LEVERAGING REMOTE SENSING FOR GLOBAL RIVER MONITORING

COLBY K. FISHER

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
CIVIL AND ENVIRONMENTAL ENGINEERING
ADVISER: ERIC F. WOOD

JUNE 2018
Abstract

The availability and distribution of fresh water resources has always been an item of great global interest, not only because it is a critical resource for human consumption, agriculture, and industry, but also because the hazards posed by hydrologic extremes (both excesses and shortages) can have lasting global impacts. With respect to these hazards, observations of surface waters can aid in monitoring reservoir conditions, providing early warning for possible flooding conditions, or predicting areas which may become susceptible to hydrologic drought. Traditionally, these observations were done through discharge gauging stations; however, the global availability of these has declined in recent years. As such, there is a need for alternative methods and data sources to supplement these observations. The goal of this dissertation is to leverage some of the remote sensing observations available to derive new methods for monitoring global water resources. In Chapter 2, an algorithm is created to derive continuous estimates of discharge from limited in-situ gauges. This spatiotemporal interpolation method is tested in a set of synthetic experiments illustrating the potential for basin wide discharge reconstruction from limited observations. Chapter 3 builds upon this work, applying the spatiotemporal interpolation method in the context of the upcoming Surface Water and Ocean Topography mission that will provide increased spatial coverage compared to current in-situ gauge networks but will have significant temporal gaps in observations. The SWOT mission is further explored in Chapter 4, where the proposed mission orbit is examined in relation to the amount of information it might be able to provide about global river basins. As a result of the orbital pattern and the uniqueness of individual river basins, careful consideration will be required to maximize the utility of SWOT. Finally, in Chapter 5 an alternative source of observations is utilized to provide rapid predictions of inundated areas as flooding occurs. Using a machine learning algorithm and passive microwave brightness temperature
observations, high resolution estimates of surface water extents are generated. In the context of global hydrology, the work presented in this dissertation provides new pathways for monitoring the availability and distribution of fresh water resources.
Acknowledgements

First and foremost, I would like to thank my advisor, Eric Wood, for all the support that he has given me throughout my time at Princeton. When I came to his group I don’t think I really understood what I was getting into; I just wanted to make cool pictures from satellites and do more programming. From the beginning though, the experience has been so much more than that. While it could often lead to some frustrating moments, Eric allowed me to wander through a number of interesting projects and ideas, enabling me to develop a body of work that I could take complete ownership of. As part of these wanderings, I was given countless opportunities to travel to conferences and workshops in locations like Niamey, Niger or Iguazu Falls, Brazil. More than just being fun trips, these excursions helped me to understand my research in a global context and to see that there was a very practical application of what I did while sitting behind a computer screen. This has been a truly rewarding experience that I do not think I would have had anywhere else and for that I am eternally grateful.

I would also like to thank Ming Pan for all the support he has given me throughout the last 5 years. It was an idea and a paper from Ming that set me on the path of exploring hydrologic data assimilation, ultimately resulting in the work presented here. He was also critical to helping me hone my Linux skills, which I now consider to be a great asset. I am also grateful to all of the members of the Terrestrial Hydrology research group, past and present, who provided ideas and feedback throughout my studies. Along with my colleagues at Princeton, I would also like to thank the members of the SWOT Science Team community for their support. Thanks also to the members of my PhD committee who provided valuable critiques of my work: Jim Smith, Justin Sheffield, and Mike Durand. A special thanks also goes to Dennis Lettenmaier who served as a reader of this dissertation at the last minute.
My friends at Princeton have been an invaluable source of support and companionship, which every graduate student needs in their solitary pursuit of some obscure topic. In particular, I would like to thank Julio Herrera-Estrada, whom I met on my second day at Princeton and quickly came to know as a great friend. From the late night coding sessions to the never ending office arguments about random current events, there is nothing better that I could have asked for in a colleague and friend. I would also like to thank Nate and Allison Chaney who have been true friends throughout my time at Princeton and have always been there for me, even when I wasn’t super fun to have over for dinner. Finally, thanks to Niko Wanders and Marjolein van Huijgevoort who taught me that there was plenty of time for fun in the life of a graduate student. They made sure that I went on plenty of hikes and bike rides, so without them I guess I might be out of shape...

It goes without saying that I am extremely grateful for all the support that I have received from my parents, grandparents and other family members over the years. Without the resources and freedoms they gave to me, I do not think I would have made it to this level. Unfortunately, my grandfather, who was my engineering mentor and biggest supporter, passed away earlier this year, before I was able to complete this dissertation; however, I know that he was extremely proud of the work that I have done, which makes it all worth it in the end.

Finally, I want to thank my two best friends, roommates and companions for the support they have given me as worked my way through this process. First, there is Maisy, who has shown me that whenever there is work to be done, it is always the perfect time to come over, plop down in my lap and lick my computer screen. Similarly, any time is apparently the right time to chase a squirrel, regardless of the commotion it may cause (whether or not that has a deeper philosophical meaning, I don’t yet know). Without the chaos of her crazy persona and the peacefulness of her lazy one, I’m not sure I would be where I am today. Last, but certainly not
least, I would like to thank Amy for being my partner in all the ups and downs of this road. When she first moved down to Princeton, I don’t think either of us really understood what the future would look like, just that we would have some good times together. While there have been plenty of good times, it is the support she has given me in the not so fun moments that I value the most. There were countless times where she helped me put this PhD train back on the tracks and get it running again, a task I don’t believe anyone else could have done. My success in this endeavour is a testament to her unwavering persistence and patience in all things, which has provided me with the motivation I needed. For this, I am more grateful than I can ever begin to express. Thank you.
For my grandfather,

Thanks for all the help along the way.
# Contents

Abstract ................................................................. iii  
Acknowledgements ....................................................... v  
List of Figures .......................................................... xii  
List of Tables ........................................................... xx  

1 Introduction ......................................................... 1  
1.1 Dissertation Goals and Overview ................................. 3  
1.2 Publications ......................................................... 6  

2 Spatiotemporal Assimilation/Interpolation of Discharge Records 
through Inverse Streamflow Routing .................................. 7  
2.1 Introduction ........................................................ 7  
2.2 Methods ............................................................ 10  
2.2.1 Routing Model Formulation .................................... 12  
2.2.2 Inverse Streamflow Routing Model ............................. 14  
2.2.3 Experimental Design and Study Area ........................... 16  
2.2.4 Data ............................................................ 17  
2.3 Results .............................................................. 20  
2.3.1 Synthetic Discharge Interpolation .............................. 21  
2.3.2 USGS Gauge Interpolation ...................................... 24  
2.4 Discussion .......................................................... 26
5 Developing High-Resolution Inundation Estimates through a Classification of Passive Microwave Brightness Temperature Observations 87

5.1 Introduction ................................................. 87

5.2 Methods ..................................................... 90
  5.2.1 Data ................................................... 91
  5.2.2 Classification Method ................................. 95
  5.2.3 Evaluation ............................................. 96

5.3 Results ....................................................... 97
  5.3.1 Classification ......................................... 97
  5.3.2 Feature Selection ..................................... 99

5.4 Discussion .................................................. 101

5.5 Conclusions ............................................... 105

6 Summary and Conclusions ............................... 107
List of Figures

2.1 Two-sweep procedure for spatiotemporal assimilation/interpolation of discharge records. The first sweep (lower left) propagates observed information collected at gauging points upstream and backward in time following the Inverse Streamflow Routing method developed in Pan and Wood (2013) and derives continuous runoff fields (lateral influx at furthest possible upstream). The second sweep (lower right) propagates information downstream and forward in time (regular routing) to create continuous discharge values everywhere. The stacked spatial maps at the top illustrate how the observed information at a single point in time/space propagates backward in time/space (upper left) and how discharge is reconstructed from the integrated runoff (upper right).
2.2 a) Overall process flow diagram for the synthetic experiments with the ISR model, where \( p \) is the input precipitation for the VIC LSM distributed over the study domain and study period, \( r \) is the runoff fields distributed over the same space, and \( q \) is discharge at discrete points during the study period. The superscripts “Init” and “Syn” represent the initial guess and the synthetic truth, respectively, while the “Inv” superscript indicates the products resulting from the model. 

b) Overall process flow diagram for the ISR model, substituting actual USGS discharge observations for the synthetic truth of the previous experiments.

2.3 The Ohio River basin modelled at 0.125 degree resolution and the distribution of the 75 USGS gauge sites used in the creation of pseudo gauges for assimilation. Blue dots represent those gauges used in the assimilation and interpolation while the red dots represent those gauges which will be reconstructed for evaluation. The background shading indicates the travel time from each grid cell to the outlet of the basin.

2.4 Reconstructed discharge time series during 2009 for 2 of the 25 evaluation sites when the ISR model was run using the climatological initial guess of runoff conditions. The blue line represents the synthetic truth discharge used in the inverse routing, the black line illustrates the discharge derived from our initial guess of runoff, and the red line illustrates the reconstructed discharge. NSE values are given for the initial guess and the reconstruction in relation to the synthetic truth.
2.5 Distributions of NSE values for the three synthetic experiments with varied initial conditions of runoff. These daily initial conditions are: a) Null (uniform mean runoff over the entire basin), b) Climatology (average daily runoff over the entire period from NLDAS), and c) TMPA (runoff derived from TMPA precipitation and VIC LSM). In each plot, the red bars illustrate the distribution of NSE values for discharge generated from the initial guess of runoff and the blue bars indicate the same distribution after reconstruction with the inverse routing method.

2.6 Distribution of NSE values for each of the evaluation sites versus the size of the upstream area for each gauge. The ordering and experiment names are the same as those in Figure 2.5.

2.7 Reconstructed discharge time series during 2009 for two of the 25 evaluation sites when the ISR model was run using the climatological initial guess of runoff conditions. Here, the “true” discharge data are USGS observations.

2.8 Distributions of NSE values for the three synthetic experiments with varied initial conditions of runoff and USGS observations as the synthetic truth. The ordering is the same as that in Figure 2.5.

2.9 Box plots of NSE values at 25 validation sites for 100 random configurations of the gauge network. Experiments are divided into two categories: the entirely synthetic discharge reconstructions (a, c, e) and the discharge reconstructions from USGS observations (b, d, f). These experiments are further differentiated by the three initial runoff conditions used: Null (a, b), Climatology (c, d), and TMPA derived (e, f). The yellow box plots represent the gauge network configuration used for the results presented in this study.
3.1 Schematic of the model process for the spatiotemporal interpolation of discharge records with the ISR model. In this diagram, p is the input precipitation distributed over the study domain that is used to force the VIC LSM, r is the runoff fields, and q is the discharge observed at our pseudo gauges during the study period. The left side of the diagram (with superscript “Init”) creates the initial guess for the ISR method, while the right side of the diagram (with superscript “Obs”) illustrates the synthetic truth.

3.2 The Ohio River basin as modeled at 0.125 degree resolution. The shading of the basin indicates the travel time from each grid cell to the outlet of the basin. The 75 USGS gauge sites used in the creation of pseudo gauges for assimilation are also shown. Blue dots represent pseudo gauges used in the assimilation while the orange diamonds represent those gauges which will be used for evaluation.

3.3 Representation of the theoretical SWOT sampling of pseudo gauges in the Ohio River basin. a) Aggregated observation frequencies within the 21 day orbit indicating the total number of times each pseudo gauge may be observed. B) Temporal pattern of the 21 day orbit crossings where each colored block represents the day on which a specific gauge will be observed.
3.4 Spatiotemporal reconstruction of discharge time series during 2009 for two of the 25 validation pseudo gauges. These results represent the interpolation of SWOT observations with a climatological initial runoff guess. The black line represents the discharge resulting from this initial guess, while the blue line represents the synthetic truth we are trying to reconstruct. The red line illustrates the discharge derived from our interpolation method. The red stars along the bottom of each plot illustrate the times at which SWOT may observe this location.

3.5 Distribution of NSE results at the 25 validation pseudo gauges for the 9 synthetic experiments performed in this study. The three columns represent the three categories of experiments performed and the three rows represent the three initial runoff conditions used. For each plot, the red bars represent the distribution of NSE values for discharge generated from initial runoff conditions used and the blue bars represent the NSE distribution after reconstruction. NSE values of less than 0 were aggregated into the 0 bin for visual clarity.

3.6 Reconstructed discharges for 2 of the 25 validation pseudo gauges using SWOT sampled discharge observations as well as complete observation time series for 5 of the pseudo gauges. In this case, the model was provided with a climatological initial guess of runoff.

3.7 Resampled crossing cycle over the Ohio River basin using 217 pseudo gauge sites that may potentially be observed by SWOT. The structure of (a) and (b) are the same as that presented in Figure 3.3.
3.8 Time series of discharge reconstructions at one gauge using: a) the original 75 pseudo gauges sampled by the SWOT orbit and b) the optimized 217 pseudo gauges which SWOT may potential observe given the orbit and river widths. Red boxes illustrate regions in which there is a noticeable performance increase in the discharge reconstruction using the additional observations. 59

4.1 The 23 global river basins used in this study. 70

4.2 Flow chart of the spatiotemporal discharge interpolation process using Inverse Streamflow Routing (ISR). 72

4.3 Number of potential SWOT observations during the 21 day orbit cycles for 8 representative basins. These observation frequencies are given as the number of times SWOT will cross over each model grid cell. The black outlines indicate the extent of the basin domains as modeled. 75

4.4 Crossing frequencies during one orbit cycle for a) the Danube River and b) the Nile River as sampled by SWOT at the generated pseudo gauge sites. 76

4.5 Crossing cycles for the 23 basins used in the study. For each basin, the shaded grids indicate the day in the 21 day orbit cycle in which each pseudo gauge will be observed. Gauges are ordered and numbered from the SW corner of the basin to the NE corner. 78
4.6 Example of discharge reconstruction for two of the evaluation pseudo
gauge sites along the Danube River. The black line illustrates the
discharge generated from the initial estimate of runoff while the blue
line represents the discharge generated from the “true” runoff fields.
The discharge generated from our reconstruction is illustrated with
the red line. The red stars along the bottom of each plot illustrate the
times at which SWOT will cross over these locations.  . . . . . . . . . . 80

4.7 Nash-Sutcliffe Efficiency (NSE) values for the 25 reconstructed gauge
discharge time series in 8 representative basins. Red bars indicate
the performance of the initial conditions only and blue bars represent
the reconstructed gauges. An initial guess of climatological runoff was
used for these simulations and makes use of only the theoretical SWOT
observations for the assimilation.  . . . . . . . . . . . . . . . . . . . . . . . . 81

4.8 Relative information content of observations in each of the 8 represen-
tative basins. This metric represents the number of times in which the
entire basin area is observed as a result of the information contained
in each SWOT observation. A value of 1 indicates 100% of the basin is
observed and values increase beyond this due to multiple overlapping
observations.  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 83

5.1 Overall process flow for the inundation classification model. Covariates
(or features) are used to produce training and testing datasets which
are labeled using model output. The random forest is then trained on
the data and used to predict inundation for a series of daily images.  . 91

5.2 Observed AMSR-E Brightness Temperatures in Kelvin for the Upper
Mississippi River Basin on June 08, 2008. The cooler colored areas
indicate regions with more water where flooding is likely to have occurred. 92
5.3 Elevation in meters for the Upper Mississippi River Basin as provided by the National Elevation Dataset. ........................................ 93
5.4 Model produced inundation estimates for two regions. The elevation of water in flooded regions is shown in meters. This data is converted to a binary mask where all of the colored cells are given a class of “inundated” and all other cells are “not inundated”. .................. 95
5.5 Example of fully classified image for June 2008. The red regions indicate areas that were classified as inundated. .......................... 97
5.6 Receiver Operating Characteristic and Precision-Recall curves. These curves were generated from the classification results of the full 30 day dataset of daily images. .................................................. 98
5.7 Calculated feature importances for the 15 most important training features used. ................................................................. 100
5.8 Inundation probabilities generated from the random forest estimators. Regions of probability greater than or equal to 0.5 were classified by the random forest as being inundated ................. 102
5.9 Comparison of predicted surface water extents from Tb and the multi-day composite flood extent image produced by the DFO from MODIS and Landsat. For regions A and B, yellow areas represent areas of agreement between the products. and Landsat imagery. ............ 104
List of Tables

3.1 Summary NSE values for each of the 9 experiments provided the mean
NSE for all 25 validation pseudo gauges both before and after the
discharge reconstructions were performed. . . . . . . . . . . . . . . . . 52

4.1 Location, area, and travel time characteristics of each of the 23 global
basin domains. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 71
Chapter 1

Introduction

One of the most common drivers of hydrologic studies is the desire to understand and assess the cycling of water in the coupled land-atmosphere system by measuring and modeling water in all of its different forms (Bierkens, 2015). An increase in our understanding of the global water cycle is vitally important for society, as we depend on water for all aspects of our life from agriculture to industry. In particular, there has been increased interest in the recent years to better understand how this system relates to the availability of water resources for human consumption and the potential for severe hydrologic events, such as floods and droughts, which have significant global impacts. As a result of this, we would like to be able to better understand the global hydrologic cycle and its dynamics (Sahoo et al., 2011). Through the processes of measurement, modeling, and forecasting, we are able to garner a greater understanding of the global water cycle and how it relates to the land-atmosphere system as a whole.

A common framework for addressing this issue is to estimate the components of the water cycle over the globe through observations and to then determine if these measurements ensure that the conservation of mass is maintained within this system.
(Pan and Wood, 2006). This balance, as shown in equation (1.1), is known as the global terrestrial water balance (or water budget), where $dS$ is the change in water storage, $P$ is the incoming flux of precipitation, $ET$ is the outgoing evapotranspiration, and $Q$ is the discharge from a basin or model grid cell (Lettenmaier et al., 2015).

$$dS = P - ET - Q$$  \hspace{1cm} (1.1)

A vast amount of work has already been done in the field to measure and estimate these fluxes. These efforts can be divided into: 1) observations based on in-situ measurements; 2) derived estimates based on remotely sensed observations; and 3) estimates based on land surface models (LSMs) (Oki and Kanae, 2006). While in-situ observations are vitally important to hydrologic studies, there has been a movement in recent years to supplement these observations by leveraging remote sensing as an alternative (Lettenmaier et al., 2015; McCabe et al., 2017). Numerous studies have evaluated the global water budget closure using only remote sensing products, each finding that there were still significant biases in the observations that prevent the closure of the computed water balances (Gao et al., 2010; McCabe et al., 2008; Sahoo et al., 2011; Sheffield et al., 2009).

Of the water balance components, runoff is arguably the hardest to estimate from an earth observing satellite. Estimates of runoff are of great importance due to its magnitude as well as its variability, both spatially and temporally around the globe (Sahoo et al., 2011). Therefore, we are often interested in determining spatial fields of runoff; however, spatially distributed runoff is the only component of the water budget that we cannot directly measure. As such, proxies of runoff are often used for water budget studies. These typically come in the form of either runoff estimates generated from land surface models (LSMs) or streamflow measurements (Lettenmaier et al., 2015).
Streamflow (or discharge) is related to runoff through the process of routing. Routing refers to the path by which surface runoff, resulting from saturation or infiltration excess, flows over the land surface and through river networks to become streamflow. Unfortunately, the main source of this information, in-situ streamflow gauges, has seen a dramatic decrease in global observations in the last 20-30 years (Fekete et al., 2012). Due to this decline, it has become more challenging to constrain the land surface and climate models that are used to predict the other components of the global water cycle. Thus, to properly constrain our land surface models and improve our understanding of the global water balance, there is a need for additional assimilation and modeling methods which move beyond using limited discharge information and make use of other current and future sources of information on runoff, which may vary significantly in their depictions of surface waters (Pan and Wood, 2013). This dissertation aims to contribute to this effort by developing new methods that make use of remote sensing products (and in-situ gauges) to inform our understanding of runoff throughout global river basins.

1.1 Dissertation Goals and Overview

The over-arching objective of this dissertation is to develop specific methods for the application of available data sources (either in-situ or remotely sensed), whereby we can gain a better understanding of global runoff dynamics and the closure off the global water balance can be achieved. This work is specifically focused on providing improved estimates of global discharges and surface water extents, especially for applications in regions which are currently poorly monitored. Ideally, these methods and datasets will aid not only in the scientific aspects of hydrology but also in im-
important applications, such as the monitoring of global flood hazards. Some specific science questions this dissertation will address are provided below.

- How can our understanding of the global water cycle be improved through the combined assimilation of in-situ gauge data and new remotely sensed data sources?

- What is the feasibility of applying this assimilation system globally and how will regions with very sparse networks of gauges be impacted?

- Moving beyond discharge, what other insight can we gain from remote sensing observations regarding the global distribution of surface waters?

- What potential is there for leveraging these data sources to provide enhanced monitoring and prediction capabilities for global flood hazards?

The success of this effort would enable hydrologic monitoring systems that provide stakeholders and policymakers with invaluable local information on the past, present, and future state of the hydrologic cycle. To achieve this end, this thesis addresses the application of two types of remote sensing observations to the derivation of global records of discharge and surface water occurrence. Three chapters are dedicated to the interpolation of observations of discharge that are sparse in both space and time. A fourth chapter investigates the application of passive microwave observations for rapid classification of flooding extent. The final chapter summarizes the these studies, providing a synthesis of what was achieved and suggesting a path forward for further remote sensing of global freshwater resources. A brief summary of each chapter follows.

- Chapter 2 introduces an algorithm for using records of discharge (such as those from streamflow gauges that are sparsely located throughout a river basin, to
derive records of discharge that are continuous in both space and time. This spatiotemporal interpolation method is tested in a set of synthetic experiments over the Ohio River basin. These experiments investigate the capabilities of the method under varying knowledge of the prior hydrologic conditions in the basin.

- Chapter 3 applies the previously developed algorithm to a new set of remotely sensed observations that will be provided by the future NASA/CNES Surface Water and Ocean Topography (SWOT) mission. Using the proposed orbit for the mission, theoretical observations of discharge are generated and used in the spatiotemporal interpolation process. Experiments investigate the ability of the interpolation process to improve the spatial and temporal extent of a SWOT derived discharge product.

- Chapter 4 investigates the potential for creating a global discharge product from the SWOT mission through the use of the spatiotemporal discharge interpolation method. Applying the method over 23 large river basins, reconstructed discharge records are created. The performance of this reconstruction is evaluated as well as the impact that the sampling pattern of the proposed orbit will have on global observations.

- Chapter 5 uses a machine learning based algorithm to test a method for rapidly creating inundation estimates from passive microwave observations of two sensors, AMSR-E and SMAP. The algorithm is tested for two specific flooding events where it produces a high resolution dataset of daily surface water extents. In addition, the importance of different features used in the classification algorithm is explored.
1.2 Publications

The core chapters of this dissertation (Chapters 2 through 5) either have or will be submitted for publication. Below are the full references:


Chapter 2

Spatiotemporal
Assimilation/Interpolation of
Discharge Records
through Inverse Streamflow
Routing

2.1 Introduction

In the application of water resources for human use, as well as the monitoring and prediction of global hydrologic hazards such as floods and droughts, a comprehensive understanding of globally distributed runoff and river discharge is extremely important. In many regions of the world, river flows are poorly monitored with in-situ observations and the collection of the available observations for consumption by global end-users has proven to be a difficult challenge, as evidenced by the available
records from the Global Runoff Data Center (GRDC) (Fekete et al., 2012). Streamflow records are typically most complete in the relatively developed and populous parts of the world; however, streamflow data in many regions are often considered proprietary, resulting in, among other issues, difficult problems in the management of water resources in transboundary rivers (see e.g., Biancamaria et al., 2011b; Pavelsky et al., 2014).

Besides the usefulness of near real-time global river discharge data for water management, there is a great need for observations of river discharge data to further our understanding of the global water cycle and its representation by reanalysis and climate models. While other observational sources for the terrestrial water budget have become more readily available from satellite remote sensing, the lack of comprehensive river discharge observations has resulted in a key flux (runoff) in climate models and large-scale land surface models being poorly constrained by observations over much of the global land surface (Sahoo et al., 2011). Furthermore, the amount of water stored at the land surface and its space-time variability are poorly known. To better serve the global hydrologic community, there is then a need for methods which can make further use of the currently available global discharge data sources.

These data sources can be divided into the following sets: 1) observations based on in-situ measurements (gauges); 2) estimates based on remotely sensed observations (e.g. satellite altimetry, synthetic aperture radar); and 3) estimates based on land surface models (LSMs) and routing models (Pan and Wood, 2006). Traditionally, the data provided by sets 1 and 2 can be thought of as point observations of river discharge along a river network, whereas set 3 can provide us with a spatially distributed representation of discharge throughout our basins of interest derived from modelled runoff fields. These two variables can be connected by the process of streamflow routing, where the spatially distributed runoff generated at the land surface flows
over the hillslope and through a river network to become streamflow in the river channels. From this process we can then say that the streamflow (as measured at specific points in space and time) is the integrated response to the runoff through a subset of time and space. Due to this, all studies using streamflow as representative of basin runoff are limited to applications where the temporal differences can be ignored or accounted for (e.g. for long time scale studies the aggregation of the runoff data allows us to ignore the temporal differences) (Sahoo et al., 2011; Sheffield et al., 2009; Pan et al., 2012). There is then a need for new methods that are able to derive spatially and temporally continuous records of runoff and river discharge from the available data sources.

The goal of this study is to present a methodology by which we can integrate and use the point scale observations of river discharge to derive a product that is spatially and temporally continuous. One possible method is the combination of point observations with spatially distributed model estimates through assimilation. Due to the integrated nature of the streamflow generation process, any discharge assimilation must be able to propagate information throughout the range of influence in a basin for any given gauge. Additionally, we must be able to assimilate all available observations in a basin simultaneously in time and space, to resolve conflicts due to observational errors. There have been a number of recent studies to investigate the potential for such discharge assimilation using a wide variety of methods (see e.g. Andreadis et al. 2007; Biancamaria et al. 2011b; Paiva et al. 2013; Pan and Wood 2013). While these methods are often robust and comprehensive, the large computational burden, particularly for those using Ensemble Kalman Filters such as that used by Andreadis et al. (2007), limits their potential for rapid global application. In addition, many of these methods simply adjust discharge in a forward sense and do not fully account for the upstream spatial and temporal correlations of
the streamflow generation process. Alternative methods have focused on the use of kriging based statistical techniques to derive spatially distributed estimates of river discharge (see e.g. Paiva et al. 2015; Yoon et al. 2013). These statistical methods often show good agreement for the reconstructed discharge with little computational cost but are highly dependent on the formulation of the covariance matrix for each river system. Here we propose the use of an assimilation and interpolation scheme for creating spatially complete and temporally continuous river discharge records from point observations based on the Inverse Streamflow Routing (ISR) model previously used by Pan and Wood (2013) for the generation of spatially distributed runoff fields to be used in land surface modelling applications, such as the calibration of model parameters. The new approach maintains the important structure of the streamflow generation process with a relatively low computational burden and guarantees an exhaustive propagation of observed information to all reachable times and locations across the river network. The new approach can be more effective compared to the assimilation and interpolation methods discussed previously, which perform an assimilation by adjusting the state variables (e.g. water height/volume, flow rate, etc.) and propagate the observed information much less exhaustively.

2.2 Methods

In short, the proposed method aims to propagate the observed discharge information across all reachable parts of the river network (up/downstream from gauging point) and all reachable times (before/after observation time) using a two-sweep procedure that first propagates information backward in time to the furthest possible upstream location (inverse routing) and then propagates it forward in time to the furthest possible downstream (forward routing). Figure 2.1 provides a detailed illustration of the
proposed scheme. The first sweep of the procedure, known as the Inverse Streamflow Routing (ISR) model developed by Pan and Wood (2013) to generate spatially distributed runoff fields, plays the key role here (left side of Fig. 2.1). The ISR helps to guarantee an exhaustive propagation of observed information by updating the boundary influx (runoff) at pixel level (the furthest possible upstream) throughout the entire spatial and temporal domains. The second sweep simply re-runs the same routing model forward using the runoff fields derived from the first sweep to reconstruct continuous discharge values everywhere (right side of Fig. 2.1). Since ISR does not require an initial guess of discharge from the routing model (Pan and Wood, 2013), the proposed method works for both data assimilation (if an initial guess exists) and pure interpolation of observations (without an initial guess). Since the discharge records are ultimately created by a routing model, this approach preserves all the physical consistencies embodied by the chosen routing model and its parameters such as the flow confluence relationship on the river network and the resulting mass balance, wave velocity and diffusivity (if a diffusive wave routing model is used). Such a strong physical consistency can hardly be implemented by methods based on statistical correlations between different gauging points or different state variables in the routing model, for example, in the river kriging method (Paiva et al., 2013). When used as an interpolator, the proposed method can also exactly reproduce the input observations at gauging locations/times (Pan and Wood, 2013). The mathematical formulation of this method is described below.
Figure 2.1: Two-sweep procedure for spatiotemporal assimilation/interpolation of discharge records. The first sweep (lower left) propagates observed information collected at gauging points upstream and backward in time following the Inverse Streamflow Routing method developed in Pan and Wood (2013) and derives continuous runoff fields (lateral influx at furthest possible upstream). The second sweep (lower right) propagates information downstream and forward in time (regular routing) to create continuous discharge values everywhere. The stacked spatial maps at the top illustrate how the observed information at a single point in time/space propagates backward in time/space (upper left) and how discharge is reconstructed from the integrated runoff (upper right).

2.2.1 Routing Model Formulation

The basic routing model selected for this work is the University of Washington (UW) routing model (Lohmann et al., 1996; Nijssen et al., 2001), which provides a simple linear routing scheme that is commonly coupled with LSMs. This model routes runoff
through two processes. The first of these is the drainage of the runoff water within a grid cell to the outlet of the grid cell as governed by a known unit hydrograph function (UHF). This is given by equation (2.1) below, where $u(t)$ is the UHF, $r(t)$ is the pixel runoff, and $o(t)$ is the pixel outflow.

$$o(t) = \int_0^t r(t - \tau)u(\tau) \, d\tau$$  \hspace{1cm} (2.1)

The second process then governs the travel of water in channels between pixels through the one-dimensional diffusive wave equation. This is given by equation (2.2), where $q$ is the streamflow generated by the pixel outflow at a distance $x$ downstream, $C$ is the channel wave velocity and $D$ is diffusivity.

$$\frac{dq}{dt} = D \frac{\delta^2 q}{\delta x^2} - C \frac{\delta q}{\delta x}$$  \hspace{1cm} (2.2)

This model is linear as long as the parameters $C$ and $D$ are assumed not to be a function of the streamflow, i.e., retention effects such as lakes and reservoirs as well as human management are not considered and thus it is a good candidate for our inversion. These two stages of the routing process are then solved together using the form presented by equation (2.3) below, where $i(x, t)$ is the impulse response function as defined by equation (2.4).

$$q(x, t) = \int_0^t r(t - \tau)u(t - \tau)i(x, t) \, d\tau$$  \hspace{1cm} (2.3)

$$i(x, t) = \frac{x}{2t\sqrt{\pi}Dt} \exp\left\{-\frac{(Ct - x)^2}{4Dt}\right\}$$  \hspace{1cm} (2.4)
By integrating equation (2.3) for all upstream pixels, denoted as \( \text{all}(g) \), for any given gauge \( g \) at discretized time steps, we can determine the streamflow at any gauge location, \( Q(g,t) \) as shown in equation (2.5).

\[
Q(g,t) = \sum_{all(G)} q(x,t)
\]  

(2.5)

This formulation serves as the basic routing model for inversion and for the final reconstruction of discharge from the inverted runoff fields.

### 2.2.2 Inverse Streamflow Routing Model

Using the routing model presented above, the fixed interval Kalman smoother can now be established for the inversion process following Pan and Wood (2013). First, the routing model must be written in a linear state space form as a function of input states as seen in equation (2.6).

\[
y_t = H_0 x_t + H_1 x_{t-1} + \cdots + H_k x_{t-k} + \epsilon_t
\]  

(2.6)

In this form, \( y_t \) is a vector of the discharges at a number of gauges in the basin and \( x_t \) is a matrix of the runoff for all cells at time \( t \). Because of the integration to determine flow at each gauge, the model requires runoff information up to a lag time of \( k + 1 \) steps, which is the travel time of the basin. Finally, \( H_t \) represents the measurement operator matrix, whose elements represent the amount of runoff that one specific cell will contribute to each gauge at a given time. These values are calculated from the impulse response function. As a result of the integration described above, there is a need for the solution of this inverse problem for multiple time steps at once, which gives rise to the fixed interval smoothing component of this inversion. Through a
time augmentation, the model can ultimately be written in the Kalman filter form as shown in equation (2.7) below (Pan and Wood, 2010).

\[
\begin{align*}
\widehat{x}''_{t} &= \widehat{x}'_{t} + K_{t}(y'_{t} - H'\widehat{x}'_{t} - L'\widehat{x}'_{t-k}) \\
\end{align*}
\]  

(2.7)

Here, \(\widehat{x}'_{t}\) is the initial guess of the time augmented runoff fields, \(y'_{t}\) is the time augmented streamflow measurements, \(H'\) and \(L'\) are time augmented measurement operators, \(K_{t}\) is the Kalman gain as given by equation (2.8) and \(\widehat{x}''_{t}\) is the updated estimate of the runoff fields.

\[
K_{t} = P_{t}H'^{T}(H'P_{t}H'^{T} + R_{t})^{-1}
\]

(2.8)

The Kalman gain represents a weighting of the update to the runoff fields and is controlled by \(P_{t}\), which represents the error covariance matrix of the initial forecast for the runoff, and \(R_{t}\), which is the error covariance matrix of the gauge measurements. For this study, we perform a set of idealized experiments in which we set \(R_{t}\) equal to 0, such that the inversion process provides a maximum correction to the initial runoff guess. The error covariance \((P_{t})\) is defined as a diagonal matrix of the long-term mean runoff error variance. In practical applications this error term will be derived from the error utilized in the particular form of the discharge observations. It should be noted that this method can function without an estimate of the initial runoff conditions (a null field) and thus, it also works for streamflow interpolation in which river discharge is reconstructed purely from observations. With the first sweep completed through ISR, the second sweep of flow reconstruction is done by running the same routing model in a forward sense with the new runoff influxes.
2.2.3 Experimental Design and Study Area

For this study we perform two sets of streamflow interpolation experiments over the Ohio River basin. The Ohio River basin, along with the Tennessee River in the southern part of the basin, is a large basin covering an area of approximately 490000 km$^2$. This basin contains a wide variety of river sizes that drain a mix of developed, undeveloped and agricultural areas, all of which are monitored by the United States Geological Survey (USGS) with a dense network of gauges. This monitoring network makes the basin a good study area for these streamflow interpolation experiments, as we will be able to use the extensive USGS observations as another set of data inputs.

The first of these experiments performs the inversion using synthetically created streamflow values as a proof of concept for the method. The goal of this experiment is to see if the streamflow interpolation method can generate the true discharge, given varying levels of information about the prior runoff conditions in the basin. The second experiment is the same as the previous, except that the synthetic gauge data is replaced with actual USGS gauge data. In this experiment, the performance of the ISR method is evaluated under “real world” conditions given that the routing model does not account for the effects of human management and will produce streamflows that are likely different from the observed streamflows. Flow charts of these two experiment sets can be seen in Fig. 2.2. Each of these experiments were run for the entire year of 2009. This period was selected because the daily discharge characteristics were representative of the climatology, with some individual high flow events. Based on the previous work of Pan and Wood (2013), the wave velocity parameter and the smoothing window for the Ohio River basin were set at 1.4 m/s and 70 days, respectively.
Figure 2.2: a) Overall process flow diagram for the synthetic experiments with the ISR model, where $p$ is the input precipitation for the VIC LSM distributed over the study domain and study period, $r$ is the runoff fields distributed over the same space, and $q$ is discharge at discrete points during the study period. The superscripts “Init” and “Syn” represent the initial guess and the synthetic truth, respectively, while the “Inv” superscript indicates the products resulting from the model. b) Overall process flow diagram for the ISR model, substituting actual USGS discharge observations for the synthetic truth of the previous experiments.

2.2.4 Data

In each of the experiments, the NLDAS 0.125 degree meteorological dataset (Cosgrove et al., 2003) is used to force the Variable Infiltration Capacity (VIC) LSM (Liang et al., 1994, 1996) to produce runoff fields that are considered the “true” runoff. The NLDAS precipitation forcings were chosen for this experiment as they combine hourly radar analyses and daily gauge observations and are considered to provide a comprehensive and reliable set of forcings over the United States (Pan et al., 2010).
This NLDAS derived runoff is then used with the routing model described previously to generate synthetic streamflow values at set evaluation sites (“pseudo gauges”) for the study period. These 75 sites are the routing model grid cells in which an actual USGS gauge is located. Of these, 25 sites are designated as validation sites and the remaining 50 sites provide river discharge time series to be assimilated in the ISR model. The selection of these gauge sites was based solely on finding gauges within the basin that had relatively complete discharge records (>95% days available) for the year 2009 and the distribution of validation sites was random. The use of these USGS gauge based sites for validation of the synthetic model results also allows for later experiments and comparisons with the actual USGS observations. The distribution of these pseudo gauge stations as well as a representation of the routing model river basin can be seen in Fig. 2.3.
Figure 2.3: The Ohio River basin modelled at 0.125 degree resolution and the distribution of the 75 USGS gauge sites used in the creation of pseudo gauges for assimilation. Blue dots represent those gauges used in the assimilation and interpolation while the red dots represent those gauges which will be reconstructed for evaluation. The background shading indicates the travel time from each grid cell to the outlet of the basin.

The generated synthetic streamflows are considered the “true” observations and are used in the streamflow interpolation process to correct an initial estimate of river discharge (derived from an initial estimate of daily runoff that is also routed using the Lohmann routing model). To investigate the impact of this initial runoff estimate, we perform each synthetic experiment with three daily initial conditions. These are: (1) a long term mean value of runoff applied over the entire basin (same value in every grid cell for every day in the study period), (2) a daily climatology of runoff at each grid cell, derived from the NLDAS forced VIC LSM, and (3) daily runoff values derived from the VIC LSM forced with the real-time TRMM Multi-Satellite Precipitation Analysis (TMPA) version 3B42RT (Huffman et al., 2007) precipitation product. The
TMPA product was selected for this experiment as it is globally available between 60 N and 60 S at a 3-hour temporal and 0.25 spatial resolution. The product was interpolated to 0.125 to force the VIC simulations (Pan et al., 2010). While this product is not as accurate as the ground observation based NLDAS product, it is globally available and can be used along with the VIC LSM to provide us with a realistic initial forecast of runoff even when ground observations do not exist (Pan et al., 2010). The results of these three purely synthetic experiments and three USGS observation-based experiments are presented in Section 2.3.

2.3 Results

Following the above methodology, six discharge interpolation (reconstruction) experiments were performed. To evaluate the performance of the interpolation in each of these experiments we compute the Nash-Sutcliffe Efficiency (NSE) at each of the 25 pseudo gauges designated for validation in Fig. 3. The NSE is a measure of model performance and is defined in equation (2.9), where $Q_o$ is the mean of observed discharges, $Q^t_m$ is modeled discharge at time $t$, and $Q^t_o$ is observed discharge at time $t$ (Nash and Sutcliffe, 1970).

$$NSE = \frac{\sum_{t=1}^{T} (Q^t_m - Q^t_o)^2}{\sum_{t=1}^{T} (Q^t_o - Q_o)^2}$$

(2.9)

The NSE may range from $-\infty$ to 1, with an efficiency of 1 meaning that there is a perfect match between the modeled discharge and the observations (or the synthetic truth). An efficiency of 0 indicates that the model is just as accurate as the mean of the observations and a value less than 0 indicates that the mean would be a better predictor than the model.
2.3.1 Synthetic Discharge Interpolation

For the first set of experiments we follow the procedure outlined by Fig. 2.2a. An example of the discharge interpolation can be seen in Fig. 2.4, where the time series of discharge are shown for 2 of the 25 validation gauges. The runoff initial conditions for this set of reconstructions was the climatological daily runoff. Figure 2.4a, which represents a downstream gauge with a large upstream area, shows good performance for the ISR method. We find that the NSE increased from 0.527 to 0.995, indicating a large increase in the model performance through assimilation. By examining the overall time series, we can see that the assimilation was able to correct for a majority of the conditions imposed by the initial guess of runoff. For example, between days 50 and 100 we can see that the initial guess had significantly higher flows compared to the synthetic truth, where these high flows centered around day 50, and the assimilation was able to reconstruct this quite well. Figure 2.4b shows the same results for an upstream gauge with a smaller contributing area, where we observe an increase in NSE from 0.049 to 0.986 after the discharge reconstruction. Similar to the previous example, this performance is quite good, indicating that the ISR methodology can be effective for reconstructing spatially and temporally continuous discharge records. Despite this, we find that for some gauges with the smallest upstream areas, which potentially contain fewer assimilated gauges than others, the reconstructions will occasionally miss the temporal dynamics of the synthetic truth, such as between days 150 and 200 in Fig. 2.4b.
Figure 2.4: Reconstructed discharge time series during 2009 for 2 of the 25 evaluation sites when the ISR model was run using the climatological initial guess of runoff conditions. The blue line represents the synthetic truth discharge used in the inverse routing, the black line illustrates the discharge derived from our initial guess of runoff, and the red line illustrates the reconstructed discharge. NSE values are given for the initial guess and the reconstruction in relation to the synthetic truth.

Figure 2.5 illustrates the evaluation of the NSE values across all of the validation sites for each of the three initial runoff conditions. By plotting the distribution of NSE values in the validation gauge set for the initial guess and the reconstructed discharge, we can see the performance improvement from the streamflow interpolation method. Across all of the initial conditions, there is an increase in performance for many of the gauges with a shift in the NSE values towards 1. In particular, we find that the null initial guess of runoff performs the best (Fig. 2.5a). This is likely because we are not imposing any temporal or spatial dynamics on the runoff, just a mean value that allows for the interpolation to adequately reconstruct the temporal dynamics of the synthetic truth. In contrast to this, we found that the experiment with initial conditions based on the TMPA observed precipitation performed the worst, as there were often differences in when events such as high flows started or
the magnitude of these events, for which the assimilation was not able to completely correct. Despite these differences, we find that the ISR method is able to do a good job of creating discharge records across all initial conditions, with a noticeable increase in performance in each case.

Figure 2.5: Distributions of NSE values for the three synthetic experiments with varied initial conditions of runoff. These daily initial conditions are: a) Null (uniform mean runoff over the entire basin), b) Climatology (average daily runoff over the entire period from NLDAS), and c) TMPA (runoff derived from TMPA precipitation and VIC LSM). In each plot, the red bars illustrate the distribution of NSE values for discharge generated from the initial guess of runoff and the blue bars indicate the same distribution after reconstruction with the inverse routing method.

To further illustrate the impact of upstream area and gauge density on the performance of the interpolation, we plot the upstream area of each validation gauge versus the NSE for reconstructed discharge in Fig. 2.6. For each experiment we can see the same pattern, with a wide variety of NSE values for gauges with upstream areas less than $10^4$ km$^2$ while basins larger than this have NSE values consistently between 0.9 and 1. These larger sub-basins incorporate the information of other upstream gauges assimilated, allowing for a more accurate reconstruction of discharge. In addition to this, the integrative nature of the routing and smoothing procedure dampens many of the short high flow events, allowing the larger sub-basins to exhibit consistently better performance given reliable upstream observations.
2.3.2 USGS Gauge Interpolation

In addition to these purely synthetic experiments, we evaluated the performance of the streamflow interpolations under real world conditions by substituting daily observed USGS river discharge values for the synthetic truth used previously. Here we present the results of these three USGS based experiments, varying the initial runoff conditions in the same manner as the previous experiments. Figure 2.7 illustrates the performance of the ISR model for discharge reconstruction when assimilating these in-situ river discharge observations. Again, we find that the method works well for the two evaluation gauges presented, with the larger basin (Fig. 2.7a) improving the NSE from 0.166 to 0.862 and the smaller basin (Fig. 2.7b) improving from -0.061 to 0.942. Comparing these results to those from the purely synthetic experiment presented in Fig. 2.4, we see that the use of the USGS data degrades the performance of the reconstruction. This is likely due to the non-linear components of flow, such as reservoirs, dams, or backwater effects that are present in this basin and can significantly alter the flow from what this linear routing model predicts. Additionally, during some of the peak flow periods (such as days 100 to 150 in Fig. 2.7b), we can
see instances where the reconstructed discharge is greater than the synthetic truth. This is a result of a numerical correction done in the model where physically unrealistic negative runoff values resulting from each Kalman smoother update are reset to a value of zero. The effect of this correction is more apparent in the assimilation of these USGS observations than in the synthetic experiments.

Figure 2.7: Reconstructed discharge time series during 2009 for two of the 25 evaluation sites when the ISR model was run using the climatological initial guess of runoff conditions. Here, the “true” discharge data are USGS observations.

Figure 2.8 presents the overall results of these experiments, again displaying the distributions of NSE values resulting from the initial guess and the discharge reconstruction. For all initial conditions, the ISR model is able to create some improvement in the reconstructed discharge values; however, the degree of improvement is noticeably less. In contrast to the synthetic experiments, the initial guess of runoff derived from the TMPA precipitation resulted in the best performance for the interpolation while the null and climatological initial guesses performed similarly, exhibiting a smaller shift in the NSE values for all the evaluation gauges. In part, this is due to the non-linear flow characteristics in the USGS observations that we are not repre-
senting, as there are often conflicting estimates of the spatial distribution of discharge between the USGS observations and the TMPA precipitation-based discharge, which lead to a decrease in the innovation term provided by the Kalman smoother. Another potential cause of this decreased performance could be errors in the river discharge observations themselves. In-situ observations are likely to have errors of varying magnitudes; however, we are treating these observations as error free for the purposes of model evaluation. As a result, any potential errors in these observations will then be transferred to errors in the final reconstructed discharge estimates.

Figure 2.8: Distributions of NSE values for the three synthetic experiments with varied initial conditions of runoff and USGS observations as the synthetic truth. The ordering is the same as that in Figure 2.5.

### 2.4 Discussion

The need for global discharge and runoff observations and estimates is not new and there have been a number of recent studies that have taken different approaches to generating spatially and temporally distributed discharge from point observations (Andreadis et al., 2007; Biancamaria et al., 2011b; Paiva et al., 2013, 2015). The ISR method is an alternative approach to these methods, which allows for the creation of spatially distributed discharge fields that are not only spatially consistent but
are also consistent through time, due to the application of a Kalman Smoother. The results of these experiments have shown that the ISR method can produce an accurate representation of discharge throughout a basin river network given a wide variety of initial conditions. In particular, the interpolation of USGS observations is promising, as we are able to generate a very close representation of the discharge conditions throughout the basin with little to no prior information about the specific distribution of runoff present. This indicates that the ISR method may be able to extend the usefulness of observations in basins with sparse gauge networks, such as many underdeveloped regions of the world. It is also important to note that the ISR method produces fields of runoff that are consistent with the observed discharges, which may prove beneficial for the calibration and optimization of land surface model processes in poorly gauged basins.

Although these experiments have illustrated the potential for the ISR method to be used for river discharge interpolation in global basins, it is important to acknowledge that these experiments are idealized and thus, do not contain all of the potential errors and uncertainties that would be present in a real-world application. As discussed previously, our method is limited by the lack of a non-linear routing model, the presence of error free observations and the overall parameterization of the routing model (static parameters for the wave velocity and the diffusivity). With regards to errors in the observations, Pan and Wood (2013) tested the impact of these errors on the ISR models’ ability to reconstruct runoff fields and found that these errors could potentially be significant enough to remove any positive improvements from the assimilation procedure. In real world applications we will need to carefully consider the error characteristics of the data sources to be assimilated, as these will have a significant impact on the quality of the final discharge product.
Another important variable in the performance of this method is the availability and selection of gauges (or pseudo gauges in synthetic experiments) for assimilation and evaluation. In this study we present the results of assimilating one specific configuration of available gauging sites as illustrated in Fig. 2.3. To better understand how the results of these experiments might change if the network of gauges were configured differently, we performed a sensitivity study by generating 100 random configurations of gauges to assimilate and evaluate from the total set of 75. These gauge networks were then used to reconstruct discharge in each of the previous 6 experiments, evaluating the NSE at each of the 25 evaluation sites, for each possible network configuration. The results of these simulations are illustrated in Fig. 2.9, where the gauge configurations for each experiment are ranked according to the median NSE of reconstructed discharge. In addition to this, the box and whisker plots illustrate the spread of performance for each configuration, as well as any potential outliers. Finally, the yellow box in each experiment represents the specific simulation results that are presented in this paper.

Focusing first on the results of the synthetic experiments with a null initial condition of runoff (Fig. 2.9a), we can see that there is a significant amount of variation in the specific distributions of NSE; however, there is little change in the median value or the lower limit of NSE values across all configurations. This indicates that regardless of the network configuration chosen, we are able to reasonably reconstruct spatially and temporally distributed discharge within the basin river network. Looking across the three initial conditions for the synthetic experiments (Fig. 2.9a, c, e) shows results comparable to those presented previously, with the null and climatological initial runoff conditions providing relatively similar performance. The TMPA derived initial conditions show a distribution of median NSEs that is slightly lower than the prior experiments. It is also interesting to note that with increasing information in the
Figure 2.9: Box plots of NSE values at 25 validation sites for 100 random configurations of the gauge network. Experiments are divided into two categories: the entirely synthetic discharge reconstructions (a, c, e) and the discharge reconstructions from USGS observations (b, d, f). These experiments are further differentiated by the three initial runoff conditions used: Null (a, b), Climatology (c, d), and TMPA derived (e, f). The yellow box plots represent the gauge network configuration used for the results presented in this study.
initial conditions, the spread of the model performance increases considerably. This further illustrates that large differences between the initial guess and true conditions can degrade the effectiveness of the ISR model for generating discharge throughout an entire basin.

Finally, the results of these 100 random configurations for the experiments using in-situ USGS discharge observations are shown in Fig. 2.9b, d, and f. Overall, we can observe the pattern of decreasing performance from null to TMPA derived initial conditions is still present. Across all three experiments, the range of median NSE values is larger than that for the synthetic experiments, indicating that the selection of gauges for assimilation in this real-world scenario has a more significant impact on the overall performance. We also find that the spread of the NSE distributions is greater than those in the synthetic experiments, further reinforcing the influence of the previously discussed error and uncertainty sources in the ISR method. Understanding and constraining these errors will be critical to future applications of ISR.

2.5 Conclusions

In this study we have developed a two-sweep method for reconstructing spatially and temporally continuous discharge records from discrete observations of discharge, in which the first sweep applies the ISR method (Pan and Wood, 2013) to propagate observed information backward in time/space and the second sweep re-runs the same routing model to propagate information forward in time/space. The new formulation is expected to offer more complete propagation of observed information in time and space (thus a better performance) and a better physical consistency than existing approaches. The core algorithm of this method is formulated as a Kalman Smoother, allowing for an assimilation/interpolation of discharge from all available observations
of discharge in a basin. By assimilating and validating against synthetic and real observations at 75 gauging sites in the Ohio River basin, the new approach has illustrated good performance across all experiments. In particular, the results of the discharge reconstructions given a null initial runoff condition are promising, as they illustrate the ability of the streamflow assimilation/interpolation methodology to create continuous discharge records in a basin where we do not have a good climatology or a calibrated hydrologic model.

The performance of this method will be limited by the availability and quality of gauge data, the specific initial conditions chosen, the parameterization of the routing model and the exclusion of non-linear features such as dams (Yin et al., 2016a,b). Further work is needed to determine how this method will perform as the density of the gauge network is reduced or as the amount of days missing from a gauge’s discharge record is increased, as would be the case in many of the global basins which do not currently have robust observation networks. Temporally sparse observations are particularly challenging for this type of assimilation/interpolation, as specific extreme events could be missed entirely, or the method may not have enough data to maintain a correction through time from the initial guess. At a minimum, ISR can be used to reconstruct a distributed representation of discharge from one or a few in-situ gauge observations; however, the more information that can be provided for the assimilation, the more likely we are to produce an accurate estimate of the discharge conditions in that basin.

To improve upon the density of observations in sparsely gauged regions, this methodology could be extended to perform interpolations from remotely sensed river discharge products, such as those from current generation satellite altimetry, or the upcoming NASA Surface Water and Ocean Topography (SWOT) mission (Alsdorf and Lettenmaier, 2003; Durand et al., 2010; Pavelsky et al., 2014). The SWOT
mission, scheduled to launch in 2021 is of particular interest, as this will be a swath altimeter designed to provide global observations of water surface elevation and slope, from which river discharge can be estimated. Within the 21-day repeat cycle, a river reach will be observed 2-4 times, on average (Biancamaria et al., 2010). The prospects for such a space borne sensor are great, especially with respect to the global coverage; however, due to the inclination of the orbit these observations are not evenly distributed in time or space and thus they will not be as complete as the USGS observations used here. In general, we believe that this form of streamflow interpolation using the ISR method could serve as a framework for creating spatially and temporally continuous discharge records from sparse observations like the future SWOT mission. Careful consideration will be required to account for the gaps in observations and the unique error characteristics of these remotely sensed discharge observations.
Chapter 3

Deriving Continuous Discharge Records from Future SWOT Observations

3.1 Introduction

With increasing interest in global water resources there is a continual need for observations and datasets of runoff and river discharge that can be used to monitor the distribution of global fresh water resources. In addition to this, similar observations are required to monitor and protect against natural hazards, such as droughts or flood events, which can have significant impacts globally. Ideally, we would have a comprehensive set of observations of discharge (and runoff) from in-situ gauges. These observations are not only important for issues related to consumption and hazard mitigation, but records of global runoff and river discharge are also very important for furthering our overall understanding of the global water cycle (Alsdorf and Lettenmaier, 2003). Global runoff and discharge estimates are essential to constraining
the climate and land surface models (LSM) that we use to monitor and predict the global water cycle. These models can often be poorly and inconsistently parameterized or constrained due to this lack of data runoff in many regions of the world (Sahoo et al., 2011). Despite their importance, river flows are often poorly monitored in many regions of the world, especially outside of North America and Europe. Additionally, the collection and standardization of the available observations, such as the records from the Global Runoff Data Center (GRDC), has proven to be a difficult task (Fekete et al., 2012). In addition to this, many streamflow records that exist in various regions are considered proprietary for government or industry use, which can cause difficulties in monitoring and managing the water resources of transboundary rivers (see e.g., Biancamaria et al., 2011b; Pavelsky et al., 2014).

In recent years, there has been a concerted effort to develop methods of monitoring surface waters using a number of different satellites and sensors (Alsdorf and Lettenmaier, 2003; Lee et al., 2011; Prigent et al., 2007; Smith and Pavelsky, 2008). These satellite platforms have the advantage of a wider spatial coverage around the globe, but they are limited by the decreased temporal coverage compared to traditional streamflow gauges. Historically, these sensors and platforms for monitoring surface waters have fallen into one of two categories. The first of these is focused on the observation of the surface water extents using visible, microwave and synthetic aperture radar based imaging sensors (Brakenridge et al., 2005; Schumann and Domeneghetti, 2016). The second type of sensors make use of nadir radar altimetry, such as those used by ENVISAT and JASON, and have been utilized to derive global records of water surface elevations around the globe (Birkett et al., 2002; Kouraev et al., 2004; Paiva et al., 2013). While many of these sensors have proved to be very beneficial for specific studies, they are limited in part by their relatively coarse spatial and temporal frequencies, which are often quite different between satellites, as well
as the need to calibrate them with in-situ observations. As such, there is still a need for more comprehensive observations to adequately monitor global surface waters.

The future Surface Water and Ocean Topography (SWOT) satellite mission aims to provide a new source of data to the field through the use of a wide swath altimeter (Biancamaria et al., 2015). The proposed sensor for this mission is designed to provide images of water surface elevations, as well as river widths and slopes. These images will come in two parallel swaths, as opposed to a nadir looking altimeter which is only able to observe a single track of elevation. These swath observations will allow SWOT to estimate discharge at the reach scale on rivers wider 100 m to within 30% of the true discharge (Pavelsky et al., 2014; Rodriguez, 2016) With regards to the orbit configuration, SWOT is proposed to have an orbit repeat interval of approximately 21 days, at an altitude of 890 km and an inclination of 77.6 degrees (Biancamaria et al., 2015). This orbital configuration will allow SWOT to have an almost complete coverage of global rivers, but the temporal sampling will be much coarser than what an in-situ gauge is able to provide. During the 21 day cycle, most areas will be observed between 2 and 8 times, with the lowest frequency of observations occurring in the middle latitudes (Biancamaria et al., 2010).

As a result of this orbital pattern, the observations of SWOT in any given river basin will be sparse in both time and space. In order to provide a continuous estimate of discharge from these sparse SWOT observations throughout a river network, an assimilation and/or interpolation technique is needed. A number of methods have been recently proposed to achieve this goal for the SWOT mission and generally, they can be classified into one of two categories: assimilation methods and statistical interpolation methods.

An assimilation based technique for creating a spatially and temporally continuous discharge product is advantageous due to the integrative nature of the streamflow
generation process. In any assimilation or interpolation of SWOT observations, it is desirable to propagate the information contained in each observation throughout the entire basin and through all relevant times. A number of recent studies have investigated the potential for the assimilation of SWOT observations using a wide variety of methods (see e.g. Andreadis et al. 2007; Biancamaria et al. 2011a; Durand et al. 2008; Paiva et al. 2013; Pan and Wood 2013). While these methods have been shown to be robust and comprehensive, they are limited in the sense that many adjust discharge in the forward sense only. By not fully accounting for the upstream spatial and temporal correlations of the streamflow generation process, a significant amount of the information contained in a streamflow observation is ignored. Finally, these methods require a significant amount of prior modeling and calibration to provide adequate initial conditions for the assimilation. The second category of methods for creating a spatially and temporally continuous SWOT discharge product are focused on kriging based statistical techniques, (see e.g. Paiva et al. 2015; Yoon et al. 2013). These methods generate estimates of discharge in discrete reaches along a river network and have been shown to produce relatively accurate interpolated discharge values. Kriging based methods are beneficial in that they have a smaller computational cost than assimilation based methods, but they are highly dependent on the formulation of the covariance matrix for each river system, which requires an additional effort to determine (Paiva et al., 2015).

As an alternative to these approaches, we propose to use the spatiotemporal discharge interpolation method presented by Fisher et al. (2018b) to derive spatially and temporally continuous records of discharge from the sparse observations that may be provided by the SWOT mission. This interpolation technique makes use of a specific fixed interval Kalman smoother, namely, the Inverse Streamflow Routing (ISR) method as described by Pan and Wood (2013), to propagate observed information
(such as that from an in-situ streamflow gauge or a SWOT overpass) across all parts of the river network (both up and downstream from the observation) and all times before and after the specific observation (Fisher et al., 2018b). In addition, it was shown that this method is capable of reconstructing discharge even when given no prior information about the initial conditions of runoff, allowing us to see this method as a form of discharge interpolation. This method was tested in the Ohio River basin using both synthetic and real in-situ gauge observations and demonstrated accurate reconstructions of discharge records across the basin. Despite this, the experiments performed relied on having a complete time series of observations at each gauge and did not investigate the impact of intermittent data in the observation records. In this study we aim to use the same ISR based spatiotemporal interpolation method to create a continuous discharge product from the sparse discharge observations which may be generated from the future SWOT mission.

3.2 Methods

The interpolation of theoretical SWOT observations presented in this study builds upon the method for the spatiotemporal interpolation of discharge records described by Fisher et al. (2018b), which propagates the observed discharge information across all reachable parts of the river network (up and downstream from the gauging point) and all reachable times (before and after the observation time). This is achieved through the use of a two-sweep procedure that begins by propagating the discharge information to the furthest upstream contributing areas and backwards in time. This first sweep uses the Inverse Streamflow Routing (ISR) model proposed by Pan and Wood (2013) to derive a set of spatially distributed runoff maps for all time steps. The second sweep in this method propagates the updated runoff fields forward in time.
to the furthest downstream areas using a traditional routing model. This method of spatiotemporal interpolation was shown to accurately perform a reconstruction of discharge throughout the Ohio River basin when tested with a set of discrete gauge observations from both idealized synthetic sources and observations from the U.S. Geological Survey (USGS) stream gauge network (Fisher et al., 2018b). In this study, we will expand upon the prior work by investigating the effect of spatially and temporally sparse observations, which we expect will be retrieved from the future SWOT mission, on our ability to perform this spatiotemporal discharge interpolation for reconstructing streamflows throughout a basin river network. Through a series of idealized experiments, we will illustrate the potential information that may be gained from assimilating these synthetic SWOT observations. An overview of the two-sweep interpolation method, as well as the study area and experimental design, are presented in the following sections.

3.2.1 Spatiotemporal Discharge Interpolation with ISR

The interpolation and assimilation method used here is composed of two processes which propagate the information contained in point observations of discharge, such as those from streamflow gauges or from swath altimetry observations, both backwards and forwards in space and time. This is achieved through the use of a two-sweep process which begins with the Inverse Streamflow Routing model developed by Pan and Wood (2013).

The ISR method is a Kalman smoother constructed with a particular routing model, which allows us to invert observed discharges and produce spatially and temporally distributed fields of runoff that would in turn generate the observed discharges. For this study, a basic routing model, the University of Washington (UW) routing model (Lohmann et al., 1996; Nijssen et al., 2001), serves as the basis of the method.
This routing model provides a simple linear routing procedure and has typically been used for routing LSM outputs in large basins (Lohmann et al., 2004). The specific form of this equations is represented in equations (3.1) and (3.2):

\begin{align*}
q(x, t) &= \int_0^t r(t - \tau) u(t - \tau) i(x, t) \, d\tau \tag{3.1} \\
\ q(x, t) &= \frac{x}{2t \sqrt{\pi D}} \exp \left\{ - \frac{(Ct - x)^2}{4Dt} \right\} \tag{3.2}
\end{align*}

where \( q \) is the streamflow generated by the pixel outflow at a distance \( x \) downstream at any time \( t \), \( u(t) \) is a known unit hydrograph function (UHF) and \( i(x, t) \) is the impulse response function. This impulse response function is derived from the one-dimensional diffusive wave equation, where \( C \) is the channel wave velocity and \( D \) is the diffusivity. To derive discharge at any given point in a basin (or gauge), we simply integrate equation (3.1) for all upstream grid cells during any defined time period. It should be noted that this routing model does not take into account any nonlinear effects which may be present in a river system, such as reservoirs, dams, or backwater effects, which can limit its ability to predict the truly observed streamflows in a given river system; however, for the simplicity of later inversion calculations in large river basins, such as the Ohio, we believe that this model is adequate for testing the potential of discharge interpolation using ISR.

The inversion process of ISR is then constructed from this routing model and a fixed interval Kalman smoother (Fisher et al., 2018b; Pan and Wood, 2013). Because streamflow generation is an integrative process, the routing model requires runoff information up to a certain lag time, which is the longest travel time of the basin, to determine flow at any location in the basin. As a result of this, the inversion problem of deriving runoff from discharge is one in which we must solve for runoff
fields at multiple time steps all at once, ultimately leading to our use of a fixed interval smoother. The general form of the Kalman smoother is given in equation (3.3) below.

\[
\hat{x}_t'' = \hat{x}_t' + K_t (y_t' - H' \hat{x}_t' - L' \hat{x}_{t-k})
\]

(3.3)

This smoother is solved to provide an updated estimate of runoff fields throughout the basin (\(\hat{x}_t''\)) from \(\hat{x}_t'\), which is the initial forecast of the runoff fields, \(y_t'\), which is the streamflow measurements, and \(H'\) and \(L'\) which are the time augmented measurement operators derived from the routing model in equations (3.1) and (3.2). It is important to note that prior work has illustrated the ability of the ISR model to function without an initial forecast of the runoff conditions (null fields) (Fisher et al., 2018b; Pan and Wood, 2013). As a result, this method is able to reconstruct distributed river discharges purely from observations, allowing us to classify this method as not only assimilation method but also a form of streamflow interpolation.

\(K_t\) in equation (3.3), is known as the Kalman gain and is further defined according to equation (3.4) below.

\[
K_t = P_t H'^T (H' P_t H'^T + R_t)^{-1}
\]

(3.4)

In this formulation, the Kalman gain serves as a weighting function for the update to the runoff fields. In addition to the measurement operators described previously, this weighting is derived from \(P_t\), which is the error covariance matrix of the initial forecast for the runoff defined as the long-term mean runoff error variance, and \(R_t\), which is the error covariance matrix of the gauge measurements. Following the previous work of Fisher et al. (2018b), we have chosen to set \(R_t\) equal to 0 for the idealized experiments presented in this study. Doing so will ensure that the inversion process imposes the maximum correction to the initial runoff conditions, allowing for a direct
validation of our model outputs against the “true” runoff and discharges we are trying to reconstruct. Once the first sweep has been completed with ISR, streamflow is reconstructed throughout the basin with a second sweep. This consists of running the routing model forward in time and space with the updated runoff inputs from ISR to derive a reconstructed network of streamflow for all time steps. Using this two-sweep method we are able to generate the necessary updated runoff inputs with or without an initial forecast of the runoff, resulting in a reconstructed and interpolated network of streamflows derived from the observations made on that network. Further details of this methodology can be found in Fisher et al. (2018b) and Pan and Wood (2013).

3.2.2 Study Area

In order to compare the results of these interpolation experiments with those derived from the synthetic experiments presented by Fisher et al. (2018b), these streamflow interpolations were performed over the Ohio River basin. This basin, which covers an area of roughly 500,000 km² and also contains the Tennessee River in the southern section, provides an interesting test case for the interpolation methodology primarily due to the availability of a dense network of streamflow gauges maintained by the USGS. In addition, the basin contains a wide distribution of river sizes and types, many of which will be observable by the SWOT mission.

3.2.3 Experimental Design and Data

To evaluate the feasibility of reconstructing basin-wide discharge from future SWOT observations, three sets of streamflow interpolation experiments were performed. The primary set of experiments performs the reconstruction using synthetically created SWOT observations of streamflow to evaluate the general performance of the method.
Given varying levels of detail regarding the initial runoff conditions in the basin, we will determine the ability of the streamflow interpolation method to reconstruct the true discharge. Further details of the generation of the synthetic SWOT observations can be found in the following section. The second set of experiments replicates the work done by Fisher et al. (2018b) to reconstruct discharge from synthetic observations with complete time series. The results of these experiments serve as a baseline for evaluating the performance of the reconstructions from SWOT observations. The final set of experiments performed uses the same methods as the previous experiments; however, we use a combined record of observations that includes both the previously discussed SWOT observations as well as observations from 5 pseudo gauges located throughout the basin, which provide a complete time series of discharge comparable to observations from a USGS stream gauging site. Figure 3.1 illustrates the model flow for the SWOT sampled interpolation experiment. To allow comparison with the results of the prior study, each of these experiments were run for the entire year of 2009.
Figure 3.1: Schematic of the model process for the spatiotemporal interpolation of discharge records with the ISR model. In this diagram, \( p \) is the input precipitation distributed over the study domain that is used to force the VIC LSM, \( r \) is the runoff fields, and \( q \) is the discharge observed at our pseudo gauges during the study period. The left side of the diagram (with superscript “Init”) creates the initial guess for the ISR method, while the right side of the diagram (with superscript “Obs”) illustrates the synthetic truth.

For each of these experiments, the “true” runoff fields are derived from the Variable Infiltration Capacity (VIC) LSM (Liang et al., 1994, 1996) forced with the NLDAS 0.125 degree meteorological dataset (Cosgrove et al., 2003), which provides a comprehensive and reliable set of forcings over the United States (Pan and Wood, 2010). This “true” runoff is routed with the linear routing model to generate synthetic streamflow values at a set of evaluation sites (“pseudo gauges”). For the Ohio River basin, the wave velocity parameter of this routing model was set to 1.4 m/s and the smoothing window was set at 70 days (Pan and Wood, 2013).

Following Fisher et al. (2018b), a set of 75 pseudo gauges are located at the routing model grid cells in which an actual USGS gauge is located. These pseudo
gauges are also located on rivers which may be observable by the SWOT mission. Of the entire pseudo gauge set, 25 of the gauges are reserved as validation sites, while the remaining 50 gauges are used to provide river discharge time series to the ISR model for assimilation. For the experiments where we will evaluate the interpolation from SWOT observations, the discharge time series at each of these 50 pseudo gauging sites will be sampled according to the future SWOT mission orbit to produce synthetic SWOT derived discharge observations. Further details of this sampling are provided in the following section. The Ohio River basin, as modeled in this study, can be seen in Figure 3.2 along with the locations of the 75 pseudo gauge stations.

Figure 3.2: The Ohio River basin as modeled at 0.125 degree resolution. The shading of the basin indicates the travel time from each grid cell to the outlet of the basin. The 75 USGS gauge sites used in the creation of pseudo gauges for assimilation are also shown. Blue dots represent pseudo gauges used in the assimilation while the orange diamonds represent those gauges which will be used for evaluation.

Although the streamflow interpolation method using ISR is capable of reconstructing discharge from observations only (i.e. no prior knowledge of the runoff in the basin), we are also interested in investigating the influence of the initial forecasts
of runoff and river discharge on the ISR method to determine whether more prior information will enhance or degrade the model performance. To do this, we will apply three daily initial conditions of runoff to the two-sweep interpolation method. The first of these is a long term mean value of runoff applied to every grid cell in the basin. This represents the null case, indicating that we have no information about the distribution of runoff throughout the basin. Next, we provide a daily climatology of runoff derived from the NLDAS forced VIC LSM run for each grid cell. Finally, we provide runoff fields derived from the VIC LSM forced with the daily real-time precipitation product from the TRMM Multi-Satellite Precipitation Analysis (TMPA), version 3B42RT (Huffman et al., 2007). This precipitation product was selected for these experiments due to its globally availability between 60 N and 60 S at a 3-hour temporal and 0.25 spatial resolution. Given this global coverage, we believe that although this product may not be as accurate as the NLDAS product over the US, it can provide reliable global information about the prior state of runoff in most basins, especially where in-situ meteorological observations are not available. To maintain consistency with the “true” runoff fields described previously, this product was interpolated to 0.125 to force the VIC simulations (Pan et al., 2010).

### 3.2.4 SWOT Observation Sampling

In order to produce a set of discharge time series representative of the potential spatial and temporal sampling patterns of SWOT, the “true” modeled discharge must be sampled appropriately so as to not include extra discharge information beyond what the satellite may see. The future SWOT mission is proposed to have a non sun-synchronous orbit with a repeat period of approximately 21 days (Biancamaria et al., 2015). The wide swath altimeter proposed will be able to observe water surface elevations and slopes, with errors of approximately 10 cm and 25%, respectively, for
most rivers greater than 100 m in width (Rodriguez, 2016). A vector-based model of this theoretical 21 day orbit (Biancamaria et al., 2015), indicating the satellite track as well as the locations of the swath measurements, was intersected with our routing model network to determine the network grid cells that would be crossed by the satellite and the days on which these crossings would occur within each cycle. This record of satellite crossings was then used to determine when each of our 75 pseudo gauging sites will be observed in the 21 day orbit. Finally, we ensured that all pseudo gauging sites were SWOT observable, meaning that they are on rivers with widths of at least 100m.

The spatial and temporal patterns of these crossings are presented in Figure 3.3. Figure 3.3a illustrates the spatial distribution of the pseudo gauge sites along with the aggregate number of observations that may occur in any given cycle of the satellites orbit. We can see that approximately half of the sites will be observed 2-3 days per cycle, while the remaining sites will be observed once or not at all. Figure 3.3b illustrates the temporal distribution of these observations. In this plot, the pseudo gauges were ordered and numbered from the SW corner of the basin to the NE corner, and the day of the orbit cycle on which each gauge will be observed is marked with a filled box. From this, we can determine that the observations of the Ohio River basin from the proposed SWOT orbit are restricted to two distinct time bands, between days 5-9 and 15-19, despite the fact that the pseudo gauge sites are well distributed geographically throughout the basin. As a result, we can only expect to get observations of the rivers during these windows, potentially limiting our ability to accurately reconstruct a complete record of the “true” discharge. These sampling bands are a function of the general latitude of the Ohio River basin and the proposed SWOT orbit, with basins confined to a small range of latitudes receiving observations during these discrete windows (Biancamaria et al., 2015).
Using this 21 day cycle of potential observations, we then sampled the 50 “true” discharge time series to be assimilated. This sampling was done by simply removing all discharge records from the time series when SWOT would not be able to make an observation and then repeating the 21 day cycle throughout the study period. This resulted in 50 synthetic time series of SWOT discharge observations for use in the ISR method as described previously.

One important caveat with this methodology for generating synthetic observations is that we are assuming the mission will create not just a water surface elevation product but also a discharge product for each river crossed by the swath. The generation of these discharge products at each river crossing has been the focus of a large amount of ongoing research (see e.g. Durand et al. 2014; Garambois and Monnier 2015; Gleason and Smith 2014. A recent comparison study by Durand et al. (2016), presented a suite of potential algorithms for this retrieval, many of which were able to produce discharge estimates on rivers to within 35% of the true discharge, which is close to
the 30% error tolerance described as part of the SWOT mission science requirements (Rodriguez, 2016). Although the authors of this comparison study conclude that further experimentation and tuning is required to produce good discharge estimates for most river crossings, we believe it is reasonable to assume that the SWOT mission will be able to produce the river discharge estimates that we synthesize in this study.

### 3.3 Results

Using the ISR spatiotemporal interpolation method described previously, a total of nine synthetic discharge experiments were performed. In each of these experiments, the reconstructed discharge at each of the 25 pseudo gauge sites left out of the assimilation were validated by comparing the reconstructed discharge to the “true” discharge. This comparison was done through the computation of the Nash-Sutcliffe Efficiency (NSE), which is measure of the model performance (Nash and Sutcliffe, 1970). For any comparison, the NSE may range from $-\infty$ to 1, where a value of 1 indicates a perfect match between the modeled discharge and the observations (the synthetic truth in this study). A NSE value of 0 indicates that the model is as accurate as the mean of the observations and a value below this indicates that the mean would be a better predictor than the model.

#### 3.3.1 Discharge Reconstruction from SWOT

The primary set of synthetic experiments was performed using the methodology illustrated in Figure 3.1. The results of these experiments were validated by comparing the reconstructed discharge time series to the synthetic truth at each of the 25 validation pseudo gauges. Two of these comparisons can be seen in Figure 3.4, which illustrates the results for reconstructed pseudo gauges with different contributing
areas. These reconstructions were performed with the climatological initial runoff conditions. Figure 3.4a illustrates a gauge near the outlet of the basin, with a large contributing area. We can see that the NSE increases from 0.496 to 0.609, indicating that the interpolation method is capable of performing an adequate reconstruction of the discharge time series. Further inspection reveals that there are number of periods where the reconstructed discharge reverted to the initial guess, such as the period centered around day 80 or around day 150. It is also important to note that the day in which SWOT would have crossed over this site is denoted with a red star at the bottom of each plot. We can see that these observation periods generally correlate with the strong correction of the reconstructed discharge to the synthetic truth. This is expected given that we are not adding any errors to the assimilation. During the periods when no observations occur, we can see that the model tends to revert back to the provided initial guess. The same comparison is presented for a pseudo gauge with a small contributing area in Figure 3.4b. We can again observe that the NSE increases from 0.044 to 0.255, indicating only a slight improvement over the initial guess. On a majority of days, the algorithm was not able to provide a significant correction for the reconstruction and thus the resulting time series resides very close to the initial guess, except when a SWOT overpass occurs.
Figure 3.4: Spatiotemporal reconstruction of discharge time series during 2009 for two of the 25 validation pseudo gauges. These results represent the interpolation of SWOT observations with a climatological initial runoff guess. The black line represents the discharge resulting from this initial guess, while the blue line represents the synthetic truth we are trying to reconstruct. The red line illustrates the discharge derived from our interpolation method. The red stars along the bottom of each plot illustrate the times at which SWOT may observe this location.

A graphical summary of the results for these experiments is presented in Figure 3.5, where we plot the NSE values of both the initial guess of discharge (illustrated with red bars) and the reconstructed discharge (illustrated with blue bars) for each of the 25 validation gauges. Figures 3.5a, d, and g show the performance of the discharge reconstruction at the validation pseudo gauges using the theoretical SWOT discharge observations only. These results are further classified by the three initial runoff forecasts. In each individual experiment there is a noticeable shift in the NSE values following reconstruction, which indicates that the method is capable of recreating some of the “true” runoff conditions throughout the basin; however, it is not capable of capturing all of the dynamics from the SWOT observations alone. The most noticeable improvement occurs in the null initial condition experiment (Figure
3.5a), where we observe that all of the initial discharges exhibited a NSE of 0 or less. This reconstruction highlights the ability of the method to interpolate discharge from only the spatially and temporally sparse observations from SWOT. The TMPA driven initial conditions proved to be the least beneficial for reconstruction, a result that was similarly found by the previous study (Fisher et al., 2018b).

Figure 3.5: Distribution of NSE results at the 25 validation pseudo gauges for the 9 synthetic experiments performed in this study. The three columns represent the three categories of experiments performed and the three rows represent the three initial runoff conditions used. For each plot, the red bars represent the distribution of NSE values for discharge generated from initial runoff conditions used and the blue bars represent the NSE distribution after reconstruction. NSE values of less than 0 were aggregated into the 0 bin for visual clarity.

To compare these results to the ideal reconstruction, achieved using a full time series of discharge information at all assimilated gauges, we performed a second set of experiments to replicate the results of Fisher et al. (2018b). Figures 3.5b, e, and h present the results of these discharge reconstructions for the same three initial runoff
conditions. Despite the identical nature of the initial conditions for each experiment, there is a noticeable difference in the reconstruction performance when utilizing the complete observations provided by the in-situ pseudo gauges. Table 3.1 provides the mean NSE value of all 25 validation gauges for each experiment. These mean NSE values further illustrate the performance gained when incorporating more data into this interpolation method.

Table 3.1: Summary NSE values for each of the 9 experiments provided the mean NSE for all 25 validation pseudo gauges both before and after the discharge reconstructions were performed.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Initial Runoff</th>
<th>Initial Mean NSE</th>
<th>Reconstructed Mean NSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWOT Obs.</td>
<td>Null</td>
<td>-0.203</td>
<td>0.420</td>
</tr>
<tr>
<td></td>
<td>Climatology</td>
<td>-0.070</td>
<td>0.368</td>
</tr>
<tr>
<td></td>
<td>TMPA</td>
<td>-1.286</td>
<td>-0.288</td>
</tr>
<tr>
<td>In-Situ Gauge Obs.</td>
<td>Null</td>
<td>-0.203</td>
<td>0.837</td>
</tr>
<tr>
<td></td>
<td>Climatology</td>
<td>-0.070</td>
<td>0.751</td>
</tr>
<tr>
<td></td>
<td>TMPA</td>
<td>-1.286</td>
<td>0.584</td>
</tr>
<tr>
<td>SWOT and Gauge Obs.</td>
<td>Null</td>
<td>-0.203</td>
<td>0.558</td>
</tr>
<tr>
<td></td>
<td>Climatology</td>
<td>-0.070</td>
<td>0.433</td>
</tr>
<tr>
<td></td>
<td>TMPA</td>
<td>-1.286</td>
<td>0.095</td>
</tr>
</tbody>
</table>

3.3.2 Incorporating Limited In-situ Observations

To further investigate the impact that the quantity of observations can have on the performance of the ISR spatiotemporal interpolation method, we performed a third set of experiments that combines the two presented previously. In this experiment we combine the theoretical SWOT observations with the observations of 5 of pseudo gauge sites. These additional observations are located on large river branches throughout the basin and provide the interpolation method with complete time series of dis-
charge throughout the study period. This set of experiments represents what may be a very common application of SWOT observations, where the infrequent satellite overpasses are used to supplement a small set of spatially limited in-situ gauges that can monitor the rivers at a much higher temporal frequency.

An example of the discharge reconstructions from this set of experiments is shown in Figure 3.6, with the same gauges as those depicted in Figure 3.3. We can see that by supplementing the SWOT observations with an additional data source, the reconstruction performance was increased from an NSE value of 0.524 to 0.985 for the larger basin and from 0.048 to 0.820 for the smaller basin. Looking at the reconstructed time series in Figure 3.6a, we can see that the spatiotemporal interpolation method is now able to maintain the correction through periods of limited SWOT observations without reverting back to the initial guess, as was seen previously. The results for all pseudo gauge comparisons are shown in Figures 3.5c, f, and i. The mean statistics for these experiments are also presented in Table 3.1. Comparing to the results from the SWOT only experiments, it is clear that there is a large increase in the performance of these discharge reconstructions a limited set of additional observations is provided. Despite this, the best performing experiment, which used the null initial runoff conditions, only resulted in a mean NSE of 0.558, seemingly low compared to the results achieved by the in-situ gauge information experiments.
Figure 3.6: Reconstructed discharges for 2 of the 25 validation pseudo gauges using SWOT sampled discharge observations as well as complete observation time series for 5 of the pseudo gauges. In this case, the model was provided with a climatological initial guess of runoff.

### 3.4 Discussion

As mentioned previously, there is a strong interest among the remote sensing and hydrology communities to find methods that can make use of the wealth of potential surface water observations provided by current and future satellite missions like SWOT (Alsdorf et al., 2007; Schumann and Domineghetti, 2016). The Inverse Streamflow Routing based spatiotemporal method for interpolating discharge records, as presented in this study, can be seen as one approach for taking advantage of this future global data source. The results of the experiments described previously illustrate the potential for this method to produce continuous records of streamflow from limited observations; however, our reconstructions of discharge using only SWOT observations did not perform nearly as well as those using in-situ observations. This is due in part to the observation pattern of SWOT over the basin, as shown in Figure 3.3b,
where there is a 5-6 day gap in observations throughout the entire basin. When this gap occurs during a high flow period, such as the period around day 50 in Figure 3.4a for example, we are unable to recreate any of the runoff in the basin that produced this event. As a result of this, it is difficult for the interpolation method to consistently make a correction to the initial guess. These observational gaps will be particularly important for the future use of SWOT observations as many basins around the globe will be observed with very different patterns depending on position and latitude (Biancamaria et al., 2015).

With regards to the runoff conditions used for the ISR model, we found that the null conditions tended to perform the best across all experiments, while the TMPA driven initial conditions performed the worst. This is primarily due to the fact that the TMPA observations show differences in timing of large rainfall events as well as the total volume of water for these events. Previous studies have illustrated some of the uncertainties present in the global TMPA (Sahoo et al., 2011). When the input precipitation forcing is vastly different from the “true” forcing, we find that this also manifests in the discharges and the ISR method is not able to fully correct our initial guess, even when a large number of discharge observations are given. These initial conditions could be altered by running the ISR model in a forecast mode, where the assimilation and smoothing done for each Kalman window creates a new set of initial conditions for the TMPA driven LSM model run. This would provide a dynamic initial condition for the reconstruction and may increase the overall performance of this experiment set. Running the ISR model in this mode may also prove to be useful for forecasting applications with SWOT, where we are interested in generating daily initial conditions given newly available SWOT observations.

The results of our third set of experiments illustrated one potential method to improve upon the discharge interpolation of SWOT observations by incorporating
a second source of in-situ observations into the method. Additional data in a few locations throughout the basin can allow the interpolation method to maintain a relatively accurate spatial distribution of runoff throughout the entire interpolation period. Information obtained from the SWOT observations is then used to make a further correction throughout the wide network of pseudo gauge sites each time an overpass occurs. This effect can be seen in the time series comparisons shown in Figure 3.6, where we are much better able to match the temporal dynamics of the synthetic truth with our reconstructions, but we are not able to perfectly recreate the discharges of the synthetic truth. The pattern is altered when each SWOT observation occurs, allowing the reconstructed discharge series to make a complete correction to the synthetic truth, given that our theoretical SWOT observations do not have any errors associated with them.

Although this is one effective method for reconstructing a distributed discharge product from SWOT observations, it will not help in an ungauged basin. An alternative to this would be to increase the density of pseudo gauge sites that are observed by SWOT. The 75 gauges chosen for this study were ideal given they were distributed widely around the Ohio River basin, they reduced the computation burden of the inversion and they exhibited good performance for discharge reconstructions in prior studies. In reality, the SWOT mission has the potential to make a denser network of observations than the one that was utilized here (Pavelsky et al., 2014). Due to the integrative nature of the streamflow generation process, we would expect that all observations made by SWOT will contain some information about the prior runoff conditions upstream as well as information about the future downstream discharge. By sampling more sites, we aim to capture more information that can be used to inform our discharge interpolation. To determine if such a network of observations would be beneficial, we performed another synthetic experiment with a new network
of pseudo gauges that better represents the potential wealth of observations SWOT may be able to provide.

To construct this new network of theoretical SWOT pseudo gauges, we augmented the previous gauge set by adding pseudo gauges for river locations in the Ohio River basin that would be SWOT observable. This was done by first filtering out all sections of the model river network that had widths less than 100 m and then sampling these such that we would not have observations in grid cells directly adjacent to our validation gauges. Finally, we reduced the size of this new pseudo gauge set to make it more computationally efficient by removing all SWOT observable sites with less than two observations in any given cycle. The resulting network of 217 pseudo gauges is shown in Figure 3.7. The temporal structure of these observations, as shown in Figure 3.7b, is largely the same as our previous network due to the latitudinal location of the basin; however, on any given day we now have significantly more observations.

Figure 3.7: Resampled crossing cycle over the Ohio River basin using 217 pseudo gauge sites that may potentially be observed by SWOT. The structure of (a) and (b) are the same as that presented in Figure 3.3.

Using this new network to perform the spatiotemporal discharge interpolation experiment using SWOT observations only, we were able to reconstruct a new set of discharge records. The output of this experiment is shown in Figure 3.8 where
we compare the reconstructions for one specific validation site using the original network of 75 pseudo gauges (Figure 3.8a) and the new network of 217 pseudo gauges (Figure 3.8b). For this particular site, we find that the NSE for the reconstruction using the denser network is 0.921 compared to the previous value of 0.592, which represents a significant improvement. In particular, we can see that in many of the time periods between SWOT observations (outlined in red in Figure 3.8) the discharge reconstruction method is able to maintain a correction, providing a closer representation of the synthetic truth. The idea that a denser network of observations may help an assimilation problem such as this is not a new; however, we believe that it is of particular importance to the SWOT mission. Given that the general methods for developing a global discharge observation product from the future SWOT observations are still under development (Durand et al., 2016), we believe that it is advantageous for these algorithms to generate discharges in as many locations as possible, understanding that there are many rivers which may not be observed due to the specific geography, layover effects, or lack of model parameters to derive discharge. From an assimilation perspective, the density of these observations will be critical to creating a truly global and continuous discharge product from SWOT.
Figure 3.8: Time series of discharge reconstructions at one gauge using: a) the original 75 pseudo gauges sampled by the SWOT orbit and b) the optimized 217 pseudo gauges which SWOT may potential observe given the orbit and river widths. Red boxes illustrate regions in which there is a noticeable performance increase in the discharge reconstruction using the additional observations.

In addition to the density of the observations, it is important for assimilation and interpolation methods, such as the one presented here, to consider the potential errors in the future SWOT observations. In this study, we have chosen to ignore these errors as we wished to test the validity of our method without including any other uncertainties. Other assimilation studies have incorporated or investigated these errors and have shown that they will have a significant impact on the observations (Paiva et al., 2015; Yoon et al., 2016). Although the exact parameters of these errors are not known, the science requirements for the mission specify a 30% error tolerance in discharge, while the current discharge algorithms being developed have produced results with an average error of around 35% (Rodriguez, 2016; Durand et al., 2016). Further work is needed to investigate the impact of these rather substantial errors on our ability to produce a continuous discharge product from discrete SWOT observations.
3.5 Conclusions

The future Surface Water and Ocean Topography mission has the potential to provide an entirely new set of remotely sensed surface water observations to the suite currently available. By providing estimates of river width, elevation and slope, we anticipate that the mission will be able to provide discharge observations for all global rivers that are observable (Durand et al., 2010). In this study, we have implemented a two-sweep method for performing a spatiotemporal interpolation of these future observations to provide a discharge product that is continuous in both space and time. The particular advantage of this method, which makes use of the Inverse Streamflow Routing model developed by Pan and Wood (2013), is that it accounts for the continuity of discharge both backwards and forwards in space and time, allowing for a comprehensive discharge record to be produced. The results of this study indicate that the method can produce an interpolated discharge product that is representative of the true conditions. To achieve this, careful consideration is needed to ensure that the maximum amount of information is extracted from the available SWOT observations or the sparse observations must be supplemented with alternative sources of information such as in-situ stream gauges. Pavelsky et al. (2014) made a similar recommendation, finding that SWOT would likely serve as complement to the available in-situ gauge networks rather than as a replacement.

As an alternative to simply assimilating discharge observations into another discharge product, we have shown that the ISR based method is capable of performing an interpolation of discharge observations given no prior information regarding the distribution of runoff within the basin. This finding is particularly important for the SWOT mission, as it aims to provide data for and improve our understanding of basins which currently have poor in-situ gauge networks or none at all (Alsdorf et al., 2007). By creating an interpolated discharge product from discrete observations, we
are able to gain some understanding of the overall flow conditions in any basin, which will be useful for parameterizing and calibrating hydrologic processes in both land surface models and climate models.

Further work is needed to better understand the impact that the proposed SWOT orbit will have on our ability to monitor rivers around the globe. The observation cycle of any basin will generally be determined by its general latitude (Biancamaria et al., 2010); however, the orientation of each river in relation to the satellite track will have an impact on the number of observations which may be made and the quality of these observations. As we described previously, detailed studies of the potential SWOT observations are important for developing a global SWOT discharge product through assimilation or interpolation, as the amount of useful information gained from SWOT observations in any river network will depend on all of these parameters combined. Overall, we believe that the Inverse Streamflow Routing based discharge interpolation method presented here has the potential to produce spatially and temporally continuous discharge records from the future SWOT mission. Further investigation is needed to determine how the method will perform when applied globally and how significantly the potential discharge errors from SWOT will impact the quality of this global discharge product.
Chapter 4

Assessing the Potential for Creating Global River Discharge Records from Future SWOT Observations

4.1 Introduction

The availability and distribution of fresh water resources has long been an item of great global interest, not only because it is critical as a resource for human consumption, agriculture, and industry, but also because the hazards posed by water extremes (both excesses and shortages) can have lasting impacts on society. With respect to these hazards, observations of surface waters can aide in monitoring reservoir conditions, providing early warning for possible flooding conditions, or predicting areas which may become susceptible to hydrologic drought, for example. Traditionally, the critical observations of these surface waters were done through discharge gauging
stations. While these stations tend to provide relatively reliable and consistent monitoring when available, they are usually only located on large single channel rivers that can be easily measured via the river stage and a known bathymetry (Bates et al., 2013). Dense gauging networks with long historical records exist mainly in developed countries and tend to be quite limited in other regions of the world (Pavelsky et al., 2014). Efforts have been made to collect and maintain records of these global observations; however, in recent years the number of global gauges available has decreased significantly (Fekete et al., 2012; Vörösmarty, 2002). Along with this decline, the availability of discharge data in many regions of the world has also been restricted by political issues for transboundary rivers and also by industrial and commercial interests (Biancamaria et al., 2011b). Finally, it is also important to note that these traditional gauging stations are more prone to errors under high flow conditions (Di Baldassarre and Montanari, 2009). During these high flow periods, a significant fraction of the total river discharge may be conveyed in floodplains and wetlands, areas that are very rarely gauged (Prigent et al., 2007).

For applications in water management and hazard monitoring, a wide range of satellites have become available to provide critical information about rivers and lakes around the world. These sensors are not capable of measuring discharge directly, but rather they use estimates of either surface water extents or elevations with other physical parameters of the river to estimate the discharge Bates et al. (2013). Surface water extent can be monitored using a variety of optical sensors, passive microwave sensors, or synthetic aperture radars (e.g., Landsat, AMSR-E, RADARSAT). Surface water elevations are mostly measured with nadir viewing altimeters, which are either radar (e.g. JASON, Topex–Poseidon) or laser (ICESat) based (Schumann and Domeneghetti, 2016). Each of these sensor types has specific observation capabili-
ties and limitations but none can serve as a complete replacement for in-situ gauge observations (Fekete et al., 2012).

Despite the wealth of available satellites, there is still a need for more observations. The proposed Surface Water and Ocean Topography (SWOT) mission aims to fill this gap by providing an improved set of observations of terrestrial surface waters (Biancamaria et al., 2015). SWOT is different from the previously mentioned missions in that it will use a wide swath altimeter that is designed to provide images of not only water surface elevations but also of river widths and slopes. This is achieved through the observation of water heights in two parallel swaths, as opposed to the single-track observations of a nadir altimeter. The proposed sensor design will allow SWOT to make reach scale estimates of discharge to within 30% of the true value for rivers wider than ~100 m (Pavelsky et al., 2014; Rodriguez, 2016). SWOT will also have better temporal coverage than many current generation, with a proposed orbit repeat interval of approximately 21 days, at an altitude of 890 km and an inclination of 77.6 degrees (Biancamaria et al., 2015). Within each orbital cycle, most areas of the globe will be observed between 2 and 8 times, depending primarily on the general latitude of the river reach (Biancamaria et al., 2010). While this orbital configuration will allow for an almost complete coverage of global rivers, it is important to note that compared to the temporal availability of observations that a quality in-situ gauge will be able to provide, SWOT will be much coarser (Pavelsky et al., 2014).

The key to all of these earth-observing systems, with respect to surface waters, is the availability of the observations in both time and space. Many satellites missions, such as those providing altimetry observations, have limited spatial coverage and a relatively large revisit time (usually on the order of 1-2 weeks). For applications related to water hazards, these observation gaps can have significant consequences for end-users. In the case of flooding where we would like an early warning of impactful
events, it is possible that an altimeter may not observe river levels in a basin during the window in which the flood wave begins, and thus, there will be no capacity for the mission to generate any timely warnings or predictions. To resolve some of these issues, a significant amount of research has been done to derive combined global products that make use of multiple sensors (see e.g., Aires et al. 2014; Prigent et al. 2007; Schumann et al. 2016). While these products are useful, they are often aggregated in time to resolve the temporal difference between missions and thus they may be too coarse for monitoring specific flood hazard events.

Ideally, we would like to derive a continuous estimate of discharge, in both space and time, from the observations of one sensor, despite the intermittent nature of the observations. The creation of such a product is possible for discharge, as the streamflow generation process is integrative in both space and time, meaning that each individual observation actually contains information about prior and future discharge in other parts of the river network (Pan and Wood, 2013). In the context of SWOT, this product has been the focus of a large amount of effort in preparation for the mission launch. Some studies have focused on the assimilation of the observed water surface elevations into complex hydrodynamic models to derive regional estimates of discharge methods (see e.g. Andreadis et al. 2007; Biancamaria et al. 2011a; Durand et al. 2008; Paiva et al. 2013). While these methods generally produce good discharge products, they can be computationally intensive and as of writing have only been tested for small reaches or single basins. Other methods have focused on interpolation of the discharge observations through statistical methods, such as kriging (see e.g. Paiva et al. 2015; Yoon et al. 2012). While these methods are computationally inexpensive and have been applied successfully, they are reliant on detailed knowledge of each basin to generate the required covariance functions, are...
limited to reach based applications (typically coarse) and have also not been tested for global application.

Pan and Wood (2013) proposed an assimilation-based model that can serve as an alternative to those described previously. Known as Inverse Streamflow Routing (ISR), this model was demonstrated by Fisher et al. (2018b,a) to be capable of performing the spatiotemporal interpolation of discharge records from both complete and sparse records. One main advantage of this method over those described previously is that it makes use of a specific fixed interval Kalman smoother (the ISR model) to propagate observed information (such as that from a SWOT overpass) to every part of the river network (both up and downstream) over a period of time, both before and after the observation was made (Fisher et al., 2018b). While this spatiotemporal discharge interpolation method was previously tested over only the Ohio River basin, it is important to determine how it may perform over a wide variety of global river basins, including those where there may not be a reliable or dense network of in-situ gauges. These interpolations experiments will not only investigate the potential for creating a global discharge product from intermittent SWOT observations, but it will also illustrate some of the observation patterns that we can expect to see from the mission.

4.2 Methods

This study aims to investigate the impact of future global SWOT observations of water surface elevation (and thus discharge), as well as the proposed mission orbit, on our ability to create a continuous global discharge record using an interpolation/assimilation technique. To investigate this, we perform a series of spatiotemporal discharge interpolation experiments for a set of large global river basins. The inter-
polation method used here is based on the Inverse Streamflow Routing model (Pan and Wood, 2013), which has been shown to be capable of producing such discharge records from intermittent observations (Fisher et al., 2018a). The following sections describe the interpolation methods, the study domains and the experimental setup.

4.2.1 Spatiotemporal Interpolation Method

River discharge is a unique hydrologic variable in the sense that it represents an integration of all runoff originating upstream, which results in observations at discrete points (gauges, satellite crossings etc.) that are representative of the location where the measurement was taken, as well as a larger upstream region (Fekete et al., 2012). As such, it is important that any assimilation or interpolation method for discharge take this into account, so as to fully capitalize on the information contained in each observation. In this study we use a two-sweep method for discharge interpolation that begins with the ISR model developed by Pan and Wood (2013). The ISR method is essentially a Kalman smoother constructed with a particular routing model, which allows us to invert and assimilate observed discharges to produce fields of runoff that would produce the observed discharges when routed again. Following previous applications of this method (Fisher et al., 2018b,a), the University of Washington (UW) routing model (Lohmann et al., 1996; Nijssen et al., 2001) was selected for use with the ISR model. This routing model was designed for routing coarse resolution, gridded land surface model (LSM) outputs in large basins, making it relatively simple computationally while still adequately representing the streamflow generation process for these initial global experiments (Lohmann et al., 2004). The specific form of this equations is represented in equations 4.1 and 4.2:

\[
q(x, t) = \int_0^t r(t - \tau) u(t - \tau) i(x, t) d\tau
\]  (4.1)
\begin{equation}
  i(x, t) = \frac{x}{2t\sqrt{4D}} \exp \left\{ -\frac{(Ct - x)^2}{4Dt} \right\} 
\end{equation}

where \( q \) is the streamflow generated by the pixel outflow at a distance \( x \) downstream at any time \( t \) and \( u(t) \) is a known unit hydrograph function (UHF). The impulse response function, \( i(x, t) \), as shown in equation 4.2 is derived from the one-dimensional diffusive wave equation, where \( C \) is the channel wave velocity and \( D \) is the diffusivity. The calculation of discharge at any location along the river network is done by integrating equation 4.1 for all upstream grid cells during any defined time period. Given that we are applying this routing model to a wide variety of global basins, it should be noted that this model does not take into account any non-linear components of flow that are present in many regions such as tidal effects, reservoir impoundments, or other human alterations to flow. As such, the model is limited in its ability to predict the discharge conditions that actual in-situ gauges would observe; however, for the purposes of performing a first set of global synthetic interpolation experiments and examining the impacts of the proposed SWOT orbit, we believe that it is sufficient.

The first sweep of the interpolation process is constructed from this routing model and a fixed interval Kalman smoother (Fisher et al., 2018b; Pan and Wood, 2013). A fixed-interval smoother was used to maintain the integrative nature of the streamflow generation process, in which the routing model determines flow in the basin by aggregating runoff inputs up to a certain lag time. The general form of the Kalman smoother is given in equation 4.3 below.

\begin{equation}
  \hat{x}_{t}'' = \hat{x}_{t}' + K_{t}(y_t' - H'\hat{x}_{t}' - L'\hat{z}_{t-k}') \tag{4.3}
\end{equation}

Equation 4.3 is used to provide an updated estimate of runoff fields throughout the basin (\( \hat{x}_{t}'' \)) from \( \hat{x}_{t}' \), which is the initial forecast of the runoff fields, \( y_t' \), which is the
streamflow measurements, and $H'$ and $L'$ which are the time augmented measurement operators derived from the routing model in equations 4.1 and 4.2. The reason for classifying this method as a form of interpolation, as well as assimilation, is that the ISR model is able to function without an initial forecast of the runoff conditions (null fields), allowing for the creation of a discharge product informed only by the discharge observations made (Fisher et al., 2018b; Pan and Wood, 2013).

The Kalman gain, $K_t$ in equation 4.3, is further defined according to equation 4.4 below.

$$\begin{align*}
K_t &= P_t H' T (H' P_t H' T + R_t)^{-1} \\
\end{align*}$$

(4.4)

In this formulation, the Kalman gain serves as a weighting function for the update to the runoff fields. In addition to the measurement operators described previously, this weighting is derived from $P_t$, which is the error covariance matrix of the initial forecast for the runoff defined as the long-term mean runoff error variance, and $R_t$, which is the error covariance matrix of the observations. Prior studies using ISR for discharge interpolation set $R_t$ equal to 0 so as to perform an idealized experiment in which the interpolation method should completely reconstruct the observed discharges. For this study we have chosen to add an error term to the observations. These errors are discussed in the following section along with the generation process for the synthetic SWOT observations. Once the first sweep has been completed with ISR, streamflow is reconstructed throughout the basin with a second sweep. By running the routing model forward in time and space with the updated runoff fields we are able to derive a reconstructed network of streamflow continuous in both time and space. Further details of this methodology can be found in Fisher et al. (2018b) and Pan and Wood (2013).
4.2.2 Study Areas

In order to determine the global potential for SWOT observations, we have selected a set of 23 global basins in which we performed synthetic experiments of discharge interpolation (Fig. 4.1). Table 4.1 provides the outlet location of each of these basins as defined in this study, as well as the basin area and maximum travel time (from the longest flow path). These basins were chosen to represent a variety of river types globally, differing not only in their locations but also in their lengths and shapes, which will have an impact on the availability of SWOT observations.

Figure 4.1: The 23 global river basins used in this study.
Table 4.1: Location, area, and travel time characteristics of each of the 23 global basin domains.

<table>
<thead>
<tr>
<th>River Basin</th>
<th>Outlet Lat</th>
<th>Outlet Lon</th>
<th>Area (10^3 km²)</th>
<th>Max Travel Time (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amazon</td>
<td>-1.9192</td>
<td>-55.5131</td>
<td>4680.00</td>
<td>40</td>
</tr>
<tr>
<td>Amur</td>
<td>50.63</td>
<td>137.12</td>
<td>1730.00</td>
<td>34</td>
</tr>
<tr>
<td>Columbia</td>
<td>45.6073</td>
<td>-121.1734</td>
<td>613.83</td>
<td>15</td>
</tr>
<tr>
<td>Congo</td>
<td>-4.27</td>
<td>15.32</td>
<td>3475.00</td>
<td>38</td>
</tr>
<tr>
<td>Danube</td>
<td>45.2167</td>
<td>28.7167</td>
<td>807.00</td>
<td>22</td>
</tr>
<tr>
<td>Dneper</td>
<td>46.770</td>
<td>33.180</td>
<td>482.00</td>
<td>15</td>
</tr>
<tr>
<td>Don</td>
<td>47.541</td>
<td>40.647</td>
<td>378.00</td>
<td>13</td>
</tr>
<tr>
<td>Indigirka</td>
<td>69.58</td>
<td>147.35</td>
<td>305.00</td>
<td>11</td>
</tr>
<tr>
<td>Kolyma</td>
<td>67.37</td>
<td>153.67</td>
<td>361.00</td>
<td>14</td>
</tr>
<tr>
<td>Lena</td>
<td>70.7</td>
<td>127.65</td>
<td>2430.00</td>
<td>28</td>
</tr>
<tr>
<td>Mackenzie</td>
<td>67.4583</td>
<td>-133.7447</td>
<td>1660.00</td>
<td>31</td>
</tr>
<tr>
<td>Mekong</td>
<td>15.1167</td>
<td>105.8</td>
<td>545.00</td>
<td>15</td>
</tr>
<tr>
<td>Mississippi</td>
<td>32.315</td>
<td>-90.9058</td>
<td>2964.26</td>
<td>29</td>
</tr>
<tr>
<td>Murray</td>
<td>-34.1916</td>
<td>141.5979</td>
<td>991.00</td>
<td>15</td>
</tr>
<tr>
<td>Northern Dvina</td>
<td>64.150</td>
<td>41.917</td>
<td>348.00</td>
<td>13</td>
</tr>
<tr>
<td>Nile</td>
<td>29.7</td>
<td>31.28</td>
<td>2900.00</td>
<td>51</td>
</tr>
<tr>
<td>Ob</td>
<td>66.57</td>
<td>66.53</td>
<td>2950.00</td>
<td>39</td>
</tr>
<tr>
<td>Parana</td>
<td>-27.97</td>
<td>-58.85</td>
<td>1950.00</td>
<td>22</td>
</tr>
<tr>
<td>Senegal</td>
<td>14.9</td>
<td>-12.45</td>
<td>218.00</td>
<td>12</td>
</tr>
<tr>
<td>Ural</td>
<td>50.85</td>
<td>51.28</td>
<td>190.00</td>
<td>11</td>
</tr>
<tr>
<td>Xi Jiang</td>
<td>23.48</td>
<td>111.3</td>
<td>329.71</td>
<td>10</td>
</tr>
<tr>
<td>Yangtze</td>
<td>30.66</td>
<td>111.23</td>
<td>1010.00</td>
<td>19</td>
</tr>
<tr>
<td>Yenisei</td>
<td>67.48</td>
<td>86.5</td>
<td>2440.00</td>
<td>40</td>
</tr>
</tbody>
</table>

4.2.3 Synthetic Interpolation Experiments

Using the two-sweep interpolation method described previously, a set of spatiotemporal discharge experiments were performed using synthetic SWOT observations. A schematic of one of these experiments can be seen in Figure 4.2. Prior to performing the interpolation, we generate an initial forecast of discharge conditions in the basin from prior information of precipitation and runoff. This prior information can either
come from actual global observations of precipitation (such as the TMPA 3B42RT product) or a daily climatology of precipitation that is used to force the Variable Infiltration Capacity (VIC) LSM (Liang et al., 1994, 1996). It also possible that no prior information is known, meaning that we have a null field of runoff. Previous studies have illustrated the difference that these prior conditions can have on the performance of the interpolation method, with the climatological or null conditions performing similarly well (Fisher et al., 2018a).

Figure 4.2: Flow chart of the spatiotemporal discharge interpolation process using Inverse Streamflow Routing (ISR).

For the experiments presented here we will use only the climatological runoff prior. This daily global climatology is derived from the daily TRMM Multi-Satellite Precipitation Analysis (TMPA), version 3B42RT product at a 0.25 spatial resolution
(Huffman et al., 2007). Following the flow of Figure 4.2, this prior is used to force the VIC LSM, whose runoff output is then routed to produce an initial network of discharge at a 0.25 resolution. A second set of discharge observations is needed to generate the synthetic SWOT observations for interpolation. These discharges represent the “true” conditions in each basin that we are attempting to reconstruct. Following the same flow described previously, the daily 0.25 precipitation product from the Princeton Global Forcing (PGF) dataset (Sheffield et al., 2006) is used to force the LSM and routing model. The resulting discharges are then sampled according to the future SWOT orbit (as described in section 4.2.4) to derive a set of synthetic SWOT observations. Each of these simulations was done for the entire year of 2009, giving us a 1 year dataset for our synthetic experiments.

For each basin, a network of pseudo gauge sites at which we make SWOT observations was generated (described in section 4.2.4). Of these sites, 25 are removed from the interpolation process to be used for later validation. Following the generation of the two discharge pseudo gauge datasets, the two-sweep interpolation method is executed, producing both a network of reconstructed discharge and the runoff fields which produced those discharges. The routing model used was tuned to each individual basin, with the wave velocity parameter being set based on the basin properties and the smoothing window set at double the longest travel time in the basin (as shown in 4.1).

4.2.4 Generating Synthetic SWOT Observations

In order to produce a set of discharge observations that are indicative of the future spatial and temporal sampling patterns of SWOT, the “true” synthetic discharges must be sampled. This ensures that we do not include any extra discharge information in the reconstruction process beyond what the satellite may see. The SWOT mission
is proposed to have a 21-day orbit during the nominal phase, which will provide swath measurements of rivers greater than 100 meters in width. A vector-based model of this theoretical orbit (Biancamaria et al., 2015), indicating the satellite track as well as the ground tracks of the swath observations, was rasterized and intersected with our routing model network to determine the network grid cells that would be crossed by the satellite and the days on which these crossings would occur within each cycle. The results of this process can be seen in Figure 4.3 for 8 representative basins. From these plots we can see that the number of observations per cycles varies widely across basins, as well as within each basin. This observation pattern is dictated by the orbit, which can overlap in some areas and miss others completely, such as in the Amazon (Figure 4.3a). Northern latitude basins such as the Lena (Figure 4.3e) and the Yenisei (Figure 4.3h) will have a high number of observations, due to the continual overlap of the orbit in the polar regions, while regions in the middle and lower latitude regions will experience continually fewer observations with decreasing latitude.

Applying these observation patterns to the river model network, we then created a set of pseudo gauge sites for each basin. These sites represent the locations at which we will generate discharge observations for interpolation. By first filtering out all regions of the model river network that had widths less than 100 m and then sampling these to ensure there were no observations in grid cells directly adjacent to other sites, a dense network of potential SWOT observations was created. To reduce the computational complexity of the inversion process, this observation network was further constrained to 150 pseudo gauge sites by randomly sampling the dense network. For some basins the number of SWOT observable sites was less than 150, in which case we did not perform this random sampling. Two examples of these networks are shown in Figure 4.4, along with the frequency of observations at each site. For the Danube River (Figure 4.4a), we find that most of the observations are concentrated along the main
Figure 4.3: Number of potential SWOT observations during the 21 day orbit cycles for 8 representative basins. These observation frequencies are given as the number of times SWOT will cross over each model grid cell. The black outlines indicate the extent of the basin domains as modeled.
stems of the river network, primarily oriented east to west, and show a range of observation frequencies from 1 to 4 per cycle. Figure 4.4b illustrates a similar pattern for the Nile River basin; however, we can see that the dominant orientation of the observations along the river is north-south and on average the observations are less frequent compared to a higher latitude basin, such as the Danube.

Figure 4.4: Crossing frequencies during one orbit cycle for a) the Danube River and b) the Nile River as sampled by SWOT at the generated pseudo gauge sites.

At each of the pseudo gauge sites within a basin, the temporal pattern of SWOT observations was extracted from the orbit for use in sampling the “true” modeled discharge values described previously. An illustration of these observation cycles for each of the 23 study basins can be seen in Figure 4.5. For each day in the orbit, the presence of an observation is recorded with a filled box for the corresponding gauge number, where these pseudo gauges were ordered and numbered from the SW corner of the basin to the NE corner. It is clear from these plots that the observation of global river basins will not be uniform or consistent from one basin to another. Some basins, such as the Columbia, Dneper, Don, Mekong and Nile exhibit banding patterns in the observations, with large data gaps occurring. Referring back to Figure 4.1, we can see that these basins are more limited in terms of longitudinal extent, and
thus they are observed only during specific segments of the orbit. Others, such as the Amazon, Congo, Lena, and Mississippi basins, exhibit a more distributed pattern of observations. In general, the main observable rivers of these basins cover a wider range of longitudes and latitudes, allowing for more spatially and temporally different observations to be made.

With these potential observation cycles for each of the pseudo gauge sites in each of the global basins, we then sampled the 125 “true” discharge time series to be assimilated from the PGF derived conditions described previously. The remaining 25 pseudo gauge sites were removed for use in validation and were not sampled. Sampling of each discharge time series was done by simply removing all discharge records from the time series when SWOT would not make an observation during each consecutive 21 day cycle of the study period. Along with the temporal sampling, these discharge observations were assigned errors for use in the ISR method. A recent comparison study by Durand et al. (2016), detailed a number of potential algorithms for the retrieval of discharge from the SWOT surface water observations. Given the potential errors in SWOT observed heights, many of the algorithms were able to derive discharge estimates to within 35% of the true discharge, close to the 30% error tolerance described as part of the SWOT mission science requirements (Rodriguez, 2016). As such, we have decided to assign a 30% error to the SWOT observed discharges for use in equation 4.4. These discharge observations were then used in the ISR model to interpolate the continuous discharge records (as illustrated in Figure 4.2). Validation was done by comparing the discharge time series at each reconstructed pseudo gauge to the “true” discharge records. The Nash Sutcliffe Efficiency (NSE) was used as a model performance metric at each of the 25 validation sites, following the previous methods of Fisher et al. (2018a). The NSE is a commonly used normalized statistic for hydrologic model performance which determines the relative magnitude of the
Figure 4.5: Crossing cycles for the 23 basins used in the study. For each basin, the shaded grids indicate the day in the 21 day orbit cycle in which each pseudo gauge will be observed. Gauges are ordered and numbered from the SW corner of the basin to the NE corner.
residual variance compared to the variance of the measured data (Nash and Sutcliffe, 1970).

### 4.3 Discharge Interpolation Results

For each of the global basins, the previously described synthetic experiments were carried out over a one-year period and evaluated against the “true” discharge conditions at 25 validation pseudo gauge sites. An example of these results can be seen for the Danube river in Figure 4.6, where we present the discharge reconstruction time series for two specific gauges. Figure 4.6a illustrates a pseudo gauge with a smaller upstream area. Overall, we can see a general shift in the discharge towards the “true” conditions from the more constrained initial conditions. Evaluation of the NSE for these time series shows an increase in performance following the discharge reconstruction from 0.54 to 0.891. Figure 4.6b illustrates these time series for a pseudo gauge located near the outlet of the basin with a larger upstream area that the previous. Analysis of the NSE values indicates a similar increase in performance, from 0.545 to 0.860. Qualitatively, though, we can say that the interpolation effects are more noticeable than the previous example, with the reconstruction providing significant corrections of up to 2500 m$^3$/s at times. Furthermore, we find the results to be consistent with those of the ISR based discharge interpolation studies done previously (Fisher et al., 2018a,b), in which the reconstructed discharge tends to revert towards the initial guess in periods of limited observations and then makes a strong correction when SWOT observations are made. This effect is more apparent for the pseudo gauge with the smaller upstream area (Fig. 4.6a). For discharge reconstruction at gauges with larger upstream areas, we find that the interpolated product is able to
better maintain the “true” discharge conditions, due to more data being collected and integrated from upstream areas, throughout the entire smoothing window.

Figure 4.6: Example of discharge reconstruction for two of the evaluation pseudo gauge sites along the Danube River. The black line illustrates the discharge generated from the initial estimate of runoff while the blue line represents the discharge generated from the “true” runoff fields. The discharge generated from our reconstruction is illustrated with the red line. The red stars along the bottom of each plot illustrate the times at which SWOT will cross over these locations.

The overall results of these experiments are presented in Figure 4.7 for the same 8 representative gauges shown previously. For each basin, the distribution of NSE values at the 25 validation sites are shown for both the initial forecast (red bars) and the reconstructed discharges (blue bars). For all basins, there is a shift in the performance (to higher values of NSE) following the reconstruction. This indicates that for each of the basins, the spatiotemporal discharge interpolation method using the ISR model is able to perform a reasonable reconstruction of discharge, recreating some of the “true” runoff conditions throughout the basin; however, it is not capable of capturing all of the dynamics from the SWOT observations alone. For example, the Danube, Nile and Yangtze basins exhibit large shifts in the overall NSE but still have some pseudo gauge sites where the reconstruction is no better than that of the initial
conditions. Other basins, such as the Amazon and the Congo, which are both located in the lower latitudes, exhibit very little change in the overall NSE distribution.

Figure 4.7: Nash-Sutcliffe Efficiency (NSE) values for the 25 reconstructed gauge discharge time series in 8 representative basins. Red bars indicate the performance of the initial conditions only and blue bars represent the reconstructed gauges. An initial guess of climatological runoff was used for these simulations and makes use of only the theoretical SWOT observations for the assimilation.

4.4 SWOT Orbit Impacts

While the previous work of this study has focused on the ability of the spatiotemporal discharge interpolation method to generate accurate continuous discharge records, it is important to consider these results in the context of the mission itself. The performance of the interpolation method in each basin is mainly driven by the availability of observations and the temporal and spatial patterns with which these observations will be made. The SWOT crossing cycle plots of Figure 4.5 illustrate one component
of these basin specific factors. As was previously mentioned, these plots illustrate the wide variety of temporal patterns which may exist throughout the global river basins, from distinct windows of observation 2-3 days in length (e.g. the Don) to multiple observations throughout the basin on every day (e.g. the Amur). These gaps are driven in part by the latitudinal location of the basin, with higher latitude basins receiving more frequent observations (Biancamaria et al., 2010). Given that each basin is relatively unique, we would expect the river orientation, length and width to be different, which means each river system will be observed differently by SWOT.

To better understand the information content of the available SWOT observations and how it can be interpolated, we compute the number of total basin “observations” (TBO) for each day in the orbit cycle. This metric was computed by first determining the upstream area of each site where a SWOT observation could be made. Given that these observations contain information about the prior upstream conditions (which the ISR method will make use of), we can then consider these upstream areas to also be observed by SWOT. For each day, these areas are then aggregated over the entire basin and weighted by the total basin area. This produces a metric where a value of 1 indicates that the available SWOT observations on that day contained information about an area equivalent to 100% of the entire basin. When more observations are made on a given day, this number will increase. For subsequent days in the orbit cycle, the value may increase or decrease depending on the orbit track. A summary of this metric for each of the 23 basins can be seen in Figure 4.8.

As an example, if we consider the Amazon river, which is a relatively large basin located in the lower latitudes, we can see from Figure 4.3 that the aggregate number of observations in each cycle will be relatively low (between 0 and 2). Looking at the temporal pattern of these observations, as shown in Figure 4.5, it is clear that despite the limited number of total observations, these are distributed evenly throughout
Figure 4.8: Relative information content of observations in each of the 8 representative basins. This metric represents the number of times in which the entire basin area is observed as a result of the information contained in each SWOT observation. A value of 1 indicates 100% of the basin is observed and values increase beyond this due to multiple overlapping observations.
the cycle, ensuring that there will always be some information about the discharge conditions within the basin. These characteristics correlate well with the TBO metric presented in Figure 4.5. For the Amazon, we see that the TBO maintains a relatively constant value of approximately 5 throughout the cycle, with some small fluctuations from day to day. Although we would expect that these temporally dense observations would provide a good discharge reconstruction, this was not the case (Figure 4.7a). This is likely due to the overall size of this basin and the attenuation of flow along the observable river segments, which allows for the climatological initial guess of runoff to produce a good representation of the “true” discharges before performing any interpolation. Contrasting this with a higher latitude basin of comparable size, such as the Lena River that will be observed a maximum of 8 times in each cycle (Figure 4.3e) in similarly uniform temporal pattern (Figure 4.5), we can see that the TBO metric produces a very different result. Due to the uniform temporal sampling, the height of the TBO is relatively constant but it is significantly higher than that of the Amazon. This is a result of the increased total number of observations and the size of the upstream areas that these observations represent. From this simple comparison we can see that the TBO metric provides a relatively straightforward interpretation of how each basin will be observed by the future SWOT mission in both space and time.

One interesting basin in this analysis is the Nile River basin, which has the longest maximum travel time of the 23 rivers studies. The Nile is oriented such that the main stem of the river, and most of the SWOT observable reaches, runs essentially S-N (Figure 4.4b). As a result of this orientation in relation to the SWOT orbit, the SWOT observable pseudo gauges are crossed predominantly in two 4-day windows. From Figure 4.8 we can see that during these periods of observation a significant amount of TBO is generated, while during the gap periods the TBO drops to 0.
Despite these gaps in observations, the discharge reconstruction results (Fig. 4.7f) show that we were still able to make some use of the information for the interpolation, improving the discharge records at some of the validation sites.

With regards to the overall information which may be available from SWOT observations, it is important to note that this analysis does not take into account the ability of the mission to generate discharge values from the water surface elevations. Because the mission will be primarily observing water surface elevations, a retrieval of the discharge is required. Depending on the capabilities of these retrieval algorithms, the global availability of SWOT observations may be further limited beyond what we have described in this study. A significant amount of research has focused on the creation of these algorithms and their application globally through variations of Manning’s equation (see e.g. Durand et al. 2014; Garambois and Monnier 2015; Gleason and Smith 2014). The challenge of universally applying these methods is that they often require prior estimates of the river bathymetry or slope, which may limit the available SWOT observations to regions where these prior parameters are known or can be easily estimated.

4.5 Conclusions

In order to investigate the potential availability and usefulness of discharge observations derived from the future Surface Water and Ocean Topography mission, we have utilized a spatiotemporal discharge interpolation algorithm to derive discharge records in 23 global basins. Overall, the results indicate that this interpolation method is able to reproduce the “true” discharge records relatively well when given intermittent SWOT observations as the only source of information. Although we saw a general benefit from the interpolation, it was evident that the interpolation was more effective
in some basins than others. These differences are due in part to the wide variety of
global river basin types and also to the specific orbit of the SWOT mission. By exam-
ining the proposed mission orbit as it relates to this global interpolation/assimilation
problem we have shown that there a number of factors which will govern the amount
of information available from future observations. By computing the total amount of
basin area observed during each day (TBO), we illustrate the many types of observa-
tion patterns which may be seen globally.

To maximize the usefulness of the SWOT mission for the generation of a spatially
and temporally consistent global discharge product, it will be important that we con-
sider these patterns for each specific basin, ensuring that we extract the maximum
amount of information possible. Furthermore, it is important to note that a combi-
nation of the future SWOT observations and a small set of in-situ gauge observations
can also be very useful. Remote sensing observations and in-situ gauges can be seen
as complementary sources of information for discharge (Pavelsky et al., 2014). It is
often the case that regions of the world which prove too challenging for in-situ river
gauging are ideal for remote sensing (Neal et al., 2009). Further investigation into
the use of these observations, including the assimilation with other discharge products
(either remotely sensed or in-situ) will prove to be beneficial not only to the SWOT
mission products but also for global hydrologic studies in general.
Chapter 5

Developing High-Resolution Inundation Estimates through a Classification of Passive Microwave Brightness Temperature Observations

5.1 Introduction

Globally, flooding is one of the most devastating natural hazards, occurring on a regular basis. These floods can have an impact on human lives and can cause severe economic damage to communities. Numerous studies have recently illustrated that under the future potential for climate change, the risks of these floods occurring will not decrease, but rather there may be an increase in flood intensity and frequency (Khan et al., 2012; Wing et al., 2018). In order understand the ways in which these
future risks may evolve and to accurately inform decision makers, reliable datasets on the spatial and temporal characteristics of floods are needed. Despite this need, there is generally a lack of reliable, continuous, and consistent information on the extent and dynamics of flooding events (Prigent et al., 2015). While traditional methods of observation, such as in-situ streamflow gauges, can provide us with general information on river conditions, these observations become inadequate at the onset of a major flooding event where the surface waters extend beyond the river channel (Bates et al., 2013). Moving beyond these observations, the utilization of satellite remote sensing data can help to monitor river levels and observe the extent of surface waters during a flooding event.

Although the previous chapters of this dissertation have focused on the derivation of distributed discharge records and the potential for generating these records from the future Surface Water and Ocean Topography mission, information regarding the presence of water on the land surface can also be gathered from a different subset of earth observing satellites that are capable of observing and estimating the areal extents of surface waters. In many regions, it is possible to examine the extent of historical and real-time inundation occurrence from visible and infrared imagery provided by sensors such as MODIS or the Landsat TM (see e.g Khan et al. 2011; Li et al. 2013; Tarpanelli et al. 2017). In addition, alternative methods use other complementary remotely sensed data sources, such as available microwave brightness temperature (Tb) observations (e.g., SMAP, SMOS, AMSR2, AMSR-E, and GMI), to aid in the estimation of global flooding (see e.g. Revilla-Romero et al. 2015; Sun et al. 2011; Khan et al. 2012).

The derivation of water extent maps from remotely sensed imagery has been the subject of a significant amount of research in recent years. Pekel et al. (2016) derived a long term global dataset of surface water presence using the Landsat 5, 7 and
8 imagery at a resolution of 30 meters from 1984 to 2015. Such observations are incredibly useful for looking at seasonal and long term changes in the distribution of surface waters, but due to the sample frequencies of the satellites this product was temporally aggregated to provide a consistent monthly product. In the case of flood monitoring and early warning systems, these revisit times of these satellites make it such that continuous observations may not be made (Schumann et al., 2016). In addition, flood classifications from this type of imagery are more complicated in regions of dense vegetation or persistent cloud cover (which is also likely to exist precipitation and flooding are occurring) (Du et al., 2016).

To overcome some of these issues, Prigent et al. (2007) developed a global dataset of inundation extents and dynamics derived not only from visible imagery, but also from passive and active microwave observations. These microwave observations are advantageous in that they are sensitive to surface water presence and are also insensitive to the signal degrading effects of varying levels of solar illumination and atmosphere contamination (Du et al., 2016). Compared to the visible or IR imagery products, microwave based estimates of surface water areas are significantly coarser in terms of spatial resolution. One common approach to resolve this issue is to provide a measure of the fraction of water in each observed grid cell; however, the development of these fractional water conversion factors is highly dependent on observations from visible imagery or ground surveys (Prigent et al., 2015). Instead of performing these water fraction retrievals for each sensor using a specifically fit set of parameters, we propose to use general machine learning based algorithm to derive high-resolution estimates of water surface extents from flooding. In doing so, we aim to take full advantage of complementary remotely sensed data sources, such as SMAP, SMOS, AMSR2, AMSR-E, and GMI, as well as globally available high-resolution elevation datasets to aid in the estimation of global flooding on a daily basis.
5.2 Methods

The main goal of this study is to test the application of a generalized machine learning algorithm to derive estimates of flooding from passive microwave observations. Machine learning, a term which has become increasingly popular in the fields of hydrology and remote sensing, is simply the use of generalized methods to gain insight from datasets without prescribing a specific statistical model to the analysis (Chaney et al., 2016). Random forests were chosen as the best algorithm for an initial classification system to derive estimates of inundation from passive microwave observations. These are an ensemble method in which a number of decision trees are built using bootstrap samples of the original training data to perform an unsupervised classification of the input dataset (Murphy, 2012). An overview of this modeling system is shown in Figure 5.1. In order to predict inundation using a random forest classifier, a number of data covariates were used to inform the prediction. These datasets are described in the following section. All data was collected for the Upper Mississippi River Basin at the native resolution of each dataset. For this study we will focus on the time period of May and June of 2008, when a large flood occurred in the areas around Cedar Rapids, Iowa.
Figure 5.1: Overall process flow for the inundation classification model. Covariates (or features) are used to produce training and testing datasets which are labeled using model output. The random forest is then trained on the data and used to predict inundation for a series of daily images.

5.2.1 Data

The primary source of information for this initial experiment was the brightness temperature observations provided by the AMSR-E sensor on board the NASA EOS Aqua satellite. For this classification, the land surface brightness temperatures were used at two frequency bands (the 36.5 and 89 GHz bands with both horizontal and vertical polarizations) (Kawanishi et al., 2003). These brightness temperatures are passive measurements of the emitted microwave radiation from the earth surface and vary with differing quantities of water on the land (Brakenridge et al., 2012). As such, a decrease in the brightness temperature compared to the surrounding land surface is strongly indicative of inundation or flooding on the land surface (Khan et al., 2012). The daily observations of brightness temperature for each of the bands
and polarizations (4 total) are available at a spatial resolution of approximately 25 km and are split into ascending and descending datasets based on the satellite orbit. This results in 8 time varying features that will be used as the main drivers for predicting inundation. Along with this, additional features are added to represent the spatial correlation between observations at the 8 neighboring grid cells. This results in 72 total features related to the brightness temperature observations. An example of these observations can be seen in Figure 5.2 below.

![Observed AMSR-E Brightness Temperatures in Kelvin for the Upper Mississippi River Basin on June 08, 2008. The cooler colored areas indicate regions with more water where flooding is likely to have occurred.](image)

Figure 5.2: Observed AMSR-E Brightness Temperatures in Kelvin for the Upper Mississippi River Basin on June 08, 2008. The cooler colored areas indicate regions with more water where flooding is likely to have occurred.

Additional features for the classification come from the National Elevation Dataset (NED) that provides a digital elevation map for the entire United States at a 30 meter spatial resolution (Gesch et al., 2002). From this, a number of additional indices are calculated to represent various features of the landscape that are related to the movement of water. The indices derived from the NED are listed below and result in 12 additional features for the classification. An example of the NED can be seen in Figure 5.3 below.
• Elevation (DEM)
• Upslope Area
• Aspect
• Curvature
• Slope
• Multi-resolution Index of Valley Bottom Flatness (mrvbg)
• Multi-resolution Index of the Ridge Top Flatness (mrrtf)
• Profile Curvature (profc)
• Plan Curvature (planc)
• Topographic Index (ti)
• Topographic Position Index (tpi)
• Terrain Ruggedness Index (tri)

Figure 5.3: Elevation in meters for the Upper Mississippi River Basin as provided by the National Elevation Dataset.

The final feature used for the classification was the Stage IV Radar Precipitation product produced by the National Center for Environmental Prediction (Lin and
E Mitchell, 2005). This product provides hourly estimates of rainfall at a 4 km spatial resolution, which was then averaged up to daily values. Given that large amounts of precipitation cause flooding, it is likely that the observed rainfall will be useful in predicting inundated regions.

The combined feature set for each day consists of the 85 features previously mentioned. Each dataset was downloaded and processed at its native resolution and then resampled to the 500 meter resolution desired for prediction using a box averaging method. This resulted in a domain of 2200x2340 grid cells (5148000 total) for each day. In order to construct the training and testing datasets from these features, all possible coordinate pairs were randomly sampled to select 2% of the grid cells for training. These sampled grid cells were evenly split between those that were inundated and those that were not. Over a 30 day period in which severe flooding occurred, the feature set was sampled for each day and was combined to produce a training dataset of 3294720 grid cells. The remaining 98% of cells were used for testing to determine the ability of the classification method to reproduce the modeled conditions.

Along with these features used for classification, a dataset of class labels was needed for each day over the entire domain for use in both the training and testing stages. Classes were created at a 500 meter resolution using the Variable Infiltration Capacity Land Surface Model (VIC LSM) forced by Stage IV Precipitation and routed with the Catchment-based Macro-scale Floodplain (CaMa-Flood) Model (Yamazaki et al., 2011). With CaMa-Flood we were able to model the flooding conditions and generate a mask of inundated area for each day with which our classification method could be trained. An example of this data can be seen in Figure 5.4 below.
Figure 5.4: Model produced inundation estimates for two regions. The elevation of water in flooded regions is shown in meters. This data is converted to a binary mask where all of the colored cells are given a class of “inundated” and all other cells are “not inundated”.

5.2.2 Classification Method

A random forest classification algorithm was constructed to utilize the training features of the sampled points for the entire 30 day period. This classification method was implemented from the SciKitLearn Python libraries (Pedregosa et al., 2011) using 100 trees and the Gini impurity scores as the criterion for selection at each split in the tree. The maximum number of features considered at a split and the maximum depth of the trees were not constrained. The random forest was selected over a single decision tree for this task because a regular decision tree tends to overfit the training data, leading to a model that is not robust for other testing datasets. By using multiple decision trees constructed with different samplings of the training data, we are able to make a model that is less sensitive to noise in the training data and will likely lead to better classification results (Murphy, 2012). A flow chart of the overall classification process can be seen in Figure 5.1.
5.2.3 Evaluation

In order to evaluate the performance of each classification method on the testing dataset, we used the precision and recall scores. The entire dataset was split such that only 2% of the grid cells were used for training and 98% were used for testing. The precision and recall scores are presented as a curve to illustrate how the performance changes when the threshold value is varied. The general metrics are defined in the following equations, where the number of false positives (FPs), true positives (TPs), and false negatives (FNs) are used:

\[
\text{precision} = \frac{TP}{TP + FP}, \quad \text{recall} = \frac{TP}{TP + FN}
\]  \quad (5.1)

Along with the calculation of these metrics, Receiver Operating Characteristic (ROC) curves were created for each of the methods used to evaluate the prediction quality. These plots illustrate the relationship between the true positive rate and the false positive rate. The ideal curve will have a false positive rate of zero, and a true positive rate of one. Along with this, the area under each curve was calculated, with area increasing to one for the best performing methods. Although the ROC curve is commonly used to evaluate binary problems such as this, it has been shown that they can present an overly optimistic view for skewed class distributions, such as we have in our testing dataset (Davis and Goadrich, 2006). As a result of this, the combined information from these two types of curves will be used to evaluate the model performance.
5.3 Results

5.3.1 Classification

During May and June of 2008, the Upper Mississippi River Basin experienced significant rainfall and flooding. This event was used as a case study for the classification. By training the random forest method on the training dataset described previously, we were able to create a model that could be used to classify the 30 daily images. An example of this daily classification can be seen in Figure 5.5 below. The subsetted region is the same as that in Figure 5.4a. By comparing these two regional images we can see that a large number of the inundated cells were classified correctly; however, there are still a number of grid cells that were not correctly classified.

![Figure 5.5: Example of fully classified image for June 2008. The red regions indicate areas that were classified as inundated.](image)

To better understand the performance of the classification, we can look at the ROC and Precision-Recall curves shown in Figure 5.5 below. In this analysis, a classification method that provides a good prediction will create a curve close to
the upper left corner in the ROC plot and close to the upper right corner in the Precision-Recall plot. In each case, the calculated area under each curve can be used as a performance metric, where the area should be close to 1. Overall we can see that the model performs relatively well. From the ROC curve in Figure 5.6a we can see that the model performance is significantly better than pure chance, with an area of 0.8. By examining the Precision-Recall curve in Figure 5.6b we can see similar results. Because this curve does not take into account the true negative classifications, of which there are significantly more for the testing dataset, we can get a clearer picture of the model performance than using a ROC curve alone. In this case the area under the curve is 0.7 indicating that we still have a model which performs relatively well for classifying inundation throughout the flooding event. It is interesting to note the sharp drop in precision that occurs at a recall of 0.5, indicating that the number of false positives increases significantly with the varying threshold. This threshold corresponds to the cutoff threshold of probability for the random forest classification. Further analysis of this is given in the Discussion section.

Figure 5.6: Receiver Operating Characteristic and Precision-Recall curves. These curves were generated from the classification results of the full 30 day dataset of daily images.
5.3.2 Feature Selection

Along with the performance of the classification model, the random forest can be used to extract relative feature importances for all of the covariates used. The top 15 most important features and their importances are plotted in Figure 5.7 below. It is interesting to note that the topographic indices which are static in time are the most important. The top four features are all related to the curvature and slope of the land surface and differentiate between areas that have a high slope (such as hillsides) and regions that are close to flat (such as floodplains and rivers). This result makes sense, as we would expect low lying areas that are flat to be significantly more likely to be classified as inundated than areas with high slopes and higher elevations. For the brightness temperature observations, the most important features are both polarizations of the 36.5 GHz band. Recent studies have also found this to be the case, with strong correlations observed between this band and the amount of water on the earth surface (Brakenridge et al., 2012). Finally, it is interesting to note that the precipitation feature is not found to be important in relation to all of the other features used.
From an analysis of the feature importances, we can also find that the added features of brightness temperatures in neighboring grid cells does not have a significant impact on the classification. As a result of this, it is likely that we could remove these features from the model while still achieving similar performance. This would have a significant impact on the model by reducing the amount of computational effort that it takes to train the random forest model by removing 64 features. Further work is needed to determine whether there is a better way to incorporate these features into the model, such as calculating the anomaly between the chosen and neighboring grid cells, given that a number of recent studies have found this relationship to be indicative of flooding and high discharges (Brakenridge et al., 2007, 2012; Khan et al., 2012).
5.4 Discussion

Within this study, we have performed the classification of inundated areas through the use of a random forest classifier. This classifier was trained with a feature dataset consisting of a number of elevation derived indices, precipitation, and AMSR-E brightness temperature observations. By training the method on data collected over the months of May and June 2008, a period when significant flooding occurred. The application of this method resulted in daily classified maps of inundation at a resolution of 500 meters. Given that the performance was relatively good overall, we can conclude that the classification model created is useful for predicting the extent of flooding in the defined region. This method may prove useful in further applications because it makes use of satellite observations and static elevation indices that can be determined globally, allowing us to easily make daily estimates of inundation for applications in hydrology and risk management.

Although the results of this study are generally positive, the performance evaluation revealed that there are a number of inundated grid cells that are not being correctly classified. To better understand the reason for this we looked at the probability of a cell being classified as inundated, as shown in Figure 5.8 below. These probabilities are generated by looking at the range of predictions generated by the decision trees in the random forest. Figure 5.8a corresponds to the same regional subset shown in Figure 5.5. From these two maps we can see that a number of the regions which were not correctly classified as being inundated lie in the probability range of 0.4 to 0.5, indicating that they were not classified correctly because the probability was slightly below the classification threshold used in the random forest. By adjusting this threshold, we would be able to get a higher number of true positive classifications but this may also increase the number of false positive classifications, which would not be desirable. These patterns agree well the with the performance
characteristics shown in Figure 5.6b where the precision decreases significantly above a recall of 0.5. In order to improve upon these results, further work is needed to identify time varying features that can increase the probability of inundation prediction for the truly inundated cells.

Figure 5.8: Inundation probabilities generated from the random forest estimators. Regions of probability greater than or equal to 0.5 were classified by the random forest as being inundated.

One of the main challenges in developing and validating algorithms for monitoring inundated areas is the availability of training and validation data. In this study we chose to use a large flooding event during the AMSR-E period of observation and used a synthetic dataset for training and validation. One alternative to this is to use visible imagery of flooding (when available) as the training and validation datasets. The Dartmouth Flood Observatory produces these maps for certain flooding events, combining the observations of MODIS and Landsat (Brakenridge et al., 2015). As an extension of the study described previously, we applied the same random forest algorithm to a new set of data to look at the performance over a more recent flood event (that existed in the DFO record). Because the AMSR-E mission ended in 2011 we chose to use the brightness temperature observations from the more recent
Soil Moisture Active Passive (SMAP) mission. This satellite contains a radiometer that provides L-band brightness temperature observations (at a ground resolution of approximately 36km) with a revisit time of 3 days or better (Entekhabi et al., 2010). which can be used instead of the AMSR-E. Keeping all other external data products consistent, the random forest was again fitted with the SMAP data.

This analysis was done for the floods that occurred during May and June of 2015 throughout much of Texas. The DFO produced a composite product of flooding indicating where any surface water had been observed during the two week period of flooding. A comparison of the DFO product can be seen in Figure 5.9, where the DFO product is overlaid on the predicted surface water extents from the random forest. In general we find see relatively good agreement between the two products but in some areas the extent of flooding observed extends beyond what the classification predicted. Figure 5.9A illustrates this well, with many DFO observed surface waters extending beyond the main river channel. It is possible that this is in part an artifact of the MODIS data, as it has a coarser resolution than that of the covariates used in the random forest. Overall, it is clear that there is some agreement between the two products, indicating that our random forest algorithm may be effective at predicting surface water extents from passive microwave observations. Without a true baseline product of inundation measured at a high temporal frequency, though, it is difficult to perform training and validation of either of these products.
Figure 5.9: Comparison of predicted surface water extents from Tb and the multi-day composite flood extent image produced by the DFO from MODIS and Landsat. For regions A and B, yellow areas represent areas of agreement between the products and Landsat imagery.
5.5 Conclusions

Overall, this work has shown that the use of a trained random forest classification method can produce reasonable estimates of inundation. By combining the time varying information from the AMSR-E brightness temperature observations with the static topographic features, we are able to generate estimates of inundation state at a 500 meter resolution. Given that AMSR-E observations are available globally and exist for regions that are persistently covered by clouds or vegetation, this method can provide a new source of information for monitoring and predicting the extent of flooding. We have also shown that the same algorithm will work relatively well with new satellite observations of Tb, potentially allowing for a near real-time application.

As was mentioned previously, there are a number of items that could potentially improve the results of this classification method. First, it would be useful to generate an informative feature set that relates the brightness temperature of the current grid cell to the neighboring cells. Simply using the values was shown to be ineffective in this study, but another type of feature, such as the difference between values, may prove to be more informative. In addition, further refinement is needed for a feature based on precipitation. The feature used in this study was most likely ineffective because it represented only the precipitation of the current day, while flooding is a result of the precipitation multiple days prior to the current. Adding a feature of accumulated 7-day precipitation, for example, may prove to be more useful.

With regards to feature selection and training, further refinement and testing is needed to derive the ideal training dataset. Within this study, all possible topographic indices were used to train the classifier but it was found that not all of these features were important. As such, one could vary the features selected to determine how the model performance is affected to find the optimal set. By reducing the size of the feature set we will be able to decrease the computational complexity of the model,
which will make the classification method more applicable to large spatial domains and near real-time applications. Finally, further work is needed throughout the flood observation community to develop training datasets based on real world observations of flooding extent. For this study we labeled areas of inundation using model outputs, as well as another remotely sensed product. In a real world application, we may not have these data and we would like to have a classification algorithm that is trained on maps of inundation extent for a number of real flooding scenarios. These maps should be generated from a careful analysis of satellite imagery or by water resource managers who hand map the extent of flooding for local areas. Incorporating these data sources for training will likely provide a more robust and accurate passive microwave based surface water extent classification system.
Chapter 6

Summary and Conclusions

The underlying motivation of this dissertation was to utilize new sources of hydrologic information to improve our understanding of the global water balance and to improve our ability to accurately model these systems. To that end, the work presented here has focused on the application of remote sensing observations of global surface waters to enhance the currently available suite of observations. In doing so, this work has described methods by which new datasets of global water availability can be created. A brief overview of the key findings of this work follows.

Given the current decline in globally available in-situ gauge networks, Chapter 2 focused on the generation of methods that can extend the usefulness of these limited observations to other areas. In particular, a two-sweep method for reconstructing spatially and temporally continuous discharge records from discrete observations of discharge was developed. This method represents an improvement over traditional streamflow data assimilation techniques in that it is constructed to maintain the spatial and temporal consistency throughout a river basin, both backwards and forwards in time. The results of the synthetic experiments conducted illustrate the potential for this method to reconstruct discharge records when provided with adequate obser-
vations. Additionally, it was shown that this method can function purely as a form of interpolation, requiring no hydrologic initial conditions to reconstruct basin-wide discharge.

While this spatiotemporal interpolation method was shown to produce viable discharge reconstructions from in-situ gauge records, it does not leverage any remote sensing which may be able to provide additional information. This is especially important for regions of the world with very sparse gauging or none at all, where remotely sensed discharge observations may be the only option. Chapter 3 addresses this issue in the context of the upcoming Surface Water and Ocean Topography mission, which will provide global river heights from a swath altimeter. While this mission provides near complete global coverage, the temporal sampling is relatively coarse, with some areas being observed only 1-2 times in the 21 day cycle. A series of synthetic experiments were carried out to determine how well we would be able to reconstruct continuous discharge records from SWOT. The results of these experiments illustrated the potential for the interpolation method to reconstruct basin-wide discharge; however, the performance was limited by the temporal sampling of SWOT. As a result of this, we explored alternative interpolation schemes, finding that by adding a small number of in-situ gauges we were able to perform a much better reconstruction of the true discharges. While this may not help to construct discharge records in ungauged basins, it does indicate that SWOT can serve as a compliment to in-situ gauge networks.

Moving beyond the Ohio River basin, Chapter 4 explores the potential for global SWOT observations of discharge as well as the potential for interpolating these observations to create a global discharge product. In this context, the spatiotemporal interpolation method is used to perform discharge reconstructions in 23 large global basins from synthetic SWOT observations alone. The results of these reconstructions
were generally good; however, it was noted that each basin appeared to have varying levels of success. These differences were related to the specific pattern with which SWOT will cross over the observable rivers, as defined by the general latitude of the basin, the orientation of rivers within the basin, and the overall size of the basin. These characteristics were further explored to examine the potential information that SWOT may be able to provide, illustrating the need for careful consideration of these characteristics to provide a reliable continuous global discharge product.

In-situ measurements can provide us with general information on river conditions, but these observations become inadequate at the onset of a major flooding event where the surface waters extend beyond the river channel and floodplain areas. In addition to SWOT, there are a number of other earth observing satellites that can be used to address specific problems related to surface water observations such as passive microwave measurements traditionally intended for use in monitoring soil moisture conditions. Chapter 5 describes one potential pathway to move beyond the traditional observations through the use of machine learning techniques. In this study remotely sensed passive microwave observations are used to derive high resolution estimates of inundated areas additionally informed by a set of high resolution topographic parameters. This method was shown to produce relatively accurate flood maps for two specific floods in the United States. In addition to simply producing reasonable maps of inundation, the machine learning based approach is efficient and fast, which could allow for the development of a near real-time flood monitoring product derived from the passive microwave observations of multiple satellites.

Overall, this dissertation explores the application of remote sensing generated global estimates of runoff and discharge that are greatly needed. The methods presented here have focused on a limited set of sensors and applications. In reality, the best method to provide an accurate and complete global discharge product will be
to integrate and assimilate the multitude of observations, such that we can constrain the uncertainties in each product alone. Advancements in these areas will be critical to monitoring the availability of freshwater resources for human consumption, agriculture, and industry, as well as to minimize the impacts of hazards posed by water extremes like floods and droughts.
Bibliography


Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A.,


