Summer 2017

A spatio-temporal dynamical evaluation of satellite rainfall products in hydrologic applications

Mohamed ElSaadani
University of Iowa

A Spatio-Temporal Dynamical Evaluation of Satellite Rainfall Products in Hydrologic Applications

by

Mohamed ElSaadani

A thesis submitted in partial fulfillment
of the requirements for the Doctor of Philosophy
degree in Civil and Environmental Engineering in the
Graduate College of
The University of Iowa

August 2017

Thesis Supervisor: Professor Witold F. Krajewski

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

_____

PH.D. THESIS

_____

This is to certify that the Ph.D. thesis of

Mohamed ElSaadani

has been approved by the Examining Committee for
the thesis requirement for the Doctor of Philosophy degree
in Civil and Environmental Engineering at the August 2017 graduation.

Thesis Committee: _____
Witold F. Krajewski, Thesis Supervisor

_____
Ricardo Mantilla

_____
Allen Bradley

_____
Dale L. Zimmerman

_____
Emad Habib

To my mother, Sawsan, my sisters, Rania and Marwa, and the memory of my deceased father, ElSayed.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Prof. Witold F. Krajewski, for his guidance during my studies. Working with Prof. Krajewski significantly contributed to my professional growth and improving my lifestyle. I am glad I had the opportunity to work with him. I would like to extend my gratitude to my thesis committee members, Dr. Ricardo Mantilla, Emad Habib, Allen Bradley, and Dale Zimmerman for their useful suggestions and comments.

I would also like to thank my lab mates Chad Drake, Danny Horna, Andre Zanchetta, Silvia Volpi, Tibebu Ayalew, Navid Jadidoleslam, Munsung Kim, Gabriel Perez, Felipe Quintero, and others, for their friendship and the good times. Special thanks go to staff engineer Radek Goska for his help with my research projects, and for being such a great friend.

Lastly, I would like to thank my mother for all her help and sacrifices to insure my success. This work could not have been possible without her help. No words can express how lucky, and grateful I am for having such a great mother.

ABSTRACT

In February of 2014 NASA has launched the core observatory of The Global Precipitation Measurement Mission (GPM). Since then, the mission has been providing a wealth of observation data collected by the core observatory along with other satellites belonging to the mission space constellation. One of the most important data products that GPM provides is the Level 4 (L4) rainfall data product called Integrated Multi-satellitE Retrievals for GPM (IMERG). IMERG is constructed using the raw data collected by the Microwave (MW) sensors on board the constellation satellites along with the Infrared (IR) sensors on board geostationary satellites and the advance Dual-frequency Precipitation Radar (DPR) on board the GPM core satellite. The IMERG product is available globally for all interested researchers to use. In this dissertation, I focus on the applicability of IMERG in hydrologic applications, and specifically in flood peak modeling.

In order to conduct a comprehensive evaluation of IMERG that is oriented towards hydrologic modeling. I have explored multiple hydrologic models which can be used to produce stream flow estimates using IMERG without the need of parameter calibration based on the model's inputs. The calibration free capability is essential since model parameter calibration obscures the effect of the errors associated with the rainfall input on the estimated discharges, which in turn will limit our understanding about the distribution of the errors in IMERG over space and time. The two hydrologic models we used in this study are both physically based distributed models and were setup over the domain of the state of Iowa which is located in the United States' Midwest. I also explored the performance of one of the hydrologic models' component, which is the

runoff-routing component, in order to estimate an additional portion of the errors in the discharge estimates that is not attributed to the model's input but rather to the hydrologic model itself.

A significant portion of my dissertation is concerned with identifying and using accurate methods to evaluate both IMERG and the hydrologic models' outputs in a hydrologic context that is useful for flood modeling. Several studies have evaluated other satellite rainfall products using methods that vary in complexity. Some studies used the simplest methods of evaluation, such as, mean aerial differences and standard deviation of the differences (additive or multiplicative) compared to a benchmark rainfall product. This is done without taking the spatial dependency of the errors in space into consideration. Other studies modeled the spatial dependency (correlation) between the errors in the rainfall product, however, using Euclidean distance based approaches that do not account for the hydrologic basins' shape and size. Nevertheless, it is important to realize that hydrologic models will eventually aggregate the rainfall values, along with the errors associated with them, through a stream network that is dichotomous in nature and does not comply with Euclidean distance. Thus, we employed a stream based evaluation framework, called the Spatial Stream Network (SSN) approaches, to characterize the errors in IMERG taking into account the stream distances and the stream connectivity information between evaluation sites. Although previously used in applications such as modeling water temperatures and pollutant transport, to the best of my knowledge this approach has not been used in rainfall product evaluation before this study. The SSN analysis of IMERG allowed me to answer the question, "What is the

proper basin scale which is capable of filtering out the correlated errors in IMERG by accumulating the rainfall values through the stream network?"

Finally, in order to add value to the current methods of evaluating model simulated stream flows. I proposed a time based evaluation that is capable of detecting peaks in both the observed and simulated flows and estimating the lag time of the simulated peaks. Typically, previous studies have used simple skill scores such as Root Mean Squared Errors (RMSE), correlation coefficient, and Nash-Sutcliff Efficiency (NSE) to evaluate hydrograph performance as a whole, or the difference in time to peak which involves primitive peak detection method (e.g., a moving or a defined time window). In this dissertation I propose a Continuous Wavelet Transform (CWT) based method to evaluate the peak times and shapes produced by the hydrologic model. The method is based on filtering the frequencies in the hydrograph by treating it as a signal and detecting sharp features in both the observed and time series and the phase difference between them. We also emphasized on the importance of the choice of wavelet shape used in the evaluation, and how different wavelet shapes can affect the inference about the time series.

PUBLIC ABSTRACT

Rivers provide us with one of the most essential natural resources that we need to survive, fresh water. Human populations lived alongside rivers since ancient times to use their fresh water for drinking and irrigation. This benefit comes with the cost of being exposed to the risk of flooding. Without proper monitoring technologies, floods can be unpredictable to individuals. This is because a flood can result from a rainfall that occurred hundreds of miles away, but was accumulated and transported through rivers to the affected location. In addition, floods are not caused by excessive rainfall alone. A combination of land and atmospheric conditions must happen at the same time to cause floods. For example, a dry soil can absorb large amounts of rainfall and prevent water from reaching the rivers to cause floods. Thus, it is important to collect information regarding the land and atmospheric conditions simultaneously. In recent decades, scientists produced instruments that can monitor atmospheric and land conditions over large areas. These instruments can be attached to satellites that revolve around earth and can stay in space for many years, thus providing global coverage for a long time. These instruments provide valuable information, however, we need to know how accurate they are. My Ph.D. thesis topic is to evaluate the performance of one of the instruments that monitor rainfall from space. This evaluation is intended to solve one piece of the puzzle, which includes evaluating the performance of other instruments that can monitor other natural conditions such as humidity, temperature, and soil water content, among others. Together, all these instruments can provide us with a complete picture of the existing natural conditions. This information can then be used by experts in order to warn the public in case of potential floods before they happen.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

.

CHAPTER 1: Introduction

Flooding is one of the most destructive natural hazards in the United States and many other regions across the world. Flood related damages include economic losses as well as losses in lives. During the past decade alone, the central United States experienced many floods (e.g., 2008, 2011, 2013, and 2014) that caused economic losses estimated by billions of dollars and many fatalities (Mallakpour, 2016; NCDC, 2015). The 2008 flood in particular caused about $11 billion in damages and resulted in 24 fatalities (Mallakpour, 2016; Mutel, 2010). Thus, it is crucial to continuously monitor floods in order to be able to warn the public in case of potential floods. Nevertheless, excess stream discharges which cause floods are a result of complex land and atmospheric interactions (e.g., rainfall, temperature, pressure, humidity, and soil moisture). Monitoring these process everywhere using point measurement instruments is not possible. Thus, remote sensing of environmental and atmospheric variable plays an important role in flood modeling and predictions. Remote sensing of atmospheric and surface variables is available through both ground and space based instruments. However, these instruments provide indirect measures of the variables they are monitoring based on theoretical algorithms that are prone to errors.

In this dissertation I focus on evaluating the rainfall estimates provided by earth observing satellites. Satellite rainfall estimates provide a wealth of information to interdisciplinary applications related to the water cycle. This is due to the availability of observation records that cover a period of over two decades. In addition, satellite rainfall products provide global rainfall estimates in a gridded format that is compatible with numerical weather prediction models and distributed hydrologic models. During the past

couple of decades, a variety of satellite rainfall products emerged. These products used information provided by geostationary Infrared (GEO IR) satellites and Passive Microwave (PMW) Low Earth Orbiting (LEO) satellites. However, each product has its own algorithm which incorporate the information retrieved from satellites in different ways. Moreover, each of these products have its own spatial and temporal resolution based on the purpose for which it was created. Examples of products that have been heavily used and evaluated during the past two decades include, first, the Tropical Rainfall Measurement Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) Product produced by NASA. Second, the Climate Prediction Center (CPC) which is a part of the National Oceanic and Atmospheric Agency's (NOAA) produced the CPC's Morphing Technique (CMORPH) which produces a high resolution satellite rainfall product. Finally, the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) produced by the Center for Hydrology and Hydrometeorolgy, at the University of California, Irvine.

TMPA rainfall estimates are produced in 0.25˚ x 0.25˚, three hourly, spatial and temporal resolutions respectively, with latitude coverage between 50˚ N and 50˚ S. TMPA merges the GEO IR rainfall estimates with the estimates collected by PMW sensors such as TRMM's Microwave Imager (TMI), Advanced Microwave Sunding Unit (AMSU), Special Sensor Microwave/Image (SSM/I), and th Advanced Microwave Sounding Radiometer-Earth (AMSR-E) observing system. The method used in TMPA calibrates the IR estimates using PMW data before merging them. In addition, TMPA combines satellite estimates with monthly rain gauge estimates provided by the Golobal Precipitation Climatology Project (GPCP).

CMORPH rainfall estimates are available at a higher spatial resolution of 8 km x 8 km (at the equator) and half hourly temporal resolution. The main difference between CMORPH and TMPA products is that CMORPH uses only PMW rainfall estimates. However, motion vectors from GEO IR observations are used to interpolate these rainfall estimates in space and time. This is because LEO satellite do not provide continuous earth coverage due to the nature of their orbits. Thus, interpolation is necessary to fill the gaps between satellite overpasses.

PERSIANN's method of rainfall estimation is based on artificial Neural Networks methods. This method detects clouds observed by the GEO IR and PMW satellites and then assigns rainfall estimates to these locations based on cloud shape and other characteristics that are related to rainfall rate. Throughout the process, PMW observations from TMI are used to adjust the rainfall estimation model parameters (Hong et al. 2004).

In February, 2014, NASA launched the Global Precipitation Measurement (GPM) mission's core satellite. This satellite is considered as a replacement of the recently retired TRMM core satellite, with improved instruments on board. The major improvement to this mission is the addition of the Dual-frequency Precipitation Radar (DPR). The radar has two bands, first, the Ku-band which is similar to that incorporated in the PR on board TRMM satellite, second, the DPR has a higher frequency band called the Ka-band which is useful in detecting snow and light rain. This is applicable in the case of the GPM mission since the orbit of the new satellite has higher inclination compared to that of the TRMM satellite which allows it to pass over cold regions. The rainfall estimates by the GPM mission were recently released to the public. The magnitude and nature of the errors associated with the new product is yet to be

understood. My product of interest is the recommended level 4 research product called Integrated Multi-satellitE Retrievals for GPM (IMERG). This product merges the methodologies of the three older rainfall products (TMPA, CMORPH, and PERSIANN), to produce a high resolution product (half hour, 0.01˚ x 0.01˚). However, my evaluation of this important new product was hydrology oriented, since I am mainly interested in flood modeling and prediction.

For this purpose, it is necessary to have hydrologic models that can convert rainfall to runoff without needing calibration based on the rainfall input. This is important because calibration will obscure the true performance of the model inputs and limit our understanding of the rainfall products as model inputs. In chapter 2, I conducted a study which aims to evaluate hydrologic models that match this criterion. The main goal of this study was to investigate the performance of independent hydrologic models components, mainly, the routing component of the distributed hydrologic models. I conducted a comparison between two routing components that are based on linear and non-linear approaches, and proved that the routing component can have a significant effect on the performance of the overall hydrologic model. The results of this study showed that using the non-linear routing component significantly improved the model performance. Thus, the two hydrologic models I later in this dissertation incorporate none-linear routing.

In chapter 3, I conducted the evaluation of IMERG rainfall estimates. The evaluation consists of two main steps. First, rainfall to rainfall comparisons. The benchmark product used in this comparison is the Multi-Radar/Multi-Sensor (MRMS) ground radar based rainfall product produced by the National Severe Storm Laboratory (NSSL). In this step I used traditional statistical skill scores such as bias, Root Mean-

Squared Error (RMSE), and correlation coefficient. Moreover, I performed spatial statistical characterization of the space-time dependency between rainfall estimates of each product separately using the spatio-temporal semi-variograms. I also constructed these semi-variograms for the discrepancies between IMERG and MRMS to explore if they are correlated in space or time. The semi-variograms showed that IMERG rainfall estimates have higher correlation in space and time in comparison to MRMS. This can be due to the interpolation method which the PMW estimates undergo during product construction. This in turn resulted in rainfall discrepancies that are also correlated in space, however, the temporal dependence of the discrepancies was negligible in time.

Now we move on the second step towards the second evaluation step. In this step I evaluate IMERG as a hydrologic model input. I forced two hydrologic models with both IMERG and our benchmark product MRMS. The combination of models and rainfall products is useful as a mean of investigating if the quality of the simulated stream discharges is related to the rainfall input or the model itself. Both models do not need to be calibrated based on the rainfall input. This is important because the calibration process obscures the effect of the errors in the model inputs by modifying the parameters. The evaluation included skill scores similar to those I used for rainfall to rainfall comparisons such as, correlation, RMSE, bias, as well as contingency table estimates (e.g., HR, FAR, CSI, Bias ratio). The latter methods were used specifically to evaluate the performance of the model in detecting peaks over certain thresholds. This is necessary for evaluating the model in flood forecasting framework. However, these measures cannot capture simulation peak time precisely. Therefore, I estimated other statistics such as time to peak which is widely used by the hydrologic community.

In chapters 2 and 3, I performed the evaluation of IMERG using traditional statistical skill scores. These skill scores provide limited insight about the performance of IMERG that is not hydrology oriented. For example, all the spatio-temporal dependency comparisons described above were done using the gridded rainfall products in IMERG resolution. Nevertheless, since we are interested in the performance of IMERG in hydrologic context, it is important to observe the effect of the accumulation which the rainfall product undergoes by the hydrologic model on the rainfall discrepancies dependency range. Ideally, if the errors are completely random they should cancel each other as they accumulate downstream. Studies such as Quintero et al., (2016) and Cunha et al., (2015) provided an evaluation framework of rainfall discrepancies that mimics the accumulation process done by the hydrologic model.

I tackle this problem in chapter 4 by providing a hydrologic based evaluation of IMERG. As a demonstration, I used this framework to evaluate IMERG over a mid-sized basin located in eastern Iowa (approximately 17,000 km^2), called the Cedar River basin. The evaluation focused on a major flooding event (second largest on record) which occurred during the last week of September, 2016. IMERG errors during this period did, to a large degree, cancel out as they accumulated downstream as we expected. However, now we are faced by a new challenge, mainly, how can we describe the spatial structure of the accumulated errors? This is because all rainfall estimates are now assigned to stream segments rather than a latitude and longitude based geographical location.

Thus, it was necessary to use a statistical approach which can utilize the stream segment based rainfall estimates and use network distances rather than the Euclidean distances connecting the rainfall estimate locations. I accomplished this task by using a

stream distance based spatial statistical modeling framework. This method is called the Spatial Stream Network (SSN) approach (Ver Hoef et al., 2006). Not only does this method used the stream distance rather than Euclidean distance, it also accounts for stream network connectivity information. This is important because the dependency of the accumulated rainfall estimates is higher between connected streams in comparison to unconnected streams. This is also the case for many other hydrologic variables (e.g., point source pollutants, fish populations, and sediment and nutrient transport) where the magnitude of a variable can be very different between two unconnected streams despite their geographical proximity.

I began investigating the new method by comparing the semi-variograms of the rainfall errors which I produced using stream distance and connectivity information to the traditional Euclidean based semi-variograms. The two semi-variograms obtained from the connected and unconnected streams are called the Torgegram (Zimmerman and Ver Hoef, 2016). The results showed that the errors in flow-connected stream are correlated for significantly longer distances (almost the double) in comparison to the errors in flow-unconnected streams. In addition to calculating the Torgegrams, SSN provide us with valid covariance models that can be used for spatial statistical modeling and prediction. Unlike the traditional statistical linear models, spatial linear models allow covariance between the model errors. One way to generate the covariance structure is by using a moving average construction method. The moving average functions used to compute the covariance can take many forms (e.g., exponential, spherical, linear) and also incorporate network connectivity information. Moving average functions are unilateral (pointing in on direction). The moving average functions which are pointing in the upstream direction

are called tail-up functions and allow covariance between flow-connected streams alone. The moving average functions which point in the downstream direction allow covariance between both flow-connect and flow-unconnected sites, and are called tail-down functions. The choice of the moving average function is based on the nature of the variable being analyzed. In our case, rainfall covariance should be allowed between both flow-connected and flow-unconnected sites. This is because some low order streams can be located within the same rainfall pixel thus have similar errors. However, it is expected that the degree of covariance between flow-connected streams is higher. Using a mixed covariance model can account for this effect (Ver Hoef et al., 2006; Zimmerman and Ver Hoef, 2016). I tried several model types using pure tail-up or pure tail-down models as well as a combination of mixed models. All mixed SSN based models outperformed the traditional Euclidean based models with different degrees. This indicated the possibility of better prediction of rainfall error fields over stream networks, which is relevant to hydrologic modeling applications.

In addition to the limitations of the skill scores used to evaluate the spatial rainfall estimates of IMERG, there are also limitations in the field of evaluating IMERG as a model input. For example, identifying and measuring the difference between modeled and observed peak times is traditionally done using simple methods that are useful, however, they can lead to false identification or unfair comparison between the modeled and observed discharges. Thus, I decided to explore other possible ways of identifying the modeled peaks that correspond to the observed peaks.

In chapter 5, I describe how I found out that by treating the hydrograph as a signal and using signal processing methods we can obtain accurate peak timing evaluation

estimates. The methods used in signal processing vary in complexity, starting with the Fourier Transform, to Fast (windowed) Fourier Transform (FFT), or Continuous Wavelet Transform (CWT). I focused on the CWT due to its advanced capabilities in identifying sharp features in the signal and distinguishing them from broad features. This is particularly useful to distinguish peaks from base flow. I will also describe the cross wavelet transform (XWT) which I used to relate the simulated peaks to the observed peaks. The phase difference between the real and imaginary components of the XWT can be converted to time difference which directly corresponds to the time difference between the CWT of the simulated and observed flows. I also emphasized on the effect of the choice of wavelet shape on the estimated time lag.

The remainder of this dissertation is organized as follows. Chapter 2 contains the performance analysis of the hydrologic models I used in IMERG evaluation. Chapter 3 contains the evaluation analysis of IMERG, this includes the rainfall to rainfall comparison with our benchmark product IMERG, and it also includes the evaluation of IMERG as model input. Chapter 4 contains the SSN analysis for characterizing IMERG errors in a network based framework. Chapter 5 contains the time based evaluation framework for hydrologic models using CWT and XWT. In Chapter 6 I summarize the findings in this dissertation and list the conclusions.

CHAPTER 2: An Investigation of Errors in Distributed Models' Stream Discharge Prediction Due to Channel Routing

2.1. Introduction

Floods result from complex interactions between rainfall and several other processes occurring in the atmosphere and landscape. Accurate quantification and prediction of stream discharges is a challenging topic that the hydrologic research community has been trying to overcome for decades. Over the summer of 2015, the National Water Center (NWC), located in Tuscaloosa, Alabama, USA, launched the National Flood Interoperability Experiment (NFIE) Summer Institute, a summer workshop that allowed participants (mainly graduate students) to experiment with research topics related to national flood estimation and forecasting, ranging from runoff generation and discharge estimation to digital inundation mapping and flood risk management. The first author participated in NFIE 2015 and evaluated the performance of the NFIE-Hydro framework, a hydrologic modeling framework that can provide nationwide stream flow estimates in the United States.

In this chapter, we build on this effort by identifying possible ways to improve the NFIE-Hydro framework (Maidment, 2016). The NFIE-Hydro framework consists of a combination of a runoff-generating Land Surface Model (LSM), Noah Multi-Parameter (Noah-MP) LSM (Niu et al., 2011), and the runoff discharge routing component called the Routing Application for Parallel Computation of Discharge (RAPID) (David et al., 2011). It is also simple to incorporate the runoff estimates from other LSM's, such as or the Variable Infiltration Capacity (VIC) (Xia et al., 2012a, b) as a substitute to the runoff estimates from Noah-MP. Previous studies evaluated the VIC-RAPID combination (e.g.,

David et al., 2013; 2015; Tavakoly et al., 2016) over multiple regions in the United States. In addition, studies (e.g., David et al., 2011; 2013) evaluated the accuracy of stream discharge estimates resulting from different LSMs (e.g., VIC, Noah LSM) when combined with RAPID. In this study, however, we restrict our focus to studying the contribution of the routing component to the quality of the discharge estimates. We do so by using the same runoff estimates from the VIC LSM to derive two routing components in a controlled environment that ensures that the differences in the stream flow estimates result from the routing components alone. The two routing components are RAPID, and the routing component developed by the Iowa Flood Center (IFC) and implemented in the full IFC in-house hydrologic model called the Hillslope-Link Model (HLM) (Small et al., 2011; Ayalew et al., 2014; Krajewski et al., 2016).

In this chapter, we do not intend to compare the routing components as software packages, which would require including other aspects (e.g., computational time and ease of implementation in different study areas etc.), however, previous studies investigated these issues (e.g., David et al., 2016 for RAPID; and Small et al., 2011 for the HLM). The routing methodologies incorporated in the routing components are, the simplified linear Muskingum routing method in the case of RAPID (Cunge, 1969; David et al., 2011), and the non-linear routing methodology based on the power law relationship between the stream discharge and mean flow velocity, and the network hydraulic geometry self-similarity assumption described in Mantilla et al., 2005 and Mantilla 2007.

We tested both routing components' performances over the Cedar River basin, an average sized basin located in eastern Iowa. The basin covers approximately 16,862 km2, and is monitored by about 11 United States Geological Survey (USGS) gauges. To

conduct a fair comparison, we used the same stream flowline and sub-catchment geometries provided by the National Hydrography Dataset Plus Version 2 (NHDPlus V2). The common runoff input obtained from the VIC model has a spatial resolution of 1/8˚x1/8˚ and a temporal resolution of 1 hour and available at (http://disc.sci.gsfc.nasa.gov/hydrology/data-holdings). We followed the same runoff-routing coupling process explained in Tavakoly et al. (2016). Our results show an improvement in the streamflow estimates, especially in the aspects of peak timing and magnitude, due to the implementation of the new routing methodology.

2.2. Methods

Available routing techniques generally belong to one of these two major routing groups. The first group of routing methodologies relies on solving the Saint-Venant equation of flow (also called hydraulic routing). These methods vary in their complexity based on the assumptions they use to solve the Saint-Venant equation and usually require accurate river field measurements (e.g., channel cross-sectional geometry, channel bed and bank roughness, etc.) Based on their level of complexity, these routing components can be computationally expensive and difficult to implement on large river networks. Due to its simplicity, the most commonly used approximation for the Saint-Venant equation is the kinematic wave approach. The kinematic wave approach presented in (e.g., Chow et al., 1988, Vergara et al., 2016; Wu et al., 2014) assumes that the only forces needed to balance the gravity forces acting on the flow are the friction forces (i.e. steady uniform flow). Solving the kinematic wave approximation can take different forms. One approach is to solve the kinematic wave with empirical approaches based on idealized representation of the channel properties (e.g., Wu et al., 2014). Another

approach is to statistically extend the relationship between the rating curve parameters at locations where discharge vs. cross-sectional area information are available to ungauged locations by relating the rating curve parameters to landscape and hydro-climatological properties, such as, relief ratio, mean rock volume percent, mean annual rainfall and temperature (e.g., Vergara et al., 2016).

The second major routing group uses storage based hydrologic routing techniques. Hydrologic routing is a simpler approximation of the one-dimensional Saint-Venant equation that involves ignoring the discharge momentum term of the equation, such as the Muskingum method described in Cunge (1969), or the Nonlinear Cascade Reservoir method described in Kim and Georgakakos (2014). On the other hand, some hydrologic routing methodologies use a simplified version of the Momentum equation such as the Muskingum-Cunge described in Cunge (1969) or the non-linear method presented in Mantilla (2007) based on scale invariance and network hydraulic geometry self-similarity.

In this chapter, we are going to compare two hydrologic routing methods, the Muskingum method and the method presented in Mantilla (2007). A unique aspect of the study presented in this chapter is that it compares two hydrologic routing methods using the same input from the widely used LSM VIC, in a controlled environment where the runoff amounts ingested into the sub-catchments of each routing component are identical. This ensures that the discrepancies in the discharge estimates obtained from the routing components are strictly due to the difference in the routing methodologies they utilize. In the next two subsections, we provide a brief description of these routing components we use in this study and the methods they utilize.

2.2.1. The RAPID stream flow routing component

The Muskingum method incorporated in RAPID is a one-dimensional diffusion wave approximation of stream discharge based on the continuity equation that ignores the momentum equation; in other words, it is a storage-based equation. The equation depends on two parameters: a storage constant k that has a dimension of time, and x, a dimensionless weighting parameter based on the relative influence of the inflow and outflow. As described in David et al., (2011a), the finite difference form of the Muskingum method is

$$q_i(t+\Delta t) = c_1\left[q_i^{up}(t+\Delta t) + q_i^e(t+\Delta t)\right] + c_2\left[q_i^{up}(t) + q_i^e(t)\right] + c_3 q_i(t) \quad (2.1)$$

where,

$$c_{1i} = \frac{\frac{\Delta t}{2} - k_i x_i}{k_i(1-x_i)+\frac{\Delta t}{2}}, \quad (2.2)$$

$$c_{2i} = \frac{\frac{\Delta t}{2} - k_i x_i}{k_i(1-x_i)+\frac{\Delta t}{2}}, \quad (2.3)$$

and

$$c_{3i} = \frac{k_i(1-x_i)-\frac{\Delta t}{2}}{k_i(1-x_i)+\frac{\Delta t}{2}}, \quad (2.4)$$

For any reach index i, $C_{1i} + C_{2i} + C_{3i} = 1$. In equation 1, q_i is the outflow of a reach i; q_i^e is the lateral flow within a reach i (surface and subsurface); q_i^{up} is the inflow to the reach from upstream links; and t, Δt are the simulation time and the time step, respectively. The output flows from RAPID are affected by the fact that the Muskingum

method assumes a uniform water surface profile between the upstream and downstream ends of a channel reach. Also, obtaining accurate k and x values is a difficult task that often includes calibration and optimization by using observations and is dependent on the calibration dataset. The RAPID routing component has an option to optimize the parameters k and x based on available USGS gauge observations. In this study, we used the calibrated k and x dataset available at Tavakoly (2017) and described in Tavakoly et al., (2016). RAPID assumes that the lateral flow q^e enters the channel instantly at the upstream junction of the stream. Lastly, the common approach for handling the runoff maps using RAPID (e.g. David et al., 2011, 2013, 2015, 2016; Tavakoly et al, 2016) lumps surface and subsurface runoff estimates from the LSM and routs them into the stream channels simultaneously.

2.2.2. The HLM routing component

The HLM routing is based on the method presented in Mantilla (2007) where the velocity in the channel element can be obtained using a nonlinear power law relationship with the upstream served area and the stream discharge (similar to that by Paik and Kumar (2004)), where the velocity corresponding to a certain discharge q can be described as

$$v(q) = v_0 {q_{link}}^{\lambda_1} A^{\lambda_2} \tag{2.5}$$

Consequently, the flow transport equation can be described as follows:

$$\frac{dq_{link}(t)}{dt} = \frac{v_0 q_{link}^{\lambda_1}(t) A^{\lambda_2}}{(1-\lambda_1) l} [a_h \left(q_{surf}(t) + q_{subsurf}(t)\right) - q_{link}(t) + q_{up}(t)] \tag{2.6}$$

where $q_{link}$and q_{up} are the output discharges from a link and inflow from upstream links at time t; A is the total upstream area draining into the hillslope; a_h and l are the hillslope area and length, respectively; and λ_1, λ_2 and v_0 are global parameters of the water velocity component in the model and are set to $0.2, -0.1$ and 0.3, respectively. More on how to estimate these parameters and how they vary across geographic locations in the United States Midwest can be found in Mantilla (2007) and Ghimire et al., (2017). In this chapter, the surface and subsurface runoff estimates, q_{surf} and $q_{subsurf}$ consequently, are lumped prior to routing in a similar fashion to RAPID. We matched our model setup to that used to prepare the RAPID parameters' values to ensure fair comparison. However, we conduct our comparisons at a mid-sized basin included in this domain because it is well monitored with USGS stream gauges across a wide range of scales. In addition, the effect of artificial storage in this basin is negligible. It is also important to note that the estimation of the HLM model parameters is independent of the rainfall-runoff model's input, Mantilla (2007).

2.2.3. Evaluation procedure

We conduct our evaluation at the Cedar River basin, which is located in eastern Iowa. The Cedar River Basin has a drainage area of 16,814 km2. Eleven USGS gauges monitor the area and cover a range of scales (from 776 km2), and are not affected by river regulation. The USGS gauge locations are shown in Figure 2.1 and the areas they serve are listed in Table 2.1. (ordered according to the labels in Figure 2.1). Our study period is the warm season (May through September) of 2014. We allowed a two-month spin-up period (runs started on March 1st) and proper initialization for both models.

Both simulated and observed discharges used in this study have an hourly temporal resolution.

We calculate three statistical skill scores at each gauge location, the percent Root Mean Square Error RMSE, correlation coefficient, and Nash-Sutcliffe efficiency (NSE). The reader is encouraged to review Moriasi et al., (2007) and Tavakoly et al., (2016) to learn more about model's performance classifications based on NSE, RMSE, and Correlation Coefficient performance categories. The equations used to calculate these scores are as follows:

$$\%\ RMSE = \sqrt{\frac{\sum_1^n (q_{sim}^t - q_{obs}^t)^2}{n}} \cdot \frac{1}{max(q_{obs})} . 100 \tag{2.7}$$

$$Corr = \frac{E[(q_{obs} - \mu_{obs})(q_{sim} - \mu_{sim})]}{\sigma_{obs}\, \sigma_{sim}} \tag{2.8}$$

$$NSE = 1 - \frac{\sum_{t=1}^n (q_{obs}^t - q_{sim}^t)^2}{\sum_1^n (q_{obs}^t - \bar{q}_{obs})^2} \tag{2.9}$$

In order to better assess the performance of both models in detecting peak times and magnitudes we calculated the hit rate (HR), false alarm ratio (FAR), and the critical success index (CSI) at each station using the two- year, five-year and ten-year return period peak discharges provided by Eash et al., (2013). Unfortunately, these records are not available at all stations due to lack of sufficient stream discharge records. Only one station out of our eleven stations (station number 7) has an insufficient peak discharge record. The formulas we used to calculate HR, FAR, and CSI are as follows:

$$HR = \frac{a}{a+c} \tag{2.10}$$

$$FAR = \frac{b}{a+b} \tag{2.11}$$

$$CSI = \frac{a}{a+b+c} \tag{2.12}$$

where a is the instances where the discharge exceeded a certain threshold in both model estimates and observed flow; b is when the model estimates exceed the threshold but the observed flow is lower than this threshold; and c is when the observed flow exceeds the threshold but the model estimates do not. All these parameters are calculated based on hourly discharges. In short, HR represents the instances when the model predicted that the flow would exceed a certain threshold correctly (ranges from zero (Poor) to one (Good)). FAR represents the instances when the model overestimated the discharge and exceeded a certain threshold (ranges from zero (good) to one (Poor)). CSI, also known as threat score, represents how well the exceeding events of the model correspond to those of the observations by accounting for the incorrect modeled exceedances in the denominator (ranges from zero (Poor) to one (Good)).

We have also calculated the difference in time to peak ΔT^p and percent difference in peak discharge ΔQ^p as defined in Vergara et al., (2016), where, as shown in (13) the difference in peak time will be negative in case the simulated peak ($T^p_{simulated}$) occurs ahead of the observed peak ($T^p_{observed}$), but positive in case the simulated peak occurs after the observed peak. Similarly, from (14) the percent difference in peak magnitude will have a positive sign in case the simulated peak ($Q^p_{simulated}$) is greater in value than the observed peak and a negative sign in case the simulated peak is smaller in value than the observed peak ($Q^p_{observed}$).

$$\Delta T^p = T^p_{simulated} - T^p_{observed} \quad (2.13)$$

$$\Delta Q^p = (Q^p_{simulated} - Q^p_{observed})/Q^p_{observed} * 100 \quad (2.14)$$

2.3. Results

We begin our comparison between the routing components by visually inspecting the hydrographs presented in Figure 2.2. The red hydrographs correspond to the VIC-HLM combination, the blue hydrographs correspond to the VIC-RAPID combination, and finally the black hydrographs correspond to the USGS observations. When available, the return period peak discharges are represented by horizontal gray lines on top of the hydrographs (two-year peak discharge is always represented by the lowest gray line above it is the five-year peak discharge and the highest gray line is the ten-year peak discharge). Two main peaks characterize the observed flow of the year 2014; the first took place around mid-June, and the second around early July. High flood levels were observed in the Cedar River basin during the year 2014 warm season (June, July and August); with peak flows often exceeding the 5-year or 10-year return periods discharge across the whole basin. A major characteristic of the simulated discharges is that they consistently underestimate the peak discharges in many of the stations, especially during the mid-June event. Since both routing components receive the same runoff estimates, and the runoff is eventually routed through the channels; the LSM VIC seems to be a major contributor to this underestimation in the overall water volume. An important feature of the VIC-RAPID hydrographs is that they exhibit a flashy behavior, which is dampened as the basin scale increases. This flashy behavior is mainly due to the linear

routing method incorporated in RAPID (simplified Muskingum) in comparison to the non-linear routing method incorporated in the HLM routing component.

In Figure 2.3 we present the statistical skill scores of the routing components spatially over the stream network. The right column shows the skill scores of the HLM routing, while the left column shows the RAPID skill scores. The top row shows the comparison of the correlation coefficients between the simulated and observed hydrographs, while the middle and bottom rows show the % RMSE and NSE comparisons, respectively. These results show that the HLM routing components resulted in an improved simulated discharge estimates across scales (higher correlation and NSE, and lower % RMSE). We included a summary of these skill scores in Table 2.2.. By looking at Table 2.2., one can infer that the non-linear routing exhibits a more stable performance across different locations (lower standard deviation, SD) with a better overall performance compared to linear routing.

So far, we have evaluated the time series of discharges as whole, without details about the performance of the routing components during extreme events. Hence, we evaluated indices such as HR, FAR and CSI at all station locations as well as the difference in time-to-peak and percent difference in peak magnitude. We start by comparing the HR, FAR and CSI scores at each location for the 2-year, 5-year and 10-year return period discharge estimates. Due to the overall underestimation in the simulated peak discharges caused by the LSM, the simulated discharges rarely reached the 10-year return period, which is not the case in the observed flow. Thus, the scores estimates for the 10-year return period are not available. The left column of Figure 2.4 shows the estimates for the 2-year return period and the right column shows the scores of

the 5-year return period. The red bars represent the HLM routing component while the blue bars represent the RAPID component represent. The HLM routing component outperformed the RAPID component by having higher 2-year HR and CSI values, and lower FAR values. Despite the overall underestimation in discharge estimates, the effect of RAPID flashy behavior is evident in the 5-year return period's FAR estimates. However, this allowed RAPID to capture a small percentage of the times when the observed flow was in fact above the 5-year return period threshold (e.g., the estimates at station number 8).

Both routing components experienced an improvement in peak timings and magnitudes across scales during the early July peak. However, by visualizing the hydrograph one can observe that the improvement is more significant in the case of RAPID. This might be due to the spatial distribution of the runoff estimates, where the locations that received a significant amount of runoff during that time are mostly located on the western side of the basin (based on observed discharges at these stations). The land cover is generally the same across the basin, and the terrain is flat in this region (not shown), hence, they do not contribute to the performance enhancement in case of RAPID. In order to have a better idea about the performance in peak time and magnitude we calculated $\Delta\, T^p$ and $\Delta\, Q^p$ for each peak at all station locations. Table 2.3. contains the outcome of this analysis where the left panel of the table compares the performance of the routing components during the mid-June peak (Event 1), and the right panel for the performance of the routing components during the early July event (Event 2). The summary of statistics at the bottom four rows of the table shows that for the first peak, although both routing components seem to have placed the peaks ahead of time, RAPID

estimates experienced a much earlier peak time when compared to the observed peak. On average, both routing components underestimated the peak magnitudes, however, with a higher variability (standard deviation) in the case of RAPID due to its flashy behavior. Both routing components performed much better in estimating the second peak timing with similar standard deviation of peak time errors. Nevertheless, the HLM underestimated the peak magnitudes and RAPID overestimated them, with greater variance in case of RAPID.

2.4. Conclusions and recommendations

The main goal of this study is to highlight the contribution of the stream discharge routing methodologies to the errors in simulated stream discharge estimates. We observed that the routing methodology have a significant impact on our hydrologic model's performance. This difference in performance generally decreases as the basin scale increases. As expected, the linear Muskingum method resulted in a flashy behavior in the simulated discharges, especially in smaller basin scales. The non-linear routing method incorporated in the HLM routing component resulted in improved discharge estimates characterized by better peak times and magnitudes. It is evident that a major contributor to the errors in the estimated discharges is the LSM runoff estimates, which in turn will have an effect on the routing components performance, since all the velocities in the streams are determined by the amount of runoff available in their served area. However, determining the amount of error caused by the LSM is not our focus in this study. Future recommendations for this study include, first, expanding the study domain to include a wider range of scales and different hydro-climatic regions. This however may require implementing reservoir operation in the routing components in order to have

a fair comparison with the observed discharges. Second, evaluating the HLM routing component with runoff estimates from a variety of LSMs. In addition, one can further investigate the effect of the spatial distribution and location of rainfall-runoff events on the peak times of RAPID. The latter recommendation arises from our observation that RAPID was able to detect a relatively close peak time in some occasions (e.g., the early July event, although the peak magnitude is inaccurate in most locations). Finally, a more comprehensive technique of estimating the peak time difference based on Continuous Wavelet Transform (CWT), which also accounts for peak magnitude and range, is described in ElSaadani and Krajewski (2017). We hope that our study will be followed by many similar investigations by others in different parts of the country and collectively will contribute to improve national flood estimation and forecasting platforms.

Table 2.1. A list of USGS gauges and their served area.

Rank	USGS ID	Area (km^2)
1	5459500	1342
2	5457700	2792
3	5458000	766
4	5463000	911
5	5458900	2213
6	5458500	4338
7	5458300	4044
8	5463500	780
9	5464000	13333
10	5464220	773
11	5464500	16862

Table 2.2. Summary of statistical skill scores for VIC-RAPID, and VIC-IFC.

Skill Score	RAPID	VIC-IFC
%RMSE		
Max	20.09	13.10
Mean	13.81	8.90
Min	9.20	6.42
SD	3.02	2.20
NSE		
Max	0.59	0.78
Mean	0.18	0.67
Min	-0.52	0.54
SD	0.29	0.09
Correlation Coeff.		
Max	0.81	0.90
Mean	0.53	0.84
Min	0.31	0.75
SD	0.15	0.05

Table 2.3. Summary of peak time differences (Δ T) in hours and percent difference in peak discharges (Δ Q%) for the mid-June (Event 1) and early-July (Event 2) peaks.

Station	Event1				Event2			
	VIC-HLM		VIC-RAPID		VIC-HLM		VIC-RAPID	
	Δ T (hour)	Δ Q%	Δ T (hour)	Δ Q%	Δ T (hour)	Δ Q%	Δ T (hour)	Δ Q%
1	1	-17.5	-10	-47.9	12	-43.7	-6	54.6
2	-6	-42.4	-46	-60.0	20	-2.2	34	146.4
3	4	-27.2	-56	19.9	2	-5.1	10	259.1
4	-73	-33.4	-100	48.4	-5	-35.4	-15	1.1
5	-22	-6.8	-108	-15.0	-3	8.3	2	100.4
6	0	-38.8	-31	-52.7	-16	-34.2	4	35.9
7	-2	-36.5	-38	-52.8	-1	-19.1	4	89.3
8	-72	-44.6	-89	12.5	0	-59.9	-12	-44.6
9	-17	-36.6	-54	-48.0	-6	-34.8	-18	-15.5
10	-6	-44.0	-21	50.4	-5	-40.0	-14	7.1
11	-10	-32.8	-41	-40.6	8	-36.8	3	-28.2
Max	4.0	-6.8	-10.0	50.4	20.0	8.3	34.0	259.1
Mean	-18.5	-32.8	-54.0	-16.9	0.5	-27.5	-0.7	55.0
Min	-73.0	-44.6	-108.0	-60.0	-16.0	-59.9	-18.0	-44.6
SD	27.8	11.7	32.1	42.3	9.8	20.5	14.9	89.8

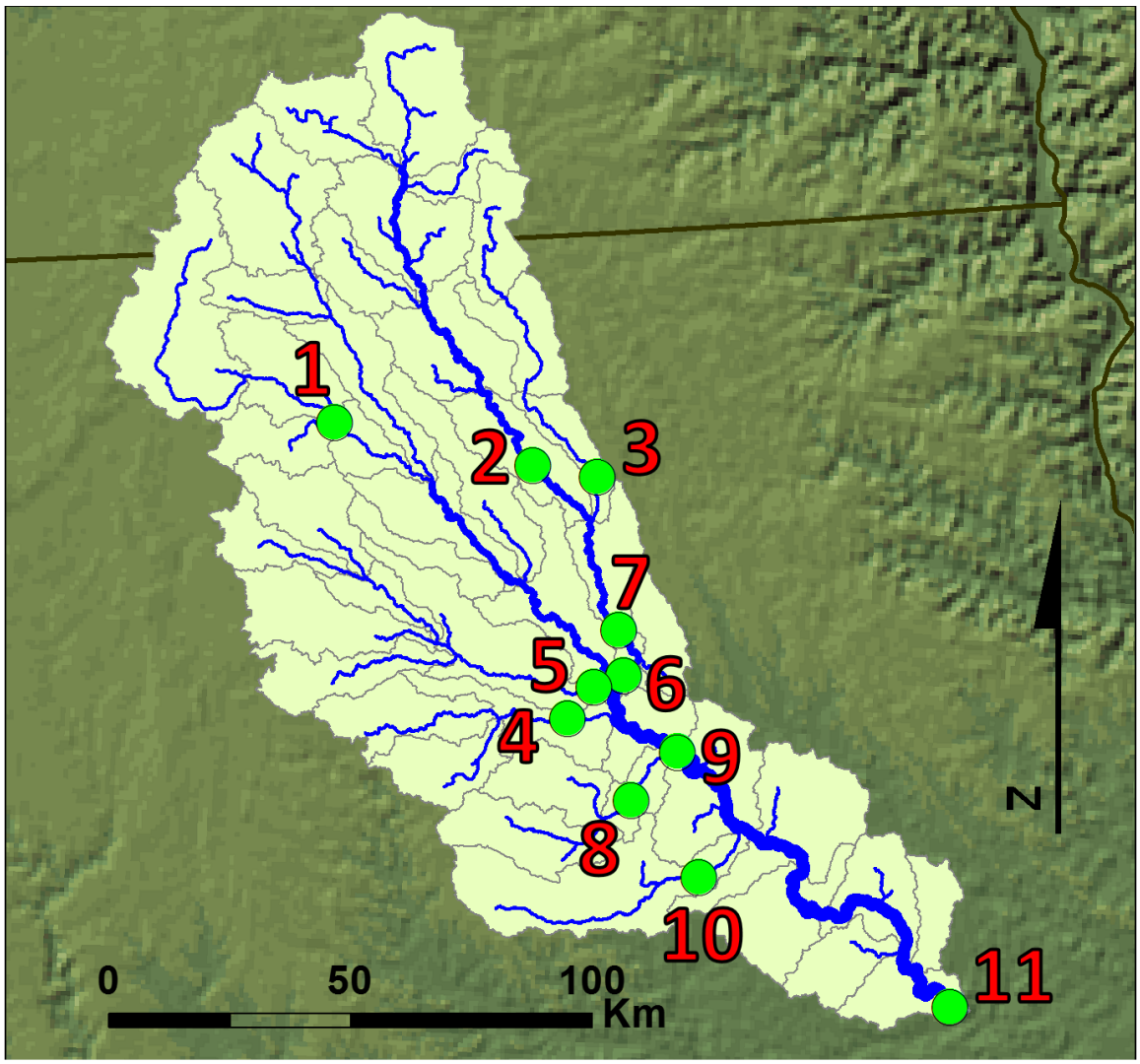

Figure 2.1. The Cedar River basin located in eastern Iowa. The USGS stream gauges located within the basin are represented by green dots. The labels of the USGS gauges correspond to the rankings in Table 2.1..

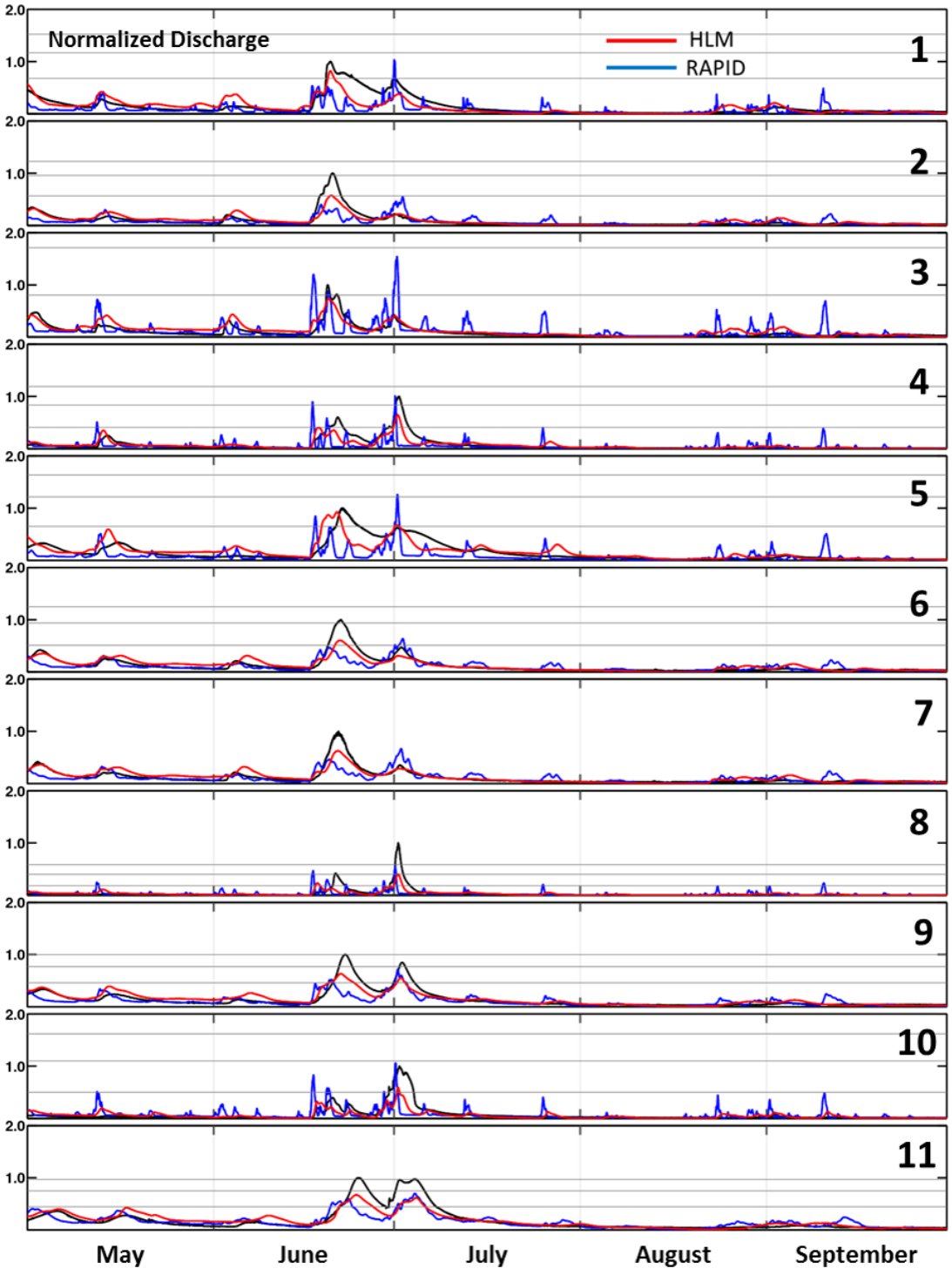

Figure 2.2. Hydrographs at the USGS station locations: the red lines represents VIC-HLM, the blue line represents VIC-RAPID, and the black lines are the observations. The numbers in the panel labels correspond to the station ranking in Table 2.1..

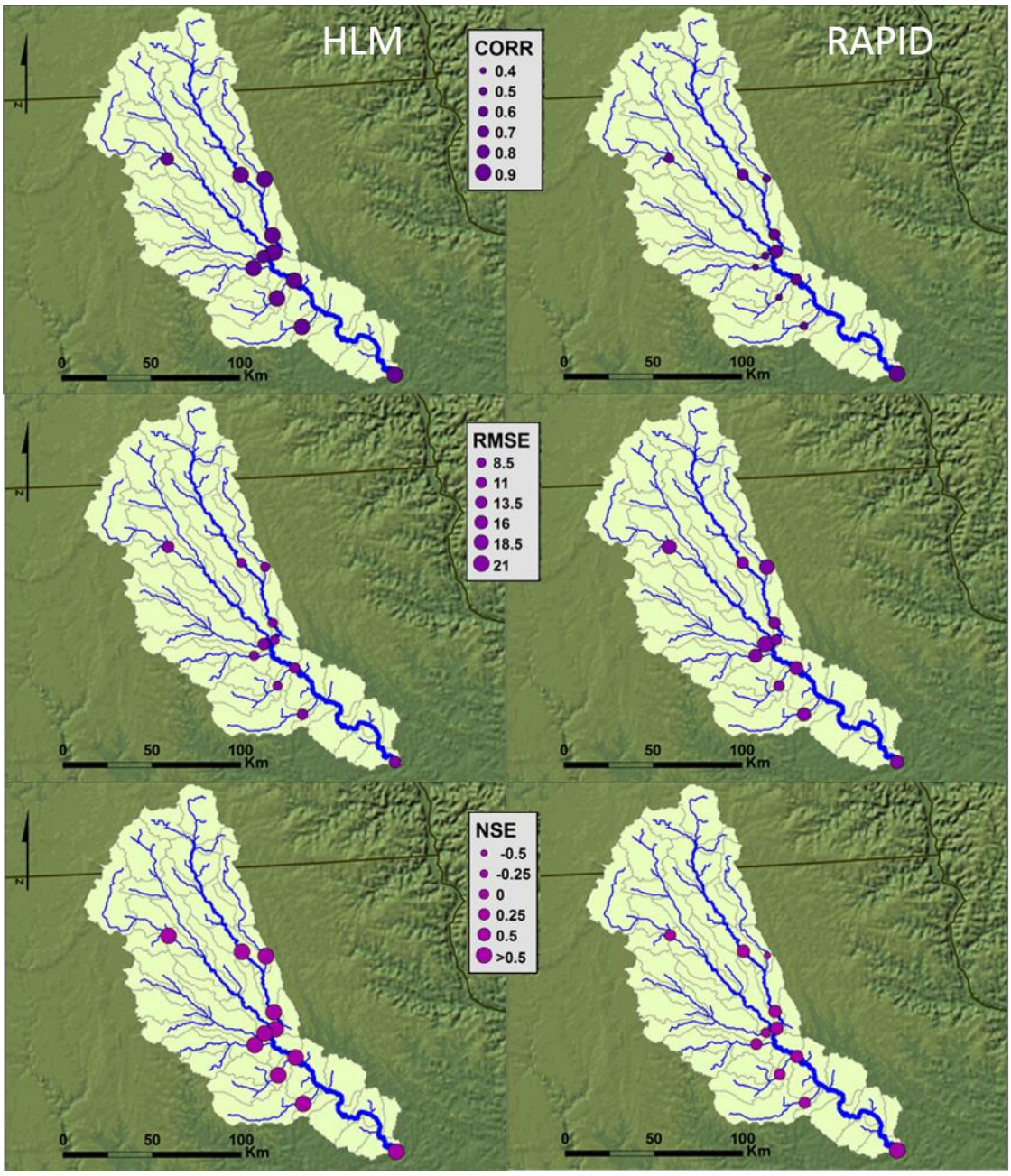

Figure 2.3. A comparison of the statistical skill scores: top row is % RMSE, the middle row is the correlation, and the bottom row is the NSE, with VIC-RAPID on left column and VIC-HLM on the right column.

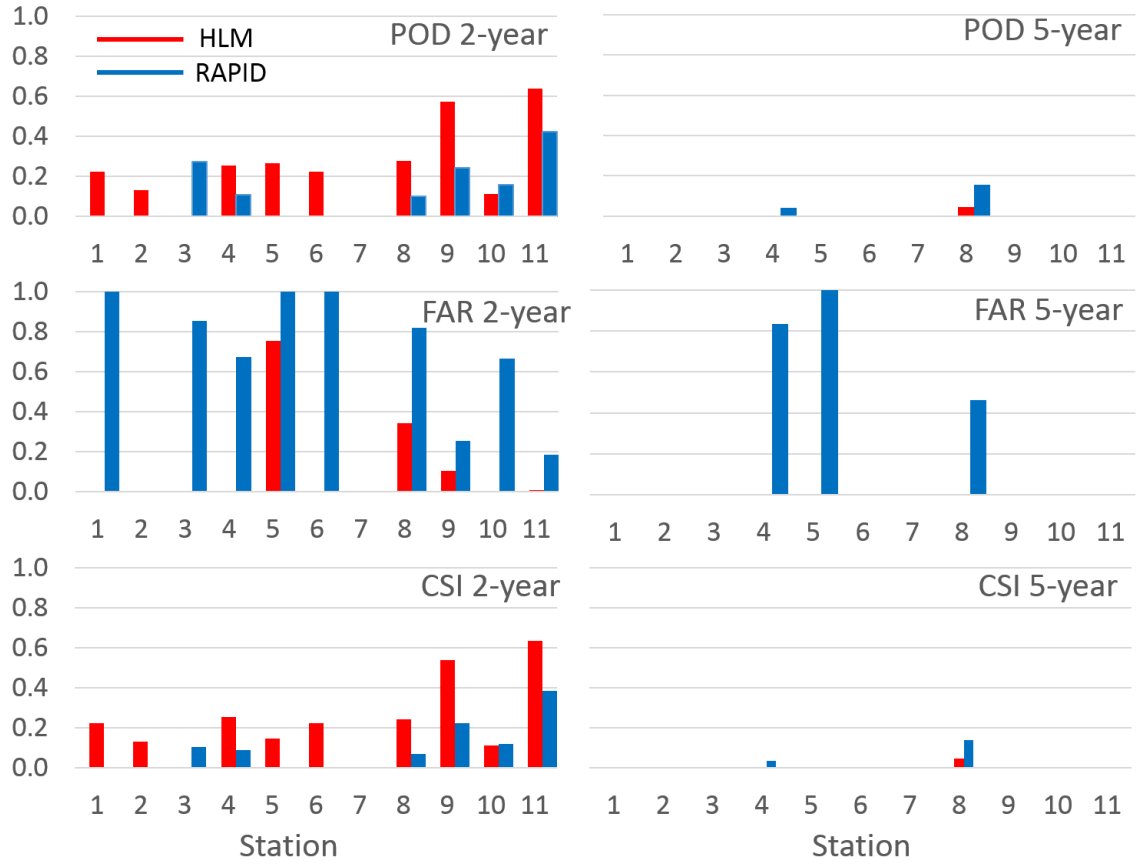

Figure 2.4. Comparison of HR left column, FAR middle column and CSI right column (y axes) at different stations (x axes) between NIFE-Hydro and IFC. The left and right columns represent the 2-year and 5-year return period flood peak thresholds respectively. Blue bars represent the VIC-RAPID scores and red bars represent VIC-HLM scores. Station numbers (x-axis) follow Table 2.1. ranks.

CHAPTER 3: A Spatio-Temporal Evaluation of IMERG Final Run and its Performance in Stream Flow Modeling over the State of Iowa

3.1. Introduction

In February 2014, NASA launched the Global Precipitation Measurement (GPM) mission's core satellite (Huffman et al., 2014). This core satellite is the successor of the retired Tropical Rainfall Measurement Mission (TRMM) core satellite which along a constellation of geostationary and Low Earth Orbiting (LEO) satellites has provided the TRMM Multi-satellite Precipitation Analysis (TMPA, references) rainfall estimates. The TMPA products were widely used and analyzed in many hydro-meteorological applications over the past three decades (Habib et al., 2012; Haile et al., 2013; Quintero et al., 2016). The GPM mission provides a verity of improved precipitation products that are obtained from different sensors (e.g., Microwave (MW) or Infrared (IR) sensors) and have different latencies (e.g., near real time or research version). In this chapter, we are interested in the recommended research version called IMERG Final Run V04A, which was released in April 2017. The product comprises of a combination of precipitation estimates that are obtained from satellite that are bias adjusted using rain gauges. The improved features of IMERG include larger spatial coverage, higher spatial and temporal resolution, and better snow detection in comparison to the TMPA products (Huffman et al., 2014). In many parts of the world, satellite rainfall products are the only available source of spatial rainfall estimates. Therefore, information regarding the accuracy of IMERG are useful for researchers who need to use it in hydro-meteorlogical applications such as land surface modeling, and flood prediction.

In this chapter, we use the Multi-Radar/ Multi-sensor (MRMS) Quantitative Precipitation Estimate (QPE) as our benchmark product to evaluate IMERG. Although we acknowledge that our benchmark product is not a perfect measure of real rainfall, it is safe to assume that it is a closer measure of it compared to satellite rainfall (Vergara et al., 2014). In addition, radar rainfall estimates can provide us with a continental scale continuous coverage and high temporal resolution that rain gauge networks cannot provide. In this study we use the latest operational version of MRMS, which is available starting from September 2015 (Zhang et al., 2015).

Our study domain is the entire state of Iowa, which is located in the US Midwest with a total area of 145,740 Km2. The state spans between the longitudes of around -96.6˚ W to -90.0˚ W and latitudes 40.0˚ N to 43.6˚ N (Figure 3.1.). In this study, we conduct a direct comparison between the time series of rainfall estimates at each location (rainfall pixel) within our domain, as well as characterizing the dependency of the rainfall estimates obtained from each product in space and time. We do so by estimating the spatiotemporal semi-variogram of the rainfall values. Naturally, rainfall estimates are correlated in space and time with magnitudes that depend on the coverage and duration of the storm which they are associated with. However, during the production of IMERG, the more reliable, but not always available, rainfall estimates from MW satellites are interpolated in space and time to fill the gaps that occur during the times when there is no satellite coverage. This is expected to alter the structure of the dependency between IMERG's rainfall estimates and affect the discrepancies between IMERG and MRMS. Vergara et al., (2014) showed that the differences in the spatial dependency between

rainfall estimates can be related to discrepancies in the discharges produced by hydrologic models.

In addition to our rainfall to rainfall comparison, we separately investigate the performance of both rainfall products in stream flow estimation. For this purpose, we used both rainfall products as the inputs of two hydrologic models. The hydrologic models are the Iowa Flood Center's (IFC) in-house hydrologic model called the Hillslope-Link Model (HLM) (Krajewski et al., 2016), and the hydrologic model presented in (ElSaadani and Krajewski, 2017) which comprises of the Noah Multi-Parameter (MP) Land Surface Model (LSM) and the routing component of the HLM. We conducted state wide hydrologic model runs and evaluated the model outputs at 130 United States Geological Survey (USGS) stream gauge locations that are within our domain and are not directly affected by artificial storage (e.g., damns or reservoirs). Both our rainfall and streamflow comparisons were conducted over the 2016 warm season (June, July, August and September).

3.2. Datasets and models

3.2.1. Rainfall products

The product we are going to evaluate in this study is the Integrated Multi-satellitE Retrievals for GPM (IMERG) is a satellite based, high resolution, global precipitation product. Precipitation estimates for IMERG are obtained through interpolating and merging precipitation estimates from satellite microwave estimates, infrared (IR) satellite estimates, and rain gauge precipitation estimates. The spatial and temporal resolutions of IMERG are (0.1˚) and (30 minutes) respectively, with global coverage from 60˚ N to 60˚ S and 180˚ E to 180˚ W. The most reliable data for IMERG are those obtained from the

GPM core observatory and the Passive Microwave (PMW) LEO satellites. The gaps between the observations from these satellites are filled by morphing, using geo-IR-based feature motion and microwave calibrated IR estimates; where the GPM core observatory serves as a calibration and evaluation tool for all other satellite products (IR and PMW) used in IMERG. The IMERG Final Run product undergoes bias correction using rain gauge rainfall estimates provided by the Global Precipitation Climatology Centre (GPCC) (Liu, 2016).

What separates the GPM mission from its predecessor TRMM is that, first, the orbit of the GPM core satellite is between 65˚ N and 65˚ S, which allows the satellite to observe a wider range of climate zones. Nevertheless, the high inclination of the orbit reduces the DPR's sampling compared to the less inclined TRMM orbit. Second, the Dual-Frequency Precipitation Radar (DPR) on board of the satellite has two active microwave channels. The Ku-band (13.6GHz) which is similar to the one available in TRMM core satellite and is good for detecting large size rain drops, and the higher frequency Ka-band (35.5 GHz) channel, which allows the satellite to detect snow and smaller rain drop sizes. In addition to the GPM's improved instruments, a combination of satellite rainfall algorithms are merged to produce IMERG. This includes the algorithms used to produce TMPA together with the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN, Ashouri et al., 2015) and the Climate Prediction Center (CPC) MORPHing technique (CMORPH, Woody et al., 2014).

Our benchmark product (ground reference) in this study is the Multi-Radar/Multi-Sensor (MRMS) Quantitative Precipitation Estimate (QPE), which was first built by the National Severe Storm Laboratory (NSSL), and now is being operationally produced by

the National Center for Environmental Prediction (NCEP). The ground based radar product is bias adjusted using hourly quality controlled rain gauge data provided by the Hydro-meteorological Automated Data System (HADS, Zhang et al., 2015). The product integrates observations from 180 operational radars (Zhang et al., 2015) and has a high spatial resolution of 0.01˚ and hourly temporal resolution. The product covers the conterminous United States (CONUS) domain and southern Canada (20˚ N – 50˚ N).

3.2.2. Atmospheric frocings

The data input requirements for the hydrologic models can be different based on the user's needs. For instance, the IFC model can use remotely-sensed or climatology-based evapotranspiration (ET) estimates. Also, users can implement their own ET module by adding the governing ET equations in ODE format to the model script. We used the climatology option based on 12 years of North American Land Data Assimilation System (Mitchell et al., 2004) for the IFC model. We chose the climatology-based ET due to its simplicity. Although this approach successfully represents the seasonal variability of ET, it does not capture the daily fluctuations nor the inter-annual variability of ET. Researchers at the IFC are currently working on producing a more sound ET estimation methodology for the IFC model; this will be the subject of future studies (Krajewski et al., 2016).

A major component in our second hydrologic model used in this study (ElSaadani and Krajewski, 2017) is the Noah-MP LSM. Noah-MP requires atmospheric forcing data that can be obtained either from the Weather Research and Forecasting (WRF) model, or retrospective atmospheric forcings such as those provided by the NASA Phase 2 North American Land Data Assimilation (NLDAS-2) dataset (Xia et al. 2012a, b). In this study,

we chose the latter option. All NLDAS-2 variables are available on the NLDAS-2 1/8˚ square grid with hourly temporal resolution. Except for precipitation, surface forcing fields (e.g. surface radiation fluxes, temperature, wind speed, humidity, and pressure) from the NLDAS-2 dataset are used to force the LSM.

3.3. Hydrologic models

3.3.1. The Hillslope Link Model (HLM)

HLM is a parsimonious model that decomposes the landscape into fine resolution hillslopes with an average area of 0.4 km2. For each hillslope, the mass conservation equations are translated into a set of ordinary differential equations and solved via a parallelized numerical solver (Small et al., 2013). This platform is flexible and allows the user to decide which processes should be included in the rainfall runoff, and runoff discharge generation and their level of complexity (complex or simple approximate equation). It also allows for the use of remotely-sensed data, retrospective data, or user-defined climatology estimates as a substitute for some of these processes e.g. ET estimates. The configuration we used for this study has three main storages in each control volume: hillslope surface S_{surf}, top soil S_{tops}, and deep soil S_{deeps}. The governing equations for the IFC model are described in Quintero et al., 2016 and Mantilla 2007.

3.3.2. Noah-MP LSM with HLM routing

The Noah-MP LSM is a one-dimensional column-based community land surface model that decomposes the landscape into a user-defined regular square grid (3 km in this study). The LSM accounts for a large number of complex natural processes that contribute to the energy and mass balance between the soil and the atmosphere, such as

vegetation growth, soil moisture, snow cover, incoming short- and long-wave radiation, and the water exchange between different soil layers. Niu et al., (2011) describes the Noah-MP model in detail. We defined the Noah-MP LSM's grid characteristics using the USGS National Elevation Dataset, the USGS 24-category land-use data, and the U.S. Department of Agriculture Natural Resources Conservation Service State Soil Geographic (STATSGO) data. The runoff estimates from the LSM are then routed using the HLM routing component described in the previous section in order to obtain stream discharges.

3.4. Results

3.4.1. Rainfall to rainfall comparisons

We begin our comparison by visually inspecting the mean hourly rainfall estimates of each product. Figure 3.2. shows the mean hourly rainfall estimates of MRMS during the months of June (top left), July (top right), August (bottom left), and September (bottom right) of 2016. Figure 3.3. shows the same estimates but for IMERG. One can see that IMERG underestimated the rainfall amounts, with a smoother and more spread rainfall domains. This is especially true for strong, clustered accumulations. The underestimation of strong events is reflected in the RMSE estimates in Figure 3.4. especially during the August and September of 2016.

In Table 3.1., we show the HR, FAR, CSI, and Bias of IMERG for various rainfall thresholds. The bias values show that IMERG has overestimated the number of hours with small amounts of rainfall (0-1 mm/hr). In addition, IMERG have significantly underestimated the number of rain instances with high rainfall values. The HR estimates point in the same direction as the bias where IMERG failed to detect instances with high

rainfall estimates while performed much better for low rainfall estimates. The FAR shows that even the instances during which IMERG detected large rainfall amounts were false alarms. In addition, IMERG has detected a fair amount of false low intensity rainfall. The CSI shows that overall, IMERG performs much better during light rainfall events than it does during heavy rainfall events.

We further investigate the spatial and temporal structure of the rainfall estimates from both products by calculating the empirical semi-variogram of rainfall. The two semi-variogram characteristics that we are interested in the most are their range and nugget. This is because we would like to relate the mean hourly rainfall distributions we observed in Figures 3.2. and 3.3., as well as the bias in heavy rainfall detection shown in Table 3.1., with the semi-variograms. For example, one should expect strong wide spread systems (e.g., tropical storms) to have larger range values while small concentrated rainfall systems (e.g., convective storms) to have a smaller range. In addition, wide spread and more uniform systems should have smaller nugget effect and less overall variance. However, if there is a group of small concentrated systems, for example, the nugget effect should be strong due to the abrupt changes in rainfall in space and in time and the overall variance should be higher as well. Since we are analyzing monthly rainfall, one should expect a combination of different storm types to occur, resulting in a mixed effect on the semi-variogram characteristics.

In Figure 3.5., we show the spatio-temporal semi-variogram of hourly rainfall for MRMS during June (top left), July (top right), August (bottom left), Spetember (bottom right). The x-axis represent the distance between the rainfall estimates in Km, while the y-axis represents the semi-variance in mm2. Each line with a different color correspond

to a certain time lag (between 1h to 10h). As expected, the semi-variograms directly relate to the observed mean hourly rainfall maps. During the June, most of the rainfall is concentrated on the top right corner of the state of Iowa and the overall maximum mean rainfall is less, compared to August and September. This is reflected in the semi-variogram as smaller range of around 100 Km (due to concentrated rainfall) and less over all variance compared to other months. In addition, the nugget to sill ratio is about 30% which is considerably higher than other months. For the month of July, the maximum mean hourly rainfall is higher than that of June but still relatively lower than August and September. In addition, July rainfall covers most of the state. This resulted in a sill that is relatively higher than that of June and a smaller nugget (around 20% of the sill value). One can also see that the more spread rainfall resulted in a range that is twice as large as the one from June. Similar behaviors are seen in the August and September semi-variograms, except with a stronger nugget due to the availability of some scattered strong systems.

Time wise, one can see that the semi-variance collapses in time quickly. During the month of June, where no strong events occurred, the semi-variance collapsed immediately at the first time lag (1 hour). However, for the other months the semi-variance of the first time lag is considerably larger and the semi-variance after the second time lag is negligible.

The same inter-monthly difference in characteristics of the semi-variograms of MRMS rainfall are observed in IMERG semi-variograms (Figure 3.6.). However, the semi-variogram ranges are considerably larger to their MRMS counterparts (all above 200 Km) with a much weaker nugget effect. In addition, the dependence in time is much

higher compared to MRMS with a span up to four hours (e.g., August). This higher auto-correlation in space and time can be due to the interpolation techniques used to generate IMERG and will in turn affect the structure of the discrepancies between IMERG and the benchmark rainfall MRMS. Ideally, the rainfall errors should be uncorrelated (random errors). However, the semi-variograms of the errors (Figure 3.7.) show otherwise. The rainfall errors of IMERG are correlated up to ranges of around 100 Km but the dependence in time collapses almost immediately at the first time lag.

3.4.2. Streamflow comparisons

We performed state wide hydrologic model runs for the 2016 wet season using the HLM and Noah-MP with HLM routing. We allowed a proper initialization period for both models and calculated all our statistics for the period between June through September, 2016. We compared the models' outputs to the observed stream flows from 130 USGS stream gauges across the state of Iowa. Gauge selection was based on record availability and minimal storage effect (e.g., dams or reservoirs).

In Figure 3.8. we show the correlation coefficient between the modeled and observed discarges. In general, the HLM model outperformed the Noah-MP with HLM routing. In addition, the MRMS product performed better than IMERG as a model input. The model selection seems to have a stronger impact on the quality of model inputs in comparison to the model rainfall input in our case. However, the spatial distribution of the performance of both models is similar. Both models performed much better in the eastern side of the state than they did in the western part of the state. This spatial change in performance can be attributed to the steeper topology of the western side of the state as well as the difference in the rainfall input quality. Similar patterns in the spatial

distribution of the performance is seen in the NSE (Figure 3.9.), and the percent normalized RMSE (Figure 3.10.).

3.5. Summary and conclusions

In this chapter, we investigated the performance of the satellite based rainfall product IMERG. The product showed an adequate performance in terms of representing the spatial and temporal occurrence of rainfall events. However, the magnitude and dependency between the rainfall estimates is different from our benchmark product MRMS. IMERG provided a reliable characterization of the monthly variability in hourly rainfall rates and rainfall distribution during our study period. However, there is an overall underestimation of the rainfall magnitudes and the product was not able to capture heavy concentrated events. In addition, IMERG showed a fair amount of false light rainfall detection due to the overestimation of the geographic extent over which rainfall occurred. This smoothing in space is clear in the empirical semi-variogram of rainfall estimates in the form of large range (almost twice as large as the ranges of MRMS) and small nugget effect. In addition, the temporal semi-variograms showed that IMERG rainfall estimates have relatively higher dependency in time (up to 4 hours) compared to the benchmark product where the dependency in time usually collapses after the first time lag.

IMERG showed an overall adequate performance in hydrologic modeling. The performance metrics obtained when IMERG is used as a rainfall input to hydrologic models were comparable to those obtained from MRMS. Generally, both hydrologic models produced better discharge estimates in the eastern side of our study area with the HLM model outperforming the Noah-MP with HLM routing combination.

Table 3.1. Contingency table for IMERG for each month.

Threshold (mm/hr)		0.1	1	5	10	20	30	40	50
HR	**Jun**	0.7	0.6	0.4	0.3	0.1	0.0	0.0	0.0
	Jul	0.6	0.6	0.5	0.4	0.1	0.0	0.0	0.0
	Aug	0.6	0.5	0.5	0.4	0.2	0.0	0.0	0.0
	Sep	0.7	0.6	0.5	0.3	0.1	0.0	0.0	0.0
FAR	**Jun**	0.5	0.5	0.7	0.8	0.9	1.0	1.0	1.0
	Jul	0.4	0.3	0.5	0.7	0.9	1.0	1.0	1.0
	Aug	0.4	0.4	0.6	0.7	0.8	0.9	0.9	1.0
	Sep	0.4	0.4	0.6	0.7	0.9	1.0	1.0	1.0
CSI	**Jun**	0.4	0.4	0.2	0.1	0.0	0.0	0.0	0.0
	Jul	0.5	0.4	0.3	0.2	0.1	0.0	0.0	0.0
	Aug	0.4	0.4	0.3	0.2	0.1	0.0	0.0	0.0
	Sep	0.5	0.4	0.3	0.2	0.0	0.0	0.0	0.0
BIAS	**Jun**	1.3	1.1	1.3	1.2	0.5	0.3	0.1	0.1
	Jul	1.0	0.8	1.1	1.2	0.8	0.4	0.2	0.1
	Aug	1.1	0.9	1.1	1.2	0.9	0.4	0.1	0.0
	Sep	1.1	0.9	1.2	1.2	0.6	0.1	0.0	0.0

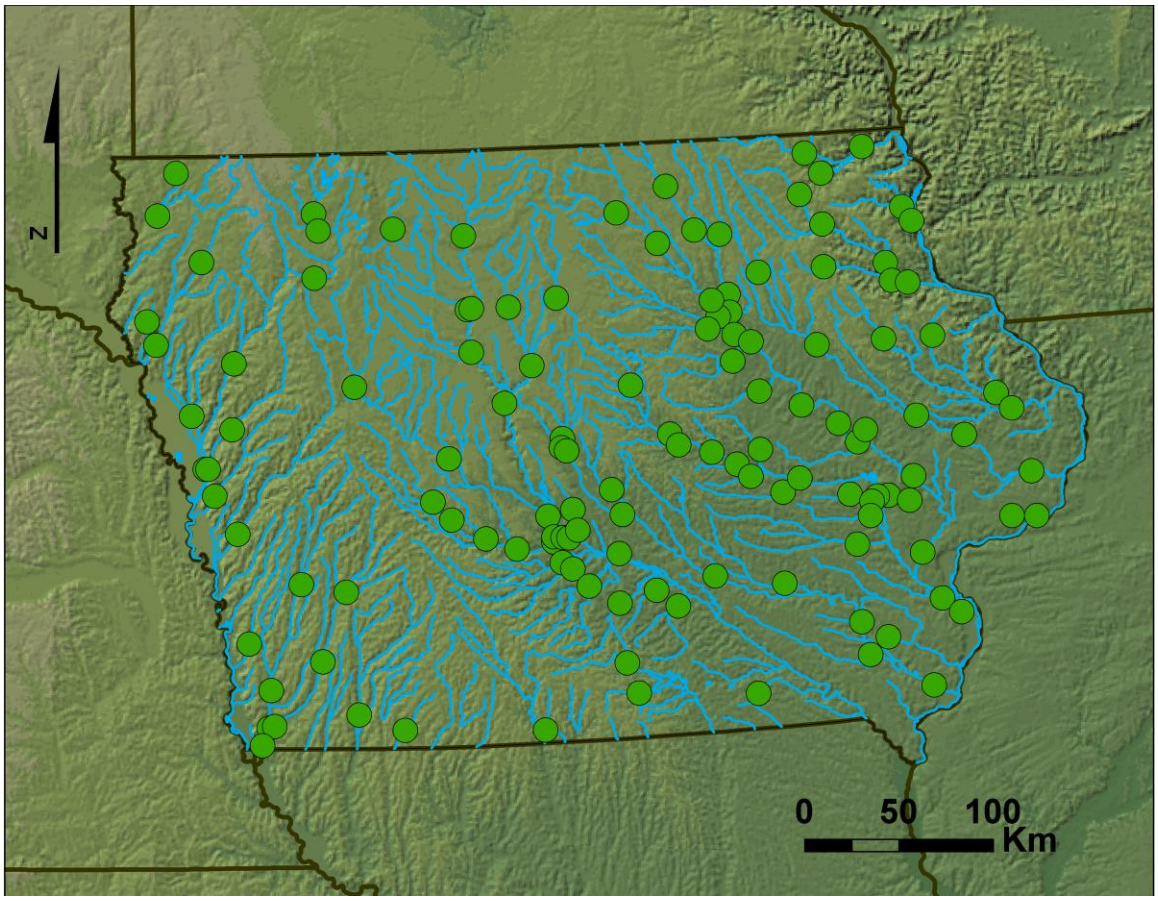

Figure 3.1. A map showing the study area (state of Iowa). The green points represent the USGS gauge locations used in the study (130 gauges).

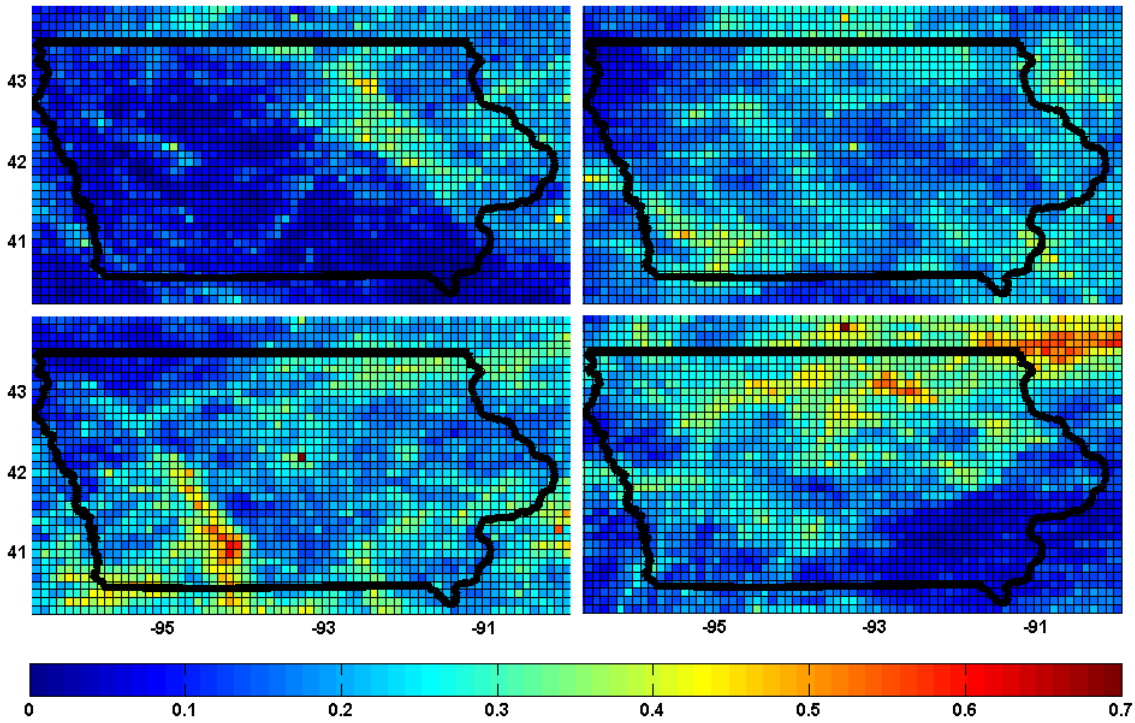

Figure 3.2. MRMS mean hourly rainfall estimates in millimeter for June (top left), July (top right), August (bottom left), and September (bottom right).

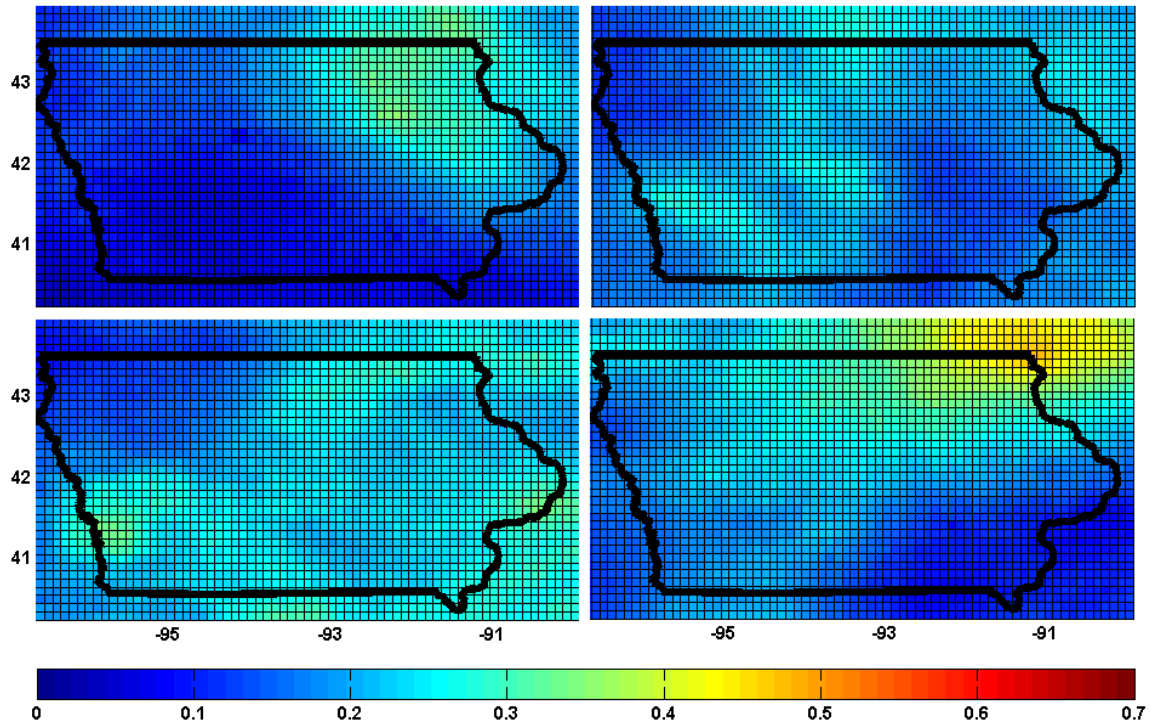

Figure 3.3. IMERG mean hourly in millimeter rainfall estimates for June (top left), July (top right), August (bottom left), and September (bottom right).

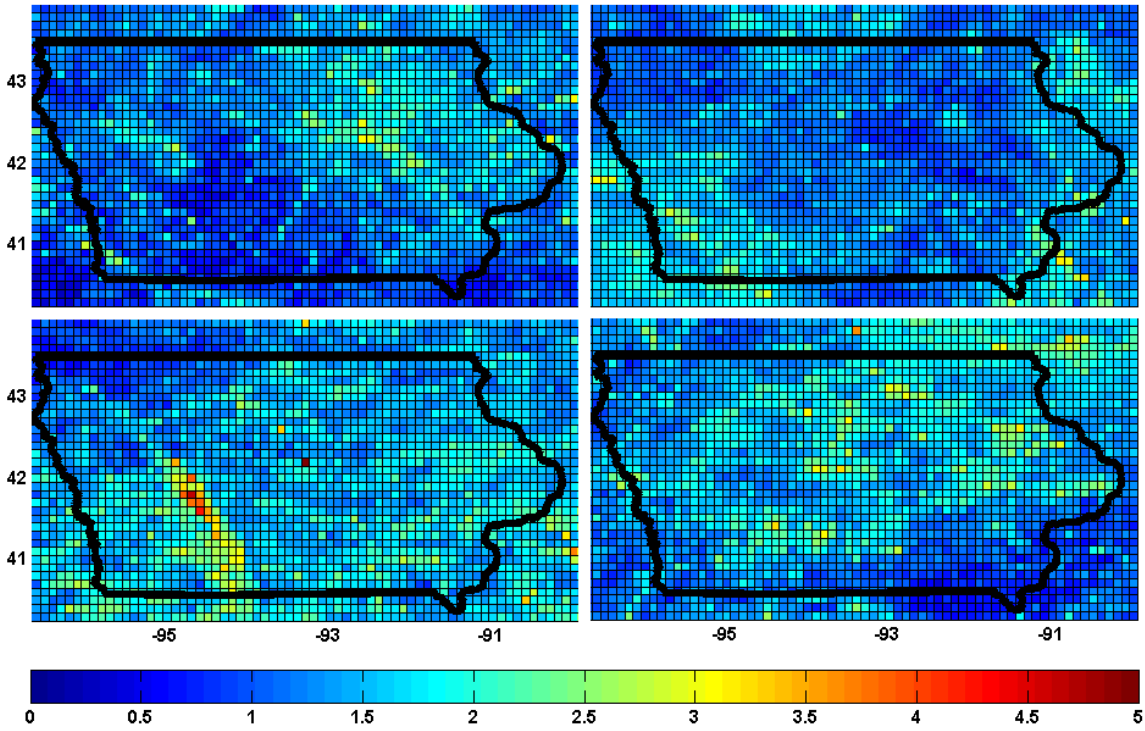

Figure 3.4. RMSE of rainfall discrepancies for June (top left), July (top right), August (bottom left), and September (bottom right).

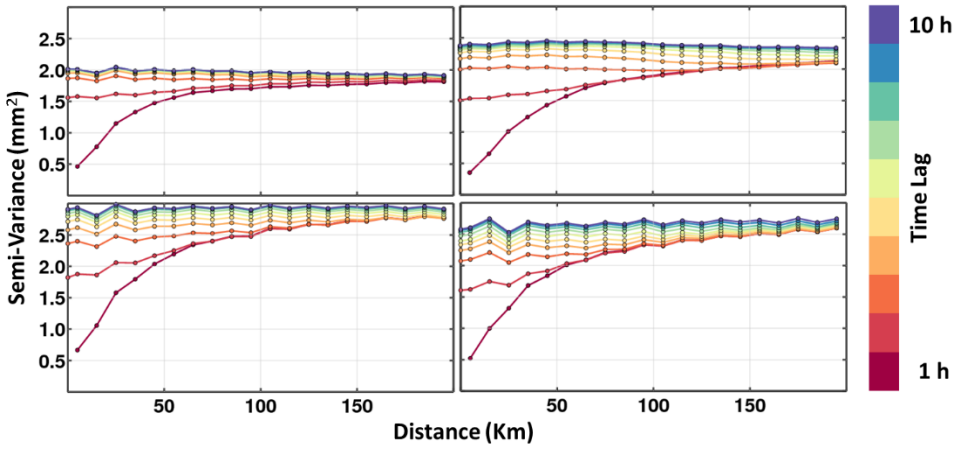

Figure 3.5. Monthly spati-otemporal semi-variograms of hourly MRMS rainfall. The top left panel represents the semi-variance of June, top right panel is for the month of July, bottom left is August, and bottom right is September. The temporal variograms are calculated for 10 lag times (colorbar on the right).

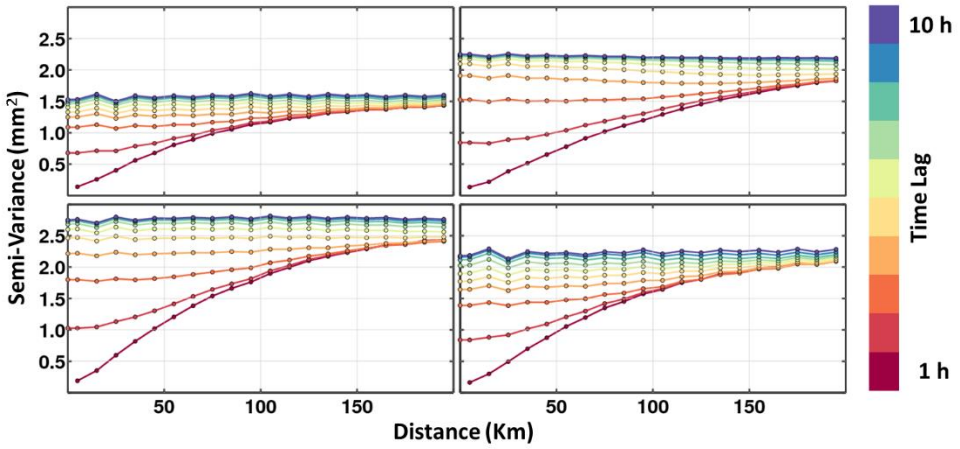

Figure 3.6. Monthly spati-otemporal semi-variograms of hourly IMERG rainfall. The top left panel represents the semi-variance of June, top right panel is for the month of July, bottom left is August, and bottom right is September. The temporal variograms are calculated for 10 lag times (colorbar on the right).

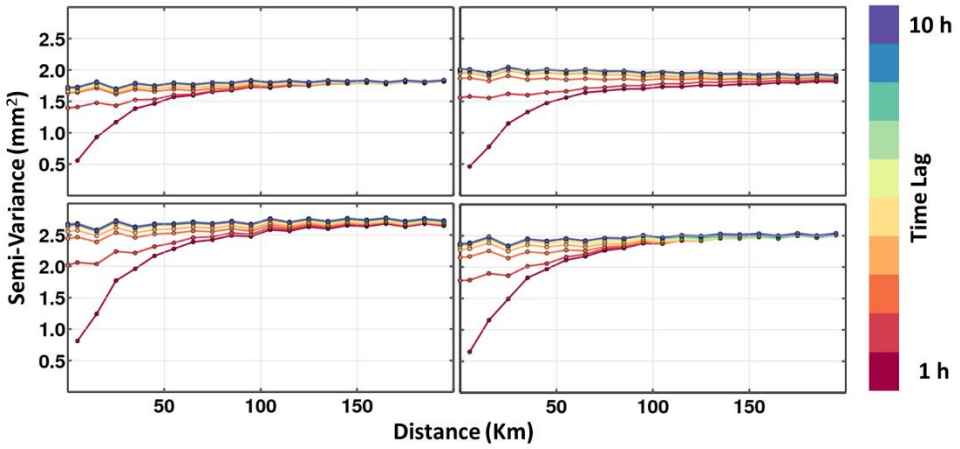

Figure 3.7. Monthly spati-otemporal semi-variograms of hourly rainfall errors. The top left panel represents the semi-variance of June, top right panel is for the month of July, bottom left is August, and bottom right is September. The temporal variograms are calculated for 10 lag times (colorbar on the right).

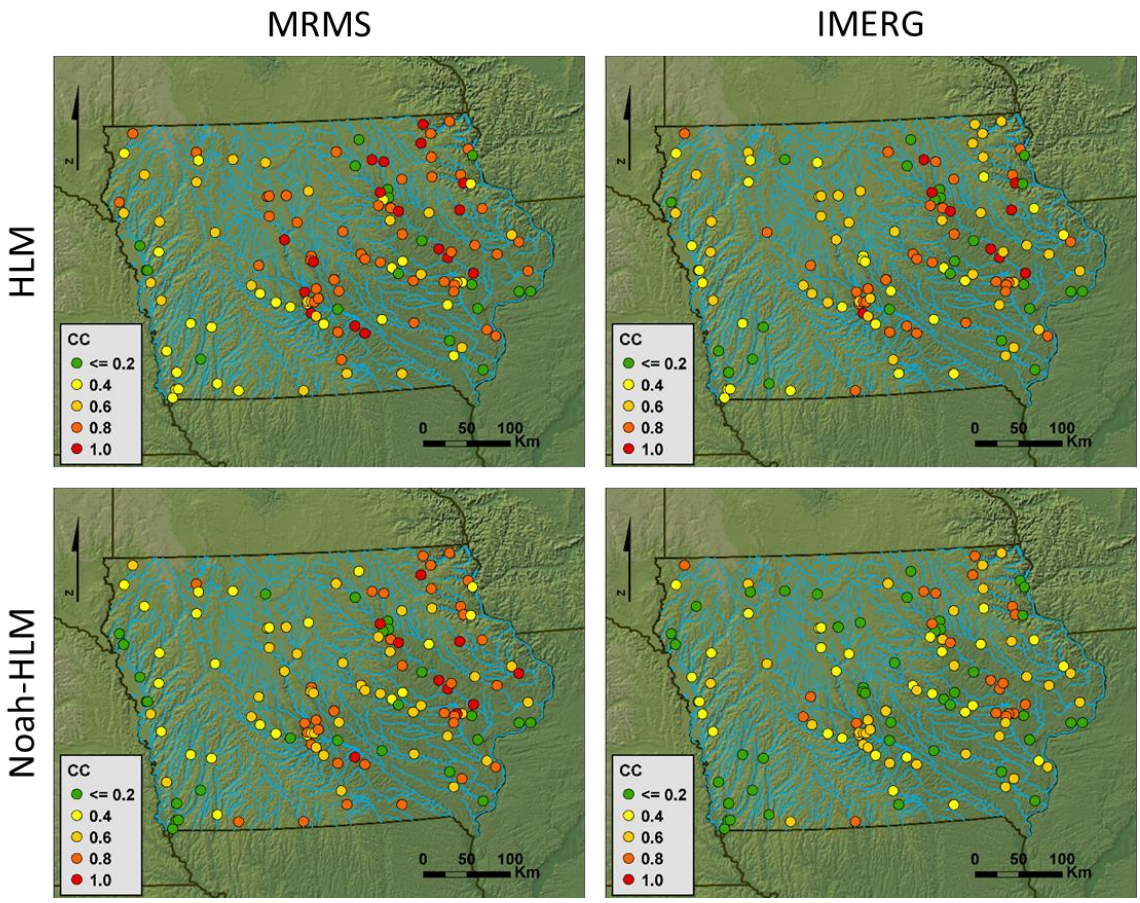

Figure 3.8. Correlation coefficient between modeled and simulated discharges. Top row is for the HLM model and bottom row is for the Noah-MP with HLM routing. Left column represents simulations with MRMS as an input while right column is for IMERG as model input.

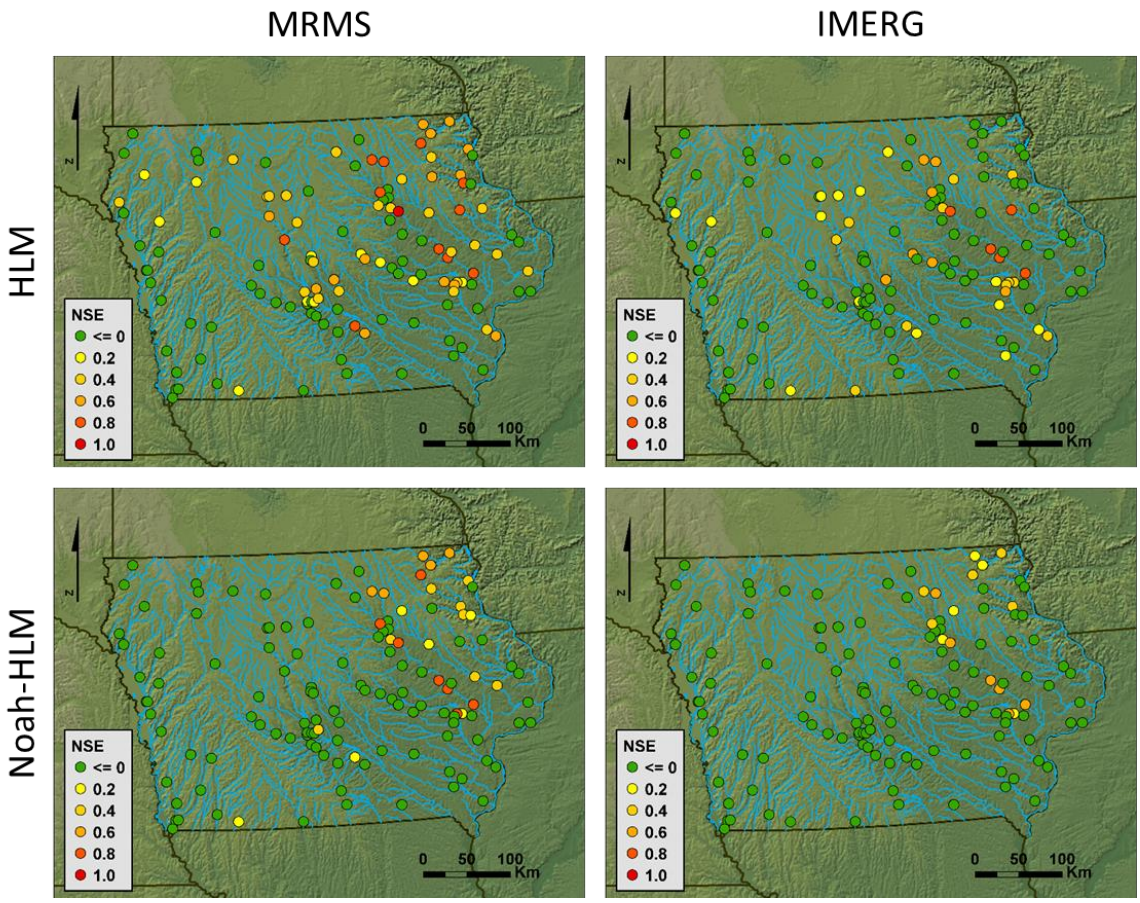

Figure 3.9. NSE between modeled and simulated discharges. Top row is for the HLM model and bottom row is for the Noah-MP with HLM routing. Left column represents simulations with MRMS as an input while right column is for IMERG as model input.

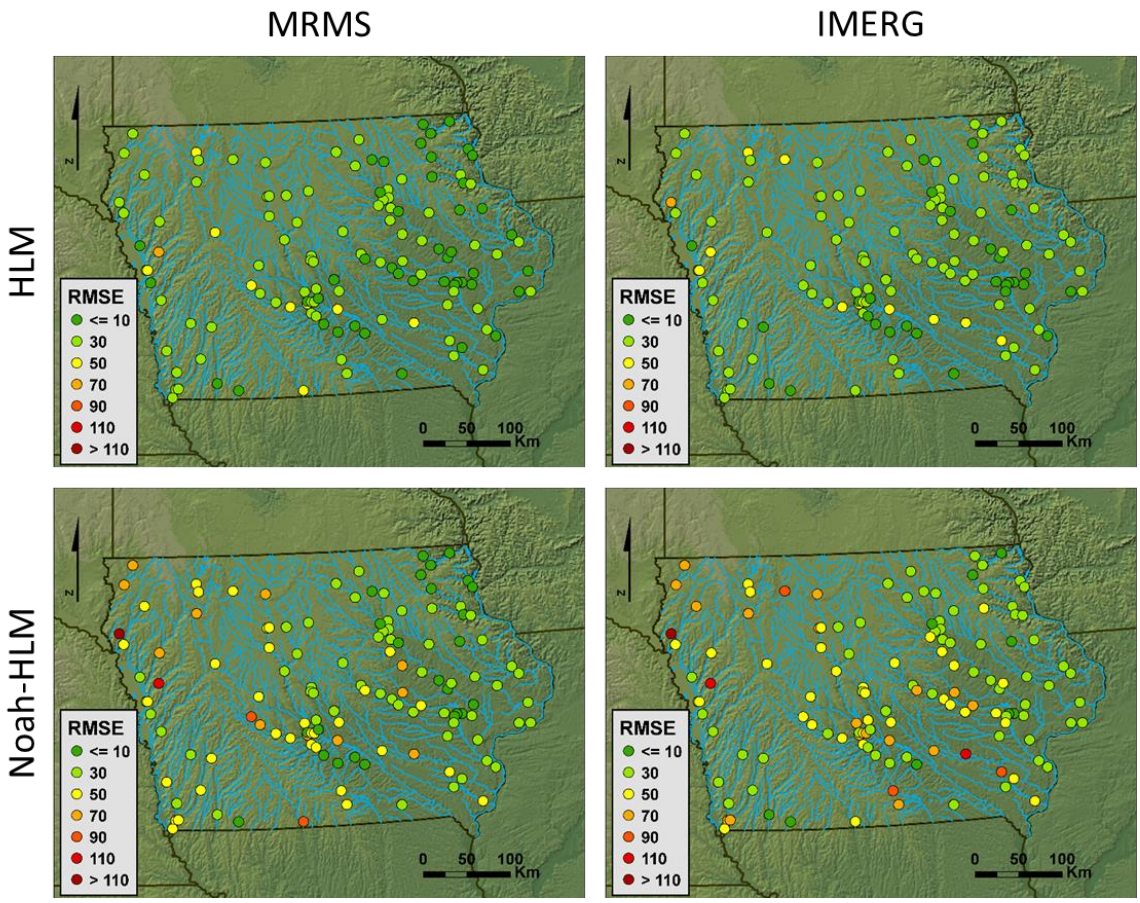

Figure 3.10. normalized RMSE% between modeled and simulated discharges. Top row is for the HLM model and bottom row is for the Noah-MP with HLM routing. Left column represents simulations with MRMS as an input while right column is for IMERG as model input.

CHAPTER 4: River Network Based Characterization of Errors in Remotely Sensed Rainfall Products in Hydrological Applications

4.1. Introduction

Many studies have aimed to evaluate or enhance the performance of satellite rainfall products in hydrologic applications (e.g., Gourley et al., 2011; Vergara et al., 2014; Habib et al., 2014). A few studies have modeled the spatial dependence of the rainfall errors from different sensors but limited the considerations to the Euclidean space (e.g., Vergara et al., 2014; Mandapaka et al., 2009). This Euclidian based characterization is suitable for many applications such as runoff generation, evapotranspiration (ET), or soil moisture content estimation since these variables are continuous in Euclidean space. However, variables such as stream discharge are confined by the stream network, which is dichotomous in nature and a fractal (e.g. Rodriguez-Iturbe and Rinaldo, 1997).

In this chapter, we present an example of river network based characterization of uncertainties in remote sensing rainfall products. River networks act as low pass filters on highly variable rainfall and its estimates, but their averaging properties are different from those of Euclidean domains. Our application is limited to rainfall only but similar considerations, with substantially more difficulties, apply to the transformed variables of runoff and streamflow and their characteristics such as peak and low flows. We characterize spatial dependence of rainfall errors using the tools of geostatistics, namely the semi-variogram (e.g. Cressie, 1993).

The method we use is described by Ver Hoef et al., (2006) and called the Spatial Stream Network (SSN). SSN accounts for the nested nature of the stream network by

using stream distances and site connectivity information. The stream distances are separated into two categories: first, flow-connected distances between sites, i.e. discharge flows directly from the upstream site location to the downstream site location; second, flow-unconnected distances between sites that are only connected through a common downstream junction in the stream network, i.e. discharge cannot flow from one site to another but rather from each site to the nearest connecting downstream junction. As a result, the complete semi-variogram produced from this analysis consists of two (or more) distinct semi-variograms called the Torgegram when combined (Zimmerman and Ver Hoef, 2016). In addition, Ver Hoef et al., (2006) developed theoretical spatial covariance models valid for stream networks.

The SSN method has received little to no attention in hydrologic modelling studies but we find it highly relevant for applications where river networks play a dominant role. In this study, our main goal is to explore the characteristics of the Torgegram when it is applied to satellite rainfall errors that are accumulated through the stream network. A secondary goal is to test an application of the Torgegram for spatial prediction. We fit multiple generalized spatial linear models to the satellite rainfall errors using both Euclidean and stream network based covariance functions, and then evaluate and compare their performance using methods such as the Akaike information criteria (AIC, Akaike 1973) and the Leave-One-Out Cross-Validation (LOOCV) test.

4.2. Methods

A good tool to estimate and visualize spatial dependence is the semi-variance. For two sites that are h distance apart, the semi-variance is described as follows,

$$\gamma(h) = \frac{1}{N(h)} \sum_{\|x_i - x_j\| \in c(h)} \left[z(x_i) - z(x_j) \right]^2, \quad (4.1)$$

where $c(h)$ is a distance bin for data locations approximately distance h apart; $\gamma(h)$ is the semi-variance for sites located inside the distance bin, $z(x_i)$ and $z(x_j)$ are the observed values at sites x_i and x_j respectively, and $N(h)$ is the number of site couples that are inside the distance bin. The result of plotting the semi-variance against distance h is called the semi-variogram. Traditionally, h is measured using Euclidean distance between the observation sites. However, for our stream network application, it is better to characterize the semi-variance in terms of stream distance rather than Euclidean distance. In addition, because some variables are impacted by the flow connectivity structure of the stream network (e.g., there is no water transport between unconnected locations as well as in the upstream direction), it is important to separate the measured distances into two categories. First, distances between flow-connected sites (e.g., the water flows directly from the upstream site location to the downstream site location) are equal to the length of the stream segments connecting the two sites. Second, distances between flow-unconnected sites are equal to the sum of the lengths of the two stream segments that connect each site to the nearest downstream connecting junction (e.g., Figures 4.5. and 4.6.). Because of this distance categorization, we obtain two semi-variograms that when plotted simultaneously are called the Torgegram. If the spatial dependence of the variable of interest is not impacted by network connectivity, the flow-connected and flow-unconnected portions of the Torgegram should be similar; otherwise, they could be quite different.

The semi-variogram and the corresponding spatial covariance function are often characterized by three parameters: sill (maximum semi-variance), range (the distance where the sill is reached), and nugget (abrupt change in semi-variance at the origin), see

for example Cressie, (1993). The covariance function can take many forms (e.g., exponential, spherical, Gaussian) as long as it satisfies certain conditions to insure positive definiteness (Ver Hoef et al., 2006). It is important to note that simply substituting Euclidean distance by stream distance does not guarantee the validity (positive definiteness) of the statistical model (Ver Hoef et al., 2006; Ver Hoef and Peterson 2010). Ver Hoef et al., (2006); Ver Hoef and Peterson (2010) and Zimmerman and Ver Hoef (2016) present groups of valid statistical covariance models for stream networks.

The SSN models are based on Moving Average (MA) constructions (Ver Hoef et al. 2006). The derivation of the models is beyond the scope of this chapter. However, we briefly summarize how the models are obtained. The MA construction is created by considering a random variable to be equal to the integral of a MA function over a white noise process on the stream network. The MA function is only non-zero in one direction, either upstream or downstream. The value of the moving average function is largest at the observation site and decreases as one moves away from the site location. The smaller the separation distance between two sites, the greater the extent of overlap of those portions of the MA functions that are relatively large. This in turn will result in higher covariance.

Covariance models for which the corresponding MA function is only positive in the upstream direction are called tail-up models and they allow only flow-connected sites to be correlated. The covariance function for the tail-up models takes the following form,

$$\boldsymbol{C_{tu}}(s_i, t_j | \{\pi_{i\,j}\}, \theta) = \begin{cases} \pi_{i\,j}\boldsymbol{C_{uw}}(s - t|\boldsymbol{\theta}), & when\ s\ and\ t\ are\ flow-connected \\ 0, & when\ s\ and\ t\ are\ flow-unconnected \end{cases}, \quad (4.2)$$

where $\boldsymbol{C_{tu}}()$ is the tail-up stream covariance function, while $\boldsymbol{C_{uw}}()$ is an unweighted covariance function valid on the real line. Here, s_i denotes a site located on the ith stream segment, a stream distance s upstream from the most downstream point on the network; t_j is defined similarly. The variables $\pi_{i\,j}$ are weights used to split the covariance function at upstream junctions' locations. Streams with larger served areas obtain larger weights and vice versa. Since there is no overlap between the MA functions of flow-unconnected sites, their covariance is equal to zero. On the other hand, the covariance function for the tail-down models takes the following form,

$$\boldsymbol{C_{td}}(s_i, t_j \,|\, \theta) =$$

$$\begin{cases} \boldsymbol{C_{fc}}(s - t|\theta), when\ s\ and\ t\ are\ flow-connected, \\ \boldsymbol{Cfu}(s - qi\,j, t - qi\,j|\theta), when\ s\ and\ t\ are\ flow-unconnected \end{cases}, \tag{4.3}$$

where $\boldsymbol{C_{td}}()$ is the tail-down stream covariance function, $\boldsymbol{C_{fc}}()$ is the covariance function between flow-connected sites, and $\boldsymbol{C_{fu}}()$ is the covariance function between flow-unconnected sites. In the flow-unconnected case the distances are measured from each site to the nearest downstream connecting junction $q_{i\,j}$ to sites s_i and t_j. $\boldsymbol{C_{fc}}()$ and $\boldsymbol{C_{fu}}()$ must be valid Euclidean covariance function of one and two variables, respectively, and are related to each other through their functional dependence on the same MA function.

One problem with the tail-down models, although useful for considering covariance between flow-unconnected sites, is that they assign the same, or an even larger, covariance value between flow-unconnected sites as they do between flow-connected sites with equivalent separation distance. This is not appropriate for many hydrologic variables since observations at flow-connected sites are usually more correlated than those at flow-unconnected sites that are separated by equivalent distances.

Thus, as described in Ver Hoef et al. (2006), it is useful to construct a mixed spatial linear model that incorporates the covariance generated by a linear combination of some or all models (e.g., tail-up, tail-down, and Euclidean) and to consider the variance of the errors to be as follows,

$$var(\boldsymbol{\varepsilon}) = \boldsymbol{\Sigma} = \boldsymbol{\sigma_{tu}}^2 \boldsymbol{R_{tu}} + \boldsymbol{\sigma_{td}}^2 \boldsymbol{R_{td}} + \boldsymbol{\sigma_E}^2 \boldsymbol{R_E} + \boldsymbol{\sigma_{nu}}^2 \mathbf{I}, \quad (4.4)$$

where $\boldsymbol{\sigma_{tu}}, \boldsymbol{\sigma_{td}}, \boldsymbol{\sigma_E}$ are the standard deviations of the errors obtained from the tail-up, tail-down, and Euclidean models respectively; and $\boldsymbol{R_{tu}}, \boldsymbol{R_{td}}, \boldsymbol{R_E}$ are the correlation matrices of the tail-up, tail-down, and Euclidean models, respectively. The final term is the variance of the nugget effect $\sigma_{nu}{}^2$ multiplied by the identity matrix I.

4.3. Rainfall products and study area

In this study, we characterize spatial uncertainties of the satellite rainfall product Integrated Multisatellite Retrievals for GPM (IMERG) Final Run (Huffman et al., 2014). IMERG combines rainfall estimates from Microwave (MW) sensors on board the Global Precipitation Measurement's (GPM) satellite constellation as well as estimates from Infrared (IR) sensors on board geostationary satellites (Hou et al., 2015). The IMERG product undergoes bias correction using the monthly rain gauge estimates provided by the Global Precipitation Climatology Centre (GPCC). The product has a half-hourly temporal resolution and is available on a 0.1˚ square grid with a global coverage between 60˚ N and 60 S˚ (Liu 2016). Our benchmark (reference) product in this study is the Multi-Radar Multi-Sensor (MRMS) Quantitative Precipitation Estimate (QPE). This product uses data from the Weather Surveillance Radar-1988 (WSR-88D) and is corrected using rain gauge data provided by the Hydrometeorological Automated Data

System (HADS) (Zhang et al., 2015). The product is available on a 0.01˚ square grid and at hourly temporal resolution.

Our study area is a mid-sized Cedar River basin in eastern Iowa, USA with an overall drainage area of approximately 17,000 Km2. The shape of the basin and the width function of the river network are illustrated in Figure 4.1. (a) and (b). The width function could be interpreted as a distribution of distances from the outlet. From Figure 4.1. (b) it is evident that the majority of the pathways to the outlet have length between 150 and 300 km. The basin experienced a heavy rainfall event in September 2016, which resulted in the second highest river water level in history on September 27, 2016 (22 feet, https://www.weather.gov/dvn/summary_09272016). We use this event as our test case for the evaluation. We use the digitized stream network provided by the National Hydrography Dataset Plus Version 2 (NHDPlus V2, http://www.horizon-systems.com/nhdplus/documentation.php) to obtain the necessary stream network information (e.g., stream distance, and stream served area). Other necessary information (e.g., connectivity, upstream distance from the basin outlet, and tail-up covariance function weights) is obtained using the STARS and SSN software packages described in Peterson and Ver Hoef (2014).

In Figure 4.2., we show the rainfall accumulations during the flood event period (September 14 through September 27, 2016) obtained from MRMS (a), IMERG (b). The differences between IMERG and MRMS as a percentage of the benchmark MRMS observed rainfall, expressed as,

$$\% \, Error = \frac{IMERG - MRMS}{MRMS} * 100, \tag{4.5}$$

are shown in Figure 4.2. (c). One can see how the satellite product generally over-estimated rainfall on the eastern side of the basin while missing a dense cluster of accumulation in the top half of the basin.

We then apply the approach described in Quintero et al., (2016) and Cunha et al., (2015) to compare our rainfall products. We convert gridded rainfall products to a more hydrologically relevant representation by accumulating the rainfall for all size basins. Normalized by the drainage area, rainfall is assigned to the river network link that drains that area. We then subtract the evaluated product from the corresponding reference product thus obtaining a river network representation of the rainfall differences (errors). Figure 4.3. shows the results of this procedure. Strictly speaking, the reference product is also subject to estimation uncertainty (e.g. Villarini and Krajewski 2007) but we ignore it here for the sake of simplicity of our illustration. Radar-rainfall estimation uncertainty is considerably smaller than that of satellite products.

An important feature of Figure 4.3. is how the magnitude of errors in rainfall estimates accumulates downstream (i.e., the errors in lower stream orders (small streams) are much higher than the errors in high order streams). This variability in performance across scales indicates that as the basin size increases, the difference in the overall performance between the benchmark product and satellite rainfall decreases. This in turn could have implications for the performance of the hydrologic models in terms of estimated stream discharges when IMERG is used as the rainfall input (i.e. better hydrologic models' performance at larger scales; Quintero et al., 2016). We use the difference values between IMERG and MRMS at approximately two thousand sites

located near the network confluences (Figure 4.1.) over the whole basin to perform our evaluation (e.g., to produce the Torgegrams and covariance models).

4.4. Results

We begin our analysis by investigating the semi-variance of the rainfall differences using the traditional Euclidean space semi-variogram in multiple directions where 0˚ is the North-South direction and 90˚ is the East-West direction (Figure 4.4.). In Figure 4.4. the diameter of the semi-variance circles is proportional to the number of observations that fall in a given distance bin. One can see that the directions with highest number of observations for longer distances are the 135˚ and the 0˚ directions because of the basin's shape and orientation. One can also see that the sill was reached at distance of around 55 Km for the 90˚ and 135˚ directions, and at around 30 Km for the 0˚ and 45˚ directions indicating little to no autocorrelation beyond these distances.

In Figure 4.5., we show the Torgegram of rainfall differences, where the green points are the semi-variances obtained from flow-unconnected sites, while the blue points are obtained from flow-connected sites. It is important to note that the distance used here is the stream distance, which should be generally greater than Euclidean distance due to the meandering nature of rivers and streams. This is illustrated in the right side of the figure where the distance between two flow-connected sites is represented by a solid line and the distance between two flow-unconnected sites is represented by a dashed line. In addition, the maximum stream connected distance should be close, but not exactly equal, to the maximum distance of the width function. As one would expect, the number of flow-unconnected sites at relatively larger distances is generally greater than the number of flow-connected sites. In addition, the flow-unconnected semi-variances extend for

longer distances in comparison with flow-connected semi-variances. The semi-variance between flow-unconnected sites is higher than the semi-variance of the connected sites at the same separation distance. The flow-unconnected semi-variogram reaches the range at around 150 Km, while the flow-connected semi-varigram stops increasing at around 200 Km. This difference in ranges indicates that rainfall errors among flow-connected sites are autocorrelated for longer distances compared to flow-unconnected sites. However, the flow unconnected semi-variogram experiences a large increase in semi-variance towards the end. This also occurs in the Euclidean semi-variogram in the 0˚ and 90˚ directions where the semi-variance keeps increasing without bound. In addition, one would expect that at very large distances the semi-variance for both flow-connected and flow-unconnected sites should be similar due to the absence of autocorrelation. This suggests trend contamination in the data (Zimmerman and Ver Hoef 2016).

To remove the underlying spatial trend, we begin with a visual inspection of the rainfall differences field. It suggests that for this particular storm, IMERG systematically overestimated the rainfall amounts on the eastern side of the basin while underestimating on the western side of the basin, thus causing differences in the mean value at different locations within the basin. In addition, the amount of accumulated rainfall error systematically decreases as the stream order increases. Detailed instructions on how to deal with this effect are available in Zimmerman and Ver Hoef (2016). Here, we follow their strategy and fit, by ordinary least squares, a linear model to the data using stream order, longitude, and latitude as regressors. Afterwards, we reproduce the Torgegram using the residuals of this model (Figure 4.6.). One can see that now the range of the flow-unconnected sites is similar to what it was for the raw data (around 125 Km) while

the monotonic increase of the flow-unconnected semi-variogram has disappeared. In addition, the range of the flow-connected semi-variogram is significantly larger than that of the flow-unconnected semi-variogram (about twice as large). This linear trend-corrected Torgegram represents the hydrologically-relevant spatial dependence of satellite rainfall errors.

Our next objective is to test the predictive ability of various models of the spatial dependence. In our example, the errors are calculated and known at some 2000 locations but one may imagine a situation where error prediction in river space is needed. Therefore, the next step is fitting spatial statistical models to our data. Given the nature of our variable, both tail-up and tail-down models should be important for error covariance modeling. This is because some flow-unconnected streams (especially low order streams) receive the same amount of rainfall differences due to their proximity. We compare the predictive performance of the stream network models with that of the Euclidean models.

For our example spatial linear model, we use the benchmark rainfall, upstream distance, and served (drainage) area as predictors. For the covariance of the errors, we produced linear models that leverage each of the tail-up, tail-down, and Euclidean covariance functions. In addition, we produced two mixed models, one with tail-up and tail-down, and another that included tail-up, tail-down, and Euclidean covariance models. We used spherical covariance functions for all models. Table 4.1. shows the comparison between all models. We fit all models using ML to be able to compare their AIC scores. The mixed model with all covariance schemes outperformed all other models and received an AIC score of 8966. The next best model is the mixed model with tail-up and

tail-down alone with AIC score of 9095, i.e. only slightly worse. All mixed models performed significantly better than the pure Euclidean model. In addition, the pure tail-up model performed worse; we expected this since the tail-up models allow nonzero correlation only between flow-connected sites, which is not a good representation of our variable (rainfall differences). In addition to the AIC comparison, Table 4.1. contains Leave-One-Out Cross Validation test (LOOCV) results. The Root Mean-Squared Error of the predicted values is in agreement with the AIC scores and shows that the mixed models outperformed the Euclidean model. In Figure 4.7., we plot the comparison between the LOOCV prediction normalized by the prediction standard error versus the observations for the Euclidean, mixed tail-up and tail-down, and mixed all models from left to right respectively. The prediction from both mixed models is better than that from the pure Euclidean model, while the difference in performance between the two mixed models is insignificant.

4.5. Conclusions and future work

The main goal of this study is to highlight the effect of the basin size and stream network configuration on the performance of the satellite rainfall product IMERG Final Run in hydrological modelling applications. The process of aggregating rainfall estimates downstream in the network helps eliminate a large portion of the differences between IMERG's rainfall estimates and the benchmark product MRMS. Nevertheless, the remaining portion is correlated for much longer ranges within the stream network. Overall, stream based spatial statistical models resulted in better prediction when compared to the Euclidean distance models. However, the choice of the covariance moving average function has a strong impact on determining the model's performance.

One should carefully choose a moving average function that is compatible with the nature of the variable being analyzed.

This approach is also applicable to other hydrologic variables, in particular to those that result from rainfall transformation, i.e. runoff and streamflow. Specifically, one could explore the nonlinear effects of system memory, via soil moisture, on the spatial dependence of errors in flood peaks. Future work that builds on our present effort can include performing hydrologic model runs and analyzing the differences in discharges and other model variables (e.g., antecedent soil moisture) and relating them to the discrepancies in rainfall estimates.

Table 4.1. Spatial statistical models performance metrics including the AIC scores and Root Mean-Squared Prediction Error RMSPE.

Covariance Model	AIC	RMSPE
Up-Down-Euclid	8966.8	4.01
Up-Down	9095.1	4.42
Tail-Up	11314.6	13.3
Tail-Down	10613	6.8
Euclidean	10481	6.8

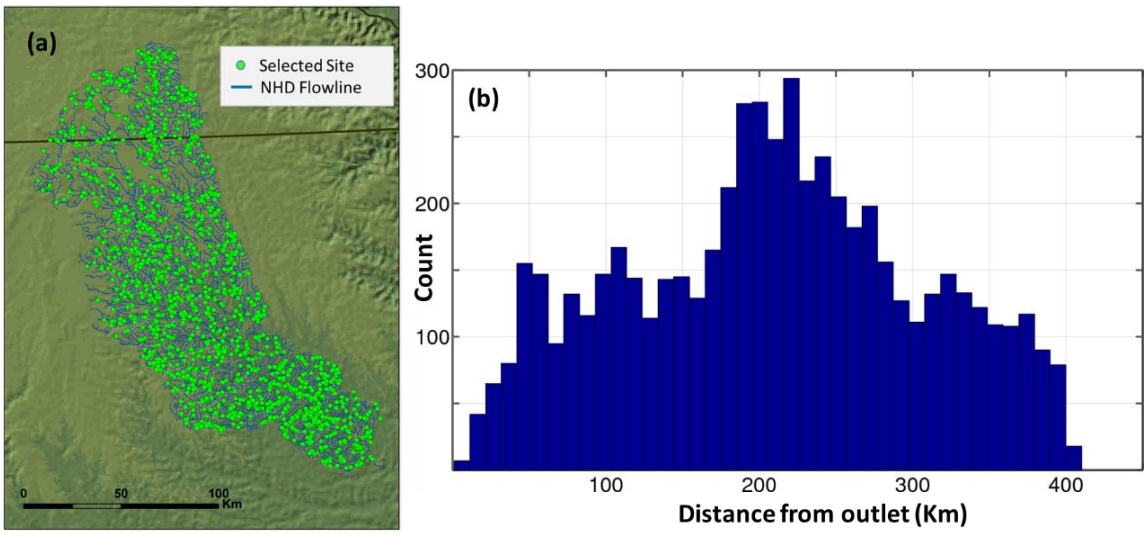

Figure 4.1. (a) The Cedar River basin located in eastern Iowa. The Blue lines represent the NHDPlus V2 stream network definition (Flowlines), while the green points represents the sample sites located near the confluences of the network. (b) The width function of the basin showing the distance from the outlet in Km (x-axis) and the number of streams at a certain distance (y-axis).

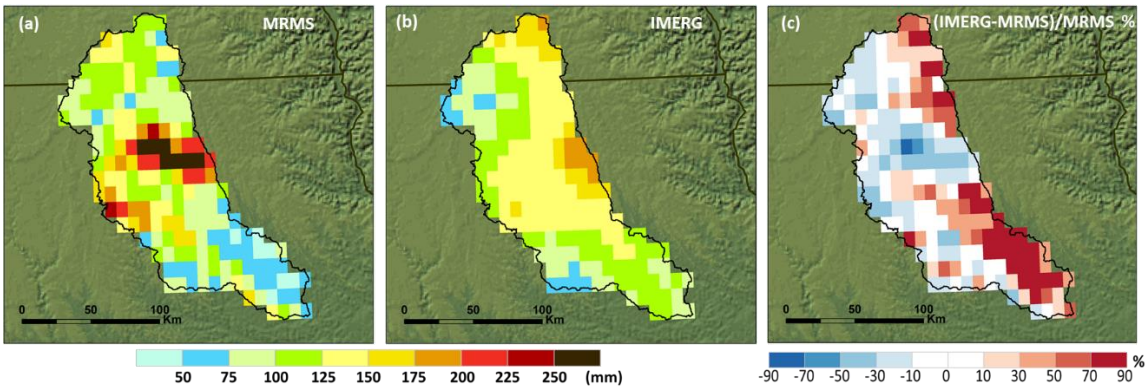

Figure 4.2. Rainfall accumulations during the period of (September 14, 2016 through September 27, 2016) for MRMS (a) and IMERG (b). MRMS has been aggregated in space in order to be compared to IMERG. The normalized differences between the two products is shown as % difference in (c).

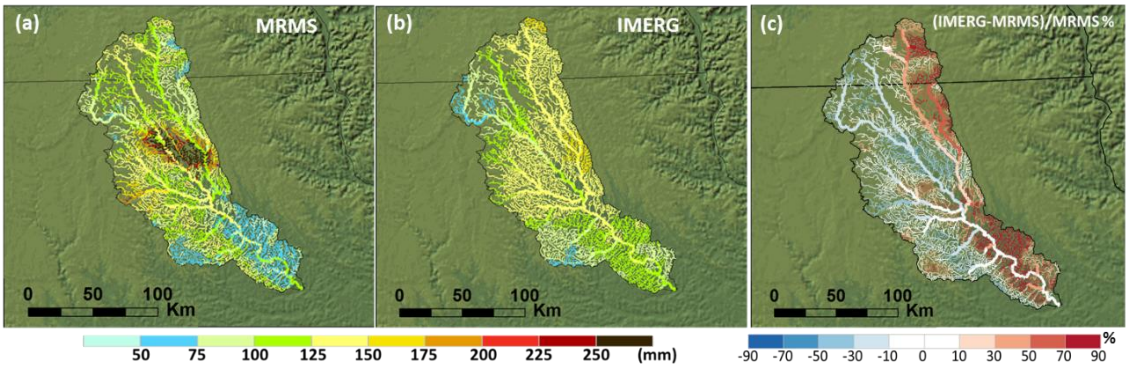

Figure 4.3. Rainfall accumulations during the period of (September 14, 2016 through September 27, 2016) for MRMS (a) and IMERG (b). Unlike Figure 2., each product was accumulated along the stream network. The normalized differences between network based accumulations of the two products is shown as % difference in (c).

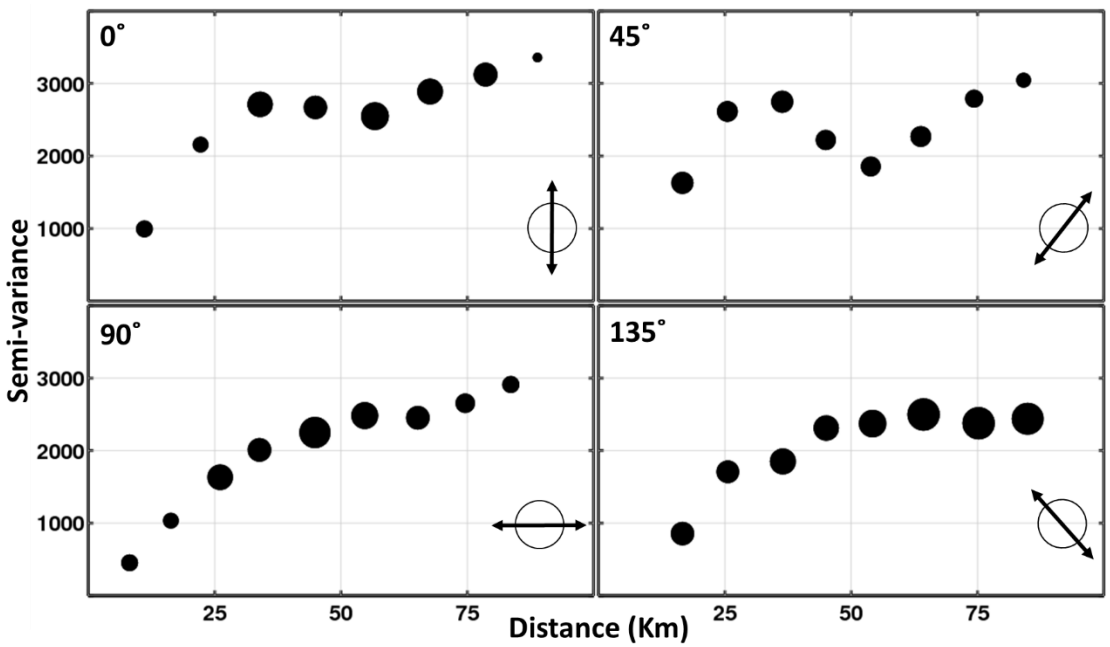

Figure 4.4. The directional semi-variogram obtained from the rainfall differences in IMERG resolution. The direction 0˚ is the North-South while 90˚ is East-West. The size of the black circles is proportional to the number of sites that are fall into a given distance bin. The circle with the arrow at the bottom right of each panel shows the variogram direction.

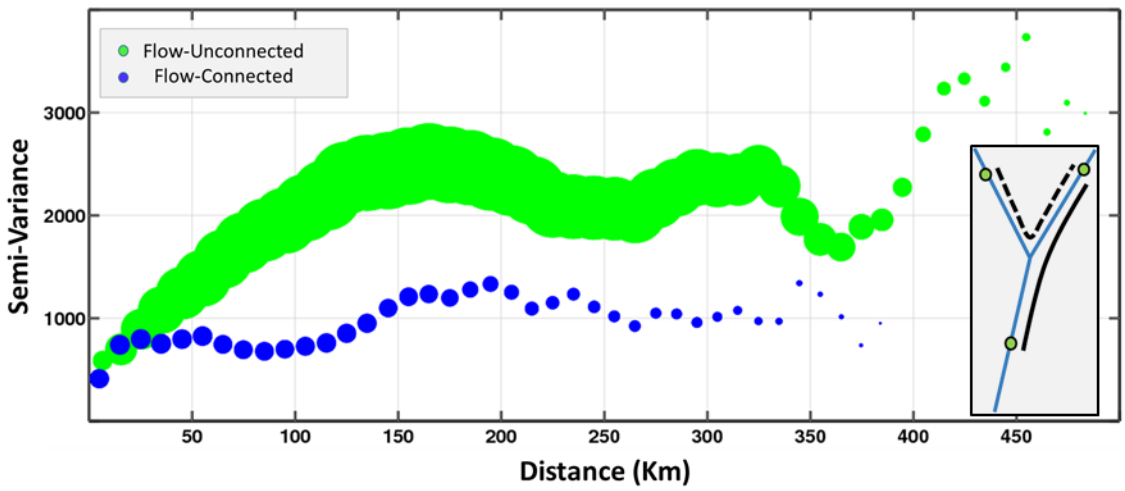

Figure 4.5. The Torgegram obtained using the differences calculated along the stream network. The green circles represent the flow-unconnected sites while the blue circles represent the flow-connected sites. The sizes of the circles are proportional to the number of sites that fall into a certain distance bin. The schematic on the right side illustrates the difference between flow-connected (solid line) distances and flow-unconnected distances (dashed lines).

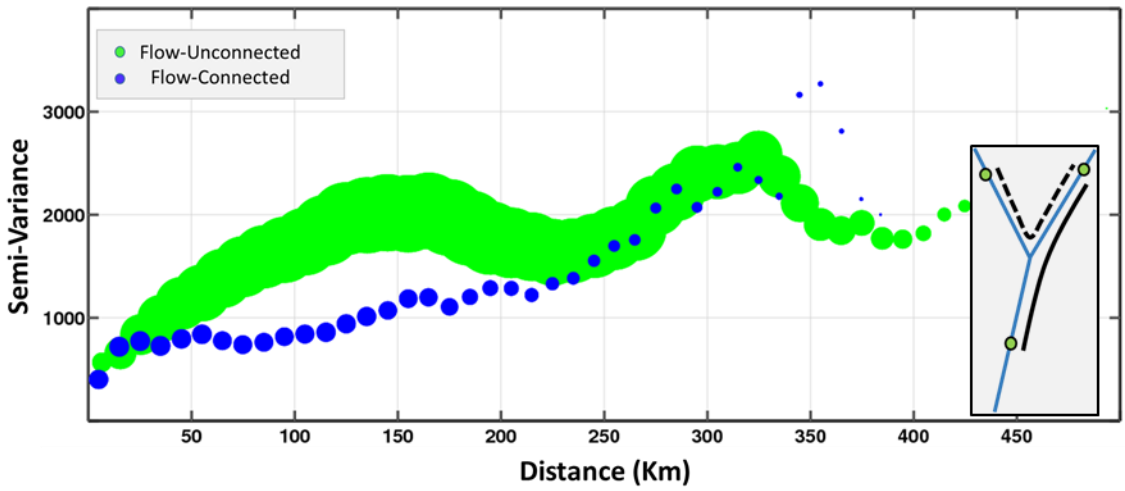

Figure 4.6. The Torgegram of the differences along the stream network after trend removal. The green circles represent the flow-unconnected sites while the blue circles represent the flow-connected sites. The sizes of the circles are proportional to the number of sites that fall into a certain distance bin. The schematic on the right side illustrates the difference between flow-connected (solid line) distances and flow-unconnected distances (dashed lines).

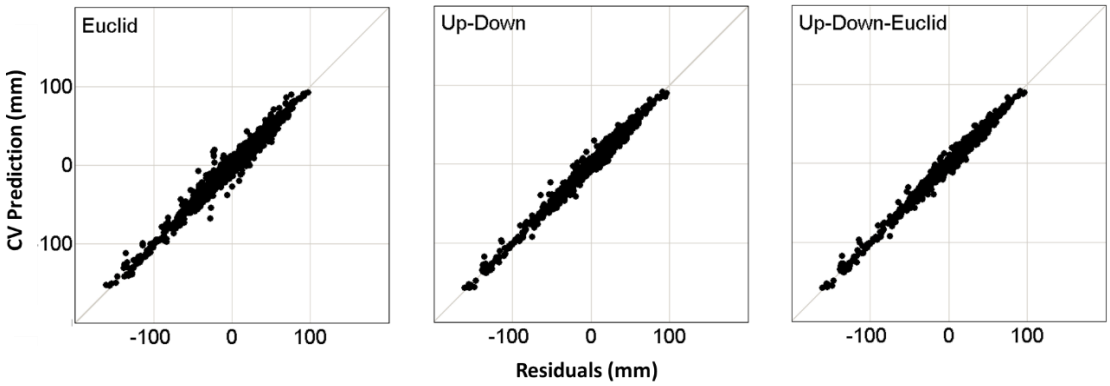

Figure 4.7. The LOOCV predictions (y-axis) plotted against the rainfall differences (x-axis). The results are shown for the pure Euclidean model (a) and two mixed models; first, tail-up and tail-down (b), second, tail-up, tail-down, and Euclidean (c).

CHAPTER 5: A Time-based Framework for Evaluating Hydrologic Routing Methodologies Using Wavelet Transform

5.1. Introduction

Distributed hydrologic models usually consist of two major components that together produce stream discharge estimates. The first component is a Land Surface Model (LSM) that decomposes the terrain into a regular (e.g. rectangular or triangular) or irregular (e.g. terrain-fit polygons outlining hillslopes) grid where the energy and mass exchange between the land and the atmosphere are modeled. This process produces estimates of the potential excess surface and subsurface runoff depths that will enter the stream channels. It also provides estimates for evapotranspiration (ET), snowmelt, and the amount of water that is stored in the soil, and would affect the discharge estimates in the short or the long term. The second component is a water routing component, which is responsible for delivering the excess runoff into the channels through surface and subsurface lateral flow motions, then transporting the runoff downstream as stream discharge (David et al., 2011; Niu et al., 2017).

In this chapter we evaluate two hydrologic routing components: first, the Routing Application for Parallel Computation of Discharge (RAPID) introduced in (David et al., 2011, 2013), which is a storage-based simplified Muskingum linear routing method. Similar to other hydrologic routing approximations, the Muskingum method relies on discharge continuity alone and does not take the momentum equation into account. The Muskingum method uses the two parameters k and x to accumulate the discharge downstream, where k is a storage parameter that has units of time, and x is a weighting parameter that is relative to the discharge inputs and outputs of the channel. The

estimation of these parameters is explained in David et al., (2011) and Snow (2015) and the estimated values can be improved through optimization. Examples of possible cost functions that can be used to optimize these parameters are found in David et al., (2011, 2013), and Tavakoly et al., (2016). The second routing component we evaluate in this study is the one implemented in the Hillslope-Link hydrologic Model (HLM) developed and used by the Iowa Flood Center (IFC). This routing component is nonlinear and accounts for the momentum equation in a simplified form (Ayalew et al., 2014; Cunha et al., 2015; and Mantilla, 2007), and the stream velocity is determined based on the nonlinear relationship between the discharge and the served area (Paik and Kumar 2004; Ghimier et al., 2017). Similar to David et al., (2011), Snow et al., (2016), and Tavakoly et al., (2016) we do not account in this study for lateral flows to the channel in our calculation for the sake of simplicity and since our sub-catchment sizes are small. This means we assume the runoff depth immediately enters the streams in their corresponding upstream junctions before the channel routing takes place.

Our study area is a moderately monitored average-sized basin called the Cedar River Basin located in the eastern part of the state of Iowa in the United States. The stream discharges of the basin are not affected by any artificial storages (e.g. dams and reservoirs) and are monitored by eleven United States Geological Survey (USGS) stream gauges that cover various connected and unconnected sub-catchments located within our basin of interest. In order conduct inter-comparisons between the observed and simulated stream discharges (hourly), we obtained the output of the two routing components at these USGS gauge locations for a whole warm season (March through October with a two month spin-up period). Both routing components were derived by the same runoff

estimates which were generated by the community Land Surface Model (LSM) called Noah Multi-Parameter (Noah-MP; Niu et al., 2011), the model inputs are discussed later in this chapter. Afterwards, the discharge is aggregated over the same stream network by each routing component; for our application we used the digital stream network called the National Hydrography Dataset Plus Version 2 (NHDPlus V2, available at http://www.horizon-systems.com/nhdplus/).

Statistical skill scores such as root mean squared error, correlation coefficient, and Nash Sutcliff efficiency index are widely used to evaluate the performance of hydrologic models. However, comparing two time series with such indices reduces the features of the complete time series to one value that does not provide information about the variation of the model's performance across the time series. One can decide to inspect individual parts of the time series separately, but this division of the time series can be either arbitrary if done using a fixed time window, or difficult to automate if the important features in the time series are to be located precisely in time. This implies that detecting features of interest in the time series is also a significant part of the evaluation problem. In this chapter to evaluate our routing components we use a method that offers solutions to both of these problems: first, detecting features of interest, and second, evaluating the model's performance during the time period when they occurred.

We perform this time-based evaluation using Continuous Wavelet Transform (CWT) and Cross-Wavelet Transform XWT (Torrence and Compo 1998; Labat, 2005). Wavelet transform analysis provides information about a given time series by filtering its frequencies in a time-localized manner. This information cannot be obtained using the traditional statistical skill scores such as RMSE and correlation coefficient due to their

generality and inability to locate time series' significant features in time. CWT provides information about the significant frequencies that exist in a time series and XWT provides information about where these detected frequencies coincide in two time series. This event matching is determined by computing the phase difference between the CWTs of the different time series. Torrence and Compo (1998) and Foufoula-Georgiou and Kumar (1995), are excellent references for a detailed description of CWT and XWT. We will briefly review the concepts behind CWT and XWT and elaborate on the use of different wavelets, and their advantages and disadvantages in our application.

Accurate peak time prediction for stream discharge is crucial in hazardous situations such as flash floods. Our method and results provide an insight regarding a more effective procedure to evaluate routing components in particular and hydrologic models in general based on their ability to forecast peak times. The evaluation framework we propose in this study should help decision makers and ground responders in making a more informed decision regarding peak time occurrence based on the performance of the hydrologic model which they receive their information from.

The rest of this chapter is organized as follows. First we introduce our study area and describe the data obtained from the two routing components. Second, we review CWT and XWT and different types of wavelets that we used in this study and why we used them. Third, we show the results of the study. Lastly, we discuss our conclusions and findings.

5.2. Study area, data inputs, and routing components

5.2.1. Study area

We apply our wavelet-based evaluation method to the stream discharge estimates from RAPID and HLM routing components at eleven USGS gauge locations in the Cedar River basin. A map of the stream gauge locations is shown in Figure 5.1. and the areas they serve are listed in Table 5.1.. As represented in Figure 5.1., our study area mostly fall in the "dryland crop" USGS land cover category, with a small area under "developed with low intensity category". There is little to no river regulation effect at our basin, this insures that there is no delay in discharge due to storage. The areas served by the USGS stream gauges range from 766 km2 at Little Cedar River near Ionia, IA (USGS ID 05458000) to 16,862 km2 covered by the station at the basin outlet Cedar River at Cedar Rapids, IA (USGS ID 05464500). In this study we used hourly stream gauge observations to compare to our hourly model outputs, and stream gauge observation were represented by green dots. The legend of the four major land cover types is shown in the top right corner where 1 is dryland crop, 2 is urban and built-up, 3 is cropland/grassland, and 4 is water body. The labels of the USGS gauges correspond to the rankings in Table 5.1..

5.2.2. Data inputs

We applied the same runoff estimate which we obtained from the WRF-Hydro hydrologic framework (Gochis et al., 2015) to both routing components. WRF-Hydro utilizes the Noah-MP LSM (Niu et al., 2011) to estimate the runoff depths. We used the Stage IV (Lin and Mitchell 2005) product as the rainfall input to WRF-Hydro; the rainfall estimates are projected on an HRAP (Reed and Maidment, 1999) grid with a spatial resolution of approximately 4 km x 4 km and a temporal resolution of 1 hour. We also

used the NLDAS V2 atmospheric forcings (Xia et al., 2012a,b) to drive WRF-Hydro; the dataset is available in a spatial resolution of 1/8˚ and hourly temporal resolution. We ran the model for the whole 2014 warm season (March through October), allowing a proper spin-up period starting from December 2013 with a proper initialization for channel states. We then ingested this runoff in both routing components allowing two months of spin-up period for routing. Consequently, our stream discharge estimates cover the period between May through October, 2014.

5.2.3. Routing components

The methodologies on which each routing component is based can be found in Ayalew et al., 2014, Cunha et al., 2015, and David et al., (2013). In short, RAPID is based on the linear Muskingum routing method, which is a storage-based method that does not account for flow momentum. The finite difference form of the Muskingum method can be written as shown in David et al., (2011). This method relies on two key parameters, $\boldsymbol{k}$ and $\boldsymbol{x}$, where $\boldsymbol{k}$ is a storage parameter and has units of time and $\boldsymbol{x}$ is a dimensionless weighting parameter. The $\boldsymbol{k}$ and $\boldsymbol{x}$ parameters were determined as described in Snow (2015).

On the other hand, the HLM is based on the methodology described in Mantilla (2007). In this method the channel velocity exhibits non-linear behavior in relationship with the upstream served area. As a simplification for the momentum equation the velocity v corresponding to a discharge $\boldsymbol{q}$ can be determined as follows,

$$\boldsymbol{v(q) = v_0 {q_{link}}^{\lambda_1} A^{\lambda_2}} \tag{5.1}$$

then the flow transport can be described as,

$$\frac{\boldsymbol{dq_{link}(t)}}{\boldsymbol{dt}} = \frac{\boldsymbol{v_0 q_{link}^{\lambda_1}(t) A^{\lambda_2}}}{\boldsymbol{(1-\lambda_1) l}} \left[\boldsymbol{a_h \left(q_{runoff}(t) + q_{baseflow}(t)\right) - q_{link}(t) + q_{up}(t)}\right] \tag{5.2}$$

where at a certain time $\boldsymbol{t}$ the corresponding surface and subsurface lateral inflows are $\boldsymbol{q_{runoff}}$ and $\boldsymbol{q_{baseflow}}$ respectively. The discharge inflow into a stream channel (link) from the upstream served area $\boldsymbol{A}$ is $\boldsymbol{q_{up}}$ and the total discharge at the outlet junction of a stream is $\boldsymbol{q_{link}}$. The parameters $\boldsymbol{v_0}$, $\boldsymbol{\lambda_1}$, $\boldsymbol{\lambda_2}$ are global parameters that are equal to 0.3, 0.2, and -0.1 respectively and correspond to the USGS hydraulic measurements and the methodology described in Paik and Kumar (2004) and Mantilla (2007).

5.3. Methodology

Wavelet transform is a useful tool to extract information about time localization of certain frequencies in a time series, unlike Fourier transform which does not provide any time localization information (Labat, 2005, Torrence and Compo 1998). One can consider wavelet transform as an extension to the Windowed Fourier Transform (WFT), where CWT extracts information about the signal structure through a filter of scaled and translated wavelet instead of an infinite sinusoid.

5.3.1. Continuous Wavelet Transform (CWT)

Decomposing a signal into its constituent frequencies can be done in multiple ways. First, one can perform the Discrete Fourier Transform, (DFT) on the signal. DFT can give us an accurate measure of all the frequencies in a discrete signal $\boldsymbol{x_n}$, as shown in Equation (3) by multiplying the signal by infinite harmonics $\boldsymbol{e^{-2\pi ikn/N}}$, which are also called sinusoid, and then calculating the integration of this multiplication from – ∞ to ∞. The result $\boldsymbol{x_k}$ is the magnitude and phase of this particular harmonic with frequency k in the time series. If a harmonic with a frequency is in fact a constituent of the signal, the integral will show its amplitude and phase in the signal. If however, this particular frequency does not contribute to the time series, the summation in Equation (5.7) will be

equal to zero. Most natural signals are comprised of waves with various frequencies, amplitudes, and phase angles, which in turn result in different integral values. The DFT equation for a discrete time series $\boldsymbol{x_n}$ is given as:

$$\boldsymbol{X_k = \sum_{n=0}^{N-1} x_n . e^{-2\pi ikn/N}} \quad (5.3)$$

where $\boldsymbol{k}$ is the frequency of the harmonic; $\boldsymbol{i}$ is the imaginary unit; $\boldsymbol{n}$ is the location index in the time series; and N is the total number of samples to be analyzed in the signal. The main disadvantage of DFT is that although it is efficient in filtering the frequencies, it does not provide any information about where this particular frequency appeared and lasted in the signal. The difference in representation between a frequency which lasted throughout the whole signal and another that contributed to the signal for a short time period is that Fourier transform calculated for the latter will be dilated in the frequency domain. In other words, DFT lacks any information regarding time localization. To overcomes this problem, researchers came up with a more advanced way of performing the DFT called the Windowed Fourier Transform (WFT), which relies on moving a box function which can take different shapes, e.g. a simple rectangle or a Gaussian window (Kaiser 1994), and calculating the transform over the location of that box as seen in Equation (5.4).

$$\boldsymbol{X[n,k] = \sum_{n=0}^{N-1} x_n[n] . w[n-\tau] . e^{-2\pi ikn/N}} \quad (5.4)$$

where $\boldsymbol{w[n-\tau]}$ is the window function shifted at time $\boldsymbol{\tau}$; and $\boldsymbol{X[n,k]}$ is the WFT at a particular location $\boldsymbol{\tau}$ and a particular frequency $\boldsymbol{k}$. However, the WFT still has its disadvantages. This is because WFT uses a fixed size window which forces us to sacrifice either detecting the accurate time localization by choosing a large window (one can think of DFT as WFT with an infinite size window); or on the other hand, one can choose a

very small window which will allow for better time localization but poor frequency detection. This brings us to the Continuous Wavelet Transform (CWT), which offers a compromise between frequency and time localization detection. Continuous wavelet transform is performed through translating scaled versions of a wavelet of a certain functional form, called "mother wavelet", over the time domain of a signal (a hydrograph in our case) and calculating the resulting power from the convolution between the discrete signal and the wavelet. In order to detect different frequencies in the signal, the wavelet is compressed or stretched through scaling. The lower the value of the scale the tighter the wavelet becomes, and this in turn amplifies the effect of high frequencies in the signal. On the other hand, the higher the value of the wavelet scale the more stretched and dilated it becomes, which is good for detecting smoother low frequencies in the signal.

The compliance of CWT with the window size uncertainty is done by offering good time localization for the important sharp features by using low scale wavelets. Similarly the mild features that cover the wide range of the signal will be detected by high scale wavelets that do not offer good time localization. Another advantage of using wavelet analysis rather than Fourier transform is the ability to use confined "mother wavelets" that can have different shapes based on their functional form, and not an infinite sinusoid. As shown in Figure 5.2., some wavelets such as the Morlet wavelet are more suitable for detecting frequency properties, while others such as the Derivative of Gaussian (DOG) wavelet or Paul wavelet are better for detecting sharp features. One of the main goals of this chapter is to stress the significance of the selected wavelet shape, and that it can lead to a different interpretation of the properties of our signal.

One can use either directional wavelets (a wavelet with an imaginary component), such as the Morlet or Paul wavelets, or a real valued wavelet such as the DOG wavelet to estimate the CWT. This is not the case for XWT since it solely relies on estimating the phase difference between two CWTs. Given a mother wavelet $\boldsymbol{\psi}$ that is equal to the original unscaled function $\boldsymbol{\psi_0}$ shown in Equations (5.6), (5.7), and (5.8), but instead normalized to have unit energy. The CWT of a discrete signal $\boldsymbol{x_{\acute{n}}}$ can be calculated as follows:

$$\boldsymbol{W_n^x(s) = \sum_{\acute{n}=1}^{N-1} x_{\acute{n}}\, \psi^* \left[\frac{(\acute{n}-n)\delta t}{s}\right]} \tag{5.5}$$

where $\boldsymbol{s}$ is the wavelet scale and $\boldsymbol{n}$ is a wavelet location parameter which determines the wavelet location during the translation across the time domain of the signal. The asterisk assigned to $\boldsymbol{\psi^*}$ denotes the complex conjugate, and $\boldsymbol{\delta t}$ is the time step; finally $\boldsymbol{W}$ is the wavelet spectrum and the wavelet power is obtained by calculating $\boldsymbol{|W^2|}$. It is important to note that in the case of a directional wavelet, power is generated from both the real and imaginary part. This means the imaginary peak would move the location of maximum power if it overlaps with a similar feature in the signal. As described in Torrence and Compo (1998), we normalize the power by the variance ($\boldsymbol{\sigma^2}$) in order to have a measure relative to that of a white noise process.

For our CWT analysis we chose three of the most commonly used wavelets. The first is the real DOG wavelet in the second order (also called the Mexican hat wavelet or Marr wavelet), given by the equation

$$\boldsymbol{\psi_0\,(\eta) = \frac{(-1)^{m+1}}{\sqrt{\Gamma\left(m+\frac{1}{2}\right)}} \frac{d^m}{d\eta^m}\left(e^{-\frac{\eta^2}{2}}\right)}, \tag{5.6}$$

where m is the derivative and is equal to 2 and $\boldsymbol{\Gamma}$ is the gamma function and $\boldsymbol{\eta}$ is a time parameter. Second is the directional Paul wavelet in with order of four,

$$\boldsymbol{\psi_0}(\boldsymbol{\eta}) = \frac{\boldsymbol{2^m i^m m!}}{\sqrt{\boldsymbol{\pi(2m)!}}}(\boldsymbol{1 - i\eta})^{-(\boldsymbol{m+1})}, \tag{5.7}$$

where m is the order and is equal to 4. Lastly is the directional Morlet wavelet,

$$\boldsymbol{\psi_0}(\boldsymbol{\eta}) = \boldsymbol{\pi^{-\frac{1}{4}}}\, \boldsymbol{e^{i\omega_0 \eta}}\, \boldsymbol{e^{\frac{-\eta^2}{2}}}, \tag{5.8}$$

where $\boldsymbol{\omega_0}$=6 to achieve wavelet admissibility.

We estimated the 5% significance levels of the wavelets power spectrum following Torrence and Compo (1998) by comparing the power spectrum to a background red noise. In short, the significance levels were estimated based on the hypothesis that our signal can be modeled as a univariate autoregressive with lag-1 AR(1) process. Before we proceed with the significance level calculation we had to test our signal's power spectrum decay against that of a theoretical red noise with the same lag-1 correlation. As an example, we present in Figure 5.3. the power spectrum decay of our signal (hydrograph) at the outlet of the basin (Cedar River at Cedar Rapids), where the blue line represents our signal and the red line represents the theoretical red noise, the lag-1 auto correlation was estimated from the data. One can see that a red noise process is a good approximation for our process; this was expected since many geophysical processes follow a red noise model. The Fourier power spectrum of the wavelets and the corresponding confidence intervals are then calculated as described in Torrence and Compo (1998).

Another important feature in our CWT calculations is the Cone of Influence (COI). Similar to Torrence and Compo (1998), in order to perform the integral in Equation (5.5) efficiently we chose to perform our calculations in the Fourier space, then

calculate the wavelet transform using a reverse Fourier transform of the convolution of the Fourier transform of the signal over the Fourier transform of the wavelet. We followed a zero padding procedure (adding zeros with a proper length to both ends of our series) since our signal is confined in time, which in turn results in artifacts in the wavelet transform at the edges of our signal. The COI helps us know where the effect of these artifacts is negligible, where all values outside of the COI are reliable. In this chapter, the COI as well as the transform-normalized powers are plotted against the Fourier periods of the wavelets and not their scales. The Fourier period λ for the same scale s is different from one wavelet shape to another; for the Morlet wavelet the Fourier period λ=1.03s, while for the Paul wavelet λ=1.4s, and finally for the Mexican hat wavelet λ=3.97s. For the sake of consistency in our analysis we are going to refer to the Fourier periods instead of the scales of the wavelets.

5.3.2. Cross-Wavelet Transform (XWT) and phase-time analysis

The next step after calculating the CWT is to relate these CWTs to each other to detect the similarities between our signals. This is done by calculating the CWT of the observed and simulated stream discharges, and then the XWT between each simulated flow and the observed flow. The XWT is calculated using Equation (5.9):

$$\mathbf{W}_{\boldsymbol{n}}^{\boldsymbol{XY}}(\boldsymbol{s}) = \mathbf{W}_{\boldsymbol{n}}^{\boldsymbol{X}}(\boldsymbol{s})\mathbf{W}_{\boldsymbol{n}}^{\boldsymbol{Y}^{*}}(\boldsymbol{s}), \tag{5.9}$$

The Cross-Wavelet power is $|\mathbf{W}_{\boldsymbol{n}}^{\boldsymbol{XY}}(\boldsymbol{s})|$, and as we mentioned earlier the XWT is complex with an amplitude and phase. Therefore, we only used the Paul and Morlet wavelets for this application. The XWT power is an indicator of the locations where both signals had similar transform behavior. This is where the wavelet shape plays an important role, since if the CWT wasn't able to detect or merely detected a certain feature

in the signal due to the wavelet shape, the XWT power will be weak at the location of this feature. The phase difference between the two signals, which is equal to the phase of the XWT, is shown in Equation (5.10):

$$\emptyset_n^{XY}(s) = \boldsymbol{tan^{-1}}\left[\frac{I(W_n^{XY}(s))}{R(W_n^{XY}(s))}\right] \tag{5.10}$$

where $\boldsymbol{I}$ denotes the imaginary component, and $\boldsymbol{R}$ is the real component. We then convert this phase difference $\boldsymbol{\varphi}$ to obtain the time difference $\boldsymbol{\Delta t}$ using the Fourier period $\boldsymbol{\lambda}$ as follows:

$$\Delta t_n^{XY}(s) = \emptyset_n^{XY}(s) * \frac{\lambda}{2\pi} \tag{5.11}$$

As shown in Equation (5.11), the time differences are associated with certain scales. However, some scales are more important than others based on the magnitude of the XWT at this scale over the locations of our interest. Similar to Liu et al., (2011), we decided to represent the time difference between the simulated and observed signals using the $\boldsymbol{\Delta t}$ at the scale of maximum power.

Naturally, we are interested in the performance of the routing components during extreme events; thus, we estimate the time difference between simulated and observed flows at peak locations. This also makes sense since as we have explained earlier, the wavelet transform offers good time localization for small scales and not large scales, where small scales are associated with sharp peaks. We also had to add more constraints to make sure we are getting reliable time difference estimates. First, we only consider the periods corresponding to the maximum power at the peak locations and within the COI. In Figure 5.4. we show both the Paul and Morlet wavelets with Fourier period 400 on top of the hydrograph at the outlet of the basin, where both the wavelets and the hydrographs are normalized to have maximum value of 1. One can see that at this Fourier period, the

wavelet is intersecting with multiple features in the hydrograph, and as the period increases the resulting power will have very poor time localization. In other words, the incoming power will result from the convolution of the wavelet with multiple features. For this reason we impose our second condition, which is limiting our search for maximum power to a maximum Fourier periods of 400 if the COI is not already covering this region. This seemed as a good limit as shown in Figure 5.4., since the peaks presented in it are the largest peaks, and all other peaks upstream must be smaller.

5.4. Results

We start our evaluation of RAPID and the HLM routing by visually comparing their CWTs against the CWT we obtained from the observed flow. In this chapter we only show the CWT for the hydrographs at the outlet of the basin in order to show the effect of the wavelet shape. This example will enable us to observe the different resulting power spectra associated with the hydrograph peaks and their time localization with regard to wavelet shape. The reasons we only show the CWT at the basin outlet are: first, CWT is only implicitly included in our main evaluation method, which is the time difference between simulated and observed flows; second, since we have eleven stations to analyze, this will result in an overwhelming amount of figures. Therefore, the details of the CWT performance across scales can be indirectly explained by our time difference analysis.

In Figure 5.5. we show the observed and simulated hydrograph at the outlet (a), the CWTs of the observed (b), HLM routing (c), and RAPID (d), using the Mexican hat wavelet. One can see in Figure 5.5. (b) that the two peaks in the hydrograph were accurately detected: the first peak ranges between periods of 128 hours to 400 hundred

hours and the second peak is wider with period ranging from about 150 to 500 hours. Starting from the period of about 500 hours, we can see that the power spectrum from both peaks becomes connected, since at this time scale both peaks are covered with one large wavelet. This is a good example to show why we enforced our second restriction by not considering periods larger than 400. The power spectrum values around 1000 hours and 2000 hours on the x axis, before and after the peaks, are caused by the negative correlation between the sinking parts of the wavelet and the peaks. If we look at the wavelet transform value alone, these regions have negative wavelet transform value. However, it is important to understand that we look at the wavelet transform power instead of the transform original value because in the case of directional wavelets there is an imaginary component that we will either have to look at separately from the real component, or otherwise combine the two components together by calculating the power. By looking at the HLM routing hydrograph, it is clear that the first peak is larger and occurred earlier in comparison to the observed flow; the second peak, however, is smaller in magnitude and duration. This is reflected in the CWT power values in Figure 5.5. (c): the first spike in power occurred earlier than that of the observation and the signal is much stronger all the way to the period of 500. On the other hand, the second peak is located approximately between the 128 and 300 periods. For RAPID, we can see that the wavelet power occurred earlier for the first peak which is in agreement with the hydrograph, and the second peak caused a larger than observation spike in the CWT power. Interestingly, because the two RAPID peaks are well separated, due to a very early first peak and generally flashy behavior, the connection between the two peaks in CWT occurred in a much larger period (about 700 instead of 500).

Figure 5.6. and Figure 5.7. are the same as Figure 5.5. but for the Morlet and Paul wavelet CWT respectvely. These figures suggest the same qualitative conclusions derived from Figure 5.5.. However, one can see that the Morlet wavelet maximum power is located between the two peaks, since the maximum correlation between the Morlet wavelet and the hydrograph occurs when two humps from the wavelet are co-located with the two hydrograph peaks. Nevertheless, this power is shifted to the left because when the wavelet is far left (e.g. the first major wavelet hump is located over the second peak), the rest of the wavelet peaks do not correlate with the rest of the hydrograph. In the case of the Paul wavelet, we can see that the peaks are well confined with a weak connection resulting from the imaginary part of the wavelet. It is also interesting to observe how the negative correlation due to the sinking part of the wavelet is not separated from the positive part happening over the peak, contrary to what happened in the Mexican hat wavelet case. This is due to the overlapping of the sinking part of the wavelet with the rising imaginary part of the wavelet. This however resulted in a wider continuous base of power values at large period. Also, the power peaks are slightly shifted to the left because the imaginary component is always ahead of the real component.

Before we proceed to the XWT calculation, it is important to keep in mind that the XWT estimates common power activity locations in both CWTs involved, regardless of how strong these activities are. So even if we focus our interests on the 95% significant values within the COI and a period less than 400, we still have quite a range of agreement between the two CWTs. In order to decide which XWT value at the peak locations is going to be used in time difference estimation, we follow the approach described in Liu et al., (2011). The challenge in choosing the appropriate Fourier period to estimate the time

difference at is the fact that the larger the wavelet is, the more power it produces. This means that if the peak was not strongly detected at the proper smaller period (e.g., due to the simulated peak looking very different from the observed peak in small periods or due to the wavelet shape), the agreement will instead be reached at a large period with bad time localization (this is why we stop our search at Fourier period of 400). In most cases agreement between the CWTs of the observed and simulated flows was reached in the proper scales, but in a few instances however we had to choose the power at period 400 or near the edge of the COI. Nevertheless, this limitation will not affect our conclusions, because we are not intending to use the time difference values to modify the simulated hydrographs, e.g. Liu et al., (2011); the values we get will still be an accurate representation of the relative performance between the two routing components.

We now focus our attention to the major peaks in the hydrograph. As shown in previous figures, the 2014 season experienced two major events at the Cedar River basin. We have estimated the XWT transform at each station using both directional wavelets. In Figure 5.8. we show the hydrographs (top), time difference estimated using Equation (5.11) (middle), and the XWT (bottom) at the basins outlet. The two vertical grey lines over the hydrograph peak locations are where we extracted the power and time information. The left column is for RAPID results and the right column represents the HLM routing results. We can see how the maximum power is located around the 256 period band for both routing components. Also, it is clear that the corresponding time difference from RAPID at this region at peak locations is much higher in RAPID that it is in HLM routing. We show a detailed cross section at peak locations in Figure 5.9.. The top panel is for the first event's cross section and the bottom panel is for the second

event's cross section. The dashed lines represent RAPID and the solid lines represent the HLM routing; we represent the power with blue color and the corresponding time difference with the red color. We see that maximum power convergence for both routing components in both events occurred at periods less than 300. Additionally, one can see how the behavior of the time (phase) difference is unstable in regions apart from the maximum power. Another important observation is how the magnitude of the power is generally larger for the first event than it is for the second event, and how the maximum power occurred in a larger period in the first event than it did in the second event. As we have mentioned earlier, the power profile indicates common activity between the observed and simulated flows; although we have only picked one point (maximum power) to estimate the time difference, it is evident that the HLM routing performed better across a wide range of periods.

We then expanded our analysis using the Morlet wavelet upstream by doing the same analysis shown in Figure 5.8. and Figure 5.9. at all ten upstream gauge locations. Figure 5.10. and Figure 5.11 show the time difference (top), Maximum power (middle), and corresponding scale (bottom) for the first and second events respectively using the Morlet wavelet at all stream gauge location. The red color represents the HLM routing and the green color represents RAPID. Although the chosen periods are the same or similar at all stations for both routing components, the time differences are sometimes significantly different with the HLM outperforming RAPID in both events at all locations. We plot the final estimates of the time differences using the Morlet wavelet spatially in Figure 5.12., where positive values indicate that the simulated flow is ahead of the observed flow and vice versa. The results show that both routing components

performed better in the second event, compared to their performance during the first event. What is curious, however, about Figure 5.12. is how the time difference values obtained for the first and the second event are not as different as one would have expected them to be by visually inspecting the hydrographs. This is of course because one can judge by looking at the hydrographs that both routing components performed much better in simulating the peak time during the second event than they did for the first event, especially in the case of RAPID. As we will see later, this is not the case with the Paul wavelet and is mainly caused by the shape of the Morlet wavelet. This is due to the fact that the Morlet wavelet at this range of periods (200 to 300) intersects with both peaks, and the phase difference at the location of one peak is affected by the performance at the other peak as well.

Figure .13. and Figure 5.14. are the same as Figure 5.8. and Figure 5.9. respectively but for the Paul wavelet. In this case we can see how the difference between the time difference estimates of the two events is enlarged. This is due to the better time localization of the Paul wavelet where the effect of one event on the XWT at the location of the other event is small. On the other hand one can also see in Figure 5.15. and Figure 5.16. that the maximum power did not occur before the period of 400 in few locations. Although the time differences are different from those obtained for the Morlet wavelet (Figure 5.10. and Figure 5.11.), generally we reach the same conclusions by using the Paul wavelet. We lay down the time difference spatially in Figure 5.17.. In this case the performance between the first and second events is distinguished properly. We can see how RAPID performed much better in the second event in comparison to the first event; similar behavior is observed for the HLM routing as well. One can also see how the

performance for the first event deteriorated when using the Paul wavelet as the contribution from the well-detected second peak diminished.

5.5. Conclusions and recommendations

In this chapter we introduced a wavelet-based evaluation method of two hydrologic routing components. By looking at the hydrographs it was clear that the simplified Muskingum-based routing component RAPID exhibits flashy behavior with sharp peaks in comparison to the non-linear HLM routing component. Traditional statistical skill scores such as RMSE and correlation coefficient can characterize the overall performance of a time series without providing any localized information in time about the significant features in the time series. In addition, these skill scores do not provide any information regarding the ability of the models to predict an accurate peak time, and peak time prediction is of great value to decision makers. Thus, we used the wavelet analysis which is widely used in the application of signal processing. CWT allowed us to filter the constituent frequencies of the hydrographs and also provided us with the locations of high frequency activities which correspond to significant peak locations. We also calculated the XWT, which provided us with the locations where both the observed and simulated discharges from each routing component experience similar power activity. The XWT also has a phase component which is the same as the phase difference between the CWTs of the observed and simulated flows. Then we calculated the time difference between the simulated and observed flows at peak locations using the Fourier period of the wavelet. A unique aspect of our study is that have analyzed the same hydrographs using different wavelet shapes. Although we have reached the same qualitative conclusion that the HLM routing outperformed RAPID in simulated peak

times, there were differences in the interpretation of the routing component performance event wise. The significance of these differences is solely dependent on the shape of the hydrograph. For example, if we only had one peak in the hydrograph or if the peaks were well separated, the choice of wavelet would have been of less importance. On the other hand if we had many closely located peaks, the results from the Morlet wavelet would have been very hard to analyze since the computed power at one peak location is affected by the performance at surrounding peaks. Nevertheless, although the Paul wavelet offers better time localization, the location and Fourier period of the maximum power does not exactly correspond to the location and width of the feature we would like to detect due to the wavelet shape (the imaginary component of the wavelet). Hence, it is recommended to always look at the XWTs from different wavelets, especially for the purpose of applications such as the time series modification performed in Liu et al., (2011).

Table 5.1. A list of USGS gauges and their served area.

Rank	USGS ID	Area (km^2)
1	5459500	1342
2	5457700	2792
3	5458000	766
4	5463000	911
5	5458900	2213
6	5458500	4338
7	5458300	4044
8	5463500	780
9	5464000	13333
10	5464220	773
11	5464500	16862

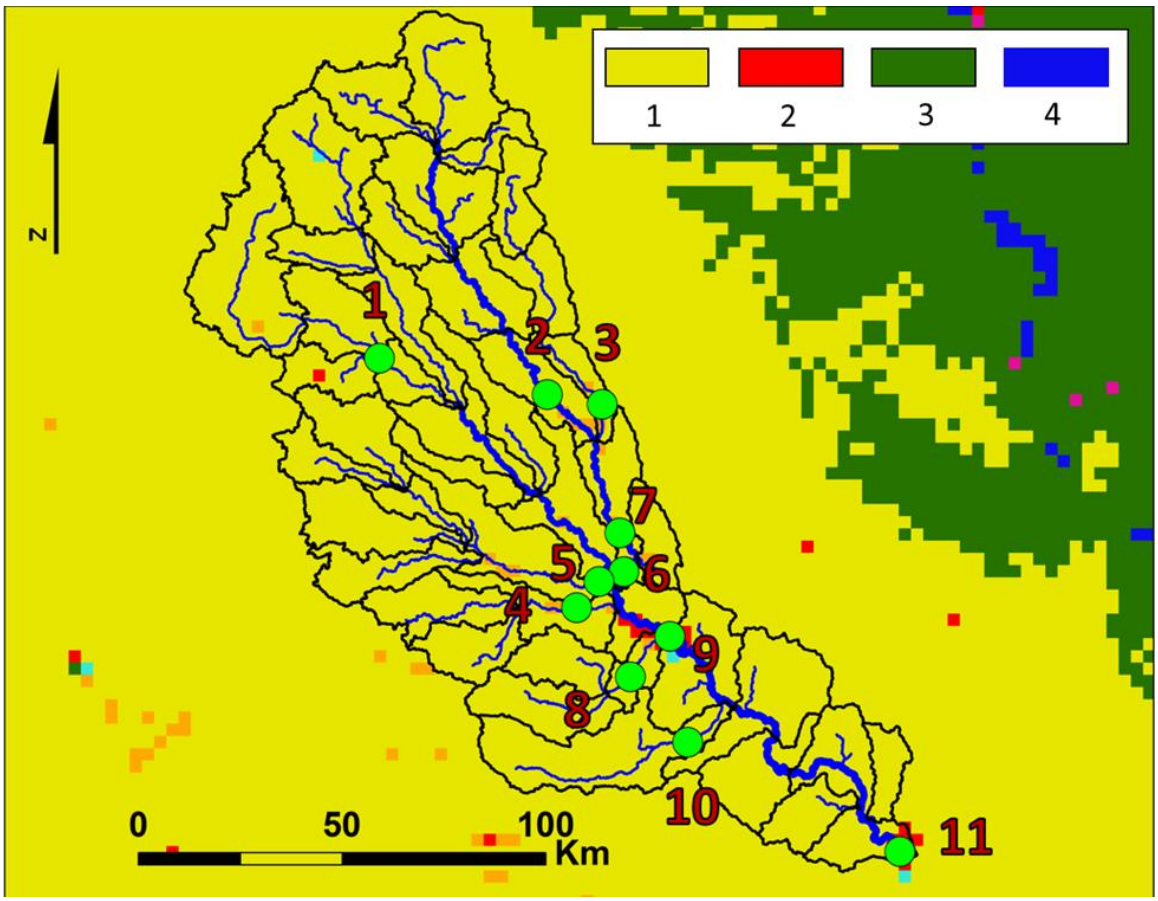

Figure 5.1. USGS Land cover is plotted over the study area. The USGS stream gauges are represented by green dots. The legend of the four major land cover types is shown in the top right corner where 1 is dryland crop, 2 is urban and built-up, 3 is cropland/grassland, and 4 is water body. The labels of the USGS gauges correspond to the rankings in Table 5.1..

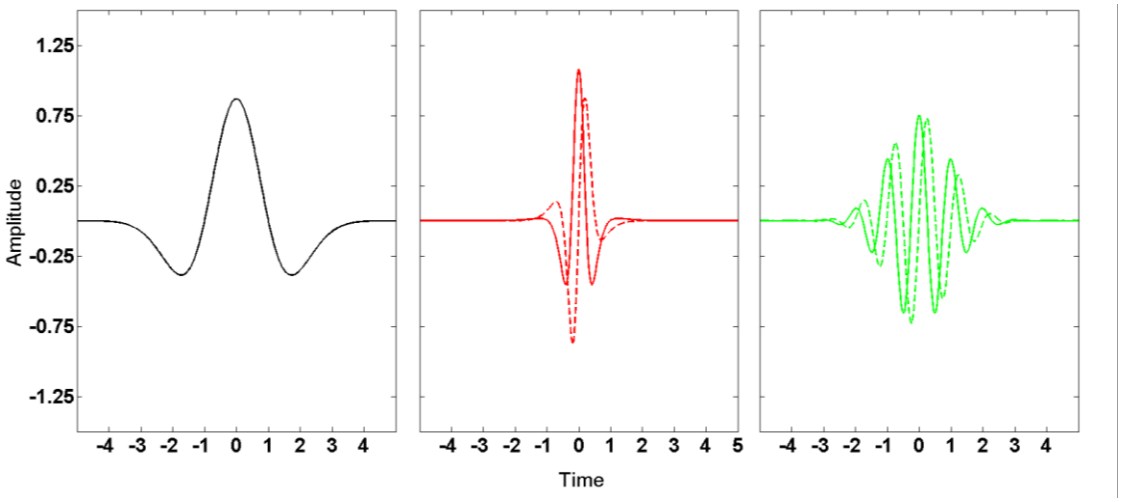

Figure 5.2. Unscaled wavelet function $\boldsymbol{\psi_0}\ (\boldsymbol{\eta})$ for the Mexican hat wavelet (left), Paul wavelet (middle), and Morlet wavelet (right). The solid lines represent the real component of the wavelets, while the dashed line represents the imaginary components (Morlet and Paul only).

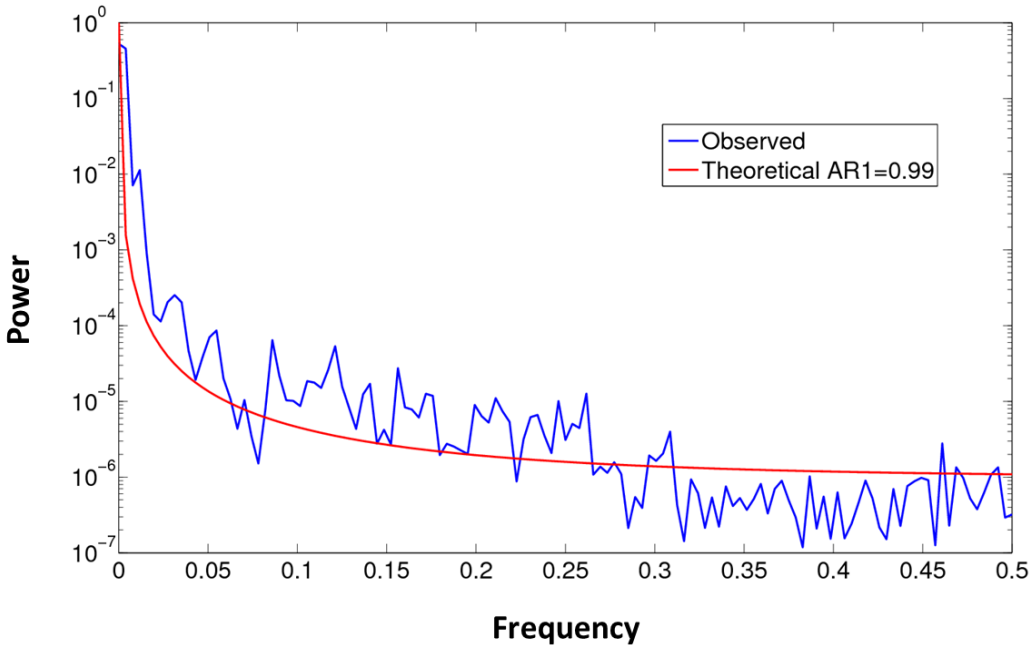

Figure 5.3. Power spectrum decay of a theoretical AR1 process (red) plotted against the power decay obtained from the data. Lag one correlation was estimated using the hourly hydrographs data.

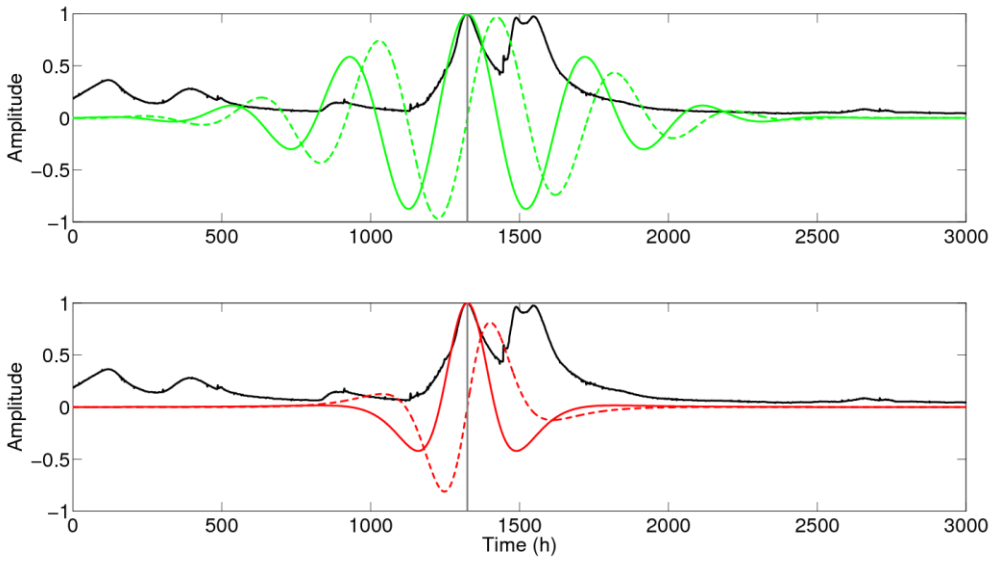

Figure 5.4. Morlet wavelet (top) and Paul wavelet (bottom) with Fourier periods of 400 hours plotted against the observed hydrograph at Cedar River at Cedar Rapids (black line). The vertical grey line shows the location of center of the wavelet, which coincides with a hydrograph peak.

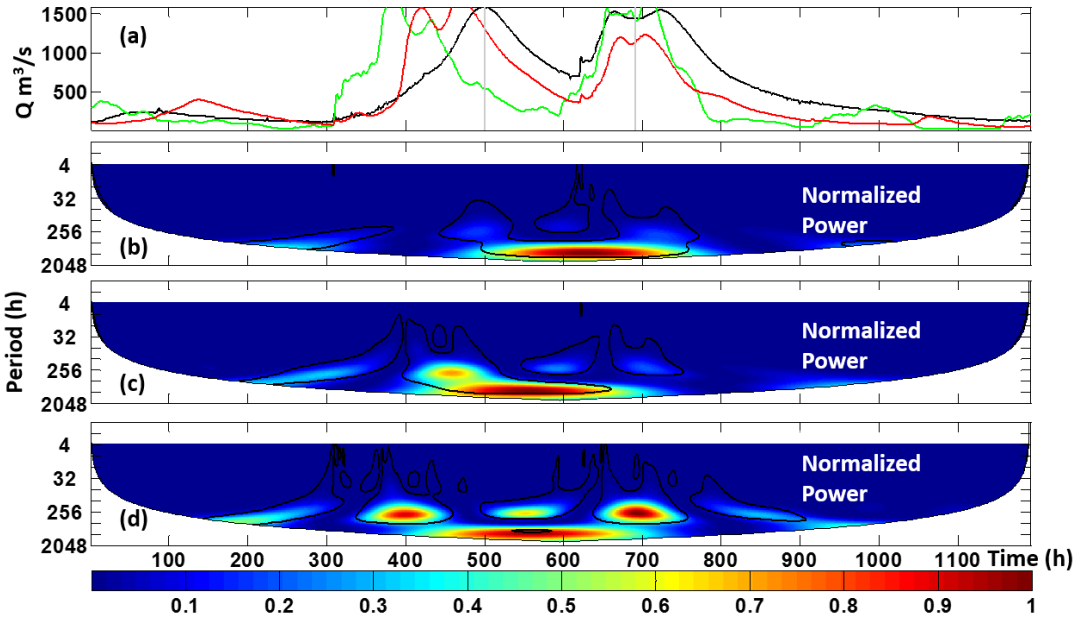

Figure 5.5. CWT power outside the COI for the Mexican hat wavelet, observed (top), HLM routing (middle), and RAPID (bottom). The x axis unites are Time in hour, and the color bar unit is normalized square power.

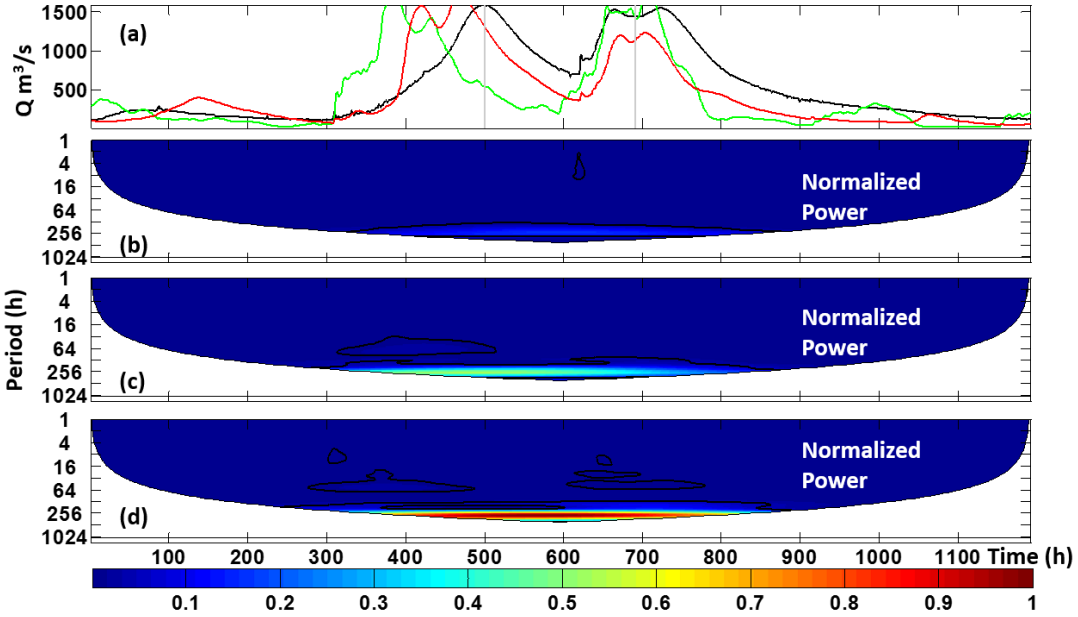

Figure 5.6. CWT power outside the COI for the Morlet wavelet, observed (top), HLM routing (middle), and RAPID (bottom). The x axis unites are Time in hour, and the color bar unit is normalized square power.

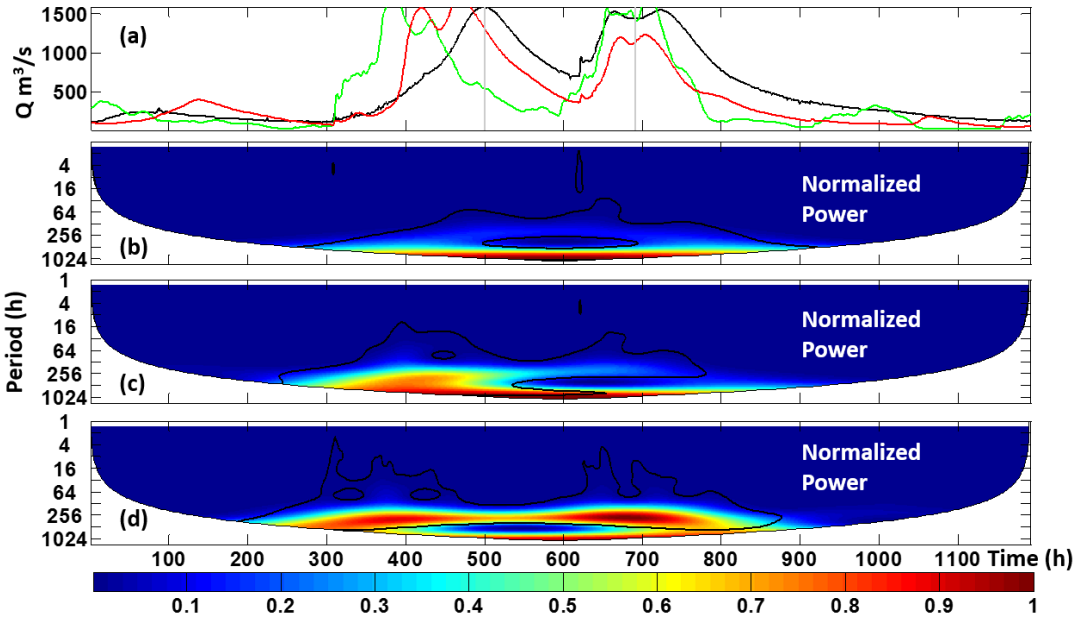

Figure 5.7. CWT power outside the COI for the Paul wavelet, observed (top), HLM routing (middle), and RAPID (bottom). The x axis unites are Time in hour, and the color bar unit is normalized square power.

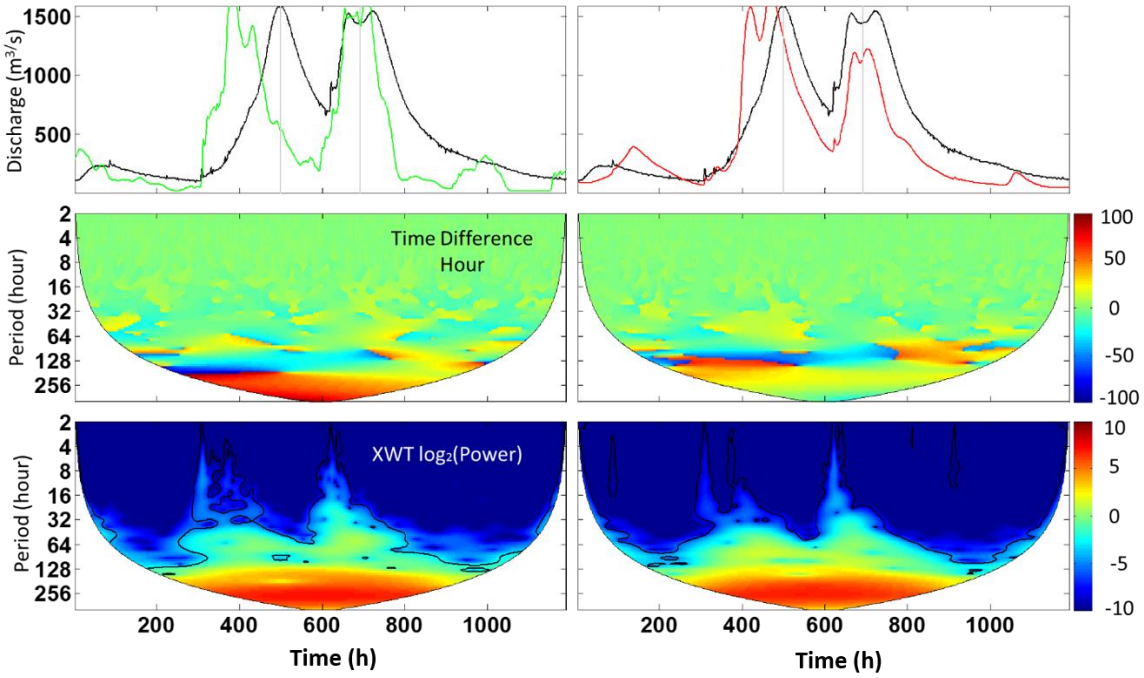

Figure 5.8. Morlet wavelet XWT analysis. The hydrographs (top) with their corresponding estimated time differences in hours (middle), and the log2 of the XWT power (bottom). The left column represents RAPID results and the right column represents the HLM routing results.

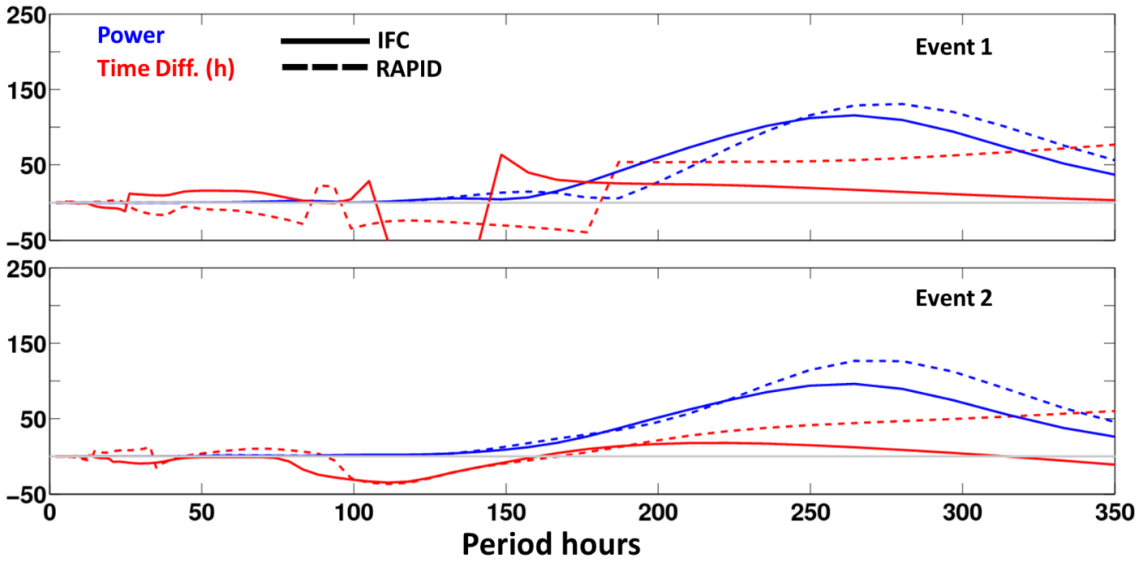

Figure 5.9. Morlet wavelet XWT analysis. Cross-section profile at the event 1 top and event 2 bottom for power (blue) and time difference (red). The solid lines represent the HLM results while the dashed lines represent RAPID results.

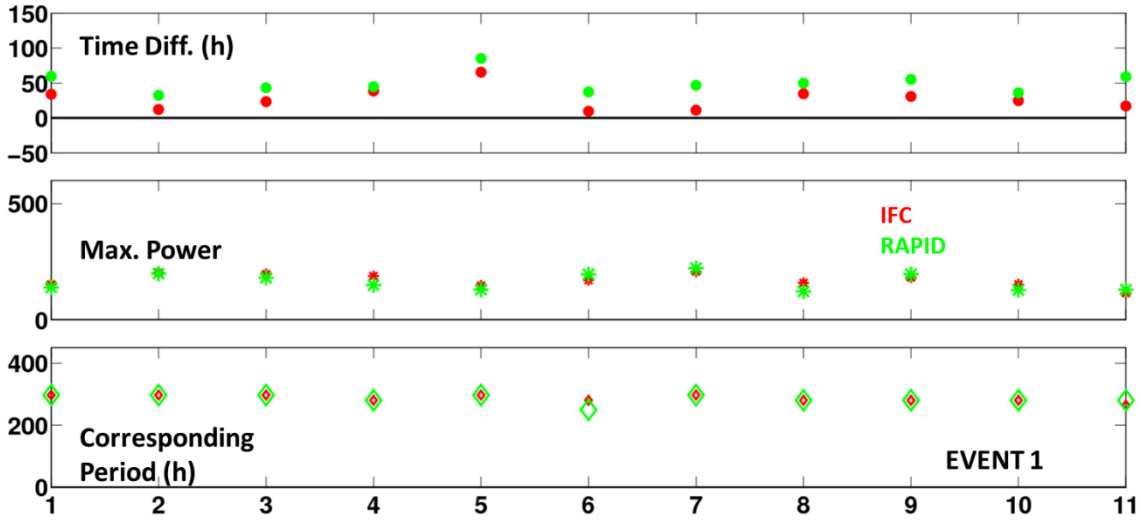

Figure 5.10. Morlet wavelet XWT analysis. Time differences in hours (top), XWT power (middle), corresponding period (bottom) for event 1 at all stations. The x-axis is the station number, and the green color represents RAPID, while the red color represents the HLM routing.

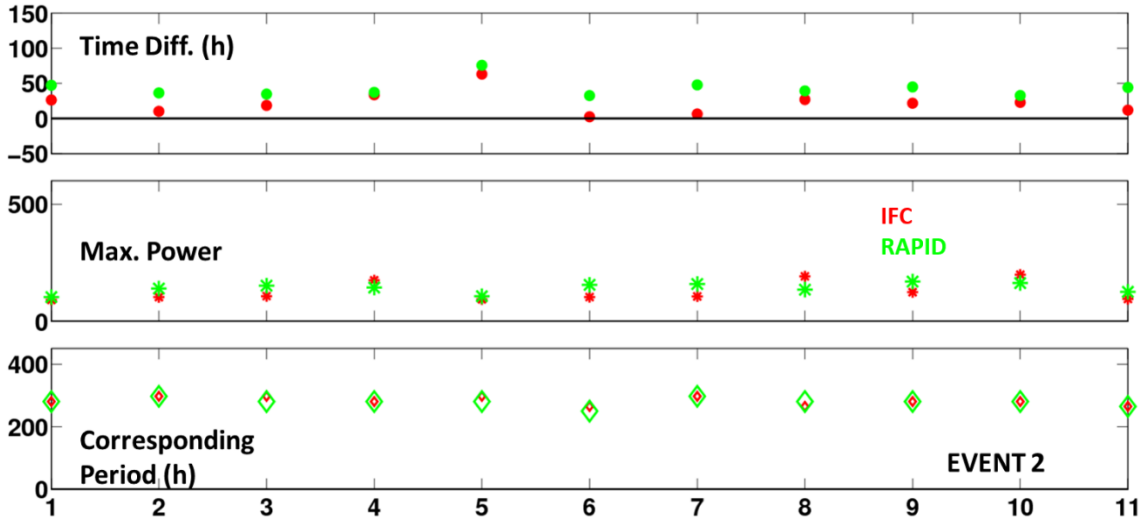

Figure 5.11. Morlet wavelet XWT analysis. Time differences in hours (top), XWT power (middle), corresponding period (bottom) for event 2 at all stations. The x-axis is the station number, and the green color represents RAPID, while the red color represents the HLM routing.

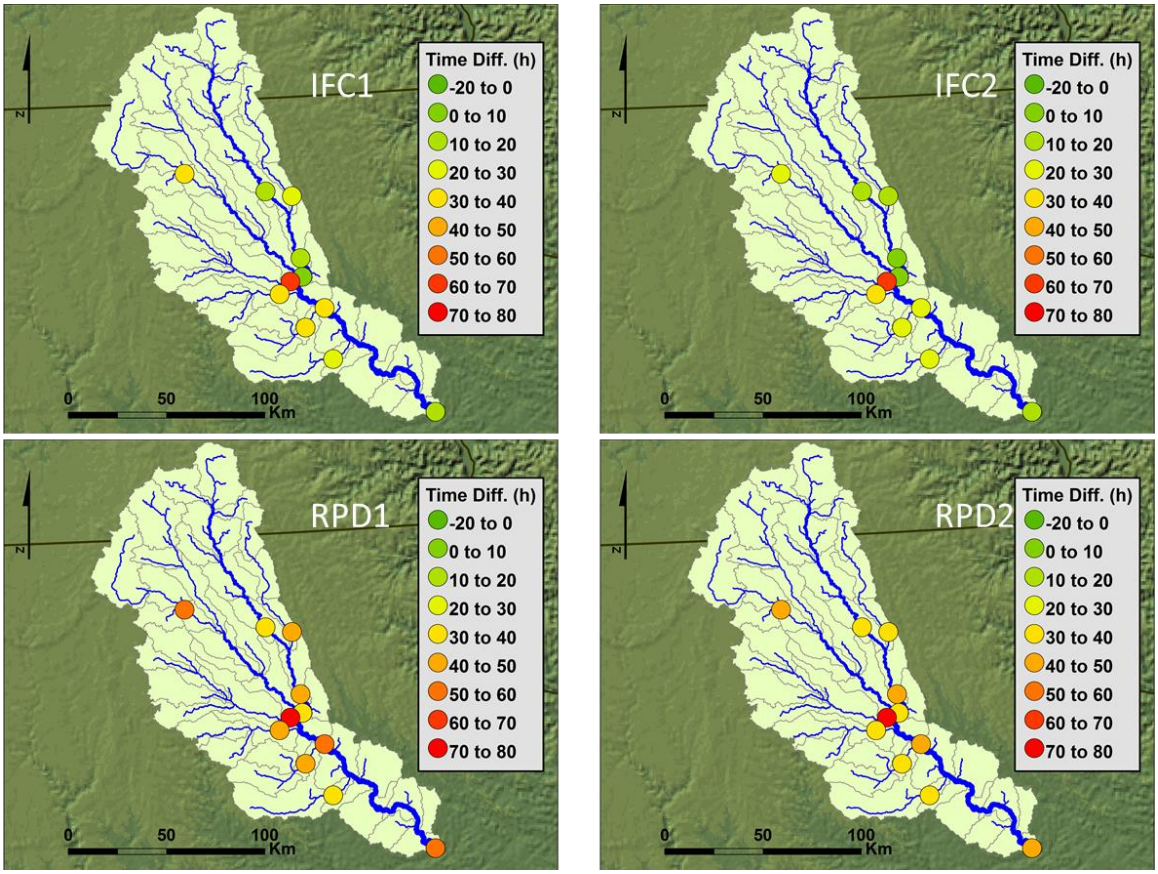

Figure 5.12. Morlet wavelet XWT analysis. Spatial plot of time differences. Top row represents the HLM routing while the bottom row represents RAPID routing. The left column shows the results of the first event while the right column is for the estimates of the second event.

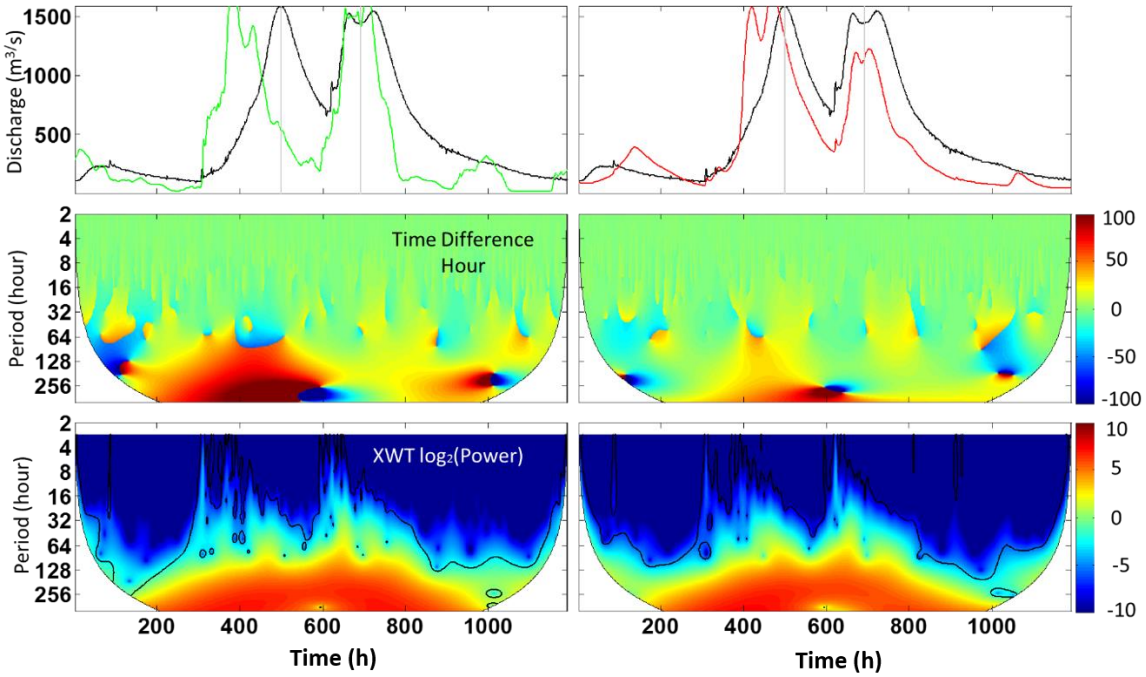

Figure 5.13. Paul wavelet XWT analysis. The hydrographs (top) with their corresponding estimated time differences in hours (middle), and the log2 of the XWT power (bottom). The left column represents RAPID results and the right column represents the HLM routing results.

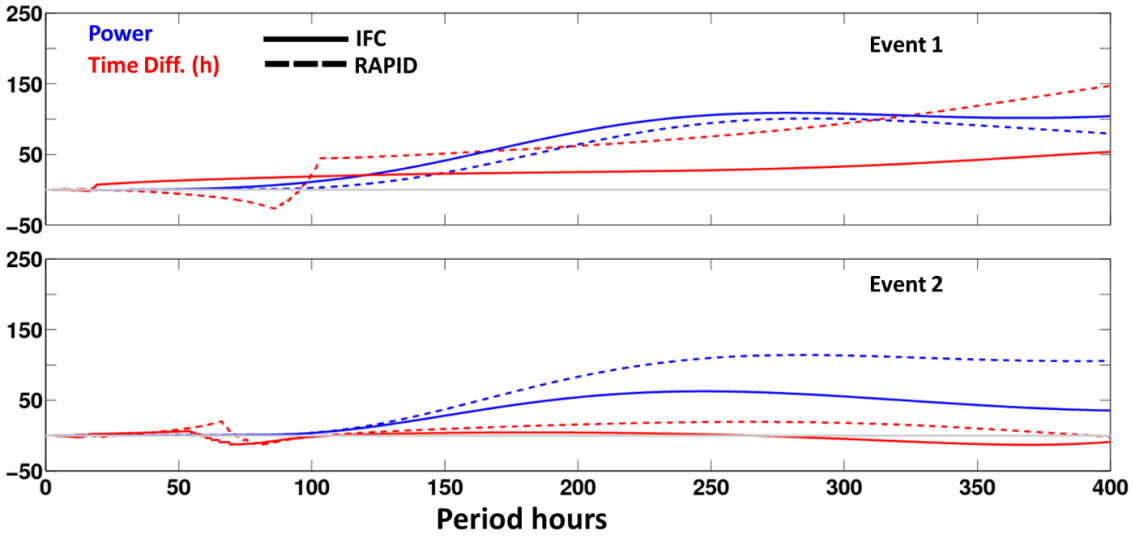

Figure 5.14. Paul wavelet XWT analysis. Cross-section profile at the event 1 top and event 2 bottom for power (blue) and time difference (red). The solid lines represent the HLM results while the dashed lines represent RAPID results.

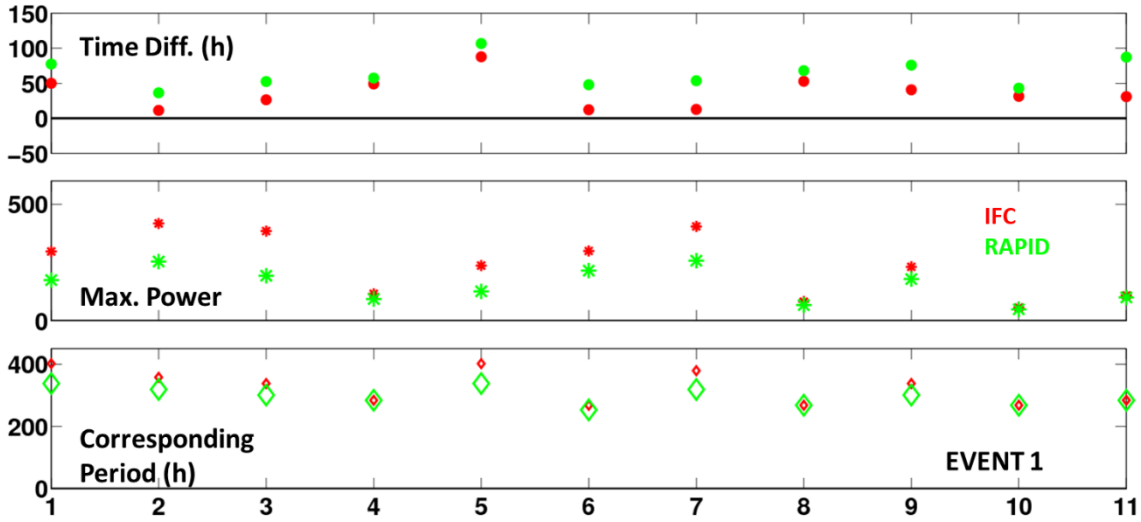

Figure 5.15. Paul wavelet XWT analysis. Time differences in hours (top), XWT power (middle), corresponding period (bottom) for event 1 at all stations. The x-axis is the station number, and the green color represents RAPID, while the red color represents the HLM routing.

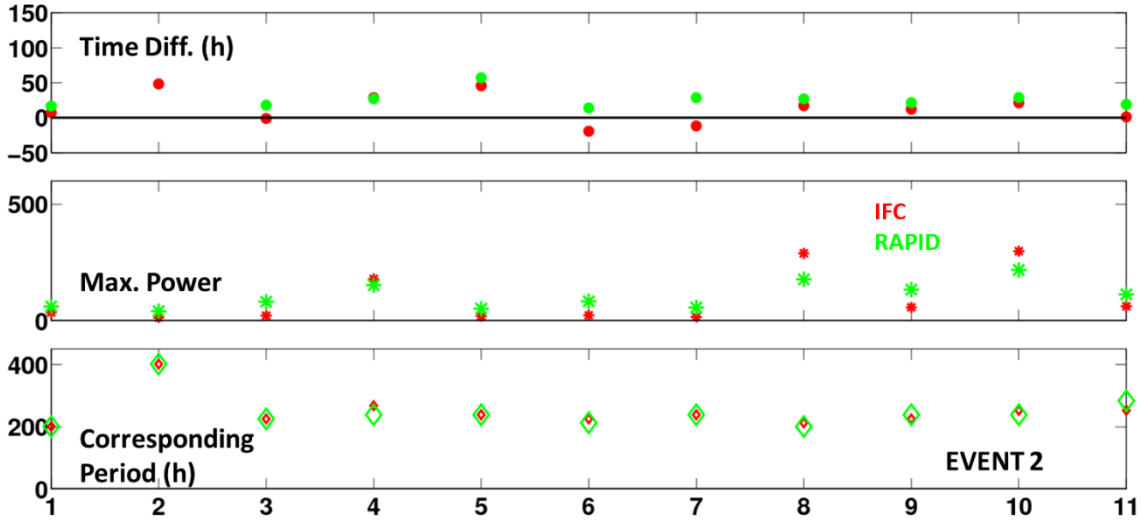

Figure 5.16. Paul wavelet XWT analysis. Time differences in hours (top), XWT power (middle), corresponding period (bottom) for event 2 at all stations. The x-axis is the station number, and the green color represents RAPID, while the red color represents the HLM routing. Paul wavelet XWT analysis. Time differences in hours (top), XWT power (middle), corresponding period (bottom) for event 2 at all stations. The x-axis is the station number, and the green color represents RAPID, while the red color represents the HLM routing.

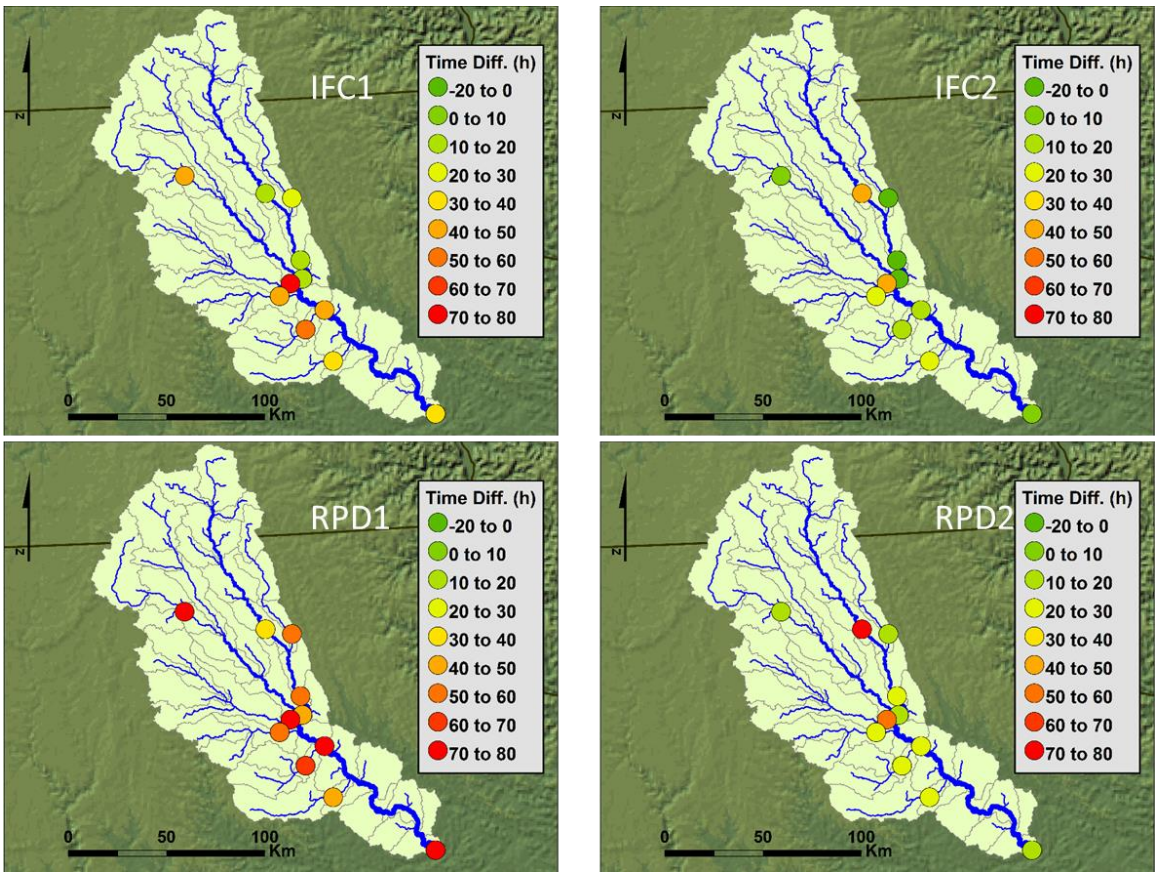

Figure 5.17. Paul wavelet XWT analysis. Spatial plot of time differences. Top row represents the HLM routing while the bottom row represents RAPID routing. The left column shows the results of the first event while the right column is for the estimates of the second event.

CHAPTER 6: Summary, Conclusions, and Future Work.

6.1. Summary

The main focus of this dissertation is to explore the efficiency of using satellite rainfall products in hydrologic modeling. My product of interest is the newly released IMERG rainfall product which is a level 4 research version product. IMERG combines data from multiple space based sensors such as the data collected by PMW sensors on board LEO satellites and IR data collected by the GEO IR satellites. It also incorporates ground based rainfall estimates from rain gauges. IMERG has a higher spatial and temporal resolution compared to its predecessor products collected by the TRMM satellite. It also provides more accurate snow and lite precipitation estimates due to the Ka band incorporated in DPR on board of the GPM core satellite. The evaluation process included both rainfall to rainfall comparisons as well as stream flow comparisons. For rainfall to rainfall comparison we used the ground radar based QPE called MRMS as our benchmark product. We aggregated MRMS to match the resolution of IMERG and compared them using various traditional skill scores and relatively complex. My study area is the entire state of Iowa, and my study period was the entire 2016 war season (June, July, August, and September)

This evaluation brought up further questions regarding conducting an evaluation framework of remotely sensed rainfall product which is hydrology oriented. This is because the hydrologic models significantly distort the original structure of rainfall estimates by assigning the rainfall values hydrologic sub-catchments and then accumulating them downstream in order to estimate discharge. First, the aggregation process propagates the errors downstream thus changing the magnitudes of the overall

errors as they accumulate downstream. Second, understanding the structure of the accumulated errors becomes a challenge since they no longer exist in the Euclidean space but rather in a stream network domain which is dichotomous in nature and affected by the connectivity information between the streams. Thus, we decided to use a network based spatial statistical framework that used stream distances rather than Euclidean distances, and is also capable of incorporating the connectivity between the streams.

Before proceeding to the evaluation of IMERG as a rainfall input. It was necessary to investigate the performance of the hydrologic models we are going to use in our study. I investigated the performance of two possible hydrologic modeling frameworks. First, the HLM model produced by the IFC. Second, a combination of a LSM with an independent routing component. I started by evaluating two possible routing methodologies, one linear and the other is non-linear. The evaluation revealed that the routing methodology can have a significant impact on the quality of the simulated stream flows. I decided to continue my studies using the non-linear routing option due to its better performance. Thus, the two final models I used for IMERG evaluation are the Noah-MP LSM with HLM routing and the full HLM hydrologic model.

I conducted the stream flow comparisons at 130 USGS locations within the state of Iowa. The gauge selection was based on record availability and to insure minimal effect of artificial storage. The evaluation period was the same as what I used for rainfall to rainfall comparison. For this purpose I used the widely used statistical skill scores for each USGS gauge location. However, again we were faced with some difficulties regarding finding a proper skill score to evaluate the performance of the hydrologic

models in estimating flood peak times. Several simplistic methods exist for this purpose, but they are either not accurate or require manual efforts by the evaluator.

I decided to use signal processing techniques to evaluate the simulated stream discharge. Signal processing techniques are fully automated and they insure that we are conducting a fair comparison by detecting the simulated peak that directly correspond to the observed peak. In addition, the difference in peak time between the simulated and observed peak flows is not only dependent on the location of the maximum value in the simulated peak but rather the overall shape and location of the entire peak "hump". The signal processing method I used is called Continuous Wavelet Transform (CWT), and provides the possibility of using different wavelet shapes as signal filters. Thus, I emphasized on the effect of the wavelet shape on the inference regarding the peak lag time.

6.2. Conclusions

In chapters 2 through 5 I conducted four independent studies that are focused on the evaluation of satellite rainfall in hydrologic applications. Each chapter is focused on one aspect of the evaluation including, hydrologic model performance and selection, IMERG comparison with ground based estimates and its performance as a model input, evaluation of IMERG in a stream network based framework, and finally a time-based evaluation method of hydrologic models' outputs. The overall conclusion of each chapter are as follows:

Chapter 2: In this chapter I investigated the effect of the routing component on the overall performance of the hydrologic models I am going to use in my studies. I conducted a detailed comparison between RAPID —a routing component based on the

simplified linear routing method called the Muskingum method— and the nonlinear routing method used in the HLM hydrologic model. The comparisons showed the using the same runoff input, HLM routing outperformed RAPID in all scales. In addition, another advantage of HLM routing is that it does not require calibration based on the rainfall input which is a capability that RAPID lacks. This is because the Muskingum method depends on 2 parameters k and x which need to be continuously calibrated. Thus, we decided to carry on the rest of our studies using the HLM routing.

Chapter 3: In this chapter I conducted the rainfall to rainfall comparisons as well as the stream flow comparisons using IMERG and the benchmark product MRMS as rainfall inputs to our hydrologic models. The hydrologic models I used in this study are the full HLM model and the Noah-MP with HLM routing. Both models incorporate the HLM routing component, however, they have different runoff generation methodologies. The rainfall to rainfall comparisons showed that IMERG overestimate the number of instances during which lite rainfall occurred and significantly underestimated the number of heavy rain instances. In addition, the space time dependence between IMERG rainfall estimates is significantly different from the MRMS. The correlation between IMERG rainfall estimates extends for ranges that are approximately twice as large as those obtained from MRMS rainfall. In addition, the dependence it time extends for about 3 time lags (in hours) which is not similar to MRMS where the dependence in time collapses almost immediately after the first time lag.

As rainfall inputs, both IMERG and MRMS performance varied based on both basin scale and geographical location. However, MRMS usually outperformed IMERG

by producing better discharge estimate. In addition, HLM model significantly outperformed the Noah-MP with HLM routing.

Chapter 4: In this chapter I proposed a new method to evaluate remotely sensed rainfall products in hydrologic context. In particular, I investigated the spatial dependence between the satellite rainfall errors in a network based domain. The investigation of the accumulated errors showed that the magnitude of the errors significantly decrease as they accumulate downstream. However, it was necessary to investigate the dependence of the accumulated errors. Using the Spatial Stream Network (SSN) method, I showed that the accumulated errors are strongly affected by the network structure and connectivity information. The correlation between the errors in the flow-connected streams extended for ranges that are twice as large as those in the flow-unconnected streams.

Chapter 5: In this chapter I proposed a time-based evaluation of the simulated stream flow hydrographs using CWT and XWT. I applied this evaluation method for the stream flow estimates produced by HLM routing and RAPID using the same runoff input. The results agreed with the evaluations I performed in Chapter 2. In addition, I showed that the shape of the wavelet used in the evaluation had an impact on the inference regarding model performance. I showed that although the Marr (Mexican Hat) wavelet directly correspond to the shape of hydrograph peaks, it is necessary to use a wavelet with imaginary component in order to be able to calculate the XWT. This shapes of these wavelet does not perfectly match the shape of hydrograph peaks, however, each of these wavelet shapes has its own advantages and disadvantages. For our application, I showed the Paul wavelet can be more useful in detecting the localized sharp peaks in the hydrograph. However, if the hydrograph has many consecutive peaks (high frequency) it

would be better to use the Morlet wavelet. I showed that using the Morlet wavelet can have negative impact on our inference if only one or two peaks exist in the hydrograph since the wavelet will use information from both peaks simultaneously due to the wide range it covers.

6.3. Future work

In this dissertation I am evaluating IMERG which was recently released by NASA, the record of this product starts in March 2014. In addition, the record of our benchmark product MRMS start in September, 2015. This small overlap limited our number of sample to few months. However, since we are using hourly rainfall rates we had a large enough sample to conduct our studies. Hopefully, in the future, the benchmark record will reveal the inter-annual variability of the IMERG performance when compared to MRMS. In addition, one can use other benchmark products. However, we chose MRMS as a benchmark since it is the product currently recognized and operationally used by the NWS. Another possible extension to this research would include applying the network based evaluation (chapter 4) to other hydrologic model state variables such as discharges, soil moisture, etc. In addition, in chapter 5 I showed the efficiency of XWT in matching and comparing peaks. The next step is to connect the XWT power estimates and timing differences to the traditional skill scores (e.g., correlation coefficient and NSE)

REFERENCES

AghaKouchak, A., A. Mehran, H. Norouzi, and A. Behrangi. 2012. "Systematic and random error components in satellite precipitation data sets." *Geophysical Research Letters*, 39: L09406, doi:10.1029/2012GL051592.

Akaike, H. 1973. "Information Theory and an Extension of the Maximum Likelihood Principle." *Second International Symposium on Information Theory*, eds. B. N. Petrov, and F. Caski, Budapest: Akademiai Kiado, 267 – 281.

Ashouri, H., K. Hsu, S. Sorooshian, D.K. Braithwaite, K.R. Knapp, L.D. Cecil, B.R. Nelson, and O.P. Prat, 2015. "PERSIANN-CDR: Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies." *Bulletin of the American Meteorological Society (BAMS)*, 96:69–83, doi:10.1175/BAMS-D-13-00068.1.

Ayalew, T. B., W. F. Krajewski, R. Mantilla, and S. J. Small, 2014. "Exploring the effects of hillslope-channel link dynamics and excess rainfall properties on the scaling structure of peak-discharge." *Advances in Water Resources*, 64:9 – 20, doi: 10.1016/j.advwatres.2013.11.010.

Chow, V. T., D. R. Maidment, and L. W. Mays, 1988. "Applied Hydrology", *McGraw-Hill, New York*.

Cunge, J. A., 1969. "On the subject of a flood propagation computation method (Muskingum method)". *Journal of Hydraulic Research*, 7(2): 205 – 230.

Cunha, L. K., J. A. Smith, W. F. Krajewski, M. L. Baeck, and B. Seo. 2015. "NEXRAD NWS Polarimetric Precipitation Product Evaluation for IFloodS." *Journal of Hydrometeorology*, 16: 1676 – 1699, doi: 10.1175/JHM-D-14-0148.1.

David, C. H., D. R. Maidment, G. -Y. Niu, Z.- L. Yang, F. Habets, and V. Eijkhout, 2011. "River Network Routing on the NHDPlus Dataset." *Journal of Hydrometeorology*, 12:913 – 934, doi: 10.1175/2011JHM1345.1.

David, C. H., Z. L. Yang, and J. S. Famiglietti, 2013. "Quantification of the upstream-to-downstream influence in the Muskingum method and implications for speedup in parallel computations of river flow." *Water Resources Research*, 49:2783 – 2800, doi: 10.1002/wrcr.20250.

David, C.H., J.S. Famiglietti, Z.-L. Yang, and V. Eijkhout, 2015. "Enhanced fixed-size parallel speedup with the Muskingum method using a trans-boundary approach and a large sub-basins approximation." *Water Resources Research*, 51(9):1 – 25, doi: 10.1002/2014WR016650.

David, C.H., J.S. Famiglietti, Z.-L. Yang, F. Habets, and D.R. Maidment, 2016. "A Dcade of RAPID – Reflections on the development of an Open Source geoscience code." *Earth and Space Science*, 3:1 – 19, doi: 10.1002/2015EA000142.

Eash, D.A., K.K. Barnes, and A.G. Veilleux, 2013. "Methods for estimating annual exceedance-probability discharges for streams in Iowa, based on data through water year 2010: U.S. Geological Survey Scientific Investigations Report 2013-5086, 63 p. with appendix, http://pubs.usgs.gov/sir/2013/5086/.

ElSaadani, M., and F.W. Krajewski, 2017. A Time-based Framework for Evaluating Hydrologic Routing Methodologies Using Wavelet Transform. *Journal of Water Resource and Protection* , 9:723 – 744, doi:10.4236/jwarp.2017.97048.

Foufoula-Georgiou, E., P. Kumar (Eds.), 1995. "Wavelets in Geophysics." *Academic Press, New York*, p. 337.

Ghimire, G.R., W.F. Krajewski, and R. Mantilla, 2017. "A power law model for river water velocity in U.S. Upper Midwestern basins." Unpublished.

Gochis, D. J., W. Yu, , and D. N. Yates, 2015. "The WRF-Hydro model technical description and user's guide, version 3.0." *NCAR Technical Document.* http://www.ral.ucar.edu/projects/WRF-Hydro/

Gourley, J. J., Y. Hong, Z. L. Flamig, J. Wang, H. Vergara, and E. N. Anagnostou. 2011. "Hydrologic Evaluation of Rainfall Estimates from Radar, Satellite, Gauge, and Combinations on Ft. Cobb Basin, Oklahoma." *Journal of Hydrometeorology*, 12: 973 – 988, doi: 10.1175/2011JHM1287.1.

Habib, E., A. T. Haile, N. Sazib, Y. Zhang, T. Rientjes. 2014. "Effect of Bias Correction of Satellite-Rainfall Estimates on Runoff Simulations at the Source of the Upper Blue Nile" *Remote Sensing*, 6: 6688 – 6708.

Habib, E., M. ElSaadani, A. T. Haile, 2012. "Climatology-focused evaluation of CMORPH and TMPA satellite rainfall products over the Nile Basin." *Journal of Applied Meteorology and Climatology*, 51:2105–2121, doi: 10.1175/JAMC-D-11-0252.1.

Hou, A. Y., R. K. Kakar, S, Neeck, A. A. Azarbarzin, C. D. Kummerow, M. Kojima, R. Oki, K. Nakamura, and T. Iguchi. 2014. "The Global Precipitation Measurement Mission." *Bulletin of the American Meteorological Society (BAMS)*, 95: 701 – 722, doi:10.1175/BAMS-D-13-00164.1.

Huffman, G. J., D. T. Bolvin, D. Braithwaite, K. Hsu, R. Joyce, C. Kidd, E. J. Nelkin, and P. Xie, 2015. "NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG). Algorithm Theoretical Basis" Doc., version 4.5, 26 pp. [Available online at http://pmm.nasa.gov/sites/ default/files/document_files/IMERG_ATBD_V4.5.pdf.]

Isaak D. J., E. E. Peterson, J. M. Ver Hoef, S. J. Wenger, J. A. Falke, C. E. Torgersen, C. Sowder, et al. 2014. "Applications of spatial statistical network models to stream data." *Wiley Interdisciplinary Reviews: Water*, 1: 277–294.

Kim, D.H., A.P. Georgakakos, 2014. "Hydrologic routing using nonlinear cascaded reservoirs." *Water Resources Research*, 50 (8):7000-7019, doi: 10.1002/2014wr015662.

Krajewski, W.F., D. Ceynar, I. Demir, R. Goska, A. Kruger, C. Langel, R. Mantilla, J. Niemeier, F. Quintero, B. Seo, S. Small, L. Weber, and N. Young, 2016. "Real-Time Flood Forecasting and Information System for the State of Iowa." *Bulletin of the American Meteorological Society*, 98:539 – 554, doi: 10.1175/BAMS-D-15-00243.1.

Laaha, G., J. O. Skøien, and G. Blöschl. 2014. "Spatial prediction on river networks: comparison of top-kriging with regional regression." *Hydrological Processes*, 28: 315–324. doi:10.1002/hyp.9578.

Labat, D., 2005. "Recent advances in wavelet analyses: Part 1. A review of concepts." *Journal of Hydrology*, 314:275 – 288. doi:10.1016/j.jhydrol.2005.04.003

Lin, Y., and K. E. Mitchell, 2005. "The NCEP Stage II/IV hourly precipitation analyses: Development and applications." *Preprints, 19th Conference on Hydrology, San Diego, CA, American Meteorological Society*, 1.2, http://ams.confex.com/ams/Annual2005/techprogram/paper_83847.htm

Liu, Y., J. D.Brown, J. Demargne, and D.-J. Seo, 2011. "A wavelet-based approach to assessing timing errors in hydrologic predictions." *Journal of Hydrology*, 397 (3-4):210 – 224. doi:10.1016/j.jhydrol.2010.11.040

Liu, Z. 2016. "Comparison of Integrated Multisatellite Retrievals for GPM (IMERG) and TRMM Multisatellite Precipitation Analysis (TMPA) Monthly Precipitation Products: Initial Results." *Journal of Hydrometeorology*, 17: 777–790, doi: 10.1175/JHM-D-15-0068.1.

Liu, Z. 2016. "Comparison of Integrated Multisatellite Retrievals for GPM (IMERG) and TRMM Multisatellite Precipitation Analysis (TMPA) Monthly Precipitation Products: Initial Results." *Journal of Hydrometeorology*, 17: 777–790, doi: 10.1175/JHM-D-15-0068.1.

Maidment, D.R., 2016. "Conceptual Framework for the National Flood Interoperability Experiment." *Journal of the American Water Resources Association (JAWRA)*, 53(2):245-257. DOI: 10.1111/1752-1688.12474.

Mallakpour, I., 2016. "The tale of flooding over the central United States." *Thesis, University of Iowa*, http://ir.uiowa.edu/etd/2115

Mandapaka,P. V., W. F. Krajewski, G. J. Ciach, G. Villarini, J. A. Smith, 2009. "Estimation of radar-rainfall error spatial correlation." *Advances in Water Resources*, 32(7): 1020–1030. Doi:10.1016/j.advwatres.2008.08.014.

Mantilla, R., 2007. "Physical Basis of Statistical Scaling in Peak Flows and Stream Flow Hydrographs for Topologic and Spatially Embedded Random Self-similar Channel Networks." *Thesis, University of Colorado.*

Mantilla, R., and V. K. Gupta, 2005. "A GIS numerical framework to study the process basis of scaling statistics in river networks." *IEEE Geoscience and Remote Sensing Letters*, 2(4):404 – 408, doi: 10.1109/LGRS.2005.853571.

Mitchell, K. E. *et al.,* 2004. "The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system." *Journal of Geophysical Research*, 109:D07S90.

Moriasi, D.N., J.G. Arnold, M.W. Van Liew, R.L. Bingner, R.D. Harmel, T.L. Veith, 2007. "Model evaluation guidelines for systematic quantification of accuracy in watershed simulations." *Transactions of the ASABE*, 50 (3):885 – 900

Mutel, C., 2010. "A Watershed Year: Anatomy of the Iowa Floods of 2008." *University of Iowa Press*, 284 pp.

National Climate Data Center (NCDC) 2015. "Billion-Dollar Weather and Climate Disasters." https://www.ncdc.noaa.gov/billions/events, Last accessed, June, 2017.

Niu, G.-Y., et al. 2011. "The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements." *Journal of Geophysical Research*, 116:D12109, doi: 10.1029/2010JD015139.

Paik, K., P. Kumar, 2004. "Hydraulic geometry and the nonlinearity of the network instantaneous response." *Water Resources Research* 40: W03602.

Peterson, E. E. and Hoef, J. M. V. (2010), A mixed-model moving-average approach to geostatistical modeling in stream networks. *Ecology*, 91: 644–651. doi:10.1890/08-1668.1.

Quintero, F., W. F. Krajewski, R. Mantilla, S. Small, and B. Seo, 2016. "A Spatial–Dynamical Framework for Evaluation of Satellite Rainfall Products for Flood Prediction." *Journal of Hydrometeorology*, 17: 2137–2154, doi: 10.1175/JHM-D-15-0195.1.

Reed, S. M. and D. R. Maidment, 1999. "Coordinate transformations for using NEXRAD data in GIS-based hydrologic modeling." *Journal of Hydrologic Engineering*, 4:174 – 182.

Rodrıguez-Iturbe I., and A. Rinaldo. 1997. "Fractal River Basins: Chance and Self-Organization", *Cambridge University Press*, Cambridge, UK.

Small S. J., L. O. Jay, R. Mantilla, R. Curtu, L. K. Cunha, M. Fonley, and W. F. Krajewski, 2013. "An asynchronous solver for systems of ODEs linked by a directed tree structure." *Advances in Water Resources*, 53:23–32, doi:10.1016/j.advwatres.2012.10.011.

Snow, A. D., 2015. "A New Global Forecasting Model to Produce High-Resolution Stream Forecasts." *Thesis*, BYU.

Snow, A.D., D. C. Scott, N. R. Swain, J. Nelson, D. P. Ames, N. L. Jones, D. Ding, N. Noman, C. H. David, F. Pappenberger, 2016. "A Cloud-Based High-Resolution National Hydrologic Forecast System Downscaled from a Global Ensemble Land Surface Model." *Journal of the American Water Resources Association (JAWRA)*, 52(4):950 – 964. doi:10.1111/1752-1688.12434

Tavakoly, A.A., 2017. "RAPID input files corresponding to the Mississippi River Basin using the NHDPlus v2 Dataset" [Data set]. Zenodo. doi: 10.5281/zenodo.322886

Tavakoly, A.A., A. D. Snow, C.H. David, M.L. Follum, D.R. Maidment, and Z.-L. Yang, 2016. "Continental-Scale River Flow Modeling of the Mississippi River Basin Using High-Resolution NHDPlus Dataset." *Journal of the American Water Resources Association (JAWRA)* 53(2):258 – 279, doi: 10.1111/1752-1688.12456.

Tavakoly, A.A., A. D. Snow, C.H. David, M.L. Follum, D.R. Maidment, and Z.-L. Yang, 2016. "Continental-Scale River Flow Modeling of the Mississippi River Basin Using

High-Resolution NHDPlus Dataset." *Journal of the American Water Resources Association (JAWRA)* 53(2):258 – 279, doi: 10.1111/1752-1688.12456.

Torrence, C., G. P. Compo, 1998. "A practical guide to wavelet analysis." *Bulletin of the American Meteorological Society*, 79 (1):61 – 78.

U.S. Geological Survey, 2015. "USGS NED 1 arc-second n42w091 1 x 1 degree ArcGrid 2015" *U.S. Geological Survey: Reston, VA*, http://ned.usgs.gov/, http://nationalmap.gov/viewer.html

Ver Hoef, J. M., E. Peterson, and D. Theobald. 2006. "Spatial statistical models that use flow and stream distance" *Environmenal and Ecological Statistics*, 13: 449. doi:10.1007/s10651-006-0022-8.

Ver Hoef, J. M., and E. E. Peterson, 2010. "A Moving Average Approach for Spatial Statistical Models of Stream Networks," *Journal of the American Statistical Association*, 105: 6–18.

Vergara, H., P.-E. Kirstetter, J.J. Gourley, Z.L. Flamig, Y. Hong, A. Arthur, and R. Kolar, 2016. "Estimating a-priori kinematic wave model parameters based on regionalization for flash flood forecasting in the Conterminous United States." *Journal of Hydrology*, 541 (Part A):421 – 433, doi: 10.1016/j.jhydrol.2016.06.011.

Vergara, H., Y. Hong, J. J. Gourley, E. N. Anagnostou, V. Maggioni, D. Stampoulis, and P. Kirstetter. 2014. "Effects of Resolution of Satellite-Based Rainfall Estimates on Hydrologic Modeling Skill at Different Scales." *Journal of Hydrometeorology*, 15: 593–613. doi: 10.1175/JHM-D-12-0113.1.

Villarini, G., and W. F. Krajewski. 2007. "Evaluation of the research-version TMPA three-hourly 0.25 ° 0.25 ° rainfall estimates over Oklahoma", *Geophysical Research Letters*, 34: L05402, doi:10.1029/2006GL029147.

Woody, J., R. Lund, and M. Gebremichael, 2014. "Tuning Extreme NEXRAD and CMORPH Precipitation Estimates." *Journal of Hydrometeorology*, 15. 1070–1077, doi:10.1175/JHM-D-13-0146.1.

Wu, H., R.F. Adler, Y. Tian, G. J.Huffman, H. Li, and J. Wang, 2014. "Real-time global flood estimationusing satellite-based precipitation anda coupled land surface and routingmodel." *Water Resources Research*, 50:2693 – 2717, doi:10.1002/2013WR014710.

Xia, Y., *et al.*, 2012a. "Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products." *Journal of Geophysical Research*,117:D03109, doi:10.1029/2011JD016048.

Xia, Y., et al. 2012b. "Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 2. Validation of model-simulated streamflow." *Journal of Geophysical Research*, 117:D03110, doi:10.1029/2011JD016051.

Zhang, Y., D-J. Seo, D. Kitzmiller, H. Lee, R. J. Kuligowski, D. Kim, and C. R. Kondragunta. 2013. "Comparative strengths of SCaMPR satellite QPEs with and without TRMM ingest vs. gridded gauge-only analyses." *Journal of Hydrometeorology*, 14: 153–170.

Zimmerman D., and J. M. Ver Hoef. 2017. "The Torgegram for fluvial variography: characterizing spatial dependence on stream networks." *Journal of Computational and Graphical Statistics*, 26(2): 253–264.