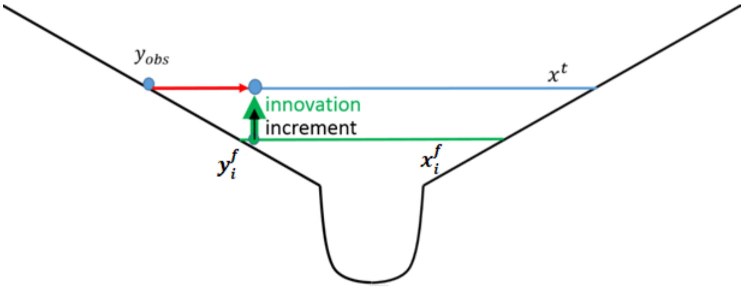


Figure 5.1: Schematic of $\mathbf{h}_{np}$ in the case of ensemble member shallower than observed.

Figure 5.1 shows how when an ensemble member predicts shallower water levels than the observation, the nearest wet pixel operator acts to effectively move the observation to the predicted flood edge, i.e. to a position at which the ensemble member predicts non-zero water depth. This allows ensemble members which are shallower than the truth to contribute information to the assimilation scheme by allowing calculation of an increment and innovation.

Finding the 'nearest wet pixel' can be difficult in practice, since is it important to find the local flood edge that corresponds to the observation. In simplified topography such as used in this study, this can be assumed to be the first wet model grid cell encountered when moving from the observation towards the centre of the river along a cross section perpendicular to the flow of the river. In situations where the topography is complex (e.g. the local direction of flow is not clear, or the river has tight meanders) finding the nearest wet pixel becomes more complicated. One approach is to require that the nearest wet pixel is in the direction of the steepest downhill descent from the observation location.

A related approach has been successfully used by Matgen et al. [2007a], Giustarini et al. [2011], Neal et al. [2009] and Matgen et al. [2010], in which it is assumed that the water level measured at a flood edge can be used to define the water level along the whole horizontal cross section of river valley perpendicular to the flow of the river. In other words, the observed water elevation at the flood edge is extrapolated across the river valley in a direction perpendicular to the flow of the river. Again, this could potentially cause problems in situations in which the local direction of flow is not clear or the river has tight meanders. There may also be problems if the observations relate to bodies of water on the floodplain that have become hydraulically separate from the river when the flood is receding; such ponding was observed in the floods of the Severn and Avon rivers near Tewkesbury, UK in 2014 (Waller et al. [2018]).

5.4.3 New observation operator, $\mathbf{h}_b$: backscatter approach

We have developed an alternative method for extracting observations from a SAR image, which directly uses SAR backscatter measurements as observations, rather than derived water elevation information. This means that the observation vector $\mathbf{y}_{obs}$ comprises p_b backscatter values at a number of selected pixel locations. The method potentially allows for more information to be used per SAR image, as information can be used from areas excluded from water elevation calculations. This could reduce the time taken to process a SAR image and produce useable observations.

The observations used in this method are measured SAR backscatter values; we follow the approach of Giustarini et al. [2016] in assuming that the backscatter values from a SAR image can be characterised as belonging to two separate probability density functions; one for wet pixels and one for dry pixels. We assume that we can create a histogram of backscatter values in the area of interest (Giustarini et al. [2016]). Two Gaussian curves are then fitted to the histogram, corresponding to the wet and dry probability density functions. These distributions represent the probability that a pixel has a particular backscatter value, given that the pixel is wet (or dry). An example of such a histogram is given in figure 5.2.

The distribution of wet pixels has a mean backscatter value m_w and variance σ_w^2. The distribution of dry pixels has mean and variance m_d and σ_d^2. Dividing the SAR image into tiles may be necessary for this to work optimally; otherwise the distribution of dry pixels is likely to dominate the histogram and make the wet pixel distribution difficult to resolve (see e.g. Chini et al. [2017]).

A new observation operator is required in order to use backscatter observations in

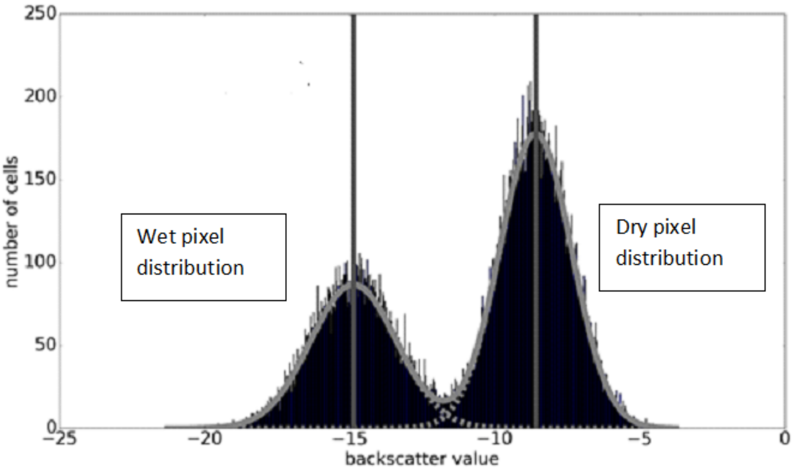


Figure 5.2: Example of backscatter histogram for flooded domain. Grey lines show fitted Gaussian distributions and black vertical lines show the mean backscatter values for each distributions

data assimilation. The operator needs to take each state vector (containing water levels in each pixel) and transform that information into model equivalent backscatter values. This could potentially be achieved using a SAR simulator to generate a synthetic SAR image, but this would be computationally expensive and would require detailed knowledge of the underlying terrain and land-use cover. Instead we take a statistical approach that makes use of the wet and dry pixel backscatter distributions obtained from a SAR image. The observation operator comprises two steps. We can describe this such that

$$\mathbf{y}_i^f = \mathbf{h}_b(\mathbf{x}_i^f) = \mathbf{h}_{b2}(\mathbf{H}_{b1}\mathbf{x}_i^f), \tag{5.4.3}$$

where $\mathbf{H}_{b1}$ is a sparse matrix, dimension $(p_b \times N)$ which extracts values corresponding to observation location positions; each row contains a 1 at positions corresponding to backscatter observation locations and all other values are zero. The non linear operator $\mathbf{h}_{b2}$ is then applied to $\mathbf{H}_{b1}\mathbf{x}_i^f \in \mathbb{R}^{p_b}$. This operation transforms each entry in the vector $\mathbf{H}_{b1}\mathbf{x}_i^f$ into m_w if water is predicted in the cell, or m_d if the cell is predicted to be dry. As for the other observation operators, interpolation will be necessary when observed backscatter cells do not correspond to the positions of model forecast information. As already mentioned, this was not necessary in our synthetic study as we used the same model to generate both the forecast values and synthetic observations; cell locations were therefore the same. The observation equivalent forecast vector is then given by

$$\overline{\mathbf{y}^f} = \frac{1}{M}\sum_{i=1}^{M} \mathbf{h}_{b2}(\mathbf{H}_{b1}\mathbf{x}_i^f). \tag{5.4.4}$$

This method potentially allows the use of more observations: in general the number of available backscatter values from a SAR image, p_b is much larger than the number of reliable flood edge observations.

A different approach to using binary-type observations in data assimilation is used by the authors of Rochoux [2014], Rochoux et al. [2014] and Rochoux et al. [2017] in an application in which the spread of wildfires is modelled. This approach uses shape recognition and front mapping; these ideas would be applicable to flood modelling but are not investigated here.

5.5 Experimental design

5.5.1 Hydrodynamic model

The inundation model used in this work is a non-linear hydrodynamic model. The model uses Clawpack code (Clawpack Development Team [2014], Mandli et al. [2016], LeVeque [2002]) to solve the two dimensional shallow water equations everywhere in the domain, in order to simulate water flowing in a channel and overtopping onto a flood plain. Clawpack solves the shallow water equations using Riemann solvers and finite volume methods, and is able to simulate the wet-dry interfaces that occur during a flood George [2008]. The software considers the domain of interest as a user-defined number of cells, N, and calculates changes in depth and velocity of the water in each cell. In our simulations the boundary condition is extrapolating (outflow) on the $y = 0$ boundary; all other boundaries are solid wall. Clawpack uses a source term in the momentum equation to model friction effects. Momentum reduction depends on a user-specified Manning's friction coefficient. Our experiments required an inflow source term to model water arriving in the river from upstream; we added this functionality to the Clawpack code, see Cooper et al. [2018b] for details. The time step for the calculations is automatically adjusted to preserve numerical stability.

5.5.2 Domain

Experiments to compare the performance of the three operators have been carried out in an idealised river valley-like domain. The use of an idealised domain is important here so that we can examine the effects of the operators under ideal conditions, without the complications of complex topography. It will also be important to understand how the operators work under real conditions, but experiments in an idealised topography are a vital first step.

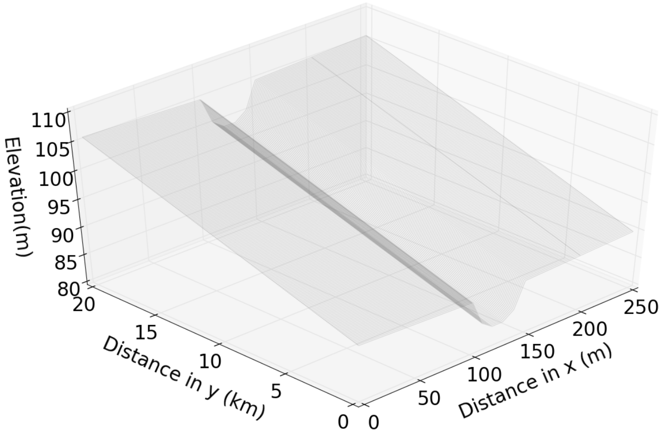


Figure 5.3: Test domain used in all assimilation experiments.

The test domain used in the experiments in this paper is the same as that used in Cooper et al. [2018b] and is shown in figure 5.3. The domain has dimensions of 20km by 250m and describes a gently sloping valley and river channel (with upstream-downstream slope of 0.08%). The domain is split into grid cells of size 10m by 10m for computation. The river channel is prescribed to be the central 5 grid cells in the x direction for all values of y and is 50m wide; the flood plain is defined as the rest of the domain. The slope of the floodplain towards the river is 0.8% based on values derived from a DTM of a stretch of the river Severn in the U.K.

5.5.3 Twin experiments

We have carried out a number of twin experiments in order to illustrate and compare how well forecasts can be corrected when using the three different observation operator approaches. The experiments use a 'truth' flood simulation and a forecast ensemble of flood realisations comprising 100 members. The forecast ensemble is updated using synthetic observations at several times during the simulation time; synthetic observations are created from the truth as described in section 5.5.4. The analysis water levels (and parameter values) can then be compared to the true water levels (and parameter values) to see how well the assimilation corrects the forecast.

In this work, the truth flood is driven by a time-varying inflow based on data taken from a gauge on the River Severn during a flood in November-December 2012. The true inflow is shown in figure 5.4; the figure also shows the inflows driving the ensemble members. All the inflows used here were also used in the experiments reported in Cooper et al. [2018b]. Inflows for each ensemble member were generated by perturbing the true inflow with additive, time correlated random errors. Time correlated errors were generated for each ensemble inflow using a first order autoregression (AR(1)) technique (Wilks [2011]) with zero mean, according to

$$\begin{aligned} e_{i,0} &= w_{i,0}, \\ e_{i,k} &= r e_{i,k-1} + (1 - r^2) w_{i,k}, \end{aligned} \tag{5.5.1}$$

where $e_{i,k}$ is the error added to the inflow at the kth timestep in the ith ensemble member. The term $w_{i,k}$ is taken from a normal distribution $\mathcal{N}(0, 0.15\times$ true inflow); i refers to ensemble member and k refers to the timestep. The autocorrelation coefficient, $r < 1$ was set to 0.997; this very high coefficient means that the errors are close to persistent in time for each ensemble member and that each inflow ensemble member is smooth. The standard deviation of the random part of the error corresponds to the value used to generate inflow errors in Garcia-Pintado et al. [2015] and results in inflows that fit within the range given in Di Baldassarre and Montanari [2009] (4% to 43%). The mean of the inflow ensemble has negligible bias relative to the true inflow. The experiments shown here all use the same inflow for the truth and the same set of perturbed inflows for the forecast ensemble. For a different true inflow and different ensemble inflow error realisations, the results obtained using the different observation operators may compare slightly differently. However, the mechanisms we describe would be the same.

Each ensemble member was run with a different value of the channel friction parameter, n_{ch}. The behaviour of flood water is highly sensitive to n_{ch} (Hostache et al. [2010], James et al. [2016]), with low channel friction parameter values leading to water travelling through and leaving the domain more quickly. This leads to shallower water levels (and less flooding) in our simple domain for a given inflow. Conversely, higher channel friction parameter values lead to water moving slowly through the domain, leading to deeper water levels in the channel and more flooding. We chose a true value of $n_{ch} = 0.04$, equal

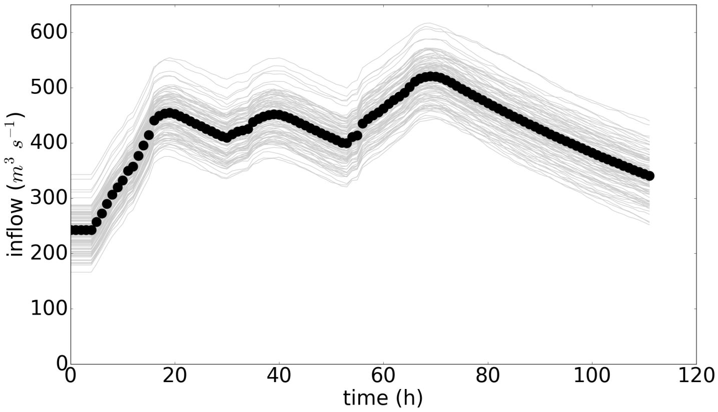


Figure 5.4: Inflows with time. True inflow values are represented with circles and ensemble inflows are shown by grey lines.

to the value for a natural stream given in Maidment and Mays [1988]. For the initial forecast step, a value of n_{ch} for each forecast ensemble member was drawn from a normal distribution with mean, μ, that is different to the true value and standard deviation σ. This imposed bias in the forecast ensemble channel friction parameter means that we can test how well data assimilation with different observation operators can correct the forecast state and parameter value towards the truth. In our state estimation experiments, the value of n_{ch} assigned to each ensemble member remained constant throughout the simulation. For joint state-parameter experiments, the values of n_{ch} were updated at each assimilation time through the ETKF equations, as described in section 5.3.2. Using an incorrectly specified channel friction parameter in the forecast is realistic, as the true value is unlikely to be known in operational situations. Initial forecast channel friction parameters are randomly drawn from a normal distribution with $\mu = 0.05$ and $\sigma = 0.01$ for experiments with positive bias in n_{ch} and with $\mu = 0.03$ and $\sigma = 0.01$ for experiments with negative bias in n_{ch}. The true value of n_{ch} falls within one standard deviation of the mean of each initial n_{ch} distribution and our choices of friction parameter values fit with the range used in Horritt and Bates [2002]. On the flood plain the value of the friction parameter is likely to be higher than n_{ch} due to the effects of vegetation. In this paper we used a true value for the flood plain friction parameter of $n_{fp} = 0.05$; the same, true value for n_{fp} was used for each ensemble member. The value of this parameter is likely to have an impact on the dynamics of a flood event, but flooding is commonly understood to be less sensitive to n_{fp} than n_{ch} (e.g. Hostache et al. [2010]). Here we focus on the ability of the observation operators to update n_{ch} only.

5.5.4 Synthetic observations

In identical twin experiments, observations are generated from a truth run; in this case the 'truth' flood simulation is described in section 5.5.3. For the two conventional observation operators we selected six synthetic observations of water elevation at the true flood edge at $y = 500m$, $700m$, $900m$, $1100m$, $1300m$, $1500m$. The elevation at these points is directly available from the state vector of water levels provided by our truth run. Each synthetic observation mimics a SAR-derived water level observation at a given cross section by locating a flood edge and using the true, calculated water elevation at this position as the observation. Here we define the flood edge WLO to be the elevation at the first 'dry' model cell encountered when moving in a perpendicular direction from the centre of the channel along one of our defined cross sections. (We use observations on the left hand side of the domain, i.e. where $x < 125m$, but since the domain and inflows are symmetrical in our simple experiments this choice is arbitrary; we could have instead used observations from the right hand side of the channel, or a combination of the two.) We added unbiased, Gaussian noise with a standard deviation of 0.25m to each observation; this is the same as the observation error used by Garcia-Pintado et al. [2015] in a case study. Observation error may be due to SAR instrument error or errors in determination of flood extent. The spacing of 200m between observations represents an optimistic best case situation, and is the same as the smallest recommended distance between thinned flood edge values for use in an assimilation system in Mason et al. [2012] (note that the other selection criteria used in the paper are not applicable here due to the use of synthetic observations). In fact, more recent work suggests a much longer correlation length scale between observation errors in a real case study (Waller et al. [2018]), in which the authors point out that part of the observation error correlation is due to the observation operator.

In order to test our backscatter observation operator we require synthetic backscatter observations; we therefore create a synthetic SAR image from our truth run, comprising backscatter values in each cell. We can then extract synthetic backscatter observations at desired locations. We have taken a very simple approach to generating a simplified synthetic SAR image in order to perform proof-of-concept experiments with our new observation operator; we will apply the method to a real case study and real SAR images at a later date. To generate a synthetic SAR image, we have taken our truth run water level output and applied a threshold water level of 5cm in each cell to determine which cells are wet and which are dry. Water levels below a threshold of a few cm are likely to be misclassified as dry in a real SAR image due to vegetation. Synthetic backscatter values are then assigned to each cell: dry cells are assigned a backscatter value drawn from $\mathcal{N}(m_d, \sigma_d^2)$ and wet cells a value from $\mathcal{N}(m_w, \sigma_w^2)$. For this, we have used values of $m_w = -14.84$, $\sigma_w = 2.25, m_d = -8.59$ and $\sigma_d = 1.53$, which are experimentally derived from a SAR image in Giustarini et al. [2016]. An example simplified synthetic SAR image, generated from the truth run at $t = 40h$, is shown in figure 5.5.

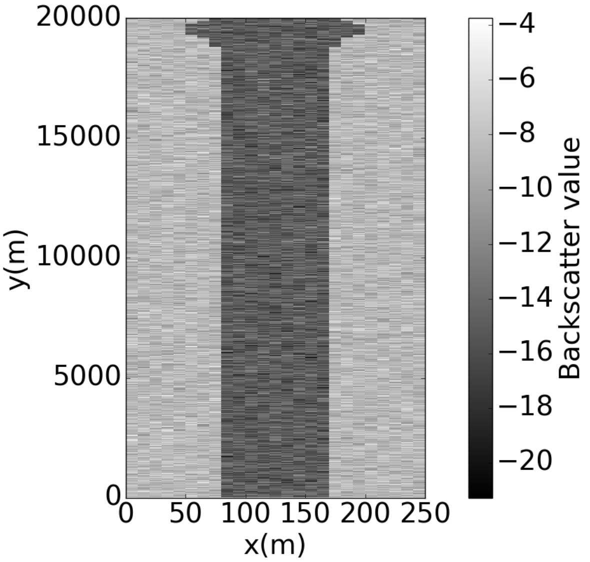


Figure 5.5: Synthetic SAR image generated from truth run water levels as described in section 5.5.4

In order to derive synthetic observations from the synthetic SAR image, the observation process is then carried out, i.e. we

- bin all the synthetic backscatter values in a histogram - see figure 5.2
- fit two Gaussian curves to the synthetic backscatter values (using python fitting algorithm scipy.optimize.curve_fit) - see figure 5.2
- extract new values of m_{w1},σ_{w1},m_{d1} and σ_{d1} from these distributions; these values are naturally very similar to the experimental values used to create the synthetic SAR image. We use a different realisation of observation error for each synthetic image (i.e. at each observation time); typical values of m_{w1},σ_{w1},m_{d1} and σ_{d1} are within 1% of m_w,σ_w,m_d and σ_d.

We then extract backscatter values to be synthetic observations. Although it would be possible to use a large number of backscatter observations in this method, for the experiments presented here we have not used all of the available synthetic observations. There are a number of reasons for limiting the number of observations. Firstly, observation errors are likely to be correlated for observations that come from positions close to each other in physical space. Some thinning of the observations is therefore necessary to meet the requirement that the observations used in the assimilation have uncorrelated errors (Mason et al. [2012]); this allows use of a diagonal observation error covariance matrix. Secondly, without ensemble localisation, using a number of observations larger than the number of ensemble members can cause the assimilation algorithm to overfit the observations (Kepert [2004]).

In this study we wish to investigate the differences in the updates generated by different observation operator approaches. We therefore use equivalent observation information for each of the operators. In the case of the water level observation operators we have used flood edge water level observations at six locations, where the flood edge location is defined as the position of the first dry model cell (see section 5.5.4). For the new operator we use two backscatter observations for each transect.

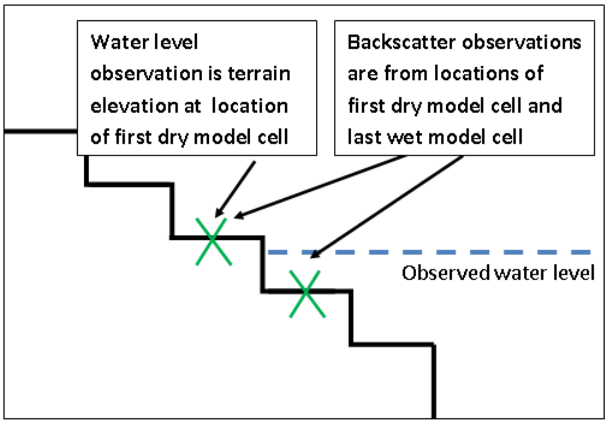


Figure 5.6: Schematic of observation locations used in this study for each transect in cross section. The thick black line shows the discretised domain elevation, the dashed blue line shows the observed flood water level. The arrows and green crosses show locations of the observations as labelled.

Figure 5.6 shows a schematic of the locations of the observations we have used in this study, relative to the edge of the flood. All observations used in this study come from transects at $y = 500m$, $700m$, $900m$, $1100m$, $1300m$ and $1500m$. In practical application of the backscatter operator, observations could be used from any location covered by the SAR image.

5.5.5 Observation error covariance matrices

It is important to specify the observation error statistics in data assimilation. In all cases we assume that our observation errors are uncorrelated so that we can use a diagonal error covariance matrix, **R**. We assume that the error in flood edge WLOs is $0.25m$. This is close to the calculated error in SAR-derived water level observations in Mason et al. [2012], and is the same value used in Cooper et al. [2018b] and Garcia-Pintado et al. [2015].

The uncertainty in each backscatter observation reflects the distribution to which it belongs (wet or dry). We assume that each entry can be set to be σ_{d1}^2 corresponding to a dry observation or σ_{w1}^2 for a wet observation.

5.5.6 Further experimental details

We present here the results from a number of data assimilation experiments, each lasting for a total simulation time of 112 hours. This includes an initial spin-up period with constant inflow for 4 hours (as shown in figure 5.4) to allow the water to reach an equilibrium state. In each experiment we use 100 forecast ensemble members. Assimilations are carried out at 12 hourly intervals. This is currently the shortest likely time between observations due to return times for satellites equipped with SAR instruments. The ETKF is used without localisation or inflation in all of the experiments as we did not encounter any spurious correlations or problematic ensemble collapse (see Petrie and Dance [2010]). This suggests that 100 ensemble members is sufficient in this particular case.

Experiments were run as follows

- State only estimation. State estimation experiments show how well data assimilation is able to correct forecast water levels at each observation time using the three different observation operators. In all of the experiments, a large bias is present in the forecast channel friction parameter values, which means that by design the error between the ensemble forecast and the truth grows quickly during each forecast step; the forecast corresponding to each of the observation operators relaxes to the same no assimilation (open loop) forecast. This allows us to examine the effect of each observation operator on the water levels in isolation at each observation/assimilation time, as the operators are each acting on very similar pre-assimilation forecasts.

  State only estimation experiments were carried out using a positive bias in the forecast channel friction parameter, which leads to forecast water levels that tend to be deeper than the truth (experiment PBSO) at any given cross section, and with a negative bias in the channel friction parameter, leading to shallower forecast water levels (experiment NBSO).

- Joint state and parameter estimation. Updating the value of n_{ch} along with water levels allows us to see the effect of the observation operators on the forecast when the large parameter bias can also be corrected by the assimilation process. Correcting the channel friction parameter in this way leads to better persistence in the forecast correction (Cooper et al. [2018b]). Experiments were again carried out using both a positively biased initial channel friction parameter distribution for the forecast ensemble (experiment PBJ) and negatively biased initial channel friction parameter distribution (experiment NBJ).

5.6 Results and discussion of update mechanisms

5.6.1 State only estimation

5.6.1.1 Positive bias in forecast ensemble channel friction parameter (PBSO)

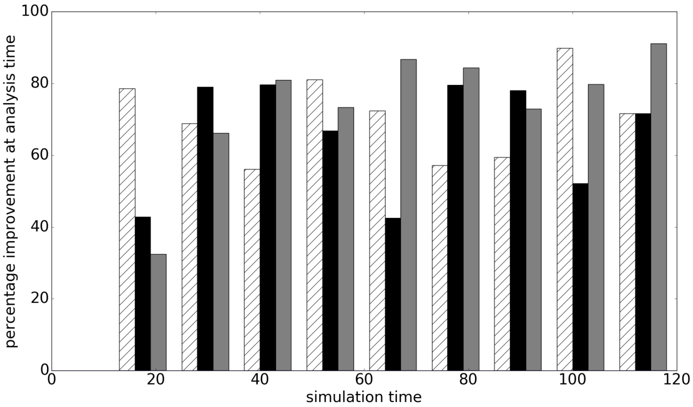


Figure 5.7: Improvement in the forecast at each assimilation time, PBSO experiment. The hatched white bars show improvement for the $\mathbf{h}_s$ operator, the black bars show improvement for the $\mathbf{h}_{np}$ observation operator and the grey bars show the improvement for the $\mathbf{h}_b$ observation operator.

Figure 5.7 shows improvement in the analysis compared to the forecast at each observation time for the PBSO experiment. Improvement is defined as

$$improvement = \frac{(\overline{\mathbf{x}^f - \mathbf{x}^t}) - (\overline{\mathbf{x}^a - \mathbf{x}^t})}{\overline{\mathbf{x}^f - \mathbf{x}^t}} \times 100, \tag{5.6.1}$$

where $\mathbf{x}^t$ is the true state of the system. This improvement measure is positive when the error in the analysis is smaller than the error in the forecast, while negative values imply a larger error in the analysis than the forecast. A perfect analysis ($\mathbf{x}^a = \mathbf{x}^t$) would result in a 100% improvement measure. Figure 5.7 shows that in the PBSO experiment all of the operators reduce the difference between the forecast mean and the truth at each observation time. We found that the error in the forecast then quickly relaxed back to the no assimilation (open loop) case for all of the observation operators. This short lived persistence in forecast improvement (less than approximately 3 hours here) when only water levels are updated is typical for such systems and is reported in many studies, including Cooper et al. [2018b], Andreadis et al. [2007], Neal et al. [2009], Garcia-Pintado et al. [2013] and Matgen et al. [2010].

5.6.1.2 Negative bias in forecast ensemble channel friction parameter (NBSO)

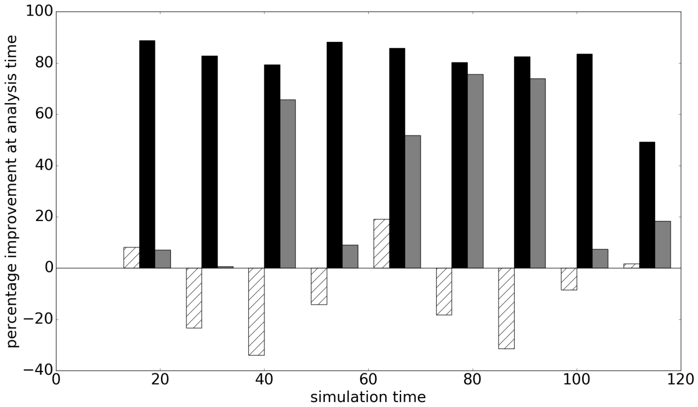
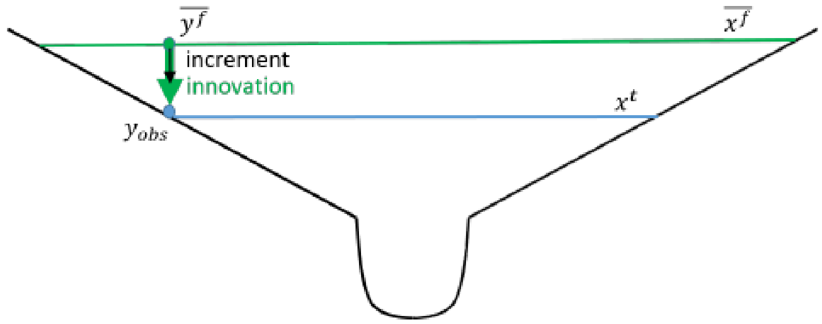
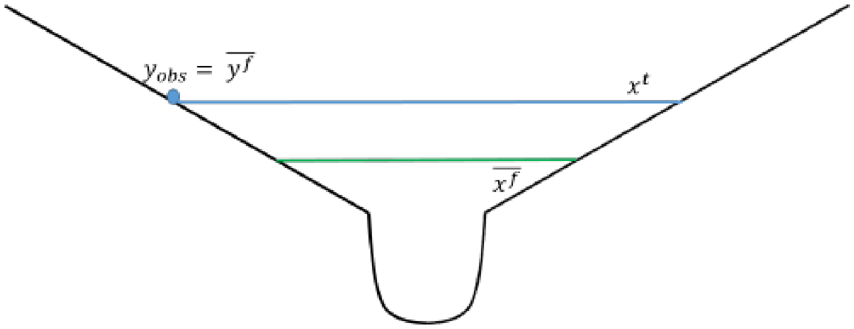


Figure 5.8: Improvement in the forecast at each assimilation time, NBSO experiment. The hatched white bars show improvement for the $\mathbf{h}_s$ operator, the black bars show improvement for the $\mathbf{h}_{np}$ observation operator and the grey bars show improvement for the $\mathbf{h}_b$ observation operator.

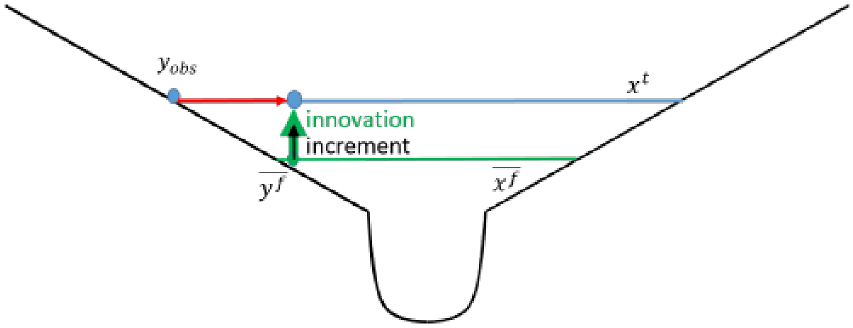
Figure 5.8 shows the improvement in the forecast at each the assimilation time for the NBSO experiment. Here, the ensemble channel friction parameters are such that the mean forecast water level tends to be shallower than the truth at any given cross section in our domain. Unlike in the PBSO experiment, the operators do not all provide a good analysis at every observation time. In fact, assimilation of flood edge observations using the simple flood edge observation operator, $\mathbf{h}_s$, makes the forecast significantly worse at many assimilation times. The reason for this is illustrated by considering the innovation produced by the simple flood edge operator when the forecast is shallower than the truth. The types of innovations produced for mean forecasts that are either deeper or shallower than the truth are shown in a schematic in figure 5.9.



(a) Mean forecast deeper than the truth



(b) Mean forecast shallower than the truth



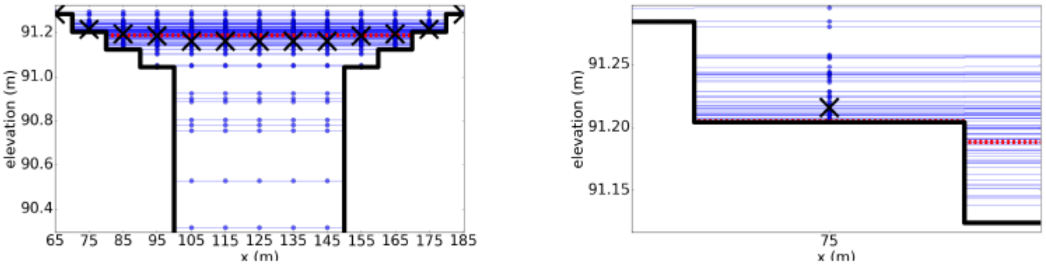
(c) Mean forecast shallower than the truth, nearest wet pixel approach

Figure 5.9: Schematic showing innovation for flood edge observation. In all cases blue lines represent the true water level and blue circles represent the corresponding flood edge observation, y_{obs}. Green lines show the mean forecast water level and green circles show the corresponding mean forecast-observation equivalent, $\overline{y^f}$. Innovations (δ) are shown with green arrows and increments by thinner black arrows - see equation (5.3.11) for definitions. The red arrow shows the difference between the observation location and the nearest wet pixel location.

Figure 5.9(a) shows a simple domain in a cross section where the mean forecast is deeper than the truth, with the innovation generated by the simple flood edge operator.

The innovation is such that the data assimilation algorithm can generate an increment and adjust the forecast water levels to be closer to the true water levels. However, as shown in figure 5.9(b), when the mean forecast is shallower than the truth, the simple flood edge assimilation method generates an innovation equal to zero. This is because the observation implies that at the flood edge, the water depth relative to the topography is zero; the ensemble forecast mean also predicts that the water depth is zero at the observation position. The increment is therefore also zero and the forecast cannot be adjusted to be closer to the truth (i.e. to shallower water levels), even though the observation clearly indicates that this is necessary. Figure 5.9(c) illustrates the way that the nearest wet pixel approach solves this problem by taking the water elevation at the observation position and extrapolating it in space (as also shown in section 5.4.2). This effectively moves the observation location to the nearest wet pixel, allowing a non-zero innovation to be calculated.

Figure 5.9 illustrates the fact that the simple flood edge operator cannot produce a useful update when the mean of the forecast ensemble is shallower than the observed water level. Figure 5.8 shows that in our experiments the simple flood operator in fact makes the forecast worse, increasing error relative to the truth at several assimilation times. The reason for this is linked to the fact that it is possible for the mean of the forecast ensemble to be deeper than the truth on the floodplain but shallower than the truth in the river channel.



(a)

(b)

Figure 5.10: Cross section of the domain showing bathymetry as a solid black line. The true water level is shown as a dotted red line, water levels predicted by each ensemble member are shown as blue circles. The mean forecast in each model cell is shown as a cross. Figure 5.10(a) shows the central part of the domain from $65 \leq x \leq 185$m. Figure 5.10(b) shows the forecast water levels and resulting forecast mean in the cell centred at 75m in greater detail. Reprinted from Cooper et al. [2018b] with permission from Elsevier

Figure 5.10 shows the domain at one cross section. In figure 5.10 we see that in the channel (e.g. at $x = 125$m) the true water level is deeper than the ensemble mean. At the edge of the flood, the true water depth is (by definition) zero relative to the topography and the majority of ensemble members also predict zero water depth in these cells. However, there are a small number of ensemble members that predict non-zero water depth at the flood edge; it follows that the ensemble mean at this location is therefore a small non-zero water depth as per equation (5.3.1). The flood edge operator therefore generates an innovation such that the mean forecast water depth at the flood edge is reduced and the analysed water depths are closer to the truth at this location. Correlations between water levels in the domain mean that the water depth in the channel is also reduced by the update step; this increases the error relative to the truth in the channel. This explains the overall increase in error at assimilation times seen in figure 5.8.

The results in figures 5.7 and 5.8 show that the new backscatter operator works well at most of the observation times. The mechanism by which the backscatter observation operator works is illustrated in figure 5.11.

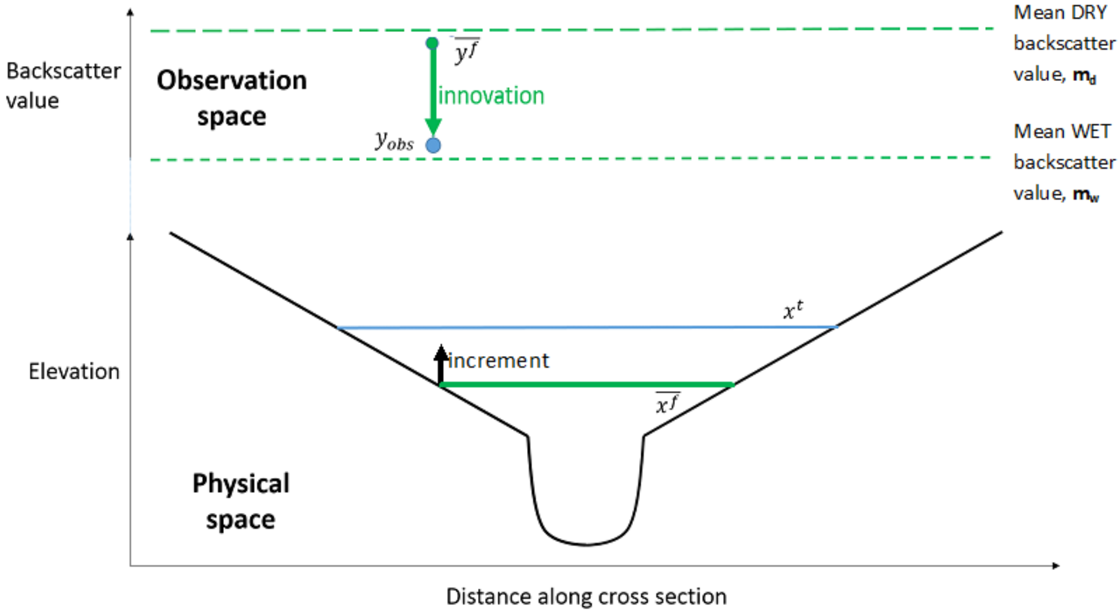


Figure 5.11: Schematic of innovation (in backscatter/observation space) and increment (in physical space) for one backscatter observation. The horizontal blue line represents the true water level and the blue circle represents a corresponding backscatter observation, y_{obs}. The solid green line shows the mean forecast water level and the green circle shows the corresponding mean forecast-observation equivalent in observation space, $\overline{y^f}$. The innovation (defined in equation (5.3.11)) is shown in observation space with a green arrow and the increment in physical space at the observation position (equation (5.3.12)) is represented by a thinner black arrow.

Figure 5.11 shows a simplified river channel in cross section. The lower part of the figure shows an example of a true and mean forecast water level, as in figure 5.9. The upper part of the figure shows the same cross section, but is a representation in observation space of an example (single) observation and equivalent mean forecast backscatter value, $\overline{\mathbf{y}^f}$. The green circle in observation space shows $\overline{\mathbf{y}^f}$ in the cell at the observation position. The value of $\overline{\mathbf{y}^f}$ is calculated using water levels forecast by all the ensemble members, through equation (5.4.3), and is essentially a measure of the proportion of ensemble members which predict that cell to be wet (or dry). The mean forecast backscatter, $\overline{\mathbf{y}^f}$, will always take a value between the mean observed wet value, m_{w1} and the mean observed dry value, m_{d1}; if half the ensemble members predict a cell to be dry and half predict it to be wet, the value of $\overline{\mathbf{y}^f}$ will lie halfway between m_{w1} and m_{d1}. If most ensemble members predict the cell to be wet (dry), the value of $\overline{\mathbf{y}^f}$ will be close to the mean observed wet (dry) backscatter value. The observed backscatter value, $\mathbf{y}_{obs}$, is shown as a blue circle in observation space.

The innovation is shown in observation space in figure 5.11. The innovation is the difference between the observed backscatter value, $\mathbf{y}_{obs}$, and the mean forecast backscatter value, $\overline{\mathbf{y}^f}$. Figure 5.11 shows that for the $\mathbf{h}_s$ and $\mathbf{h}_{np}$ the state variables and observed variables are the same. In the approach using $\mathbf{h}_b$, the observations are different to the state variables. For $\mathbf{h}_b$ the increment is the calculated difference in water level between the forecast and the analysis in metres, but this is calculated using an innovation that is a difference in backscatter value between the model and the observation. In the example shown, the mean forecast backscatter value indicates that most of the ensemble members predict the cell containing the observation position to be dry. This corresponds to the shallow mean water level prediction shown in physical space. The backscatter observation indicates that the cell is wet. The innovation is therefore large, and indicates that the cell is more likely to be wet than the forecast indicates. This maps into an increment in physical space through equation (5.3.12) such that the calculated analysis water level at the observation position is deeper than the forecast water level.

A potential problem with the backscatter operator can be illustrated through inspection of equations (5.3.13) and (5.3.10). Equation (5.3.10) shows that when the value of the Kalman gain matrix is zero, there can be no update to the forecast through assimilation of observations, even when there is a large innovation - i.e. a large difference between a model prediction and an observation. Equation (5.3.13) shows that this $\mathbf{K} = 0$ condition can be met if either $\mathbf{X} = 0$ or $\mathbf{Y} = 0$. For $\mathbf{Y} = 0$ to be true, it is only required that the ensemble members all predict the cell containing the observation to be dry, or all ensemble members predict the cell to be wet. This is because if all ensemble members predict a cell to be wet then they all give the same value of $\mathbf{y}_i^f = m_w$ through equation (5.4.4). Equation (5.4.3) then shows that the value of $\overline{\mathbf{y}^f}$ will then also be equal to m_w, and each term in $\mathbf{Y}$ must therefore be zero according to equation (5.3.9), since all the ensemble members are the same as the mean. This means that if all the ensemble members predict different but positive water depths (i.e. no non-zero water depths are predicted in the ensemble), no increment can be generated and no update made to the forecast, regardless of whether the observation indicates a wet or dry condition. For this reason, observations at or near the edge of the flood are most valuable to the data assimilation algorithm when using the backscatter observation operator, since these are locations where it is most likely that the ensemble members will predict a variety of wet/dry predictions. We did not observe any situation in which $\mathbf{Y} = 0$ in these experiments. It would in principle be possible to add a small amount of noise to each value of $\mathbf{y}_i^f$ in order to prevent $\mathbf{Y} = 0$, but this risks generating an innovation and increment such that the analysis error is larger than the forecast error.

5.6.2 Joint state-parameter estimation

The large source of error in these experiments is, by design, due to a large bias in the forecast ensemble channel friction parameter values. In this section we show the results of updating the forecast channel friction parameter values as part of the assimilation process. One way to measure the effectiveness of a data assimilation approach is to compute the root mean square error (RMSE) between the resulting forecast and the truth. Here, RMSE is defined as

$$RMSE = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(d_j^t - d_j^f)^2}, \tag{5.6.2}$$

where d_j^t is true water depth in the jth cell; d_j^f is mean forecast water depth in the same cell. As before, N represents the number of cells in the domain.

5.6.2.1 Positive bias in forecast ensemble channel friction parameter (PBJ).

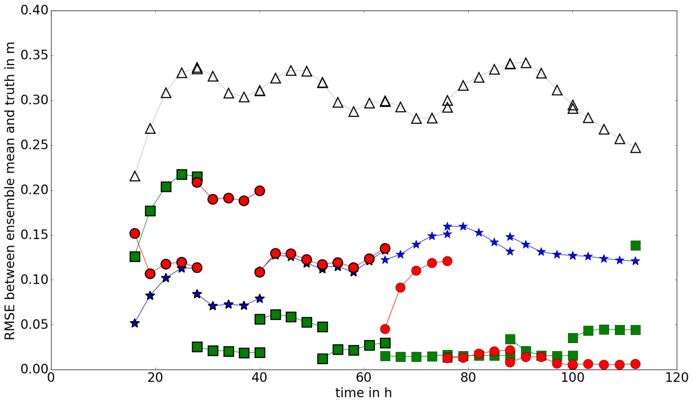


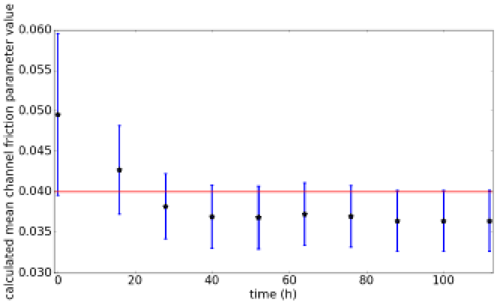
Figure 5.12: RMSE between forecast and truth, PBJ experiment. Open triangles show the RMSE between the open loop forecast and the truth. Blue stars, green squares and red circles show the RMSE between the forecast mean and the truth using the $\mathbf{h}_s$, $\mathbf{h}_{np}$ and $\mathbf{h}_b$ observation operators respectively.

Figure 5.12 shows the RMSE between the mean water levels predicted by the model and the true water levels for the PBJ experiment. The mean value of n_{ch} and the mean value of the predicted water levels are updated at 12 hourly intervals starting from 16h. At each assimilation time the RMSE for both the forecast (pre-assimilation) and analysis (post-assimilation) water levels are shown; points within a forecast step are joined with a line. The results show that the assimilation leads to a much improved forecast of water levels for all of the operators at all times. There is persistence in the improvement to the forecast, and each of the observation operators provides a better forecast than the open loop ensemble for the whole of the simulation time. The results obtained using the $\mathbf{h}_s$ operator converge to higher RMSE values than the other two operators. Use of the $\mathbf{h}_b$ operator shows a gradual reduction in RMSE over successive forecast-analysis cycles. The results for the $\mathbf{h}_{np}$ operator show faster reduction in the RMSE values, but the final analysis value (at 112h) has a much higher RMSE. This is because at 112h the inflow has reduced such that the water is well back within bank and in these conditions the assumptions used to derive water elevation observations break down; the sides of the river are too steep for the water edge position to accurately determine elevation. In an operational setting, it would be necessary to test for an in-bank condition and discard observations for the $\mathbf{h}_{np}$ operator when the river is within bank. This means that it is not possible to calibrate a hydrodynamic model on a river using SAR images when it is not in flood if water level observations are being used (i.e. with either the $\mathbf{h}_s$ or $\mathbf{h}_{np}$ observation operator).

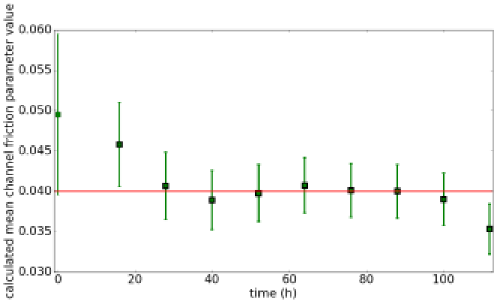
Figure 5.13 shows the calculated (analysis) mean channel friction parameter values at each assimilation time for the three observation operators. All of the operators produce values for the parameter that are closer to the truth than the initial value. The value of the channel friction parameter calculated using the $\mathbf{h}_b$ observation operator converges to a value close to the truth after 6 observations and then remains there. The value calculated using $\mathbf{h}_{np}$ converges more quickly to a value close to the truth, but the last value in the time series (at 112h) then diverges from the true value. This is because the river is now well within bank and water elevation observations cannot be reliably determined.

5.6.2.2 Negative bias in forecast ensemble channel friction parameter (NBJ)

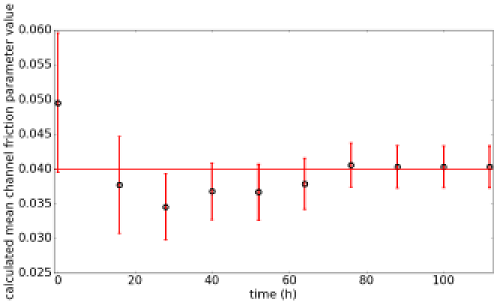
Figure 5.14 shows the RMSE between the forecast and the truth for the NBJ experiment. The nearest wet pixel approach provides a forecast which is very close to the truth for most of the simulation time. The backscatter operator performs well after the first two assimilation steps, showing a slower convergence to the true solution as in the



(a) Simple flood edge operator, $\mathbf{h}_s$



(b) Nearest wet pixel operator, $\mathbf{h}_{np}$



(c) Backscatter operator, $\mathbf{h}_b$

Figure 5.13: Calculated analysis mean channel friction parameter, PBJ experiment. Horizontal red line shows true value of channel friction parameter. Error bars show one std of ensemble parameter distribution.

PBJ experiments. The simple flood edge operator performs badly, leading to a forecast which is worse than the open loop case for most of the time. The reason for the poor performance in this particular experiment is likely due to the mechanisms outlined in section 5.6.1.2. The forecast is adjusted in the wrong direction at the first assimilation time (at 16h) such that the water levels are too shallow; the mechanism by which this can happen is demonstrated in figure 5.10. All subsequent corrections are very close to zero, due to the mechanisms illustrated in figure 5.9, so that the blue line appears to be unbroken.

Figure 5.15 shows the calculated analysis mean channel friction parameter values at

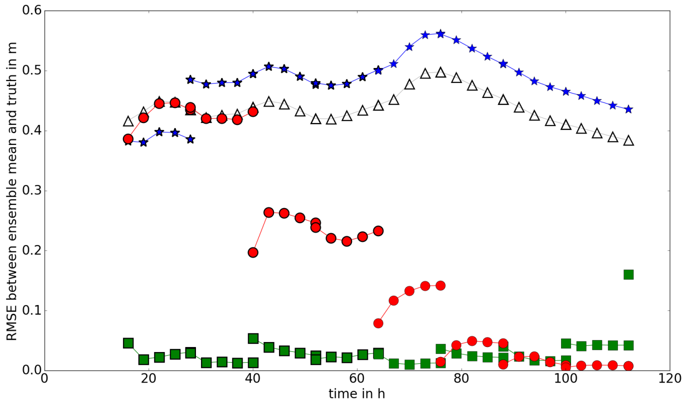
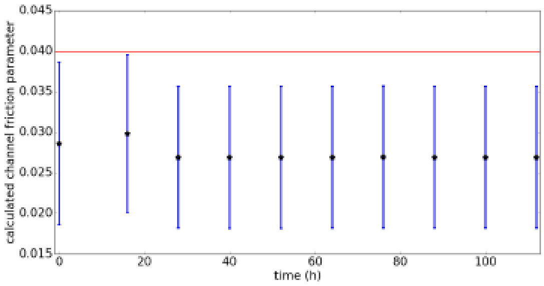


Figure 5.14: RMSE between forecast and truth, NBJ experiment. Open triangles show the RMSE between the open loop forecast and the truth. Blue stars, green squares and red circles show the RMSE between the forecast mean and the truth using the $\mathbf{h}_s$, $\mathbf{h}_{np}$ and $\mathbf{h}_b$ observation operators respectively.

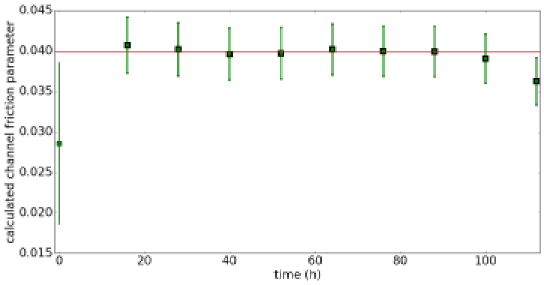
each assimilation time in the NBJ experiment for the three observation operators. The results for the simple flood edge operator support the scenario outlined above, whereby the water levels are initially adjusted in the wrong direction and then cannot be updated towards the truth. Although the details of this will depend on topography, observation error and choice of forecast inflows and parameters, this is nevertheless an important mechanism to consider when choosing an observation operator. Both the $\mathbf{h}_{np}$ and $\mathbf{h}_b$ operators do successfully correct the value of the parameter towards the truth, with the $\mathbf{h}_{np}$ operator recovering a good value in a shorter time than the $\mathbf{h}_b$ operator. Both figures 5.14 and 5.15 show that at the final assimilation time, the analysis and parameter value provided by the nearest wet pixel operator are not close to the truth. Again, this is because the river is well within bank so the flood edge observation is on ground which is too steep to provide a good observation; in operational settings observations such as these would be screened out and no update would be made with the operator.

5.7 Discussion

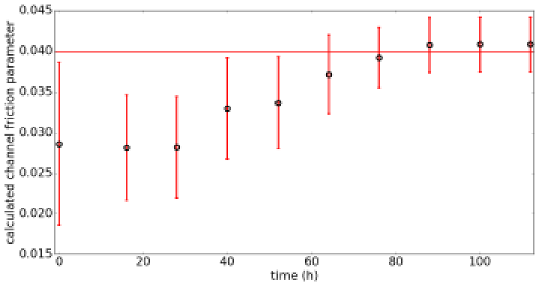
In this study we have chosen to use a small number of backscatter observations for our experiments. This allowed us to compare updates between the three observation operators when the observation operators were all given equivalent information; in this way we can draw conclusions about the physical mechanisms responsible for the different updates. In a real case, one of the major advantages of using our new backscatter observation operator is that it would be possible to use a large number of backscatter observations compared to



(a) Simple flood edge operator, h_s



(b) Nearest wet pixel operator, h_{np}



(c) Backscatter operator, h_b

Figure 5.15: Calculated analysis mean channel friction parameter, NBJ experiment. Horizontal red line shows true value of channel friction parameter. Error bars show one std of ensemble parameter distribution.

the number of water level observations which are typically available. The availability of a large number of observations may be a major strength of our new approach; in our simple experiments (not shown) we found that assimilating a larger number of observations with the backscatter operator provided a better analysis than using only a few. Another merit of the backscatter operator is that there is less processing involved in using backscatter observations directly, potentially reducing the amount of time between acquisition of a SAR image and its use to update an inundation forecast. The backscatter operator also removes the need for locating the nearest wet pixel in the model forecast, which can be computationally costly.

There are a number of potential problems with practical implementation of the backscatter operator. One is that using histograms to produce SAR-derived inundation maps can lead to errors in assigning pixels to wet/dry categories. One way to deal with this would be to use region growing techniques (see e.g. Horritt et al. [2001]) or change detection techniques (see e.g. Hostache et al. [2012]) to produce robust wet/dry maps for SAR images, and then perform a quality control procedure to discard any backscatter observations which would lead to mis-classification due to e.g. emergent vegetation. This procedure would remove the advantage of fewer processing steps for the backscatter operator, but may not be necessary. Further research is required to understand how robust the method is to the proportion of misclassified SAR pixels in a real case study. We note that the backscatter operator would not generate an update the forecast in model cells that all the ensemble members predicted to be dry (or wet) as discussed in the last paragraph of section 5.6.1.2. This means that SAR pixels far from the river wrongly classified as wet, or SAR pixels in the river channel wrongly classified as dry would not degrade the forecast through an erroneous update.

The new backscatter operator is likely to work well in cases where good separation of the wet/dry distributions can be obtained through a histogram, and less well in cases where the distributions overlap. The new observation operator does not require a digital elevation model to generate forecast-observation equivalents, although the hydrodynamic model would require topography information to generate a forecast. Water level observations cannot be accurately determined in areas with high slope, whereas backscatter observations will be unaffected. Like the other observation operators, the new operator will likely provide better results in rural settings than urban settings; double-bounce and layover effects due to buildings are potential sources of problems for all of the operators (Mason et al. [2018]).

5.8 Conclusions

We have carried out a series of experiments to test the performance of three different types of observation operators in an ETKF approach to data assimilation for fluvial inundation forecasting. Although the results are for one specific idealised domain, one realisation of true inflow and a single realisation of observation error per observation type, we believe that many of our conclusions will be applicable much more widely through the mechanisms we describe. Repeats of experiments (not reported here) with different realisations of observation error show evidence of the same behaviour in terms of the mechanisms we have described. Our experiments show that:

- Simple assimilation of flood edge water elevation observations can result in no correction to the forecast even when there is a large difference between the forecast and the observation. This happens when both the model prediction and the observation predict no flooding at the observation location. We have illustrated the physical mechanism responsible for this (figure 5.9) and shown an example in which this happens in our experiments (see assimilation times from 28h onwards in figure 5.14). The simple flood edge operator can also generate an update such that the analysis has a larger error than the forecast. This can occur when the forecast is deeper than the truth at the observation position, but shallower than the truth in the channel. In such cases the assimilation updates the water levels to shallower levels as required at the observation position, but also wrongly updates the channel water levels to be shallower. The mechanism for this is shown in figure 5.10; this is responsible for the negative improvement measures in the NBSO experiments (see figure 5.8). We have shown in our experiments that the simple flood edge operator fails in these ways when the mean ensemble channel friction parameter is negatively biased but it would also fail if, for example, the mean forecast inflow was negatively biased since errors in friction parameter and inflow are correlated (Cooper et al. [2018b]). Since in operational settings both forecast inflow and channel friction parameter values are uncertain, we conclude that the simple flood edge operator is not a good choice.

- The nearest wet pixel approach provides better assimilation accuracy than simple flood edge assimilation: in our experiments we find no evidence of negative 'improvement' scores or zero increments when the forecast and observations are very different. In our idealised system it is the best choice of observation operator in terms of better forecast accuracy in the state only experiments and in terms of rapid convergence to the true solution for both water levels and mean channel friction parameter value in the joint state-parameter experiments. However, we have shown that using water edge observations when the river is well within bank can lead to a degradation of the forecast. Also, locating the nearest wet pixel is likely to be difficult in practise for operational applications using real, more complicated topography. One way to limit the distance between the flood edge observation position and the nearest wet pixel is to locate the nearest pixel at which some threshold of ensemble members predict a positive water depth. The predicted water elevations at this location could then be used to create $\overline{\mathbf{y}^f}$. This approach balances out the need to include information from ensemble members predicting shallow water levels at the observation position with the requirement that the nearest wet pixel is not too far from the observation location.

- Our new backscatter observation operator performs well compared to more conventional options in our idealised domain using synthetic observations. The operator does not suffer from the problems of the simple flood edge operator and is able to correct the forecast for the state only assimilation cases. The backscatter operator approach also allowed the forecast to converge to the true solution for both water levels and channel friction parameter value in the joint state-parameter experiments, although in our experiments convergence was slower than for the nearest wet pixel approach. Using backscatter values operationally may speed up the time taken from image acquisition to assimilation and an improved forecast due to fewer steps in the processing. The new operator could also potentially allow the use of much more information from any given SAR image, although there is likely to be a limit to the number of backscatter observations that can be used without causing variance collapse in the channel friction parameter distribution. Tests using larger numbers of backscatter observations have not been presented here; we plan to address this question in a real case study so that the results will be more directly applicable to real world situations.

This work has shown that our novel backscatter operator has the potential to improve inundation forecasting in fluvial floods, and we believe it may have applications in other types of flooding where SAR images are available. Further work is required to test the operator against the $\mathbf{h}_{np}$ approach in a real case study, using real SAR data and real topography in order to further assess the strengths and weaknesses of the different approaches. We have explained the physical mechanisms associated with the assimilation increments for each type of observation operator; these mechanisms will also be applicable to variational data assimilation methods using similar observations. Improved understanding of these physical mechanisms provides insight into the best approaches to improve the effectiveness of assimilation of SAR data in the future.

5.9 Chapter summary

In this chapter we have designed and implemented a novel method to extract observational information from a SAR image for use in data assimilation. We have used our novel approach in synthetic experiments and showed that the new method can update modelled water levels towards the truth. We have compared the water level updates produced by our new approach to those generated by two more conventional approaches and discussed the physical mechanisms by which observation operators (including our new operator) update water levels.

Chapter 6:

New observation operator for inundation forecasting - case study.

In this chapter we address the last question from chapter 1: How does our new backscatter observation operator work with real topography and SAR data? We have shown in chapter 5 that our new backscatter observation operator can improve the forecast in an idealised domain with synthetic observations; here we want to answer the following questions:

- Can we apply our new observation operator to a real case study using real topography and SAR images?
- Does our new observation operator improve the forecast in a real case study?

The new backscatter observation operator described in chapter 5 has been applied to a flood event in the Severn catchment, U.K, in November/December of 2012. We chose this event as we have access to a set of SAR images over the course of the flood to use as observational data. The first SAR overpass happened after the start of the flooding and we show here results from assimilating data from six images. Our experiments cover a subdomain of the area used in Garcia-Pintado et al. [2015] and Garcia-Pintado et al. [2013], in which the hydraulic model LISFLOOD-FP (Neal et al. [2012a]) was used to model the same flood event. We used a subdomain in order to simplify the analysis of our new observation operator and also due to the high computational cost of running our inundation model at high resolution (see section 6.1.1).

This chapter is organised as follows: In section 6.1 we describe the application of our hydrodynamic model to part of the Severn catchment. We describe the observations used in our data assimilation experiments in section 6.2; in section 6.3 we briefly describe the experimental design of the data assimilation experiments we have carried out. In section 6.4 we discuss the results of our experiments and in section 6.5 we conclude that our new observation operator can be successfully applied to a real SAR data.

6.1 Inundation model

6.1.1 Topography and computational grid resolution

The inundation model used for the work in this chapter used the same Clawpack code as in chapters 4 and 5. The terrain elevation data for the simulations was adapted from the 75m resolution elevation data used in Garcia-Pintado et al. [2015] and Garcia-Pintado et al. [2013]; this was based on upscaling the NEXTMAP British digital terrain model (Intermap Technologies).

Figure 6.1 shows the terrain elevation in m used for our inundation simulations, with three gauge locations. Data from gauges at Knightsford Bridge and Worcester Barbourne were used to drive the model inflows; data from the gauge at Saxon's Lode were used for validation of results. Gauge data were not assimilated. The domain includes a stretch of the river Severn (traversing north-south) north of Tewkesbury, U.K. The domain also includes part of the river Teme (east-west), which is one of the Severn's larger tributaries. We used reflecting boundaries at the north, east and west boundaries of the domain, and an extrapolating outflow boundary at the southern end of the domain in order to allow water to flow out freely.

It was necessary to make changes to the elevation used in Garcia-Pintado et al. [2015] and Garcia-Pintado et al. [2013] due to differences in the way Clawpack and Lisflood-FP represent topography. Lisflood-FP uses a sophisticated sub grid parameterisation technique to represent river channels (see Neal et al. [2012a]). This gives the model the capability to use low resolution digital terrain maps while still resolving all the river channels; the 30-50m wide rivers in the Severn-Avon catchment can therefore be accurately represented using a 75m grid resolution. This sub-grid approach was used in Garcia-Pintado et al. [2015] and Garcia-Pintado et al. [2013] to model river channels in LISFLOOD-FP as rectangular cross sections, with paramaterised depths and widths of the channels chosen to match measured cross sections. Clawpack does not have sub-grid capability. Since we used the 75m terrain model, we necessarily assumed the river channels were all 75m wide and adapted the bed elevation in the river model cells in order to satisfy the requirements that:

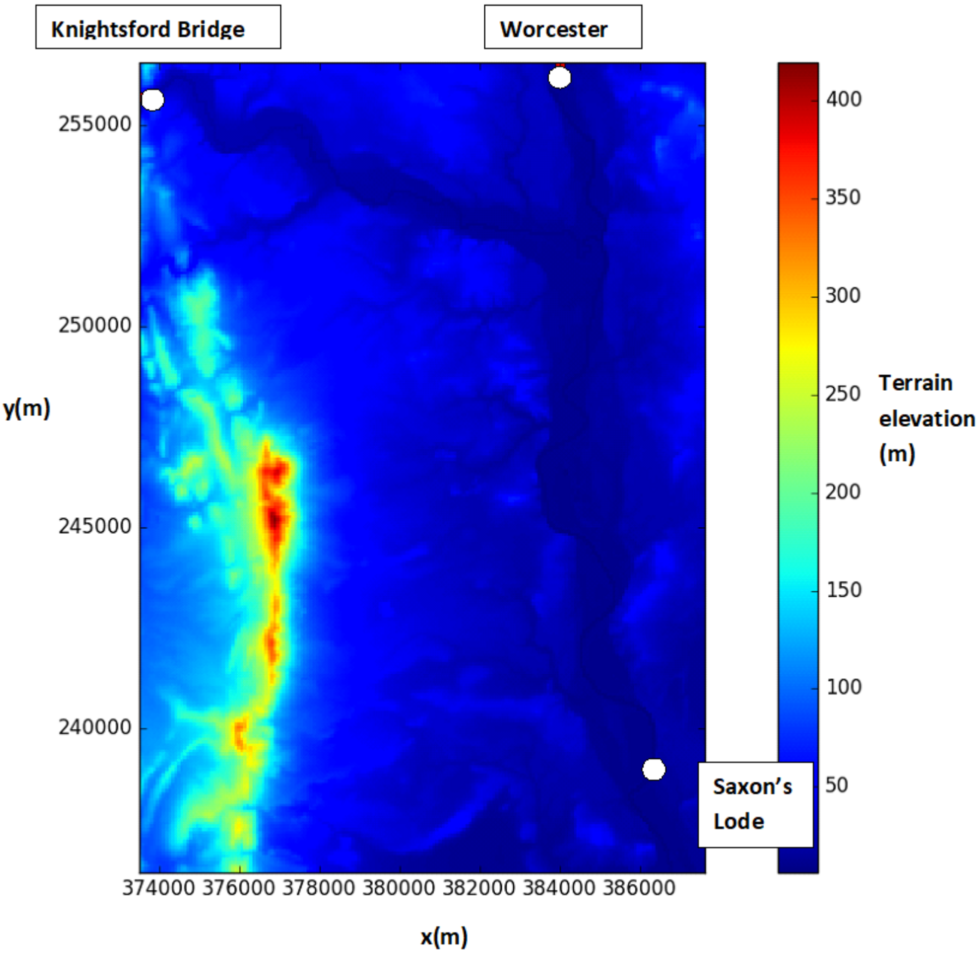


Figure 6.1: Experimental domain, elevation shown in metres. Axes co-ordinates (x and y) are OSGB 1936 British National Grid projection in m. White circles show the location of gauges. Data from Knightsford Bridge and Worcester are used for inflow generation. Data from the Saxon's Lode gauge are used for comparison with model prediction.

- the cross sectional area of the channels approximates the parameterised cross sectional area used in Garcia-Pintado et al. [2015] and Garcia-Pintado et al. [2013] and
- the bed elevations are smoothly varying.

We found that matching the cross sectional area exactly in each river grid cell meant that the bed elevations were not smooth enough to accurately represent the rivers. Some river cells had bed elevations much higher than all their neighbours, resulting in water going out of bank in unrealistic conditions (i.e. even for very low flows, and in areas that do not show flooding on the SAR images). We therefore smoothed the bed elevations to generate more realistic behaviour. All models of this type are required to make approximations and parameterisations for river and floodplain topography, resulting in flows which are not completely realistic. Nevertheless, the updates generated using our new observation operator are still of interest as we are able to examine the impact of the observations and assess to what extent the data assimilation is able to compensate for the inaccuracies of the model.

Although Clawpack does not have sub-grid capabilities, it is possible to specify a computational grid resolution that is different to the topography resolution. We therefore used a computational grid resolution of 25m with the topography grid of 75m resolution. This means that each topography grid cell was split into nine smaller cells for calculation of the solution, each with the same elevation defined by the topography map. Each 75m-wide river cross section is therefore modelled by three 25m grid cells. Using this finer resolution for the computational grid results in longer run times for the simulation experiments, but we found it was necessary in order that the simulated flow rates matched the measured flow rates at the Saxon's Lode gauging station at low flows. To illustrate this, we used reported flow rates from Knightsford Bridge and Worcester Barbourne over a (no flooding, low-flow) 10 day period before the floods as driving inflow and measured the resulting model prediction at the location of the Saxon's Lode gauge. The simulation period here runs from 0100 11th November 2012 to 0100 21st November 2012.

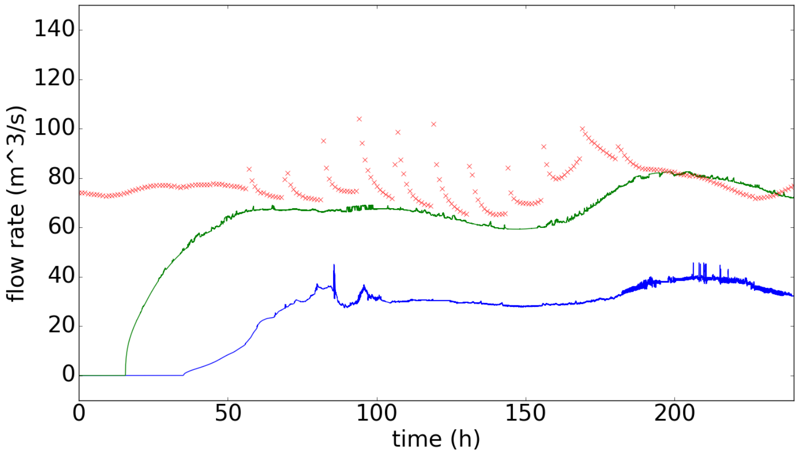


Figure 6.2: Measured and modelled flow rates at Saxon's Lode gauge location. The red crosses show measured hourly values. The blue line is modelled flow rate for 75m computational grid resolution, the green line is modelled flow rate for 25m computational grid resolution; all other experimental conditions and model parameters are the same.

Figure 6.2 shows modelled flow rates at the location of the Saxon's Lode gauging station for 25m (green line) and 75m (blue line) resolution computational grid. In both cases the Manning's friction coefficients were set to 0.02 in the river channels and 0.06 elsewhere. The red crosses show hourly flow rates reported by the Environment Agency at the gauge. The discharge spikes seen in the observed data are due to tidal backwater effects, sometimes observed at this location (Neal [personal communication, 2019]). This figure shows that our model is unable to successfully reproduce the in-bank flow rates when using the 75m resolution computational grid; this was true for a range of realistic channel friction parameter values. Using a 25m resolution grid gives results that match the reported values better. The increased computational cost of running the model at 25m resolution is the main reason for using a subdomain of the area covered by the SAR observations.

It was also necessary to switch off terrain interpolation in Clawpack for the elevation model. Clawpack is designed to take a (coarse) topography map and interpolate model cell corner elevations in order to give a good approximation for the elevations of the terrain in the centre of each computational cell. In the case of a river profile, this can have the effect of changing the cross section of the river channel. In order to use the bed elevations that were specified with no additional smoothing it was necessary to write a new 'no-interpolation method' to set up the topography exactly as specified.

6.1.2 Choice of value for Manning's friction coefficient

Unlike in the synthetic twin experiments of previous chapters, we do not have a truth run, nor a known value for the true friction parameters. In order to determine a likely range of values for n_{ch} we compared modelled flow rates at Saxon's Lode for a number of channel friction parameter values. As in section 6.1.1 we used pre-flood, low measured flow rates at the inflows from 0100 11th November 2012 to 0100 21st November 2012 to drive the model.

Figure 6.3 shows predicted flow rates for three different values of n_{ch}. Initially there is no water anywhere in the domain and in all cases the domain is the 75m topography with a 25m resolution computational grid. As expected, higher channel friction parameters result in the water taking longer times to reach the location of Saxon's Lode. After this, there are only small differences between the flow rates predicted by different channel friction parameter values. The green and blue lines are not smooth because flow rates are plotted at every model time step, which can be very small (a few seconds).

Based on our simple investigations, we randomly sampled values of n_{ch} for each ensemble member from a normal distribution centred on 0.03, with standard deviation 0.01. We used the same values in both the Severn and the Teme for each ensemble member; this follows the approach of Garcia-Pintado et al. [2015] and Garcia-Pintado

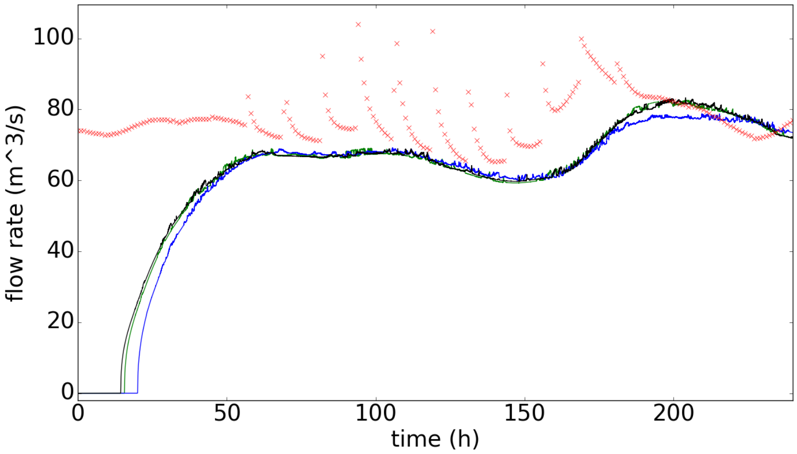


Figure 6.3: Gauged and modelled flow rates at Saxon's Lode gauge location. Red crosses show hourly gauged values. The black line shows modelled flow rate when $n_{ch} = 0.015$, the green line shows the modelled flow rate $n_{ch} = 0.02$, the blue line shows modelled flow rate when $n_{ch} = 0.04$. In all cases the friction parameter outside the channels is set to 0.06

et al. [2013]. Since the Teme is narrower than the Severn and we have represented both with 75m wide channels, it would make sense to attempt to parameterise the friction coefficients to try to compensate for this. In practice, initial investigations showed little effect on the resulting flow rates of using differing friction parameters; this is perhaps because the main contribution to the flow at Saxon's Lode comes from the river Severn.

Outside of the river we used a friction coefficient of 0.06, which is the same value used in Garcia-Pintado et al. [2015] and Garcia-Pintado et al. [2013]. It would be interesting to investigate the effects of using different values for friction coefficients in the river and flood plain, or on different branches of the river, but we did not consider this here. We carried out only brief calibration experiments for the friction coefficients because the purpose of our experiments is to assess the impact of our new observation operator; this does not rely on the model giving the best possible forecast, just a broadly realistic one.

6.1.3 Inflow ensemble generation for assimilation experiments

We used measured flow rates from gauges at Knightsford Bridge and Worcester Barbourne before and during the flood to generate inflows for the forecast ensemble. We created an ensemble of inflows for both inflow gauges by adding time correlated errors to the gauged inflows in the same way as described for the synthetic experiments in section 5.5.3. Each inflow simulation started 21:00 on 17th November 2012, approximately 10 days before the first SAR observations, to account for any spin-up effects. The simulation time period ended at the time of the final SAR overpass, at 18:14 on 2nd December 2012.

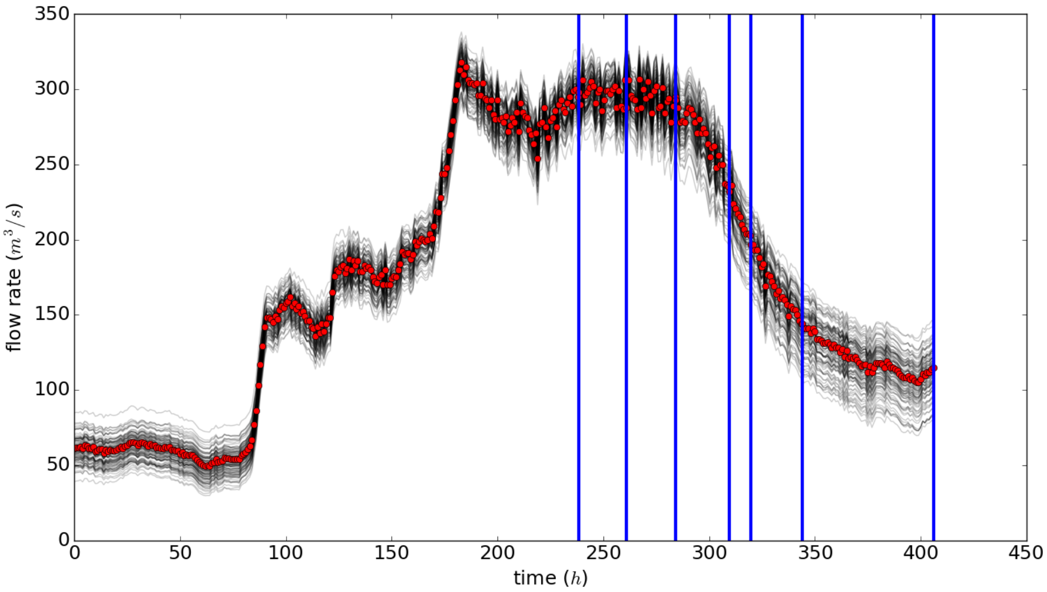


Figure 6.4: Inflow at Worcester gauging station. Red circles show measured data from gauging station, black lines show flows for ensemble members. Blue vertical lines show the times of the satellite overpasses used for observational information.

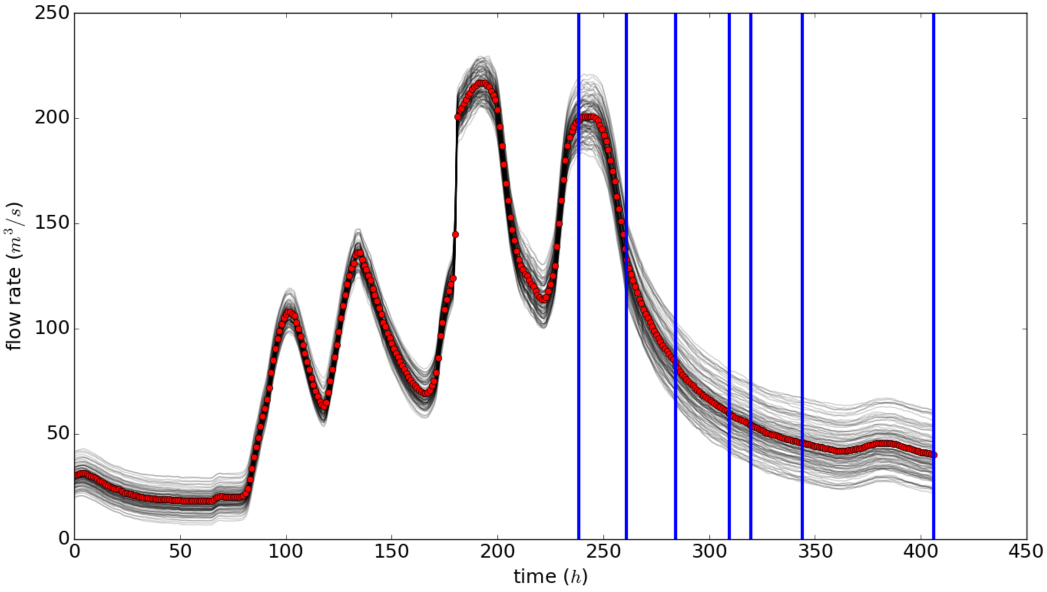


Figure 6.5: Inflow at Knightsford Bridge gauging station. Red circles show measured data from gauging station, black lines show flows for ensemble members. Blue vertical lines show the times of the satellite overpasses used for observational information

Figures 6.4 and 6.5 show the inflows used for each of the 100 ensemble members for the Worcester and Knightsford Bridge inflows respectively. Time is measured in hours from the start of the simulation at 2100h on 17th November 2012. For dates and times of the overpasses see table 6.1.

6.2 Observations

6.2.1 SAR observations

We used digital number (DN) backscatter observations from SAR images obtained by the COSMO-SkyMed (CSK) constellation at 6 times during the course of the flood. Table 6.1 shows the times and dates of the SAR observation overpasses. The SAR data we used is processed to level 1C-GEC (Garcia-Pintado et al. [2015]) and comprises digital number (DN) values. Each image covers a larger area than our subdomain, and DN measurements are at a higher resolution (2.5m) than our inundation model (25m). We therefore used ENVI image analysis software to extract backscatter DN measurements co-located with model cell centres. An alternative approach would be to take an average of the measured DN values covered by each model cell. Using cell centre DN values means that only one out of every 10 DNs in the x and y directions is used. The observations are therefore thinned at this stage, reducing the risk that there will be correlations in the observation errors unaccounted for in the data assimilation algorithm. Such correlations between observation errors are found to have longer than expected lengthscales in Waller et al. [2018].

We only used the first six of the images for our data assimilation experiments. We were unable to extract the observations needed from the seventh image as there were insufficient wet pixels. See section 6.2.2 for more details.

6.2.2 Extraction of wet/dry distributions

In order to use our backscatter operator we require histograms of wet and dry pixel DN or backscatter values for each image. In the work of Giustarini et al. [2016], backscatter values in dB are used to identify wet and dry areas. To get from DN to backscatter in dB it is necessary to first calibrate the DN values using a radiometric calibration constant; backscatter values are therefore usually a more expensive product to purchase from SAR data providers. The backscatter value for each cell in dB, b, is then given by $b = log_{10}(DN_c)$, where DN_c is the radiometrically calibrated DN value. Here, we did not apply radiometric calibration and instead used raw DN to extract histograms of wet and dry pixels. We did work with values of $log_{10}(DN)$ in order to be able to discriminate better between the values. Using the DN rather than the backscatter values means that it is not possible to compare values between images. This would be important if we were employing change detection techniques to identify wet/dry areas (as in e.g. Hostache et al. [2012]), but this is not the case here.

Figures 6.6 and 6.7 show an example of an extracted backscatter map and resulting histogram for the overpass on 28th November 2012.

Figure 6.6 shows the $log_{10}(DN)$ values obtained from the 2nd SAR satellite overpass.

Timestamp (UTC)	**Pass**
27/11/12 19:20	Descending
28/11/12 18:01	Descending
29/11/12 18:20	Descending
30/11/12 19:32	Descending
01/12/12 05:38	Ascending
02/12/12 05:56	Ascending
04/12/12 18:14	Descending

Table 6.1: Times of Cosmo-Skymed satellite overpasses.

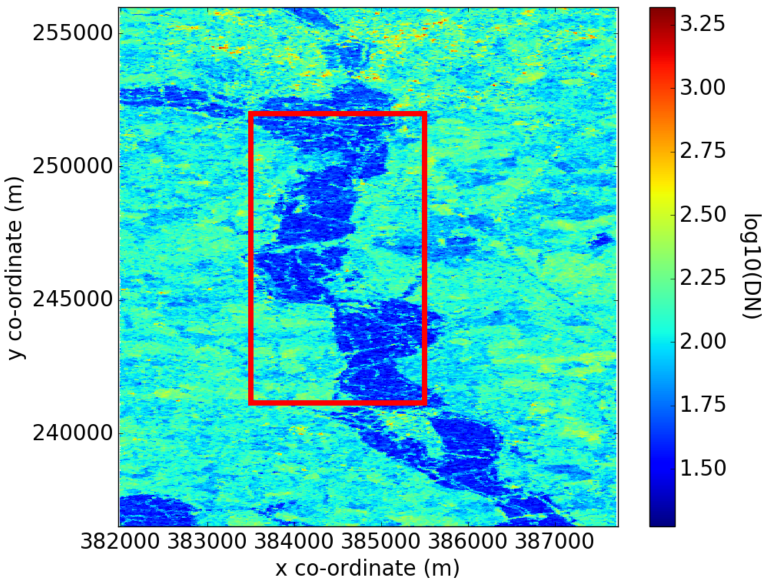
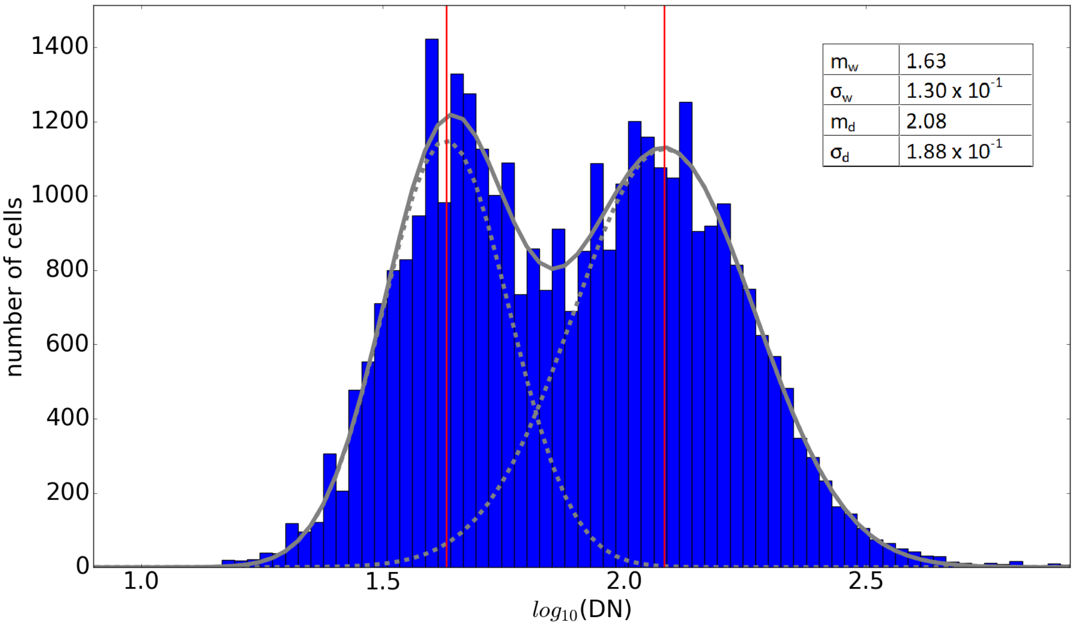


Figure 6.6: $log_{10}(DN)$ values for SAR image obtained on 28/11/2012. Red rectangle shows the area used to create a histogram of $log_{10}(DN)$ values.

The values inside the red rectangle were used to create the histogram shown in figure 6.7. Two different distributions of $log_{10}(DN)$ values can clearly be seen in figure 6.7; the dotted grey line shows two Gaussian distributions fitted to the peaks using python fitting algorithm scipy.optimize.curve_fit. The fitting algorithm uses a non-linear least squares method to generate the curves. The method was moderately sensitive to initial 'guess' data for the Gaussian distribution parameters; these were chosen by eye for each distribution. The solid grey line shows the sum of the two distributions. We use only the values inside the red rectangle in figure 6.6 so that the wet and dry distributions can be separately identified. There are many more dry pixels than wet in the whole image, meaning that a histogram of all the values shown in 6.6 is dominated by dry pixels. It is then not possible to fit a Gaussian distribution to the wet pixel distribution. We chose the area used for making the histograms in this chapter by trial and error, following Chini et al. [2017] in making sure that the Asham D criterion for separation of the Gaussian distribution peaks was met for each image. The quality of fits of the distributions would likely be better if we used a more complex method for extracting the cells to use in a histogram, as in Chini et al. [2017], since this would result in identifying a greater number of cells and therefore a larger number of values in each histogram. In these experiments we are more interested in the observation operator and data assimilation approaches.

Unfortunately in the case of the seventh image we were unable to extract a histogram



m_w	1.63
σ_w	1.30×10^{-1}
m_d	2.08
σ_d	1.88×10^{-1}

Figure 6.7: Histogram of a selection of $log_{10}(DN)$ values (see rectangle in figure 6.6) for SAR image obtained on 28/11/2012. The blue bars show the number of cells in each bin. The dotted grey lines show the two Gaussian fitted distributions, and the solid grey line shows their sum. The inset table contains the mean and standard deviation values for the wet and dry pixel distributions; m_w and m_d are the mean wet and dry $log_{10}(DN)$ values respectively (shown with vertical red lines), σ_w and σ_d are the corresponding standard deviations.

showing two peaks. This is because the water was back in bank at this point and extracting DN values at model cell centres produced only a very small number of wet cells. We therefore carried out an assimilation at the first six times detailed in table 6.1.

6.2.3 Observation thinning and quality control

As stated in section 6.2.1, the SAR image DN values are thinned to match the resolution of the inundation model. We then thinned the observations further before assimilating them. We did this so that

- we further reduce the chance of spatial correlations between observation errors
- the assimilation of the observations is less computationally costly
- we reduce the risk of overfitting to the observations in the assimilation

We thinned the observations by taking every second observation in the x direction and every 10th observation in the y direction. We thinned in this way in order to try to maximise the chance of locating flood edges associated with the Severn river. This step reduces the number of observations from $\sim 180,000$ to ~ 9000. We also reduce the number of observations of (dry) pixels far from the river, since these do not add any useful information to the assimilation. We then quality control the observations based on their $log_{10}(DN)$ value and the fitted distributions shown in 6.7, rejecting any observations which are far from either of the distribution means. We also discard observations with very high or low $log_{10}(DN)$ values as instrument outliers. We reject any observation, y_{DN} which does not satisfy

$$m_w - 3\sigma_w \leq y_{DN} \leq m_w + \sigma_w \tag{6.2.1}$$

or

$$m_d - \sigma_d \leq y_{DN} \leq m_d + 3\sigma_d \tag{6.2.2}$$

These procedures result in a set of ~ 5000 observations per image.

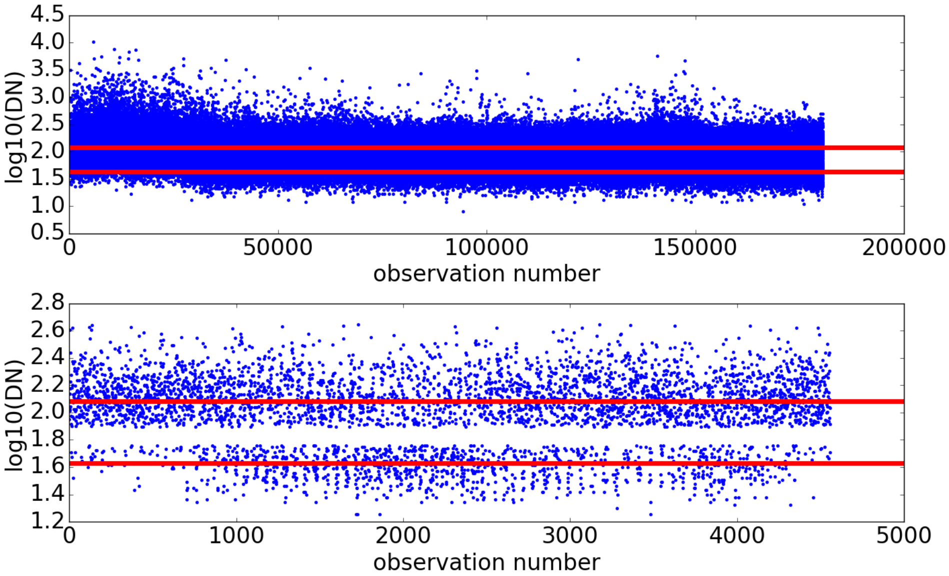


Figure 6.8: Top panel shows observation values taken from SAR image on 28th November 2012 (~ 180000); bottom panel shows the subset of observations used in assimilation after thinning and quality control (~ 5000). Horizontal lines show the values of m_w and m_d.

Figure 6.8 shows the observation values from the SAR image obtained on 28/11/2012 (top panel) with the subset of observations used in the assimilation (bottom panel). The figure shows that a large number of observations which fall between m_w and m_d in 6.7 are removed; This is desirable as these observations are the most uncertain in terms of which distribution they belong to.

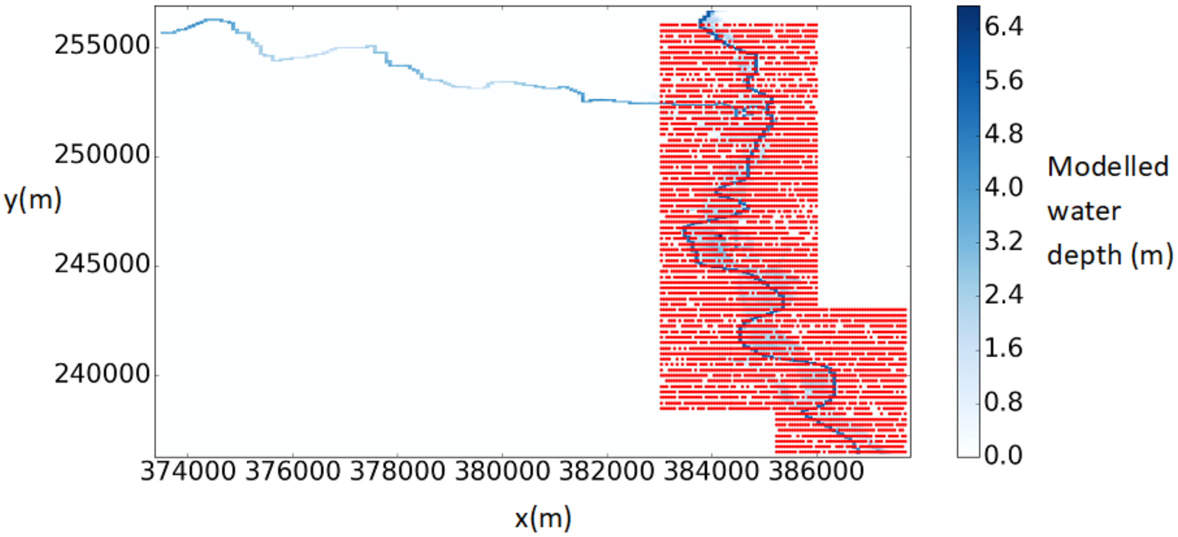


Figure 6.9: Spatial locations of thinned and quality controlled observations used in assimilation. The mean forecast (pre assimilation) water depth predictions at the time of the second overpass are shown in blue. Backscatter observation locations for the second SAR image are shown as red dots.

Figure 6.9 shows the locations of the observations used in the assimilation for the SAR image dated 28th November 2012. The image covers the same area as shown figure 6.1 but is stretched in the x direction in order to better show the locations of the observations.

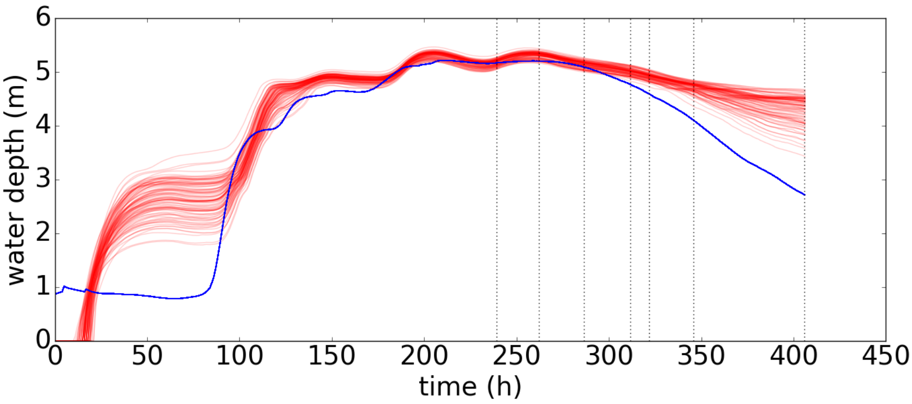
6.3 Experimental design

We ran an assimilation experiment using the inundation model, inflows and observations as described in sections 6.1 and 6.2. At the time of each available SAR image we assimilated $\sim$ 5000 DN observations using an ETKF as described in section 4.3.2.1 and used our novel observation operator as described in section 5.4.3. For each SAR image, values of m_w, m_d, σ_w and σ_d were obtained as described in section 6.2.2. Using our new observation operator, these values were applied to each ensemble member at the corresponding assimilation time in order to generate an ensemble of forecast observations. We also ran an 'open loop', no assimilation experiment for comparison with the analysis solution to allow us to assess the impact of the assimilation. Our forecast ensemble comprises 100 members in each case.

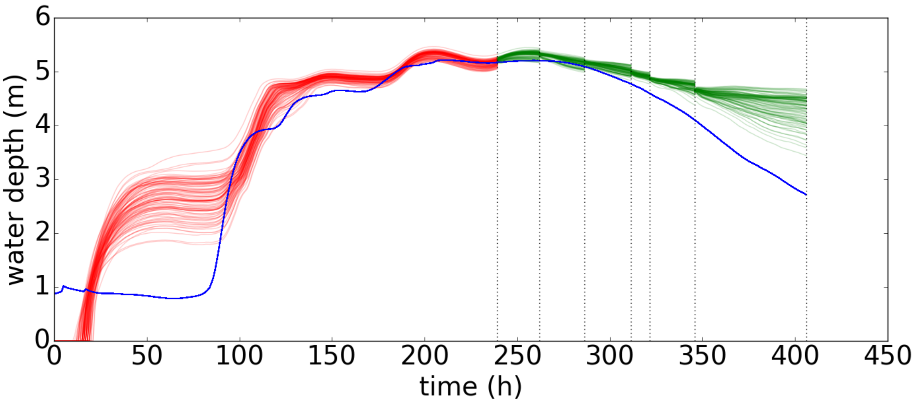
6.4 Results

In section 6.4.1 we examine the effect of the assimilation on the forecast at Saxon's Lode gauge and compare measurements of depth in the river with our model predictions. In section 6.4.2 we examine the effect of the assimilation on the spatial pattern of flooding, using the binary flood map measures described in section 2.5.

6.4.1 Depth at Saxon's Lode gauge



(a) Openloop case.

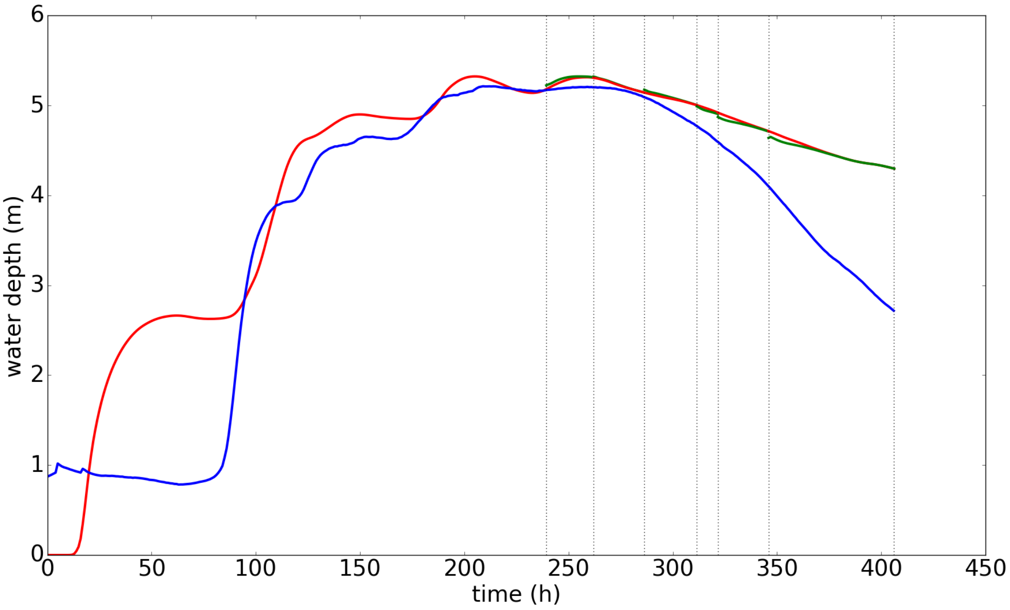


(b) Case with assimilation

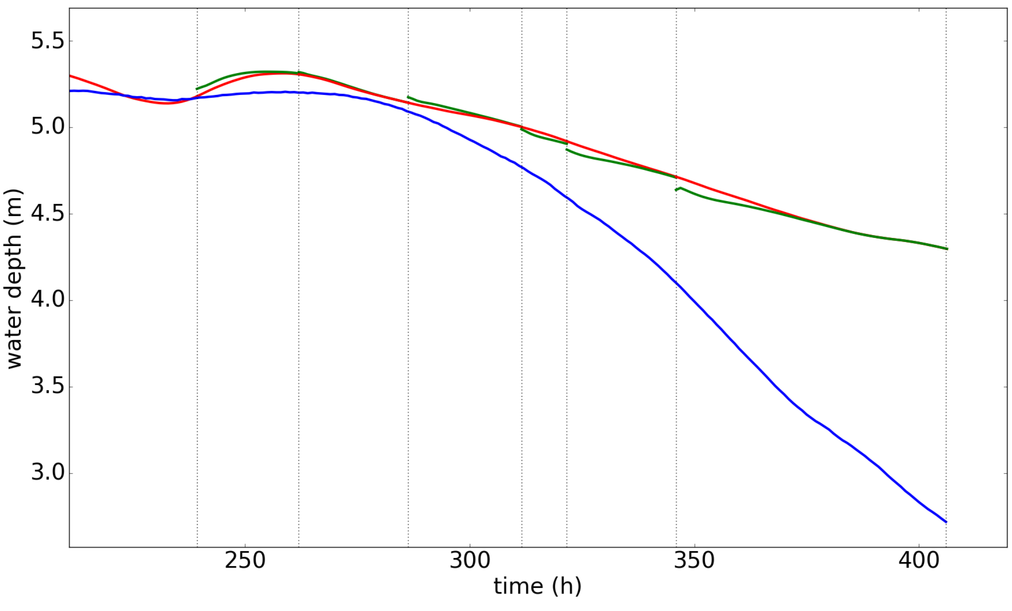
Figure 6.10: Measured and modelled water depths at Saxon's Lode. Solid blue line shows measured data from the gauging station, red and green lines show ensemble member predictions. Dotted vertical lines show times of satellite overpasses.

Figure 6.10 compares the measured depths at Saxon's Lode with depths predicted by the model at the gauge location. Figure 6.10a shows the depths predicted by the open loop (no assimilation experiment) and figure 6.10b shows the depths produced by assimilating SAR DN observations at the overpass times shown. In each case the ensemble shows a greater range of values for low, in-bank depths than for deeper water predictions. This is because once the water is out of bank the physically possible range of depths becomes very small. Figure 6.10 shows raw depths for the ensemble members; we note that the depths predicted by the model are not expected to be exactly the same as the measured values due to the parameterisation of the river channel in our hydrodynamic model. An alternative would be to compare EA reported and modelled flow rates but since the reported flow rates are calculated using a rating curve and not likely to be reliable at high flows we do not show these here.

Comparing figures 6.10a and 6.10b shows that the effect of the assimilation is small in terms of changing the depth in the river. Physically we expect only a small variation in river depth when the river is in flood; even a small change to the in-river depth can have a large effect on the resulting flood volume.



(a) Mean depths.



(b) Magnified view of mean depths.

Figure 6.11: Measured and mean predicted water depths at Saxon's Lode. Solid blue line shows measured data from the gauging station, red lines show open loop case, green lines show case with assimilation. Dotted vertical lines show times of satellite overpasses. Figure 6.11a shows the mean depths at all times, figure 6.11b is a magnified view of the depths over the time of the assimilations.

Figure 6.11 shows measured depths at Saxon's Lode with the mean depth predicted by the open loop and with-assimilation ensembles. Comparison of the open loop and assimilation mean depths shows a small and short-lived update to the mean predicted water levels. Figure 6.11b shows more clearly that at the first three overpass times the assimilation update acts to slightly increase the mean water depth at Saxon's Lode; at the second three assimilation times the mean water depth is updated to a shallower level. The fact that the update to the water depths is short-lived is expected and agrees with the work in chapters 4 and 5 for state-only estimation. It is also expected that the update is small at this location. This occurs because the physically possible range of river depths is very small once the river is out-of-bank; extra water will flow overbank to cause more flooding rather than increase river depth. There is therefore only a small range of possible water depths available to ensemble members, and a locally narrow ensemble spread leads to a small local update (see equations 5.3.12 and 5.3.13).

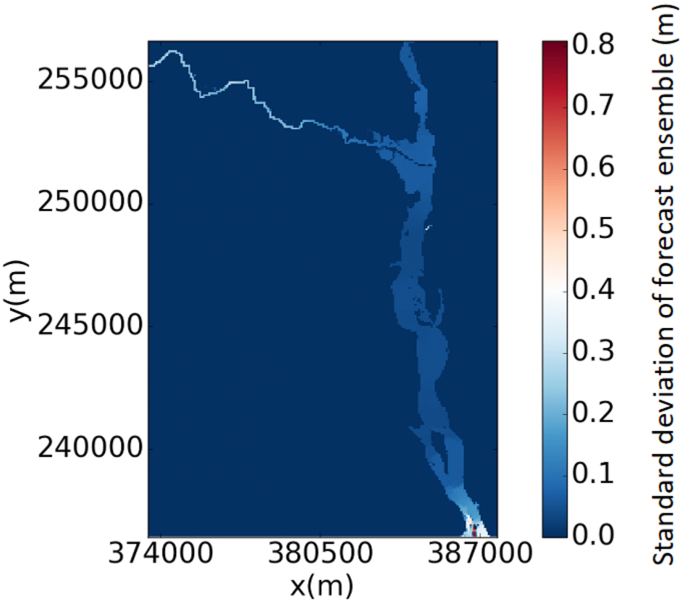


Figure 6.12: Square root of forecast variance in each cell of the domain at second overpass

Figure 6.12 shows the standard deviation of the ensemble of water depths (i.e. the square root of the variance) in the domain before the assimilation at the second overpass time. Larger values of standard deviation/variance correspond to larger spread in the ensemble, higher uncertainty in the model prediction and permit larger updates to the mean. We note that the large uncertainty at the downstream boundary is due to the fact that this is the edge of the numerical domain. An alternative approach would have been to extend the domain beyond the domain of interest in order to minimise edge effects.

We see in figure 6.12 that there is low variance (low uncertainty) in the river channels where the river is in flood, and particularly at the Saxon's Lode location (see figure 6.1), leading to small updates to water depths at these points. We note that the variances look smooth; covariance plots (not shown) also look smooth and this justifies not using any localisation in these experiments. Larger domains with larger state vectors may require covariance localisation in order to avoid spurious updates, but care would need to be taken to preserve physical structures in covariances (e.g. see Garcia-Pintado et al. [2015] in which an along-network localisation scheme is implemented).

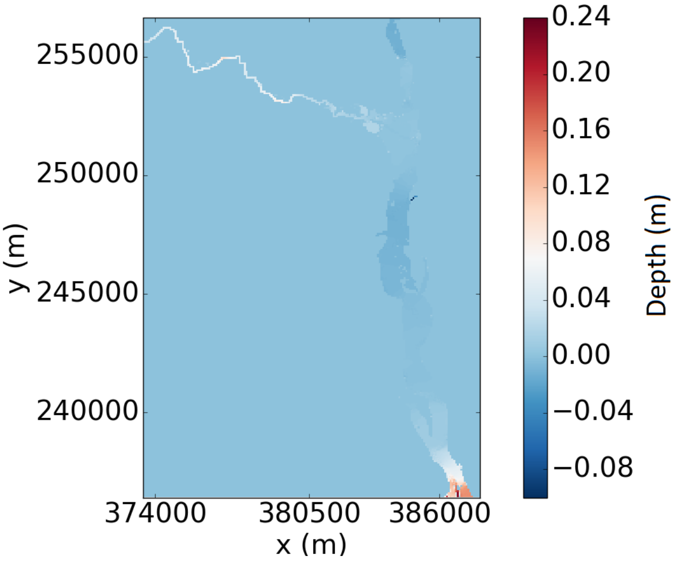


Figure 6.13: Mean analysis water depths minus mean forecast water depths at second overpass

Figure 6.13 shows the difference between the mean forecast and the mean analysis (analysis increments) at the second overpass time. The pattern of increments is clearly more complex than those shown in figure 4.7c for the simple domain used in chapters 4 and 5. This is because the more realistic topography used in this chapter leads to more complex correlations and covariances between errors in water depth at physical locations in the domain. Figure 6.13 shows updates to water levels throughout the domain, including in the channel at some distance from the assimilated observations (see figure 6.9). Updating of water levels at locations far from observations was also seen in the work in chapters 4 and 5 in the simple domain. Updates at locations far from observations can sometimes indicate that localisation is needed. Here, however, the increments are smooth, and locations with larger updates correspond to larger variances; comparing figures 6.13 and 6.12 we see that the patterns in these are very similar. We therefore conclude that the updates generated in our experiments are physically reasonable.

6.4.2 Spatial flood measures

6.4.2.1 Binary flood pattern measures

In section 2.5 we introduced contingency tables and eight binary flood pattern matching measures. Differences in these measures can be used to indicate the extent to which the assimilation of DN observations affects how well the model agrees with the observations. Here, we can only compare the goodness of fit to observations from the same images that we used in the assimilation as we only have one source of spatial flood information available; a completely independent set of observations from e.g. aerial photography would give us more confidence in our conclusions. However, when computing contingency table values (A,B,C,D) and the pattern matching measures we use a much larger set of observations than the thinned subset we used for assimilation; we use all of the observations at model cell centres rather than the thinned set of observations. We also quality controlled the validating observations slightly differently in order to reduce the number of false wet observations; these can be seen as darker blue patches/speckles in areas far from the river in figure 6.6. Such false wet observations do not affect the assimilation as there will be no update in cells that all the ensemble members predict to be dry (see the last paragraph of section 5.6.1.1). For calculating binary flood pattern matching we reject any validating observation, V_{DN}, which does not satisfy

$$m_w - 2\sigma_w \leq V_{DN} \leq m_w + \sigma_w \tag{6.4.1}$$

or

$$m_d - \sigma_d \leq V_{DN} \leq m_d + 2\sigma_d. \tag{6.4.2}$$

This results in $\sim 140,000$ observations (in a total of $180,550$ available cells) used for validation; we used ~ 5000 in the assimilation. Comparing equations 6.4.1 and 6.4.2 with equations 6.2.1 and 6.2.2 we see that we are using observations with higher uncertainty for the assimilation than we are for the validation. This is because the data assimilation can allow for uncertain observations whereas the validation measures used here assumes the observations are correct (they do not allow for uncertainty). We investigated excluding more or less observations for validation based on distance from the mean but although this resulted in the raw numbers in the contingency table being slightly different there was very little difference in terms of the resulting binary flood pattern measures.

Figures 6.14 to 6.19 show the binary flood pattern match measures, each of which is made up of a combination of A, B, C and D values as defined in section 2.5. The values of each of the eight measures are shown at each of the six assimilation times for the mean modelled forecast (pre-assimilation), mean modelled analysis (post-assimilation) and the open loop simulation. In all of the figures it is clear that the differences in these measures at each assimilation time between the forecast, analysis and open loop cases are very

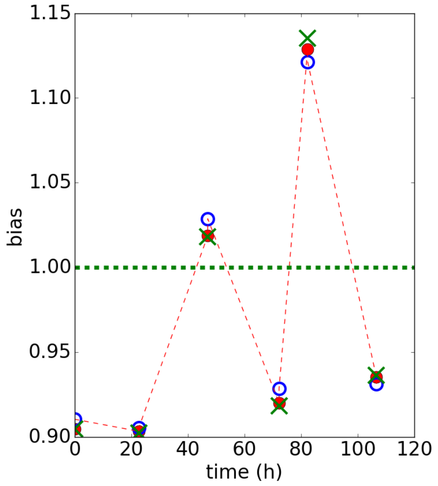


Figure 6.14: Bias at each assimilation time. Red circles are values for the mean of the forecast ensemble, blue circles are for the mean of the analysis ensemble, green crosses are for the mean of the open loop ensemble.

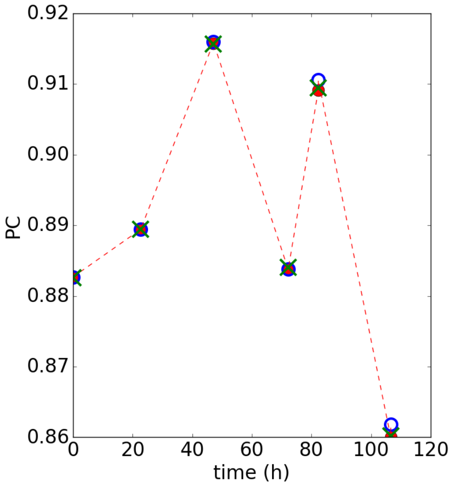


Figure 6.15: PC (proportion of cells correct) at each assimilation time. Symbols as in figure 6.14

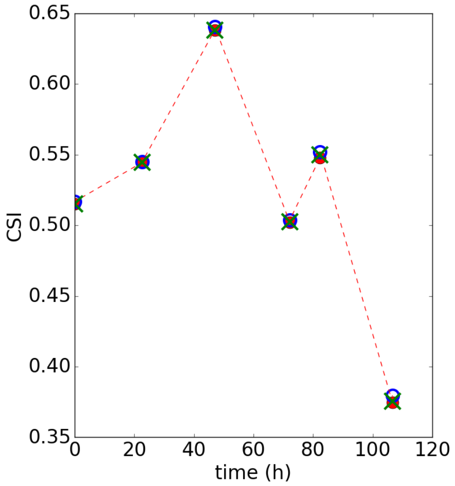


Figure 6.16: CSI (Critical Success Index) at each assimilation time. Symbols as in figure 6.14

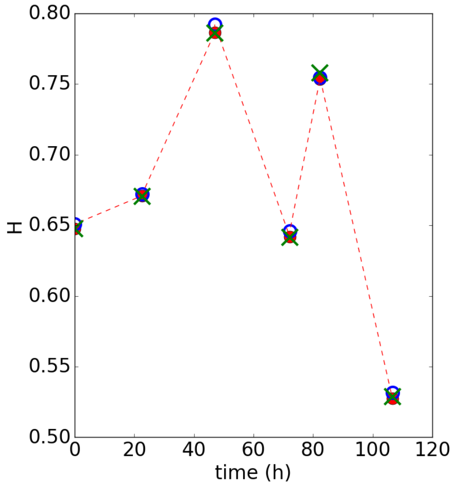


Figure 6.17: HR (hit rate for wet cells) at each assimilation time. Symbols as in figure 6.14

small. This is because there are a very large number of model/observation grid cells ($\sim 140,000$) and only a small proportion of these will change from wet to dry, or dry to wet during each assimilation provided the forecast is not very different to the observation.

Figure 6.14 shows that the forecast and open loop models underpredict the number of wet cells at two assimilation times (bias < 1), and overpredict the number of wet cells at the other two times (bias> 1). In figure 6.15 all of the PC (proportion correct) values are close to the optimal value of 1; at each assimilation time the value for the analysis prediction scores very slightly better than the forecast, indicating that the data assimilation is moving the analysis closer to the observations as required. The fact that the PC values are generally close to 1 might help to explain why we see only small

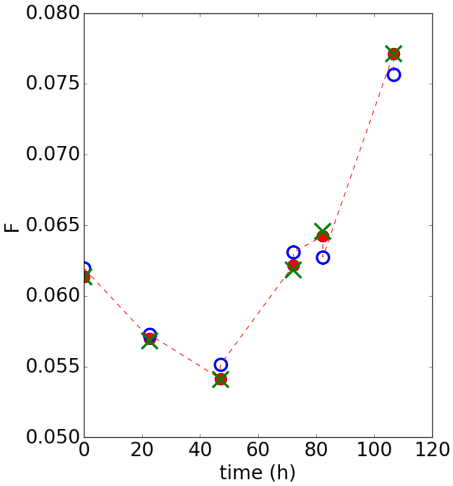


Figure 6.18: F (false alarms) at each assimilation time. Symbols as in figure 6.14

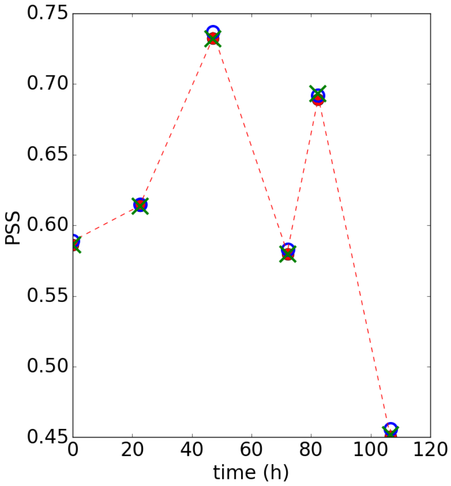


Figure 6.19: PSS (Pierce Skill Score) at each assimilation time. Symbols as in figure 6.14

improvements in the forecast at assimilation times: the match to the observations in terms of wet and dry pixels is already reasonably good. Figure 6.16 shows the CSI (critical success index) score; the CSI measures the proportion of the cells correctly modelled, excluding the D category. As expected, the CSI values are lower than the PC scores as we have excluded a large number of correctly specified dry cells in this measure. As in the case of the PC measure, the analysis CSI values score slightly higher than the forecast values at all times. The open loop score is very similar to the forecast score in each case and this agrees with our findings in chapters 4 and 5, where we saw short-lived observation impact and a rapid relaxation to the open loop case when only water levels were adjusted through the assimilation. Figure 6.17 shows the hit rate, i.e. the proportion of correctly identified wet cells. Again, all the values at each assimilation time are very similar, and the analysis scores slightly better than the forecast. The false alarm rate is the proportion of dry cells wrongly predicted to be wet; this shown in figure 6.18. The optimal value for this score is zero, and all of the values of this score are low. The open loop often has the lowest score, indicating that the assimilation is causing a small overprediction of wet cells relative to the open loop. The HR and F scores are combined in the PSS measure (Pierce Skill score), which is shown in 6.19. All of the PSS scores are greater than zero, indicating that there are more hits than misses for both the open loop and the forecast. Again we see that the analysis scores slightly better (closer to 1) than the forecast at each assimilation time, and the open loop score is similar to the forecast score, indicating short-lived observation impact. For the PSS score the open loop slightly outperforms the forecast and analysis at the 5th assimilation time; this could indicate that the long term impact of the assimilation is slightly detrimental to the forecast at some assimilation times.

Taken together, these results indicate that in general the assimilation is making a small improvement to the forecast at the time of the assimilation, i.e. the wet/dry cells predicted by the analysis are closer to all the observations (not just those used for the assimilation) than the forecast. There is no evidence that this improvement persists through the forecast step, since the (pre-assimilation) forecast does not tend to score more highly than the open loop; the open loop and forecast values are generally very similar, and in some cases the open loop scores better than the forecast. This short-lived observation impact agrees with the results shown in chapters 3 and 4.

6.4.2.2 A, B, C, D counts

In order to get a better understanding of the effect of assimilating observations, we can look directly at differences in the raw values of A,B,C and D at each assimilation time. These values are more sensitive to changes when examined alone than when combined into the measures we reported section 6.4.2.1. We can compare values of A,B,C and D between

- the mean forecast prediction and mean analysis prediction at each observation time; this will tell us whether the assimilation is moving the analysis towards the observations.
- the mean forecast and the open loop; this will tell us whether there is persistence in improvements through the forecast step.

A better forecast at given any time is likely to have larger values of A (more correctly identified wet cells). An increase in A for a given image is identical to a decrease in C as the total of $(A + C)$ is equal to the total number of cells observed to be wet and must therefore remain constant. A better forecast will also likely produce higher values of D (more correctly identified dry cells). An increase in D is the same as a decrease in B, since $(B + D)$ is constant.

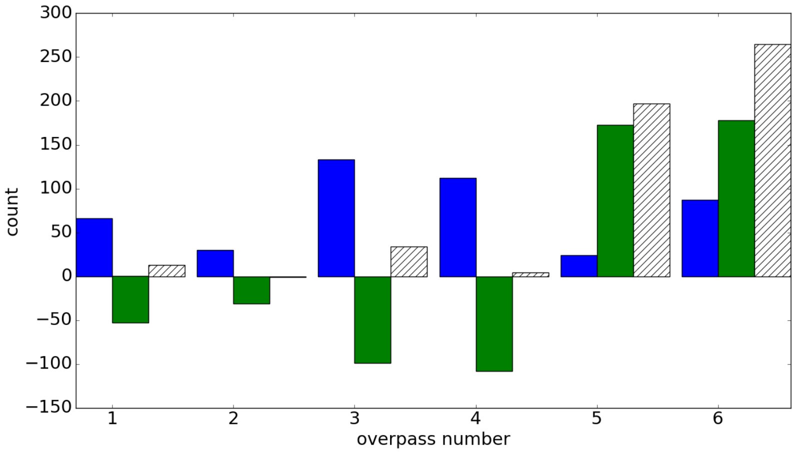


Figure 6.20: Bar chart showing differences in the values of A and D between the mean forecast and mean analysis at each assimilation time. Blue bars represent the change in A from forecast to analysis, i.e. the number of extra cells correctly identified as wet in the analysis compared to the forecast. Green bars show the change in D from forecast to analysis, i.e. the number of extra cells correctly identified as dry in the analysis compared to the forecast. The hatched white bar shows the net increase in the number of cells correctly identified by performing data assimilation (difference in A + difference in D.)

Figure 6.20 shows differences in the values of A (blue bar) and D (green bar) between the mean forecast and mean analysis at each assimilation time. Positive values represent improvement due to assimilation. When the blue bar shows a positive number, there are more cells correctly identified as wet in the analysis compared to the forecast. When the green bar shows a positive number, there are more dry cells correctly identified in the analysis then in the forecast; when the green bar is negative there are fewer correctly identified dry cells in the mean analysis compared to the mean forecast. The hatched white bar is the sum of the blue and green bars (difference in A + difference in D), we consider this to be a measure of the 'net' improvement of the analysis compared to the forecast. A positive value for the hatched white bar means that of the cells which are changed in the assimilation (wet to dry or dry to wet), more have been correctly updated than wrongly updated.

Figure 6.20 shows a modest overall improvement in the analysis compared to the forecast at overpasses 1, 3 and 4 (hatched white bar has positive values). In overpasses 1, 2, 3 and 4 the analysis correctly identifies more wet cells, but also incorrectly identifies a similar number of dry cells as wet; this gives a small net improvement for overpasses 1, 2 and 4 and a very small decrease (1 model cell) in the quality of the prediction at overpass 3. For overpasses 5 and 6 the analysis again correctly identifies more wet cells than the forecast, but also increases the number of correctly identified dry cells. The overall improvement score is therefore higher for these two times. In all cases, the percentage of observed cells in which the forecast is changed from wet to dry or vice versa by the assimilation is small. The maximum percentage of cells changed is for the sixth overpass and is approximately 0.2% of the total number of observed cells. Despite the small percentage change, this represents an area of over 165,000m^2 in which the forecast is changed from wet to dry or vice versa.

One reason that the assimilation has more of an impact at overpasses 5 and 6 may be due to the inability of the hydrodynamic model to ‘dewater’ the floodplain as the flood recedes. Pools of water on the flood plain that become hydraulically disconnected from the flood water cannot be removed by the hydrodynamic model, but could be removed in an assimilation step. This would explain the larger numbers of dry cells being correctly identified at overpasses 5 and 6.

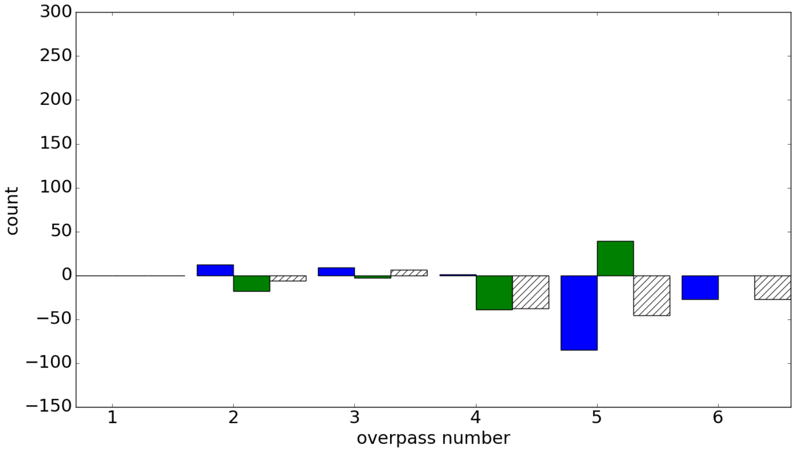
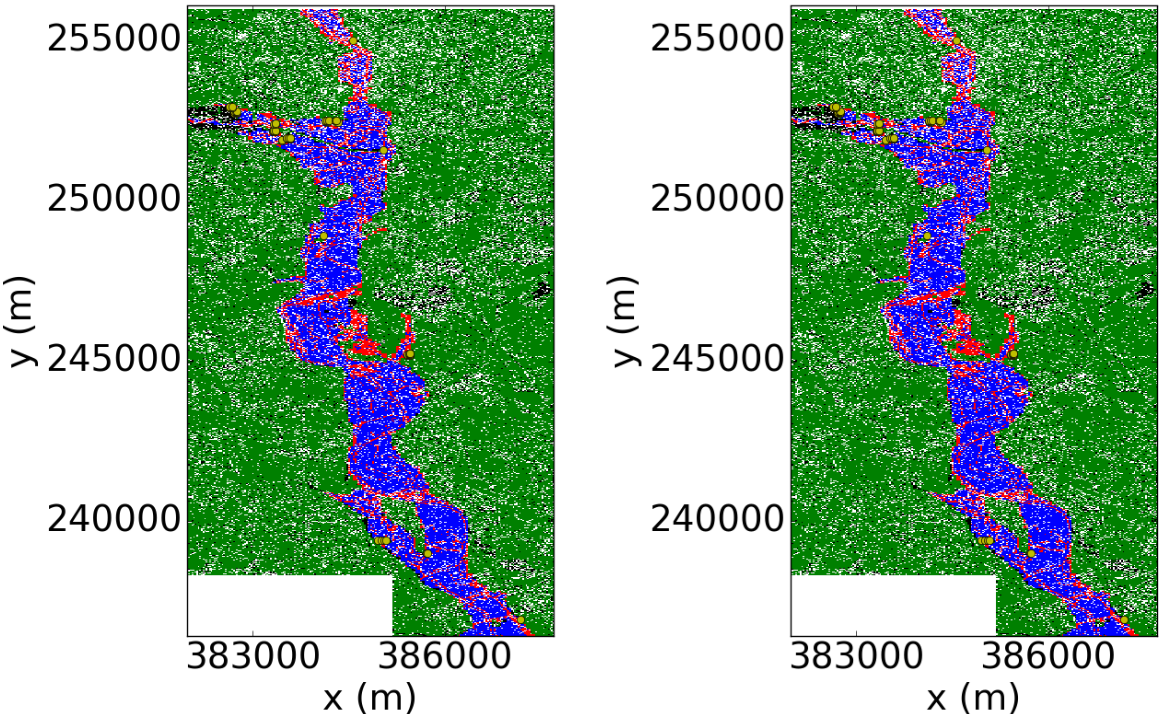


Figure 6.21: Bar chart of data comparing the mean forecast and mean open loop predictions. Blue bars represent the difference in A, green bars the difference in D and hatched white bars the net difference in the number of correctly identified cells in the (pre-assimilation) forecast compared to the open loop. When the hatched white bar is greater than zero we consider the forecast to be better than the open loop as it has a higher net number of correctly identified cells.

Figure 6.21 shows the differences in A and D between the mean open loop prediction and the mean forecast (pre-assimilation) prediction. At the first overpass time the forecast and the open loop solution are the same as no assimilation has yet been carried out; the difference at this time is therefore zero. We see that the differences between the open loop and the pre-assimilation forecast at other times are small, and much smaller than the differences between the forecast and analysis shown in figure 6.20. The percentage of observed cells which are different between the forecast and open loop cases is very small, with a maximum at the 5th overpass time of less than 0.1%. This indicates that the improvements seen at the assimilation time disappear during the forecast step. Since we are only correcting water levels, this is an expected result, and agrees with our previous results in chapters 4 and 5, where we saw the assimilated solution relax quickly back to the open loop solution following an assimilation.

At overpasses 4, 5 and 6 we see that the forecast is not as good as the open loop in terms of the difference in A and D, though the difference is small. This is particularly the case at overpass 5, where the forecast is wrongly predicting many more wet cells than the open loop. We believe that this may be due in part to the way that we have treated the terrain in our model. We have split each 75m topography cell into nine 25m x 25m cells, each of which has the same elevation. If the assimilation acts to change one of the cells in a group of nine from dry to wet, the water will naturally spread into the other eight cells once the simulation restarts, even if these cells are observed to be dry. This effect is likely to taper away throughout each forecast step as water drains back towards the river, but we note that the shortest time between overpasses is from 4 to 5 (∼ 10hours), so this might explain why the effect is more pronounced at this time.



(a) Forecast.

(b) Analysis.

Figure 6.22: Spatial plot of cells assigned to category A(blue), B(red), C(black) and D(green) at overpass 2. Yellow circles show the cells at which the assimilation changed cells from wet to dry, i.e cells which are different between figures 6.22a and 6.22b. White cells are those with no observations due high uncertainty, or in the case of the rectangle at the bottom left corner because of a lake visible in the observations but not in the model simulation.

Figure 6.22 shows an example of the spatial variation in A,B, C and D for the second overpass time. The yellow circles show the locations of clusters of cells which are updated from wet to dry or vice versa during the assimilation. They are spread throughout the domain; this spread in locations is common to all overpass time. It is notable that in both the forecast and analysis case there are cells registering as red (B) overprediction following the course of the river Severn, and this is also true for all of the overpasses. These are likely cells in which the river is really flooded but vegetation (e.g. trees and large bushes) growing at the sides of the river are causing the SAR to misclassify them as dry. The model is therefore likely correct at some of these cells rather than the observation. Dry observations in such cells, where all the ensemble members make a wet prediction, will not be updated due to a false dry observation, and this is one of the potential strengths of this operator. An alternative way to deal with such cells is to do a much more thorough wet/dry classification before assimilation and/or validation using clustering techniques (see e.g. Mason et al. [2012] and section 5.7).

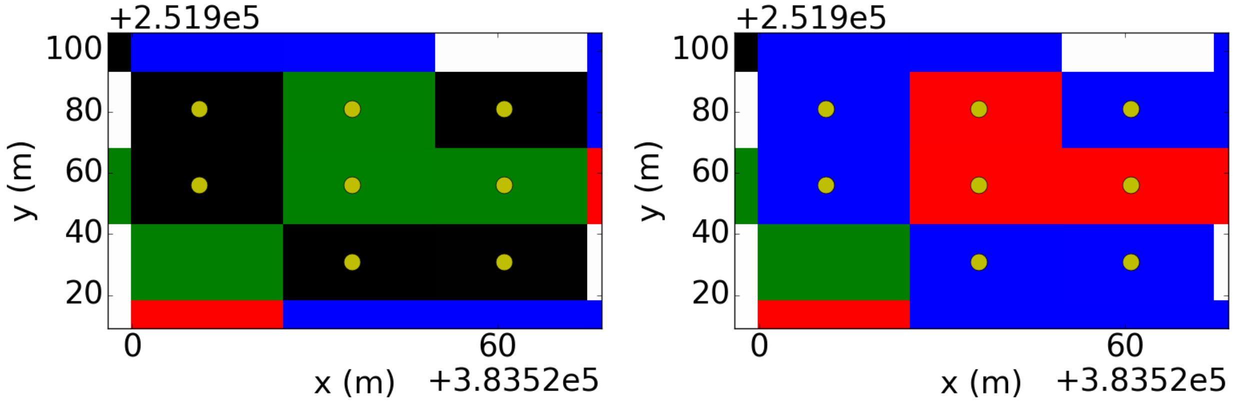


Figure 6.23: Example of update to small number of cells at step 2. Left hand panel shows forecast classifications and right hand panel shows analysis classifications. Colours as in figure 6.22.

Figure 6.23 shows a group of cells, eight of which were updated at the time of the second overpass, i.e. a subset of the cells shown in figure 6.22. The left panel shows the forecast (pre-assimilation) and the right panel the analysis (post-assimilation). The cells which were updated are marked with yellow circles. Five of the eight cells which are updated change from black (C, underprediction) to blue (A, correctly modelled as wet). However, a further three cells change from green (D, correctly modelled dry) to red (B, overprediction). This example shows that cells near each other are likely to all be moved in the same direction (wet to dry or dry to wet) by the assimilation. This results from strong spatial state error correlations between neighbouring cells and means that a desirable increase in A following an assimilation is often accompanied by an undesirable decrease in D as seen in figure 6.20.

In figure 6.22, a total of 61 cells are changed in the assimilation, i.e. are updated due to observations from the 2nd overpass and change from figure 6.22a to figure 6.22b. This is a very small proportion of the total observed cells ($\sim 0.04\%$), but still represents an area of $61 \times 25 \times 25 = 38,125m^2$ in which the prediction is changed from wet to dry or vice versa. Additionally, water depths are changed throughout the domain due to forecast error correlations between depths in different places.

6.4.3 Comparison to results from related studies

The SAR images used here were also used for assimilation experiments in Garcia-Pintado et al. [2015]. We used our new backscatter observation operator, while the authors of Garcia-Pintado et al. [2015] used water level observations (WLOs) and a 'nearest wet pixel' type approach; the domain used in our work is a subdomain of the one used in Garcia-Pintado et al. [2015] and in both cases an ETKF was used. Predicted flow rates and depths are compared with gauged values in Garcia-Pintado et al. [2015] in order to assess the effectiveness of the assimilation when estimating water depths along with combinations of inflow, friction parameters, bathymetry and downstream river bed slope. Corrections to gauged water depths are in some cases larger in Garcia-Pintado et al. [2015] than those seen in this study, but this is not necessarily due to differences in observation operator approach as there are a number of other differences between the studies. In this chapter we only update water levels, whereas in Garcia-Pintado et al. [2015] other model inputs are also updated. We use gauge reported inflows to drive our model, whereas outputs from a hydrological model are used in Garcia-Pintado et al. [2015]. Additionally, the studies use different hydrodynamic models, with different approaches to representing channel shape.

Another related study is Hostache et al. [2018], in which a different series of SAR images (from the Envisat Wide Swath Mode imagery archive) is used in assimilation experiments for four different flood events in the same spatial domain used in Garcia-Pintado et al. [2015]. In Hostache et al. [2018], the authors use the SAR images to create flood probability maps, and use these as observations. The observation operator therefore takes wet/ dry predictions from each forecast ensemble member, and combines these into a forecast-equivalent probability flooding per model cell. A particle filter technique is used to combine the observational and forecast probabilities at update times. The study in Hostache et al. [2018] uses output from a hydrological model to create inflows; perturbations in rainfall inputs to the hydrological model create an inflow ensemble as in Garcia-Pintado et al. [2015]. Since a particle filter technique is used in Hostache et al. [2018], the inflow ensembles are effectively updated at each assimilation time. This means we cannot directly compare the results with the work in this chapter (where we only update water levels), but there are nevertheless some interesting comparisons.

In terms of water depths, the results in Hostache et al. [2018] show larger assimilation increments than we see here. This could be due to the fact that the inflow is updated Hostache et al. [2018], or it could be because we are comparing different flood events and images between the studies. Comparison of our results with binary flood match measures reported for one assimilation in Hostache et al. [2018] show some similarity; in both studies assimilation results in a small, desirable increase in A (number of cells correctly predicted wet). There is an increase in A of 108 cells, ($\sim 0.3\%$ of total cells) in Hostache et al. [2018] c.f. a maximum of 133 ($\sim 0.1\%$ of total observed cells) in our study. The increase in A due to assimilation reported in Hostache et al. [2018] is associated with a smaller increase in D (cells wrongly predicted to be wet), which we also observed at several assimilation times in our experiments and attributed to situations like the one illustrated in figure 6.23. The authors of Hostache et al. [2018] report no change in the PC measure (to 3d.p.) due to assimilation, and a small increase in the CSI measure with assimilation. Our results generally show the same pattern as this. An expected difference between the results in Hostache et al. [2018] and the results in this chapter is that in the former case there is good persistence in the forecast improvement, i.e. the observation impact is longer-lived. This is due to the fact that inflow is updated in Hostache et al. [2018], whereas in this study we are updating only water levels.

6.5 Conclusions

We have presented a number of ways of measuring how the assimilation of SAR data using our novel observation operator updates model forecasts. We conclude:

- we faced a number of challenges relating to the hydrodynamic model we used. The representation of river geometry (wide, shallow channels) may have caused some issues for predicted flow rates and hence for inundation prediction and this is likely why binary measures such as the CSI give relatively low values. The lack of sub-grid river channel representation also led to the need for fine computational mesh resolution and increased computational expense. Despite this, the new observation operator was successfully applied to forecasts.

- predicted depths at the Saxon's Lode gauge show small changes at assimilation times. This is due to small ensemble spread (low uncertainty) at this position in the domain, particularly when the river is out of bank. These small changes in water depth in the river can correspond to large differences in the volume of water on the floodplain.

- a series of binary flood pattern measures show small positive changes to the forecast at assimilation times, broadly in line with results from a related study, Hostache et al. [2018].

- our results show short-lived observation impact in agreement with the work in chapters 4 and 5 when only water levels are updated; this effect has also been seen in other studies (e.g. Lai and Monnier [2009], Matgen et al. [2007b] and Schumann et al. [2009]).

- looking directly at values of A, B, C and D shows that at every overpass time, assimilation of SAR observations using our new approach increased the number of correctly identified wet cells.

- for overpasses 1,2,3 and 4 the update also increased the number of wrongly identified wet cells, i.e. generated more false positives; this effect was also observed in Hostache et al. [2018].

We conclude we have successfully applied our novel observation operator to a real case study. We have shown that our observation operator generates updates to water levels throughout the domain, and is able to change the numbers of predicted wet and dry cells. The results show small improvements to model prediction when using our new observation operator and an ETKF. We believe that the improvements would be larger in cases where the model prediction is more different to the observations, or where the topography is defined to be different in each cell. We see a short-lived improvement in the forecast; this is in agreement with our results from state-only assimilation in synthetic experiments in chapters 4 and 5. This work has shown that our new observation operator shows potential to be useful in data assimilation for inundation forecasting. Further work is required to see how the operator compares against other observation approaches when using real SAR data; it would also be interesting to test performance of the new observation operator with a different hydrodynamic model.

6.6 Chapter summary

In this chapter we have shown that we can successfully apply our novel observation operator approach to a real case study, using real SAR data. We have demonstrated a benefit to the forecast using an ETKF with our new approach.

Chapter 7:

Conclusions and Future work

7.1 Conclusions

This thesis has considered aspects of how we can best use information from satellite-based SAR instruments in data assimilation for river inundation forecasting. We have considered various ways in which we can maximise the impact of SAR observations. We have carried out synthetic data assimilation experiments, designed a new observation approach and applied our novel observation operator to a real flooding case study. The main conclusions, which answer the three questions posed in chapter 1, are:

1. **How does estimation of the channel friction parameter affect observation impact in data assimilation for inundation forecasting?**
    - In chapter 4 we showed that it was possible to retrieve the correct value for the channel friction parameter in our synthetic experiments in a simplified domain, and that correcting the parameter in this way significantly increased the impact of the observations compared to state-only estimation.
    - The results in chapter 4 showed that inflow error and channel friction parameter error are interdependent in our idealised domain, making the two sources of error difficult to separate out; this demonstrates that equifinality can be an issue for hydrodynamic models even in a simple system.
    - We discovered an initialisation shock in our idealised synthetic experiments in chapter 4 when we assumed zero velocity for all water at restart times following assimilation. We showed that using pre-assimilation velocities worked as well as updating flow rates as part of the assimilation and is computationally cheaper.

2. **Can we design and implement a new method of extracting observational information from a SAR image for use in data assimilation?**
    - In chapter 5 we designed and implemented a new observation operator which directly uses SAR backscatter values as observations in data assimilation, and showed that the method can successfully update a model forecast and the value of the channel friction parameter in synthetic experiments. Such an approach has the potential to increase the number of observations which can be used from each SAR image.
    - We showed in chapter 5 that different approaches to SAR observations and different observation operators can produce substantially different observation impacts in synthetic experiments. We explained these differences by illustrating the physical mechanisms responsible for the updates in the different approaches. Better understanding of the mechanisms which generate assimilation increments will be important in designing future approaches to inundation forecasting.
3. **Can we apply our new observation approach to a case study using real topography and SAR data?**
    - We successfully applied our novel backscatter observation operator to a real case study in chapter 6, using a series of SAR images showing flooding near Tewkesbury in the UK in the winter of 2012. We used a series of binary flood match measures to demonstrate benefit to the forecast of updating water levels in this way. The method shows great potential for use in real-time flood forecasting systems.

7.2 Future work

The work in this thesis has provided answers to the research questions posed in chapter 1. There are many interesting related topics which require further research, and we summarize some ideas for future work in the remainder of this section.

7.2.1 Extension of observation operator case study

Extension of the work in chapter 6 to compare the results already obtained with the new observation operator with those using the ‘nearest wet pixel’ approach would be extremely useful in further analysing how well the new operator can perform. Further understanding of the observation operator approach (in terms of, for example, number of observations used, location of observations) would also be interesting. Alternatively, direct comparison of results from another case study using a WLO approach (as in Garcia-Pintado et al. [2015]) and the backscatter approach would be highly beneficial in understanding the differences in the updates generated by the two approaches in realistic use. This would require running experiments using the same hydrodynamic model, model domain and SAR images in each case.

The work in chapter 6 could be extended to test whether updating the channel friction parameter in the case study in chapter 6 would increase the time over which we see observation impact, i.e. would increase persistence in the forecast improvement as seen in the twin experiments in chapters 4 and 5. There is also scope for investigating whether we could show interdependence in inflow and friction parameter errors as in the synthetic experiments in chapter 4. More broadly, it would be very useful to try to understand interdependences in errors in other model inputs including (as well as inflow and channel friction parameter), channel depth, width and shape, floodplain friction parameter, downstream slope, topography resolution and model compute grid resolution. All of these model inputs are likely to be correlated to some extent, and it would be really interesting to design and run a set of data assimilation experiments in which these correlations be explored further. This would potentially be valuable to the hydrodynamic community, and could feed back into the design and calibration of hydrodynamic models.

7.2.2 Data assimilation approaches

There are a number of questions which remain unanswered regarding the performance of different data assimilation approaches in the context of inundation forecasting. In particular, it would be interesting to compare the updates generated by an ETKF and the 3D-Var assimilation method introduced in Smith [2010], Smith et al. [2013] and Smith et al. [2009]. This hybrid 3D-Var method, which has has been shown to produce very promising results in 1D experiments, takes into account covariances between parameter and state variables, and is potentially a cheaper scheme than an ensemble method.

It would also provide new insight to directly compare results from particle filter and ETKF approaches, both of which have been used in an inundation forecasting context (e.g. Matgen et al. [2010] and Garcia-Pintado et al. [2015] respectively), in the same case study. Differences in the updates generated by different data assimilation methods would be of interest to the data assimilation and hydrology communities, as noted in Grimaldi et al. [2016].

One advantage of using a PF approach is that there is no requirement that parameter values have Gaussian errors. We note that when updating the channel friction parameter in chapters 4 and 5, the analysis parameter distributions did not always ('by eye') conform to Gaussian distributions. It would be interesting to investigate this further, and to analyse the potential consequences of non-Gaussianity of model parameters.

We did not consider model error in this thesis. In chapters 4 and 5 neglecting model error was sensible, as we had a truth run to compare forecasts to, and all the simulations came from the same model. A model error term could be added to the data assimilation scheme in chapter 6 to see what effect this would have on the results. It would also be interesting to run identical twin experiments where the truth run and the forecast ensemble had different resolutions, to see what the effect of this, and the required interpolation schemes would be.

7.2.3 SAR data

There are a number of unanswered questions regarding use of SAR data in an inundation forecasting context. One avenue of enquiry would be to investigate the optimal number of backscatter-type observations when using our new backscatter observation approach. More observations are generally beneficial to data assimilation schemes, but too many can cause filter divergence and ensemble collapse, and can also increase the time and expense of an update calculation. A better quantification of the impact of observations based on location, perhaps relative to the flood edge, would allow us to maximise the benefits of using such observations.

It would also be interesting to investigate the effect of DN vs backscatter value in data assimilation, i.e. to investigate whether the extra radiometric calibration step is beneficial in a backscatter observation approach. If the results are similar without radiometric calibration, this potentially speeds up the time taken for processing of a SAR image for use in data assimilation and reduces the cost of purchasing data (DN products are generally less expensive than backscatter products).

REFERENCES

Jeffrey L. Anderson. An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus A: Dynamic Meteorology and Oceanography*, 59(2):210–224, 2007. doi: 10.1111/j.1600-0870.2006.00216.x.

Konstantinos M. Andreadis and Guy J-P. Schumann. Estimating the impact of satellite observations on the predictability of large-scale hydraulic models. *Advances in Water Resources*, 73:44 – 54, 2014. ISSN 0309-1708. doi: /10.1016/j.advwatres.2014.06.006.

Konstantinos M. Andreadis, Elizabeth A. Clark, Dennis P. Lettenmaier, and Douglas E. Alsdorf. Prospects for river discharge and depth estimation through assimilation of swath-altimetry into a raster-based hydrodynamics model. *Geophysical Research Letters*, 34(10), 2007. ISSN 1944-8007. doi: 10.1029/2007GL029721. L10403.

Giuliano Di Baldassarre, Guy Schumann, and Paul D. Bates. A technique for the calibration of hydraulic models using uncertain satellite observations of flood extent. *Journal of Hydrology*, 367(3):276 – 282, 2009. ISSN 0022-1694. doi: /10.1016/j.jhydrol.2009.01.020.

Sbastien Barthlmy, Sophie Ricci, Etienne Le Pape, Mlanie Rochoux, Olivier Thual, Nicole Goutal, Johan Habert, A Piacentini, Gabriel Jonville, Fabrice Zaoui, and Philippe Gouin. Ensemble-based algorithm for error reduction in hydraulics in the context of flood forecasting. 7:18022, 01 2016.

Mara Bermudez, Jeffrey C. Neal, Paul D. Bates, Gemma Coxon, Jim E. Freer, Luis Cea, and Jeronimo Puertas. Quantifying local rainfall dynamics and uncertain boundary conditions into a nested regional-local flood modeling system. *Water Resources Research*, 53(4):2770–2785, 2017. ISSN 1944-7973. doi: 10.1002/2016WR019903.

Craig H. Bishop, Brian J. Etherton, and Sharanya J. Majumdar. Adaptive sampling with the ensemble transform kalman filter. part i: Theoretical aspects. *Monthly Weather Review*, 129(3):420–436, 2001. doi: 10.1175/1520-0493(2001)129$⟨$0420:ASWTET$⟩$2.0.CO;2.

Kyle M. Brown, Crispin H. Hambidge, and Jonathan M. Brownett. Progress in operational flood mapping using satellite synthetic aperture radar (SAR) and airborne light detection and ranging (LiDAR) data. *Progress in Physical Geography*, 40(2):196–214, 2016. doi: 10.1177/0309133316633570.

M. Chini, R. Hostache, L. Giustarini, and P. Matgen. A hierarchical split-based approach for parametric thresholding of SAR images: Flood inundation as a test case. *IEEE Transactions on Geoscience and Remote Sensing*, 55(12):6975–6988, Dec 2017. ISSN 0196-2892. doi: 10.1109/TGRS.2017.2737664.

Clawpack Development Team. Clawpack software, 2014. Version 5.2.2. URL http://www.clawpack.org.

E. Cooper, S. Dance, J. Garcia-Pintado, N. Nichols, and P. Smith. Improving inundation forecasting using data assimilation. *Mathematics Report Series, University of Reading*, 2016. URL `http://www.reading.ac.uk/web/FILES/maths/mathsreport.pdf`.

E. S. Cooper, S. L. Dance, J. García-Pintado, N. K. Nichols, and P. Smith. Observation operators for assimilation of satellite observations in fluvial inundation forecasting. *Hydrology and Earth System Sciences Discussions*, 2018:1–32, 2018a. doi: 10.5194/hess-2018-589.

E.S. Cooper, S.L. Dance, J. Garcia-Pintado, N.K. Nichols, and P.J. Smith. Observation impact, domain length and parameter estimation in data assimilation for flood forecasting. *Environmental Modelling & Software*, 104:199 – 214, 2018b. ISSN 1364-8152. doi: /10.1016/j.envsoft.2018.03.013.

J. Crank and P. Nicolson. A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type. *Advances in Computational Mathematics*, 6(1):207–226, 1996. ISSN 1572-9044. doi: 10.1007/BF02127704.

G. Di Baldassarre and A. Montanari. Uncertainty in river discharge observations: a quantitative analysis. *Hydrology and Earth System Sciences*, 13(6):913–921, 2009. doi: 10.5194/hess-13-913-2009.

A. Doucet, A. Smith, N. de Freitas, and N. Gordon. *Sequential Monte Carlo Methods in Practice*. Information Science and Statistics. Springer New York, 2013. ISBN 9781475734379.

G. Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical research*, 99 (C5):10142–10162, 1994.

Geir Evensen. The ensemble kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53(4):343–367, Nov 2003. ISSN 1616-7228. doi: 10.1007/s10236-003-0036-9.

Geir Evensen, Dick P. Dee, and Jens Schröter. *Parameter Estimation in Dynamical Models*, pages 373–398. Springer Netherlands, Dordrecht, 1998. ISBN 978-94-011-5096-5. doi: 10.1007/978-94-011-5096-5_16.

V. Fortin, M. Abaza, F. Anctil, and R. Turcotte. Why should ensemble spread match the RMSE of the ensemble mean? *Journal of Hydrometeorology*, 15(4):1708–1713, 2014. doi: 10.1175/JHM-D-14-0008.1.

Javier Garcia-Pintado, Jeff C. Neal, David C. Mason, Sarah L. Dance, and Paul D. Bates. Scheduling satellite-based SAR acquisition for sequential assimilation of water level observations into flood modelling. *Journal of Hydrology*, 495:252 – 266, 2013. ISSN 0022-1694. doi: /10.1016/j.jhydrol.2013.03.050.

Javier Garcia-Pintado, David Mason, Sarah L. Dance, Hannah Cloke, Jeff C. Neal, Jim Freer, and Paul D. Bates. Satellite-supported flood forecasting in river networks: a real case study. *Journal of Hydrology*, 523:706–724, 2015.

Arthur Gelb. *Applied Optimal Estimation*. The MIT Press, 1974. ISBN 0262570483, 9780262570480.

David L George. Augmented Riemann solvers for the shallow water equations over variable topography with steady states and inundation. *Journal of Computational Physics*, 227 (6):3089–3113, 2008.

L. Giustarini, P. Matgen, R. Hostache, M. Montanari, D. Plaza, V. R. N. Pauwels, G. J. M. De Lannoy, R. De Keyser, L. Pfister, L. Hoffmann, and H. H. G. Savenije. Assimilating SAR-derived water level data into a hydraulic model: a case study. *Hydrology and Earth System Sciences*, 15(7):2349–2365, 2011. doi: 10.5194/hess-15-2349-2011.

L. Giustarini, R. Hostache, D. Kavetski, M. Chini, G. Corato, S. Schlaffer, and P. Matgen. Probabilistic flood mapping using synthetic aperture radar data. *IEEE Transactions on Geoscience and Remote Sensing*, 54(12):6958–6969, Dec 2016. ISSN 0196-2892. doi: 10.1109/TGRS.2016.2592951.

Laura Giustarini, Patrick Matgen, Renaud Hostache, and Jacques Dostert. From SAR-derived flood mapping to water level data assimilation into hydraulic models. *Proc. SPIE*, 8531:85310U–85310U–12, 2012. doi: 10.1117/12.974655.

Gene H Golub and Charles F Van Loan. Matrix computations. 1996. *Johns Hopkins University, Press, Baltimore, MD, USA*, pages 374–426, 1996.

Stefania Grimaldi, Yuan Li, Valentijn R. N. Pauwels, and Jeffrey P. Walker. Remote sensing-derived water extent and level to constrain hydraulic flood forecasting models: Opportunities and challenges. *Surveys in Geophysics*, 37(5):977–1034, 2016. ISSN 1573-0956. doi: 10.1007/s10712-016-9378-y.

Thomas M. Hamill, Jeffrey S. Whitaker, and Chris Snyder. Distance-dependent filtering of background error covariance estimates in an ensemble kalman filter. *Monthly Weather Review*, 129(11):2776–2790, 2001. doi: 10.1175/1520-0493(2001)129⟨2776:DDFOBE⟩2.0.CO;2.

HEC-RAS Development Team. Hec-ras software. http://www.hec.usace.army.mil/software/hec-ras/.

J.B. Henry, P. Chastanet, K. Fellah, and Y.L. Desnos. Envisat multipolarized asar data for flood mapping. *International Journal of Remote Sensing*, 27(10):1921–1929, 2006. doi: 10.1080/01431160500486724.

Jean-Michel Hervouet. Telemac modelling system: an overview. *Hydrological Processes*, 14(13):2209–2210, 2000. ISSN 1099-1085. doi: 10.1002/1099-1085(200009)14:13⟨2209::AID-HYP23⟩3.0.CO;2-6.

Jean-Michel Hervouet. *Hydrodynamics of Free Surface Flows.* John Wiley and Sons Ltd, 2007.

Marc Honnorat, Joel Marin, Jerome Monnier, and Xijun Lai. Dassflow v1.0: a variational data assimilation software for 2D river flows. Research Report RR-6150, INRIA, 2007.

M. S. Horritt, D. C. Mason, and A. J. Luckman. Flood boundary delineation from synthetic aperture radar imagery using a statistical active contour model. *International Journal of Remote Sensing*, 22(13):2489–2507, 2001. doi: 10.1080/01431160116902.

M.S. Horritt and P.D. Bates. Evaluation of 1D and 2D numerical models for predicting river flood inundation. *Journal of Hydrology*, 268(14):87 – 99, 2002. ISSN 0022-1694. doi: /10.1016/S0022-1694(02)00121-X.

R. Hostache, P. Matgen, and W. Wagner. Change detection approaches for flood extent mapping: How to select the most adequate reference image from online archives? *International Journal of Applied Earth Observation and Geoinformation*, 19:205 – 213, 2012. ISSN 0303-2434. doi: /10.1016/j.jag.2012.05.003.

Renaud Hostache, Xijun Lai, Jrme Monnier, and Christian Puech. Assimilation of spatially distributed water levels into a shallow-water flood model. Part II: Use of a remote sensing image of mosel river. *Journal of Hydrology*, 390(34):257 – 268, 2010. ISSN 0022-1694. doi: /10.1016/j.jhydrol.2010.07.003.

Renaud Hostache, Marco Chini, Laura Giustarini, Jeffrey Neal, Dmitri Kavetski, Melissa Wood, Giovanni Corato, Ramona-Maria Pelich, and Patrick Matgen. Near real-time assimilation of sar derived flood maps for improving flood forecasts. *Water Resources Research*, 0(ja), 2018. doi: 10.1029/2017WR022205.

Neil Martin Hunter. *Development and assessment of dynamic storage cell codes for flood inundation modelling.* PhD thesis, University of Bristol, 2005.

Intermap Technologies. : Nextmap british digital terrain model dataset produced by intermap. nerc earth observation data centre. URL `http://catalogue.ceda.ac.uk/uuid/8f6e1598372c058f07b0aeac2442366d`.

Thomas Steven Savage James, Pianosi Francesca, Bates Paul, Freer Jim, and Wagener Thorsten. Quantifying the importance of spatial resolution and other factors through global sensitivity analysis of a flood inundation model. *Water Resources Research*, 52 (11), 2016.

Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(Series D):35–45, 1960.

J. D. Kepert. On ensemble representation of the observation-error covariance in the ensemble kalman filter. *Ocean Dynamics*, 54(6):561–569, Dec 2004. ISSN 1616-7228. doi: 10.1007/s10236-004-0104-9.

X. Lai and J. Monnier. Assimilation of spatially distributed water levels into a shallow-water flood model. Part I: Mathematical method and test case. *Journal of Hydrology*, 377(12):1 – 11, 2009. ISSN 0022-1694. doi: /10.1016/j.jhydrol.2009.07.058.

R. J. LeVeque. *Finite Volume Methods for Hyperbolic Problems.* Cambridge University Press, 2002.

Randall J LeVeque. Wave propagation algorithms for multidimensional hyperbolic systems. *Journal of Computational Physics*, 131(2):327–353, 1997.

David Livings. Aspects of the Kalman filter. *MSc thesis, Unversity of Reading. http://www.reading.ac.uk/web/FILES/maths/Livings.pdf*, 2005.

David M. Livings, Sarah L. Dance, and Nancy K. Nichols. Unbiased ensemble square root filters. *Physica D: Nonlinear Phenomena*, 237(8):1021 – 1028, 2008. ISSN 0167-2789. doi: /10.1016/j.physd.2008.01.005.

A. C. Lorenc, S. P. Ballard, R. S. Bell, N. B. Ingleby, P. L. F. Andrews, D. M. Barker, J. R. Bray, A. M. Clayton, T. Dalby, D. Li, T. J. Payne, and F. W. Saunders. The met. office global three-dimensional variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, 126(570):2991–3012, 2000. doi: 10.1002/qj.49712657002.

Henrik Madsen. Data assimilation in the mike 11 flood forecasting system using kalman filtering. *IAHS-AISH Publication*, pages 75–81, 01 2003.

D.R. Maidment and L.W. Mays. *Applied Hydrology.* McGraw-Hill series in water resources and environmental engineering. Tata McGraw-Hill Education, 1988. ISBN 9780070702424.

Kyle T Mandli, Aron J Ahmadia, Marsha Berger, Donna Calhoun, David L George, Yiannis Hadjimichael, David I Ketcheson, Grady I Lemoine, and Randall J LeVeque. Clawpack: building an open source ecosystem for solving hyperbolic pdes. *PeerJ Computer Science*, 2:e68, 2016. doi: 10.7717/peerj-cs.68.

D. C. Mason, I. J. Davenport, J. C. Neal, G. J. . Schumann, and P. D. Bates. Near real-time flood detection in urban and rural areas using high-resolution synthetic aperture radar images. *IEEE Transactions on Geoscience and Remote Sensing*, 50(8):3041–3052, Aug 2012. ISSN 0196-2892. doi: 10.1109/TGRS.2011.2178030.

David Mason, Javier Garcia-Pintado, Hannah Cloke, and Sarah Dance. The potential of flood forecasting using a variable-resolution global digital terrain model and flood extents from synthetic aperture radar images. *Frontiers in Earth Science*, 3:43, 2015. ISSN 2296-6463. doi: 10.3389/feart.2015.00043.

David C. Mason, Sarah L. Dance, Sanita Vetra-Carvalho, and Hannah L. Cloke. Robust algorithm for detecting floodwater in urban areas using synthetic aperture radar images. *Journal of Applied Remote Sensing*, 12:12 – 12 – 20, 2018. doi: 10.1117/1.JRS.12.045011.

D.C. Mason, P.D. Bates, and J.T. Dall Amico. Calibration of uncertain flood inundation models using remotely sensed water levels. *Journal of Hydrology*, 368(14):224 – 236, 2009. ISSN 0022-1694. doi: /10.1016/j.jhydrol.2009.02.034.

D.C. Mason, T.R. Scott, and S.L. Dance. Remote sensing of intertidal morphological change in Morecambe Bay, U.K., between 1991 and 2007. *Estuarine, Coastal and Shelf Science*, 87(3):487 – 496, 2010. ISSN 0272-7714. doi: /10.1016/j.ecss.2010.01.015.

D.C. Mason, G.J.-P. Schumann, J.C. Neal, J. Garcia-Pintado, and P.D. Bates. 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 Sensing of Environment*, 124:705 – 716, 2012. ISSN 0034-4257. doi: /10.1016/j.rse.2012.06.017.

D.C. Mason, L. Giustarini, J. Garcia-Pintado, and H.L. Cloke. Detection of flooded urban areas in high resolution synthetic aperture radar images using double scattering. *International Journal of Applied Earth Observation and Geoinformation*, 28:150 – 159, 2014. ISSN 0303-2434. doi: /10.1016/j.jag.2013.12.002.

I.B. Mason. *Ch 3, Forecast Verification: A Practitioner's Guide in Atmospheric Science*. Wiley, 2003. ISBN 9780470864418. URL `https://books.google.co.uk/books?id=cjS9kK8IWbwC`.

P. Matgen, G. Schumann, J.-B. Henry, L. Hoffmann, and L. Pfister. Integration of SAR-derived river inundation areas, high-precision topographic data and a river flow model toward near real-time flood management. *International Journal of Applied Earth Observation and Geoinformation*, 9(3):247 – 263, 2007a. ISSN 0303-2434. doi: /10.1016/j.jag.2006.03.003.

P. Matgen, G. Schumann, F. Pappenberger, and L. Pfister. Sequential assimilation of remotely sensed water stages in flood inundation models. *IAHS-AISH Symposium on remote Sensing for Environmental Monitoring and Change Detection*, (316):78–88, 2007b.

P. Matgen, M. Montanari, R. Hostache, L. Pfister, L. Hoffmann, D. Plaza, V. R. N. Pauwels, G. J. M. De Lannoy, R. De Keyser, and H. H. G. Savenije. Towards the sequential assimilation of SAR-derived water stages into hydraulic models using the particle filter: proof of concept. *Hydrology and Earth System Sciences*, 14(9):1773–1785, 2010. doi: 10.5194/hess-14-1773-2010.

P. Matgen, R. Hostache, G. Schumann, L. Pfister, L. Hoffmann, and H.H.G. Savenije. Towards an automated sar-based flood monitoring system: Lessons learned from two case studies. *Physics and Chemistry of the Earth, Parts A/B/C*, 36(7):241 – 252, 2011. ISSN 1474-7065. doi: /10.1016/j.pce.2010.12.009. Recent Advances in Mapping and Modelling Flood Processes in Lowland Areas.

Matthew J. Moye and Casey O. Diekman. Data assimilation methods for neuronal state and parameter estimation. *The Journal of Mathematical Neuroscience*, 8(1):11, Aug 2018. ISSN 2190-8567. doi: 10.1186/s13408-018-0066-8.

Mathieu Mure-Ravaud, Guillaume Binet, Michael Bracq, Jean-Jacques Perarnaud, Antonin Fradin, and Xavier Litrico. A web based tool for operational real-time flood forecasting using data assimilation to update hydraulic states. *Environmental Modelling & Software*, 84:35 – 49, 2016. ISSN 1364-8152. doi: /10.1016/j.envsoft.2016.06.002.

I.M. Navon. Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography. *Dynamics of Atmospheres and Oceans*, 27 (1):55 – 79, 1998. ISSN 0377-0265. doi: /10.1016/S0377-0265(97)00032-8.

J. Neal, personal communication, 2019.

Jeffrey Neal, Guy Schumann, Paul Bates, Wouter Buytaert, Patrick Matgen, and Florian Pappenberger. A data assimilation approach to discharge estimation from space. *Hydrological Processes*, 23(25):3641–3649, 2009. ISSN 1099-1085. doi: 10.1002/hyp.7518.

Jeffrey Neal, Guy Schumann, and Paul Bates. A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas. *Water Resources Research*, 48(11):n/a–n/a, 2012a. ISSN 1944-7973. doi: 10.1029/2012WR012514. W11506.

Jeffrey Neal, Ignacio Villanueva, Nigel Wright, Thomas Willis, Timothy Fewtrell, and Paul Bates. How much physical complexity is needed to model flood inundation? *Hydrological Processes*, 26(15):2264–2282, 2012b. doi: 10.1002/hyp.8339.

N.K. Nichols. *Mathematical Concepts of Data Assimilation, Lahoz W., Khattatov B., Menard R. (eds)*, pages 1–15. Springer, Berlin, Heidelberg, 2010.

Edward Ott, Brian R. Hunt, Istvan Szunyogh, Aleksey V. Zimin, Eric J. Kostelich, Matteo Corazza, Eugenia Kalnay, D. J. Patil, and James A. Yorke. A local ensemble kalman filter for atmospheric data assimilation. *Tellus A*, 56(5):415–428, 2004. doi: 10.1111/j.1600-0870.2004.00076.x.

H. Oubanas, I. Gejadze, P.-O. Malaterre, M. Durand, R. Wei, R. P. M. Frasson, and A. Domeneghetti. Discharge estimation in ungauged basins through variational data assimilation: The potential of the swot mission. *Water Resources Research*, 54(3): 2405–2423, 2018a. doi: 10.1002/2017WR021735.

Hind Oubanas. *Variational assimilation of satellite data into a full saint-venant based hydraulic model in the context of ungauged basins.* Theses, , January 2018.

Hind Oubanas, Igor Gejadze, Pierre-Olivier Malaterre, and Franck Mercier. River discharge estimation from synthetic swot-type observations using variational data assimilation and the full saint-venant hydraulic model. *Journal of Hydrology*, 559:638 – 647, 2018b. ISSN 0022-1694. doi: /10.1016/j.jhydrol.2018.02.004.

Ruth E. Petrie and Sarah L. Dance. Ensemble-based data assimilation and the localisation problem. *Weather*, 65(3):65–69, 2010. ISSN 1477-8696. doi: 10.1002/wea.505.

E. Pinnington, T. Quaife, and E. Black. Impact of remotely sensed soil moisture and precipitation on soil moisture prediction in a data assimilation system with the jules land surface model. *Hydrology and Earth System Sciences*, 22(4):2575–2588, 2018. doi: 10.5194/hess-22-2575-2018.

F. Rawlins, S. P. Ballard, K. J. Bovis, A. M. Clayton, D. Li, G. W. Inverarity, A. C. Lorenc, and T. J. Payne. The met office global four-dimensional variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, 133(623): 347–362, 2007. doi: 10.1002/qj.32.

Sophie Ricci, Andrea Piacentini, Olivier Thual, E Le Pape, and Gabriel Jonville. Correction of upstream flow and hydraulic state with data assimilation in the context of flood forecasting. *Hydrology and Earth System Sciences*, 15(11):3555–3575, 2011.

Catherine Rochoux, Mélanie. *Towards a more comprehensive monitoring of wildfire spread : contributions of model evaluation and data assimilation strategies.* Theses, Ecole Centrale Paris, January 2014.

M. C. Rochoux, S. Ricci, D. Lucor, B. Cuenot, and A. Trouvé. Towards predictive data-driven simulations of wildfire spread part i: Reduced-cost ensemble kalman filter based on a polynomial chaos surrogate model for parameter estimation. *Natural Hazards and Earth System Sciences*, 14(11):2951–2973, 2014. doi: 10.5194/nhess-14-2951-2014.

Mélanie Rochoux, Annabelle Collin, Cong Zhang, Arnaud Trouvé, Didier Lucor, and Philippe Moireau. Front shape similarity measure for shape-oriented sensitivity analysis and data assimilation for Eikonal equation. *ESAIM: Proceedings and Surveys*, pages 1–22, 2017.

Guy Schumann, Paul D. Bates, Matthew S. Horritt, Patrick Matgen, and Florian Pappenberger. Progress in integration of remote sensing derived flood extent and stage data and hydraulic models. *Reviews of Geophysics*, 47(4), 2009. ISSN 1944-9208. doi: 10.1029/2008RG000274. RG4001.

Guy J-P. Schumann and Paul D. Bates. The need for a high-accuracy, open-access global dem. *Frontiers in Earth Science*, 6:225, 2018. ISSN 2296-6463. doi: 10.3389/feart.2018.00225.

T R Scott, P J Smith, S L Dance, D C Mason, M J Baines, N K Nichols, K J Horsburgh, P K Sweby, and A S Lawless. Data assimilation for morphodynamic prediction and predictability. In *Coastal Engineering, Proceedings of the 31st International Conference*, pages 2386–2398. 2008. doi: 10.1142/9789814277426_0197.

T.R. Scott and D.C. Mason. Data assimilation for a coastal area morphodynamic model: Morecambe Bay. *Coastal Engineering*, 54(2):91 – 109, 2007. ISSN 0378-3839. doi: /10.1016/j.coastaleng.2006.08.008.

P. J. Smith, G. D. Thornhill, S. L. Dance, A. S. Lawless, D. C. Mason, and N. K. Nichols. Data assimilation for state and parameter estimation: application to morphodynamic modelling. *Quarterly Journal of the Royal Meteorological Society*, 139(671 Part B): 314–327, January 2013.

PJ Smith. *Joint state and parameter estimation using data assimilation with application to morphodynamic modelling.* PhD thesis, PhD Thesis, University of Reading, UK, 2010.

Polly J. Smith, Sarah L. Dance, Michael J. Baines, Nancy K. Nichols, and Tania R. Scott. Variational data assimilation for parameter estimation: application to a simple morphodynamic model. *Ocean Dynamics*, 59(5):697, 2009. ISSN 1616-7228. doi: 10.1007/s10236-009-0205-6.

Polly J. Smith, Sarah L. Dance, and Nancy K. Nichols. A hybrid data assimilation scheme for model parameter estimation: application to morphodynamic modelling. *Computers & Fluids*, 46(1):436–441, July 2011. 10th ICFD Conference Series on Numerical Methods for Fluid Dynamics (ICFD 2010).

Chris Snyder, Thomas Bengtsson, Peter Bickel, and Jeff Anderson. Obstacles to high-dimensional particle filtering. *Monthly Weather Review*, 136(12):4629–4640, 2008. doi: 10.1175/2008MWR2529.1.

Elisabeth Stephens, Guy Schumann, and Paul Bates. Problems with binary pattern measures for flood model evaluation. *Hydrological Processes*, 28(18):4928–4937, 2013. doi: 10.1002/hyp.9979.

J. Srensen and H. Madsen. Data assimilation in hydrodynamic modelling: on the treatment of non-linearity and bias. *Stochastic Environmental Research and Risk Assessment*, 2004. doi: 10.1007/s00477-004-0181-9.

J. V. T. Srensen, H. Madsen, and H. Madsen. Efficient kalman filter techniques for the assimilation of tide gauge data in three-dimensional modeling of the north sea and baltic sea system. *Journal of Geophysical Research: Oceans*, 109(C3), 2004. doi: 10.1029/2003JC002144. URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2003JC002144`.

Marion Tanguy, Karem Chokmani, Monique Bernier, Jimmy Poulin, and Sbastien Raymond. River flood mapping in urban areas combining radarsat-2 data and flood return period data. *Remote Sensing of Environment*, 198:442 – 459, 2017. ISSN 0034-4257. doi: /10.1016/j.rse.2017.06.042.

C. M. Thomas and K. Haines. Using lagged covariances in data assimilation. *Tellus A: Dynamic Meteorology and Oceanography*, 69(1):1377589, 2017. doi: 10.1080/16000870.2017.1377589.

G.D. Thornhill, D.C. Mason, S.L. Dance, A.S. Lawless, N.K. Nichols, and H.R. Forbes. Integration of a 3D variational data assimilation scheme with a coastal area morphodynamic model of Morecambe Bay. *Coastal Engineering*, 69:82 – 96, 2012. ISSN 0378-3839. doi: /10.1016/j.coastaleng.2012.05.010.

Michael K. Tippett, Jeffrey L. Anderson, Craig H. Bishop, Thomas M. Hamill, and Jeffrey S. Whitaker. Ensemble square root filters. *Monthly Weather Review*, 131(7): 1485–1490, 2003. doi: 10.1175/1520-0493(2003)131⟨1485:ESRF⟩2.0.CO;2.

Peter Jan van Leeuwen. Particle filtering in geophysical systems. *Monthly Weather Review*, 137(12):4089–4114, 2009. doi: 10.1175/2009MWR2835.1.

C. Vrsmarty, A. Askew, W. Grabs, R.G. Barry, C. Birkett, P. Dll, B. Goodison, A. Hall, R. Jenne, L. Kitaev, J. Landwehr, M. Keeler, G. Leavesley, J. Schaake, K. Strzepek, S.S. Sundarvel, K. Takeuchi, and F. Webster. Global water data: A newly endangered species. *Eos*, 82(5):54+56+58, 2001. doi: 10.1029/01EO00031.

J. A. Waller, J. García-Pintado, D. C. Mason, S. L. Dance, and N. K. Nichols. Technical note: Assessment of observation quality for data assimilation in flood models. *Hydrology and Earth System Sciences*, 22(7):3983–3992, 2018. doi: 10.5194/hess-22-3983-2018.

Philip J. Ward, Erin Coughlan Perez, Francesco Dottori, Brenden Jongman, Tianyi Luo, Sahar Safaie, and Steffi Uhlemann-Elmer. *The Need for Mapping, Modeling, and Predicting Flood Hazard and Risk at the Global Scale*, chapter 1, pages 1–15. American Geophysical Union (AGU), 2018. ISBN 9781119217886. doi: 10.1002/9781119217886.ch1.

Daniel S. Wilks. *Statistical Methods in the Atmospheric Sciences*. Academic Press, 2011.

M. Wood, R. Hostache, J. Neal, T. Wagener, L. Giustarini, M. Chini, G. Corato, P. Matgen, and P. Bates. Calibration of channel depth and friction parameters in the lisflood-fp hydraulic model using medium-resolution sar data and identifiability techniques. *Hydrology and Earth System Sciences*, 20(12):4983–4997, 2016. doi: 10.5194/hess-20-4983-2016.

Melissa Wood. *Improving hydraulic model parameterization using SAR data*. PhD thesis, University of Bristol, 2016.

Dai Yamazaki, Satoshi Watanabe, and Yukiko Hirabayashi. *Global Flood Risk Modeling and Projections of Climate Change Impacts*, chapter 11, pages 185–203. American Geophysical Union (AGU), 2018. ISBN 9781119217886. doi: 10.1002/9781119217886.ch11.