Title
Development of cloud-free MODIS datasets for hydrologic applications

Author
Tran, Hoang Viet

Publication Date
2018

UNIVERSITY OF CALIFORNIA
IRVINE

Development of cloud-free MODIS datasets for hydrologic applications

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Civil and Environmental Engineering

by

Hoang Tran

Dissertation Committee:
Professor Soroosh Sorooshian, Chair
Professor Kuo-Lin Hsu
Professor Phu Nguyen

2018

Table of Contents

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First and foremost, I owe an enormous debt of gratitude to my advisor Professor Soroosh Sorooshian for his tireless support throughout my graduate education in UCI. Soroosh is the best mentor a research student could possibly hope for. He is of course an amazing academic going by his first-rate scholarly abilities, unparalleled mastery of his specialty, and profound insights on hydrology and remote sensing. But above and beyond all these, Soroosh is a most wonderful gentleman with great human qualities—he is modest, compassionate, understanding, accommodating, and possesses a witty sense of humor. I am very fortunate, very proud, and very honored to be Soroosh's student. I am grateful to my co-advisor Professor Kuo-Lin Hsu for the valuable comments and generous time he has offered to my research. Many thanks for his continuous support, patience, and scientific guidance. Last but not least, I would also like to extend my sincere gratitude to my co-advisor Professor Phu Nguyen, for his immense support during my Ph.D. research. Phu's knowledge and insight different scientific areas, specifically hydrological modeling, is exemplary. This dissertation certainly would not have been completed without their meticulous and thorough revision.

I thank my committee member: Prof. Francois Primeau. His helpful suggestions have improved this dissertation enormously.

I am indebted to the former and current friends and colleagues at CHRS that made my UCI years so much more enjoyable. I appreciate Mr. Dan Braithwaite for providing data and IT support to my research, Ms. Diane Hohnbaum for her administrative assistance and attention.

I have been infinitely lucky to have my lovely wife, Huong and my son, Brian. Without their encouragement, I would not have started this long journey, let alone finished.

Last but not least, I thank my parents and grandma for years of toiling and putting up with my “life-long” studies. I am equally indebted to my sibling Son and brother-in law Tuan for their love and constant support.

This research was partially supported by Cooperative Institute for Climate and Satellites (CICS) program (NOAA prime award #NA14NES4320003, National Science Foundation (NSF award # 1331915) and Department of Energy (DoE prime award # DE-IA0000018) and California Energy Commission (CEC Award # 300-15-005).

CURRICULUM VITAE

Hoang Tran

Department of Civil and Environmental Engineering
University of California, Irvine
Irvine, CA 92697
`hoang.tran.1899@gmail.com`

RESEARCH INTERESTS

- Hydrometeorology and Hydrology
- Hydrologic Remote Sensing
- Hydrological Rain-fall Runoff Modeling
- High Performance Computing
- Geo-spatial Analysis
- Statistics and Statistical Modeling
- Model Calibration
- Data Assimilation

TECHNICAL SKILLS

Hydrologic and Hydraulic modeling (6 years), Remote Sensing (6 years), ArcGIS/QGIS (6 years), Computer programming (5 years), SQL (4 years), ENVI (3 years), AutoCAD (8 years).

EDUCATION

University of California, Irvine - Irvine, CA, USA December 2018
Ph.D. in Civil Engineering, focus on Hydrology/Remote Sensing

Politecnico di Milano - Como, Italy October 2014
M.S. in Civil Engineering

National University of Civil Engineering - Hanoi, Vietnam March 2012
B.S. in Civil Engineering

PUBLICATIONS

Tran, H., Nguyen, P., Ombadi, M., Andreadis, K., Hsu, K., & Sorooshian, S. Improving hydrologic modeling using cloud-free MODIS flood maps. *Journal of Hydrometeorology.* (under review)

Nguyen, P., Shearer, E., **Tran, H.**, Ombadi, M., Hayatbini, N., Palacios, T., Huynh, P., Braithwaite, D., Updegraff, G., Hsu, K., Kuligowski, R., Logan, W., & Sorooshian, S. The CHRS Data Portal, an easily accessible public repository for PERSIANN global satellite precipitation data. *Scientific Data, Nature.* (2018)

Tran, H., Nguyen, P., Ombadi, M., Hsu, K., & Sorooshian, S. A cloud-free MODIS snow cover dataset for the contiguous United States from 2000 to 2017. *Scientific Data, Nature.* (2018)

Nguyen, P., Thorstensen, A., Sorooshian, S., Hsu, K., Aghakouchak, A., Ashouri, H., **Tran, H.**, & Braithwaite, D. Global Precipitation Trends across Spatial Scales Using Satellite Observations. *Bulletin of the American Meteorological Society.* (2018)

Nguyen, P., Thorstensen, A., Sorooshian, S., Zhu, Q., **Tran, H.**, Ashouri, H., Miao, C., Hsu, K., & Gao, X. Evaluation of CMIP5 Model Precipitation Using PERSIANN-CDR. *Journal of Hydrometeorology.* (2017)

Nguyen, P., Sorooshian, S., Thorstensen, A., H., **Tran, H.**, Huynh, P., Palacios, T., Braithwaite, D., Hsu, K., Aghakouchak, A., & Ashouri, H. Exploring trends through RainSphere: Research data transformed into public knowledge. *Bulletin of the American Meteorological Society.* (2016)

CONFERENCE ORAL PRESENTATION

Tran, H., Nguyen, P., Huynh, P., Palacios, T., Braithwaite, D., Hsu, K., & Sorooshian, S. UC Irvine CHRS RainSphere a new user friendly tool for analyzing global remotely sensed rainfall estimates. *Earth Science Information Partners, Summer Meeting.* July, 2018. Tucson, AZ

CONFERENCE POSTER PRESENTATIONS

Tran, H., Nguyen, P., Ombadi, M., Hsu, K., Sorooshian, S., & Andreadis, K. Developing cloud-free MODIS images to improve flood modeling. *American Geophysics Union, Fall Meeting.* December, 2018. Washington, D.C.

Tran, H., Nguyen, P., Hsu, K., & Sorooshian, S. MODIS Snow Cover Recovery Using Variational Interpolation. *American Geophysics Union, Fall Meeting.* December, 2017. New Orleans, LA

Tran, H., Nguyen, P., Huynh, P., Palacios, T., Braithwaite, D., Hsu, K., & Sorooshian, S. UC Irvine CHRS iRain - An Integrated System for Global Real-time Precipitation Observation. *American Geophysics Union, Fall Meeting.* December, 2016. San Francisco, CA

TEACHING EXPERIENCE

Teaching Assistant, Hydrology — UC Irvine September 2017 - December 2017

PROFESSIONAL EXPERIENCE

Department of Civil and Environmental Engineering – UC Irvine September 2015 - December 2018
Research Assistant *Irvine, CA*

- Constructed hydrologic and hydraulic modeling frameworks to simulate rainfall-runoff processes using various software (e.g., HEC-HMS, VIC, SAC-SMA, CREST, HEC-RAS, and LISFLOOD-FP) in multiple regions.
- Develop cloud-free global flood maps using an image processing technique (*Variational Interpolation*).
- Analyze annual maximum drought duration and intensity from Climate Research Unit global precipitation dataset.
- Analyze land use changes in Melbourne during the Millennial drought using Landsat7 images via *ENVI* software.

Laboratory of Geomatics - Politecnico di Milano September 2012 - October 2014
Research Assistant *Como, Italy*

- Analyzed hydraulic characteristics of Serio River (Italy) for weir construction to handle erosion using *HEC-RAS.*
- Supported hydraulic model study for open channels using *ArcGIS* and *HEC-RAS.*
- Simulated behaviors of lake Maggiore (Italy) to accurately reflect storage discharge relationship using *MATLAB.*

Center for Environmental Consultancy and Technology April 2010 - August 2012
Consultant Engineer *Hanoi, Vietnam*

- Designed water supply (capacity: 40,000 m^3/d) and water treatment (capacity: 1,500 m^3/d) plans for Duong River (Vietnam) using *AutoCAD.*

PROFESSIONAL LICENSE

Engineer-In-Training November 2018
National Council of Examiners for Engineering and Surveying (NCEES) USA

PROFESSIONAL ORGANIZATIONS

American Society of Civil Engineers (ASCE)	2018 - Present
American Geophysics Union (AGU)	2015 - Present
American Meteorological Society (AMS)	2015 - Present

AWARDS

UC Irvine Henry Samueli School of Engineering Research and Travel grant	2018
National Fellowship Incentive Program	2018

ABSTRACT OF THE DISSERTATION

Development of cloud-free MODIS datasets for hydrologic applications

by

Hoang Tran

Doctor of Philosophy in Civil and Environmental Engineering

University of California, Irvine, 2018

Distinguished Professor Soroosh Sorooshian, Chair

Space-based observations, emerged in the hydrology field in the last two decades, play a fundamental role in providing alternative information of hydrologic variables besides gauge measurements, especially in data scarce regions. Among all satellite products, products derived from Moderate Resolution Imaging Spectroradiometer (MODIS) Satellite are popular due to the satellite's rapid re-visit time and adequate spatial resolutions. However, cloud obscuration limits the usage of products derived from MODIS because clouds block satellites from capturing the ground state of the earth surface. This dissertation aims to (1) recover two cloud-free MODIS datasets of snow and flood using a 3-D interpolation technique, namely, Variational Interpolation (VI) and (2) demonstrate their usefulness for hydrologic applications.

In the first part of this dissertation, the computational stability of the existing VI method is improved, then, we apply the algorithm to produce a cloud-free snow dataset for CONUS from 2000 to 2017. Moreover, by taking into consideration specific assumptions about the water body characteristic, we implement VI algorithm to remove clouds from MODIS flood maps. Promising results from a validation period over the Mississippi River are presented. We also couple the elevation information to derive the cloud-free MODIS water depth maps from the MODIS water extent maps. Water level maps are important for hydrological studies and can also act as references when the future Surface Water and Ocean Topography (SWOT) direct observations of water elevation are available in 2021.

In the second part of this dissertation, we use the resulting cloud-free MODIS flood and water depth maps to improve a hydrological model by reducing model errors via calibration and data assimilation. The calibrated output inundation maps accurately reflect flood events for the Upper Mississippi River Basin in 2013 and 2014. Also, the downstream discharge via the data assimilation scheme can correctly predict flood events during the same validation period. The results indicate that the framework can be further used to monitor and forecast floods.

Chapter 1. Introduction

1.1. Remote sensing hydrologic variables from satellites

Satellites, classified into Geostationary Earth Orbit (GEO) and Low Earth Orbit (LEO) types, have played major roles in many aspects of life, including science, specifically, hydrology. Satellites provide direct and indirect observations for a wide range of hydrologic variables such as precipitation, surface water, and soil moisture. While GEO satellites travel in the same direction as the Earth's rotation, they can provide observations/estimates in a rapid timely manner. Their main disadvantage is that they are located a large distance from Earth (i.e. 35,800 kilometers above the equator) which increases their products spatial resolutions from 5 km to hundreds of km. In contrast, LEO satellites provide more flexible choices of resolution from 5 m to 1 km depending on the re-visit time of each satellite. To study changes of hydrologic variables in terrain scales, products from LEO satellites are more appropriate. In the scope of this study, we employed images from LEO satellites, in more detail the Moderate Resolution Spectroradiometer (MODIS) sensor onboard of two NASA satellites, Terra and Aqua, to study snow cover and flood extents.

Snow cover plays an essential part in hydrological and energy cycles. Due to its importance, government agencies such as NASA and NOAA have put a lot of effort in satellite missions to observe ground-snow properties. One of the missions has produced the MODIS Snow Cover Area (SCA) product (Hall, et al., 2002). Since its establishment, the products have served as a reliable source of snow measurements for hydrologic studies as well as input in climate models. Researchers used and evaluated the product in various regions including the Columbia and Missouri river basins (Maurer, et al., 2003), Austria (Parajka & Bloschl, 2006), Xinjiang, China (Wang, et al., 2009), and in Sierra Nevada (Rittger, et al., 2013)(Micheletty, et al., 2014). The datasets were also incorporated into a land surface model (Rodell & Houser, 2004) and a hydrology model (Andreadis & Lettenmaier, 2006) for data assimilation.

On the other hand, floods are among the most devastating disasters. Hence, efforts to map floods from space have been studied extensively for more than two decades (Smith, 1997)(Frappart, et al., 2006)(Karlsson & Arnberg, 2009)(Stephens, et al., 2012). Flood images derived from satellite observations are relatively straightforward due to the distinguishable reflectance characteristics of water. Again, among various satellite products, the rapid response MODIS surface reflectance products have shown their effectiveness in flood mapping and modeling (Brakenridge & Anderson, 2006). Furthermore, the products have been used as input of NASA/ Dartmouth Flood Observatory (DFO) Near Real Time (NRT) Global Flood Mapping product (Policelli & SlayBack, 2017).

1.2. Problem statement/Research motivation

Nevertheless, cloud obscuration limits the product's usage. Clouds block satellites from capturing the ground state of the earth surface (i.e. snow/water, land) (Figure 1-1).

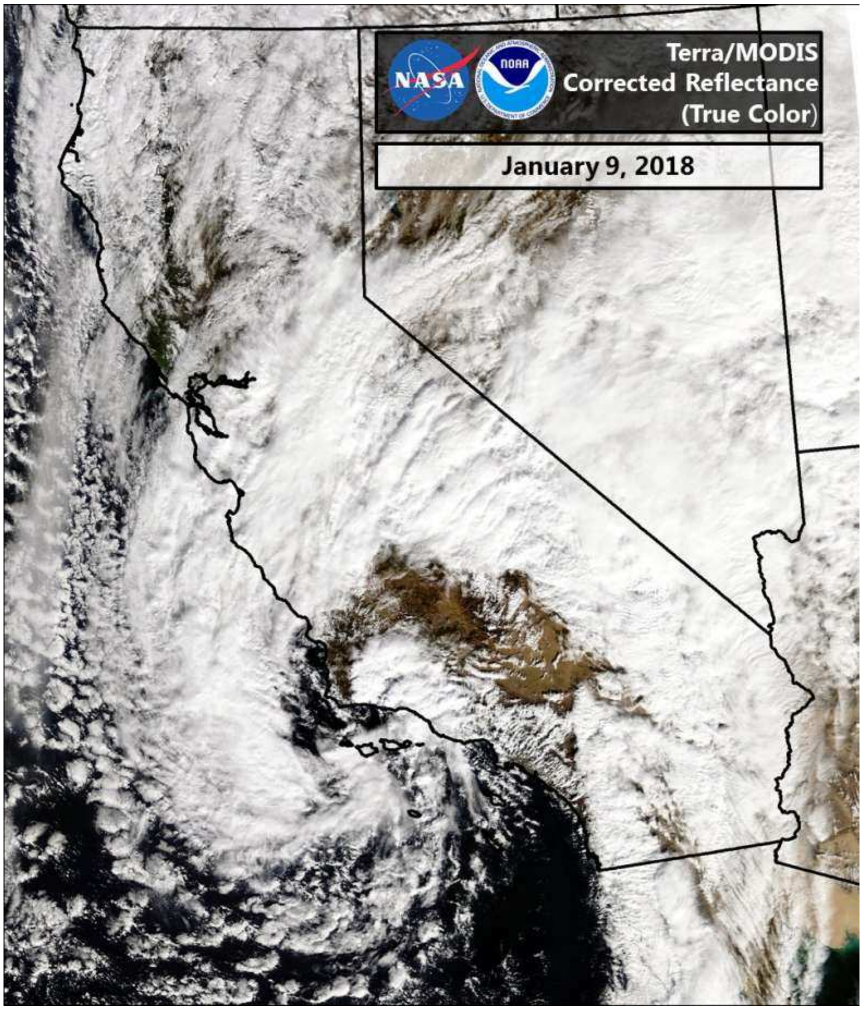


Figure 1-1. Image captured by NASA Terra satellite over California on January 9, 2018. (Image originally courtesy of NASA and NOAA)

Major concerns about MODIS SCA products polluted by clouds and snow/cloud discrimination have been repeatedly mentioned in assessment studies (Hall & Riggs, 2007)(Riggs, et al., 2017) as well as in the product user guide (Riggs & Hall, 2015). Also, clouds frequently contaminate MODIS flood products during events such as hurricanes or storms (Figure 1-2). According to (Policelli & SlayBack, 2017), their group is also struggling with this problem in producing more timely flood products.

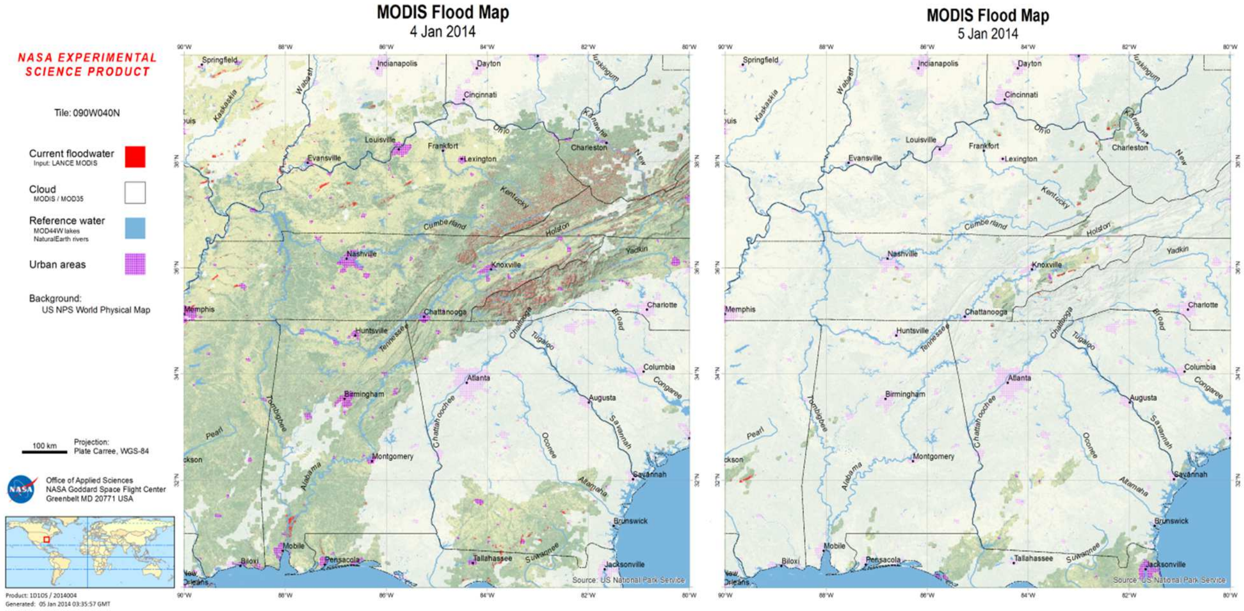


Figure 1-2. NASA MODIS Flood Map in two consecutive days. On Jan 5, 2014, clouds blocked an entire MODIS tile #090W040N

Research that focuses on the removal of clouds from satellite images has been conducted for years and can be classified into two main categories: model-driven methods and data-driven methods. Physical models are based on the relationship between snow or water extent with other factors to estimate the extent beneath clouds. Representative studies for snow could mention: (Cline & Carroll, 1999)(Barrett, 2003)(Molotch, et al., 2004)(Painter, et al., 2009). Alternatively, flood simulations by hydrological models are beneficial for both operational applications and disaster management due to the model's dynamic responses (Bates & De Roo, 2000)(Begnudelli, et al., 2008)(Nguyen, et al., 2015)(Krajewski, et al., 2017).

On the other hand, while data-driven methods to recover snow images polluted by cloud have been carried out for over a decade, studies attemping to remove clouds from MODIS flood maps while preserving the product's original resolutions are still in an early stage. Based on snow persistent characteristics, (Dozier, et al., 2008) considered snow data as a sparse space-time cube that could be filled by temporal cubic spline interpolation. Their results demonstrated that the interpolated and smoothed product has more consistent snow-covered area in the Tuolumne and Merced River basins throughout the water year 2005 than the raw, cloud cover filtered data. However, the approach of Dozier et al. is limited to 1-D interpolation primarily because of frequent zenith angles oscillation and slow computation. Later on (Gafurov & Bardossy, 2009) introduced a series of six spatiotemporal filters to mitigate cloud cover from MODIS images in the Kokcha River basin in Afghanistan; detailed information of this study will be presented in subsequent sections. (Parajka & Bloschl, 2006) proposed a regional snow-line method (SNOWL) utilizing elevation information for de-clouding; the method robustly recovered snow cover maps from clouds over Austria. (Hall, et al., 2010) suggested a cloud-gap-filled (CGF) method to produce a snow cover map with cloud-persistence count (CPC) for each grid where lower CPC snow grids are more likely to have snow. While approaches from (Parajka, et al., 2010) and (Hall, et al., 2010) are simple and proven to be suitable for use in hydrological and global models, persistent cloudy conditions during snow accumulation periods may significantly reduce their reliability. More recently, (Dong & Menzel, 2016) employed information from snow stations to estimate ground states of cloud-cover areas from MODIS in south-western

Germany. Although this approach is effective in areas with dense snow networks, it requires manually determining thresholds for each station's predicting capability of nearby snow cover based on station location and elevation.

To completely remove clouds and delineate dynamic snow boundaries, (Xia, et al., 2012) implemented the Variational Interpolation (VI) method (Turk & O'Brien, 1999) for interpolating the three-dimensional space-time cube of snow cover proposed by Dozier et al. Evaluation results in the Sierra Nevada mountain range demonstrated that the method was robust and accurate since during the accumulation period of (25 - 27) March 2007 and the melting period of (14 - 16) March 2009. However, the main drawback of the original VI method is the system instability which limits its implementation on a larger scale. (Tran, et al., 2018) improved the algorithm stability by integrating MINimum RESidual (MINRES) iterations (Paige & Saunders, 1975) instead of the traditional LU decomposition when solving linear systems.

From the previous success of the VI algorithm in removing clouds from MODIS-SCA products, we propose an approach to use the algorithm in recovering flood inundation maps from partially cloud-covered images given an assumption of water bodies persistent characteristic. The cloud-free flood maps will fill the gap of missing data caused by cloud cover in satellite images.

1.3. Objectives

The main objectives of this dissertation can be divided into three main parts:

i. Implementing the improved VI algorithm over the Contiguous United States (CONUS) to create a cloud-free snow cover dataset from 2000 to 2017.

ii. Removing clouds from MODIS images to create cloud-free flood maps. Validating those results with derived Landsat maps during flood events.

iii. Improving hydrologic modeling using the resulted cloud-free flood maps from the previous steps through calibration and data assimilation. The main goals are to calibrate an inundation model via water extent maps and assimilate water level states of a routing model via water depth maps. In other words, this study demonstrates the usefulness of the cloud-free flood maps in flood simulating and forecasting.

1.4. Dissertation outline

This dissertation consists of five main chapters. Chapter 2 presents a procedure to create the cloud-free snow cover dataset for CONUS over 17 years. Chapter 3 details an implementation of the VI algorithm to create cloud-free MODIS flood maps and related comparison and validation. Chapter 4 provides steps in coupling flood maps to improve hydrologic modeling and forecasting. Chapter 5 summarizes the findings of this study and discusses future extensions of this work.

Chapter 2. Cloud-free snow cover maps over CONUS

2.1. Review of the reference datasets: MOD10C1 and MYD10C1

2.1.1. Technical details

The main inputs in this study are products from MODIS/Terra Snow Cover Daily (MOD10C1) and MODIS/Aqua Snow Cover Daily (MYD10C1) version 6 (Hall & Riggs, 2016) released in July 2016 by the National Snow & Ice Data Center (NSIDC). These daily Climate Modeling Grid (CMG) products are in a sequence of MODIS snow product suite (Riggs & Hall, 2015), beginning with the 500 m resolution swath product (MOD10_L2).

As reported in (Riggs & Hall, 2015), the swath level snow mapping algorithm is based on the Normalized Difference Snow Index (NDSI)(Crane & Anderson, 1984)(Dozier, 1989). NDSI is calculated for Terra/MODIS using band 4 and band 6 and for Aqua/MODIS using band 4 and band 7.

$$NDSI = \frac{\text{Band } 4 - \text{Band } 6}{\text{Band } 4 + \text{band } 6} \tag{2-1}$$

The global criteria for snow is NDSI greater than 0.4 and near-infrared reflectance (band 2) greater than 0.11 and band 4 reflectance greater than 0.10. To increase snow detection sensitivity in forested landscapes, the MOD10_L2 product combines the Normalized Difference Snow Index (NDSI) range from 0.1 to 0.4 with the Normalized Difference Vegetation Index (NDVI) (Riggs & Hall, 2015). After a pixel is classified as snow, "it is subjected to a series of screens to alleviate snow commission errors and flag uncertain snow detections" (Riggs & Hall, 2015). More details about the screens can be found in the product user guide. Here, we only summarized the screens main thresholds: (1) Version 6 combines surface temperature and height screen; if snow pixels are in low elevation (< 1,300 m) and warm surfaces (> 283 K), they are reversed to no-snow. This new surface temperature screen solves the problem of detecting snow in mountain ranges during spring and summer brought up by (Rittger, et al., 2013); (2) If snow pixels have low reflectance (Very High Visible (VIS) of band 2 is ≤ 0.10 or band 4 is ≤ 0.11) or low illumination (solar zenith angles > 70°), they will be set as no-snow or night pixels; (3) Low NDSI, or unusually high Short-Wave Infrared (SWIR) reflectance snow pixels, will be converted back to no-snow pixels.

The next product in the sequence, MOD10A1, only selects one 'best' observation from all the MOD10_L2 swaths over a location using strict criteria including solar elevation, distance from nadir, and observation cover. Selecting an observation closest to nadir with maximum coverage of the cell (Hall & Riggs, 2007) could solve problems from the off nadir viewing from the MODIS reflectance product (MOD09) mentioned by (Dozier, et al., 2008).

The main input of the study, MOD10C1/MYD10C1, "maps 500 m MOD10A1 observations into 0.05^{o} CMG cells, outputs for a grid cell are determined by the percentage of counts of observations, snow or cloud, mapped in the cell" (Riggs & Hall, 2015).

It should be noted that the launch dates of the two satellites, Terra and Aqua, are different with December 18, 1999 and May 4, 2002 for Terra and Aqua respectively. In this study, the term MODIS-SCA product will be used to represent both products, MOD10C1 and MYD10C1, from Terra and Aqua whenever the text refers to 2002 or later. When referencing prior to 2002, the term only represents products from the Terra satellite.

2.1.2. Product's usage and evaluation

The daily MODIS snow product suite proves its usefulness in snow studies by an increasing number of citations in recent years. Different research employed the product to directly study change of snow cover in time from various continents of the world: Africa (Wunderle, et al., 2016), Asia (Tekeli, et al., 2016)(Yi, et al., 2016)(Li, et al., 2016)(Tang, et al., 2017)(Zhang, et al., 2017), Europe (Krajci, et al., 2016)(Harer, et al., 2018), North America (Marcil, et al., 2016)(Schneider & Molotch, 2016)(Verbyla, et al., 2017), and South America (Perez, et al., 2018)(Malmros, et al., 2018).

The product suite has "an overall accuracy of about 93%, lower accuracy is found in forested areas" (Hall & Riggs, 2007). More recent studies from (Liu, et al., 2008) and (Rittger, et al., 2013) also highlighted scenarios in which dense canopy limits the product ability to detect snow cover.

2.1.3. Problem statement

It is clear that cloud covers severely and frequently hinder MODIS snow data. Figure 2-1 shows an example of cloud cover over most of the contiguous United States (CONUS) on January 2nd, 2017. Section 2.6 below will demonstrate the effect of cloud hindrance on the original dataset.

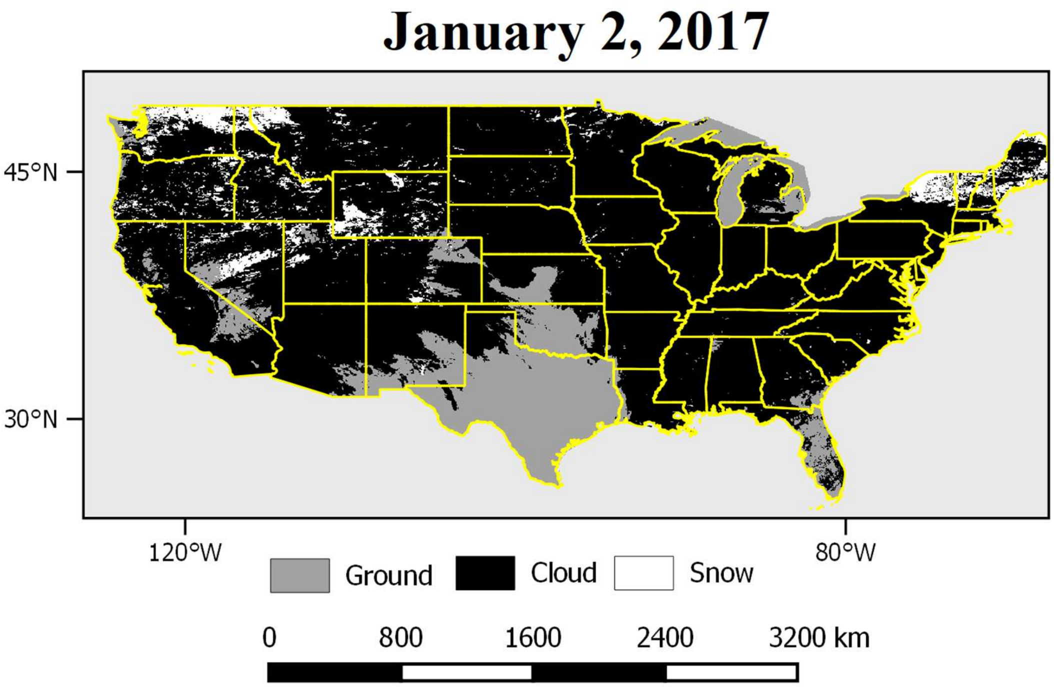


Figure 2-1. Merged snow cover image from MODIS Terra and Aqua satellites on January 2nd, 2017

Therefore, as mentioned in Section 1.2, research on removing clouds from satellite snow cover images has been carried out for decades. While approaches based on physical models can couple other information to provide accurate estimates of snow cover, it is difficult and computationally expensive to implement these models on a global scale. On the other hand, data-driven methods are more robust in the sense of using available information both in time and space to interpolate data in missing areas. (Xia, et al., 2012) introduced the Variational Interpolation (VI) method which accurately and completely removed clouds from snow cover images over the Sierra Nevada mountains. This study will continue to develop the algorithm in order to implement it on a much larger scale (i.e. CONUS).

2.2. Study domain

The spatial domain of the dataset developed in this study is the contiguous United States (CONUS) which covers about 8,080,464.3 km^2, ranges between 24^{o}30N and 49^{o}25N in latitude and from 66^{o}57W to 124^{o}46W in longitude. During winter seasons, from November to the end of February, the snow cover extent for CONUS varies from one million km^2 to four million km^2. This snow cover plays a crucial role in energy and hydrological cycles.

2.3. Methodology

Overall, the process of creating a cloud-free product from MOD10C1/MYD10C1 starts with reclassifying MODIS into one of three categories, namely snow, land (no-snow), and cloud based on the threshold of 50% fractional: if a grid has a percentage of snow greater than 50%, it is set as snow. If the sum of the snow and cloud fractions in one location is smaller than 50%, a grid is marked as land (no-snow). If neither snow nor no-snow, a grid is set as cloud. Next, the reclassified MODIS images are passed through two subsequent steps based on a series of filters and the VI algorithm. These two steps are discussed in the following subsections.

2.3.1.1. Mitigated filters

The filters are used as a first step to retrieve cloud-free snow cover images. This method has been adopted from (Gafurov & Bardossy, 2009) and it consists of five filters to mitigate cloud obstruction: (1) combining Terra and Aqua snow cover images in a same day, (2) short-term temporal filter, (3) elevation filter, (4) neighborhood spatial filter, and (5) long-term temporal filter. Figure 2-2 illustrates the flow of cloud polluted images through the filters and the functions applied by the filters. The importance of using the filters is that they provide the necessary information about snow boundaries in order for the VI algorithm to be applied.

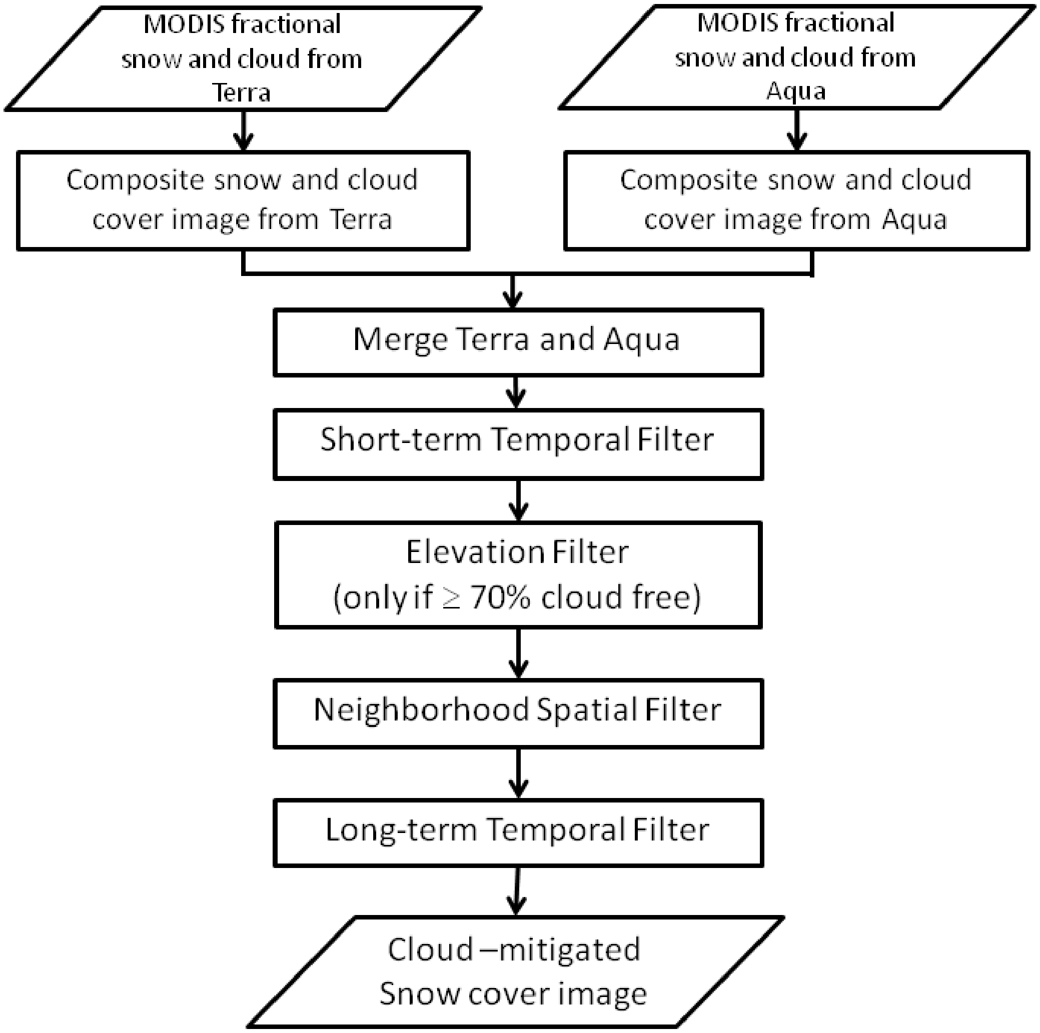


Figure 2-2. Flow chart of cloud mitigation filters. The process consists of five filters: (1) combining Terra and Aqua snow cover images in a same day, (2) short-term temporal filter, (3) elevation filter, (4) neighborhood spatial filter, and (5) long-term temporal filter.

The first filter implies an assumption that no snowmelt or snowfall occurred within two observations of MODIS in one day. Thus, as long as one satellite views a pixel as snow (or land), this ground status will be assigned to the pixel in the combined image. The formula is given as follows in equation (2-2):

$$S_{(x,y,t)} = max(S^{A}_{(x,y,t)}, S^{T}_{(x,y,t)}) \tag{2-2}$$

Where x and y are spatial (i.e. longitude and latitude) coordinates of pixel S; t is the day index of pixel S. S^A and S^T represent pixels from Aqua and Terra respectively. The second filter assigns the cloud grids as snow (or land) if the cloud covered pixel shows both snow (or land) in the preceding and succeeding days. This step is formulated as equation (2-3):

$$S_{(x,y,t)} = 1\ if\ (\mathrm{S}_{(x,y,t-1)} = 1\ and\ \mathrm{S}_{(x,y,t+1)} = 1) \qquad (2\text{-}3)$$

In the third filter, maximum and minimum elevation lines defined as the highest and lowest elevation of snow grids in the image are determined. To ensure the snow lines are correctly determined, the condition of this filter is that at least 70% of the image is cloud free. Otherwise this filter will be skipped. The filter assigns grids with a lower elevation than the minimum elevation line $H_{min}^{S}(t)$ as land. Likewise, grids with a higher elevation than the maximum elevation line $H_{max}^{S}(t)$ are assigned as snow. The formulas are given as follows in equations (2-4) and (2-5):

$$S_{(x,y,t)} = 0\ if(H_{(x,y)} < H_{min}^{S}(t)) \qquad (2\text{-}4)$$
$$S_{(x,y,t)} = 1\ if(H_{(x,y)} > H_{max}^{S}(t)) \qquad (2\text{-}5)$$

where $H_{(x,y)}$ is the elevation of a pixel (x,y) location. The last two filters apply spatial and temporal processing respectively. First, the fourth filter merges two neighborhood spatial filters of Gafurov and Bardossy into one. Specifically, if three out of four direct "side-bordering" pixels of the cloudy pixel indicate snow/land, the cloudy pixel will be set as snow/land. Second, when considering all eight neighboring pixels, if any pixel both has lower elevation than the center elevation and shows snow, the center pixel will also be assigned as snow. This step is formulated as equation (2-6):

$$S_{(x,y,t)} = 1\, if\ (\mathrm{S}_{(x+k,y+k,t)} = 1\, and\ \mathrm{H}_{(x+k,y+k)(k\in(-1,1))} < \mathrm{H}_{(x,y)}) \qquad (2\text{-}6)$$

In the last filter, a new long-term temporal filter was developed based on the fact that the annual snow-status time series can be separated into three types of periods: snow, land, and transition periods. For a snow period, the grids either show snow or cloud, therefore, cloudy pixels would be assigned as snow. On the other hand, for a land period, the grids either show land or cloud, therefore, cloudy pixels would be assigned as land. The selection of length of the period is subjective, but must be long enough to avoid phase change or long-lasting cloudy periods. For this study, we chose a 30-day window period. Interested readers should refer to (Gafurov & Bardossy, 2009) for detailed information about the filters.

2.3.1.2. Variational Interpolation algorithm

The time-varying snow cover boundaries resulted from previous filters are modeled by the VI algorithm (Turk & O'Brien, 1999) using a three-dimensional implicit function formulated as:

$$f(\vec{x})\begin{cases} > 0 & inside\ snow\ cover \\ = 0 & at\ snow\ boundaries \\ < 0 & outside\ snow\ cover \end{cases} \qquad (2\text{-}7)$$

where $\vec{x} = (x_1 \quad x_2 \quad t)^T \in R^3$, x_1 and x_2 are spatial coordinates on the projection plane, and t is the time. In three spatial dimensions, implicit functions deliver relatively simple techniques to generate complicated but useful surfaces (Gomes, et al., 2009). Once the snow cover implicit surface in space and time is determined, cloud-free images from selected days can then be obtained through cross sections of the surface(Xia, et al., 2012).

One thing to note is that interpolation from implicit surfaces depends heavily on the surface smoothness. In order to apply VI for snow cover, we have to make a hypothesis about the dynamic property of snow cover boundaries. Numerous theories in Physics such as the Principle of Least Action (De Maupertuis, 1744), Principle of Least Forcing (Gauss, 1829), and the Variational Principle (Lanczos, 1970) proved that a natural process always operates in its most efficient way. As energy cost is one of the most crucial factors of efficiency, a natural surface should hold the minimum energy cost. Hence, it can be represented as a linear combination of the radial-basis function established at selected constraint points on the surface according to the following equation (Xia, et al., 2012)(Duchon, 1977):

$$f(\vec{x}) = 0 \Rightarrow \sum_{i=1}^{N} w_i R(\vec{x} - \vec{x_i}) = 0 \tag{2-8}$$

where $\mathbf{w}$ is a set of N weights and $R(\vec{x} - \vec{x_i})$ is a selected radial-basis function established at N constraints points. We decided to use the thin plate function $R(.) = r^2 \log r$ with $r = \|\vec{x} - \vec{x_i}\|$ to present the radial-basis function. With constraint points collected on snow boundaries in discrete times, the weights of those points can be computed by solving the linear system to create the implicit surface.

To provide VI with necessary constraint points, the Douglas-Peucker algorithm has been used (Douglas & Peucker, 1973). This one parameter method is simple and widely used in vector graphics simplification and cartographic generalization. Given a relative distance dimension ε $(0 < \varepsilon < 1)$ and a starting curve of an ordered set of points, the algorithm recursively divides the curve to discard points closer than ε to line segments. The larger ε, the less points will be kept (Douglas & Peucker, 1973). After experimenting with a wide range of ε, a value of $\varepsilon = 0.2$ has been chosen to preserve shapes of snow boundaries and reduce the number of points fed into the VI algorithm.

2.3.1.3. Improving the system stability for the VI algorithm

The performance of the original VI algorithm (Xia, et al., 2012) is unstable when a massive amount of constraint points is collected. The default method for solving the linear system (2-8) is LU (lower-upper) decomposition introduced by Tadeusz Banachiewicz in 1938(Schwarzenberg-Czerny, 1995). According to (Schwarzenberg-Czerny, 1995), LU decomposition, used to solve linear equations by factoring a matrix as the product of a lower triangular matrix and an upper triangular matrix, which is subjected to singular symmetric systems. For example, if constraint points were collected for the CONUS region in January 2009 (i.e., ~130,000 points), the linear system (2-8) will become singular and the whole system will break down. A proposed solution for this problem is using MINRES algorithm by (Paige & Saunders, 1975). MINRES is a Krylov subspace method for solving large symmetric systems. When applied to an inconsistent system (i.e., a singular symmetric linear problem), (Paige, et al., 1995) and (Choi, 2006) reported that MINRES maintains the system stability and provides a least-squares solution.

Paige et al. (1995) analyzed the convergence behavior of the MINRES method in singular systems and concluded its residual monotonically decreases toward the origin satisfying several convergence properties.

In our implementation of the MINRES algorithm, to ensure both system stability and accuracy, besides setting the tolerance threshold to 1e-08 (i.e. the smaller the threshold is, the more reliable the results become), we also specified the maximum number of iterations to be 1000. Hence, when either these conditions were reached, the algorithm terminated. In a small experiment, we compared the system stability of VI using both traditional LU decomposition and the MINRES method over the CONUS region in January 2009. The performance was measured by the time elapsed over the increment of the interpolation period. Both methods started by using five consecutive days as a calculation unit. However, as more points were collected, the ordinary LU decomposition resulted in a system breakdown (Figure 2-3). In contrast, the MINRES method demonstrated its superiority by maintaining the system stability and accelerating the computation time. It is worth noting that, in all singular cases, the

MINRES method terminated before reaching the maximum number of iterations which indicated that system accuracy was also guaranteed.

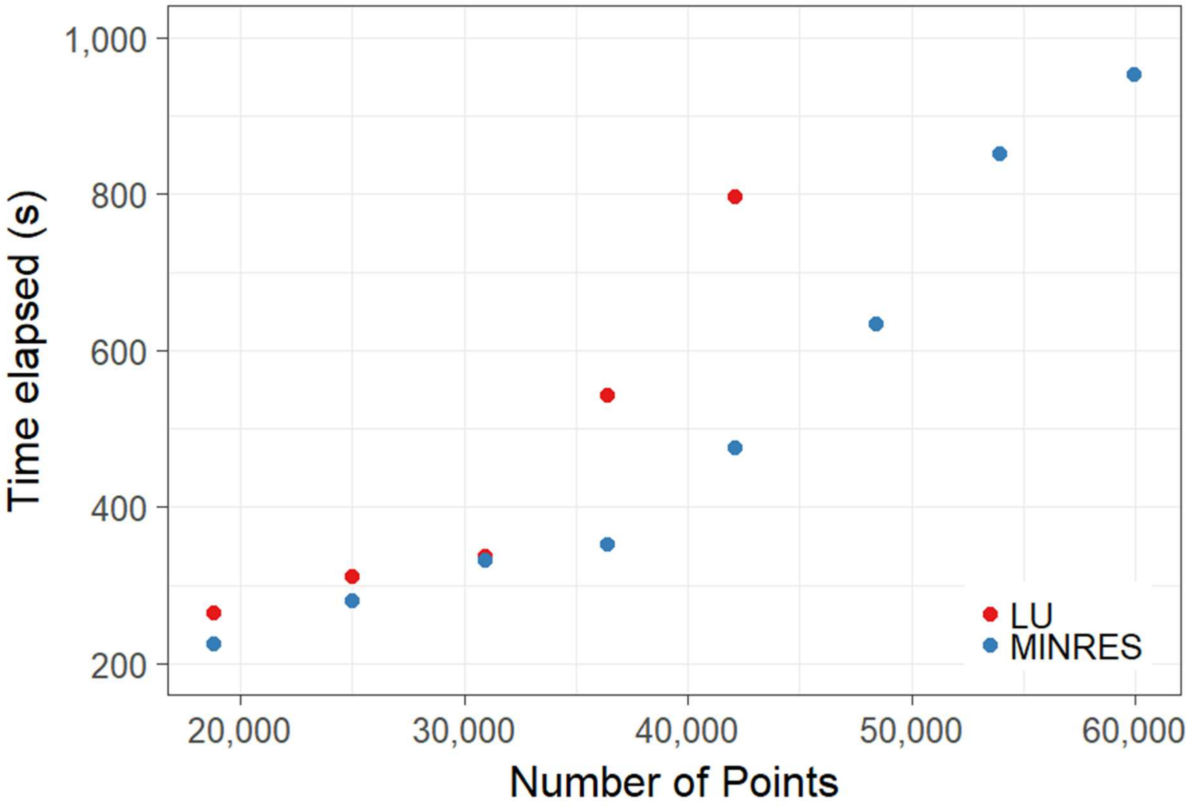


Figure 2-3. Results of computational efficiency comparison between the VI algorithm integrated with LU decomposition and the MINRES iteration. Computation time to solve a linear equation using LU decomposition (red) and MINRES iteration (blue) regarding the number

2.3.1.4. Discovering the most efficient interpolation period

(Xia, et al., 2012) claimed that five-day time series expose enough snow cover boundaries as inputs for the VI algorithm and are efficient for the computation. However, since the integration of the MINRES method to the VI algorithm, computational efficiency has greatly improved, the experiment showed that longer time series have the tendency in better accuracies. Nevertheless, once again, we have to balance between accuracy and cost efficiency: how many more days need to be included to increase the POD by 0.1, or is it worth it to have a much higher computational cost in order to make a small improvement in accuracy.

The experiment was carried out by choosing results of only the middle day of interpolation periods as it contained the most information of snow cover boundaries. Four different numbers of consecutive days were taken into account, namely, 9 days, 19 days, 29 days (about one month), and 59 days (about two months). The results were evaluated in two separate winter periods (in 2014-2015 winter and 2015-2016 winter) by the SNOTEL network using three categorical verification scores of FAR, POD, and HK.

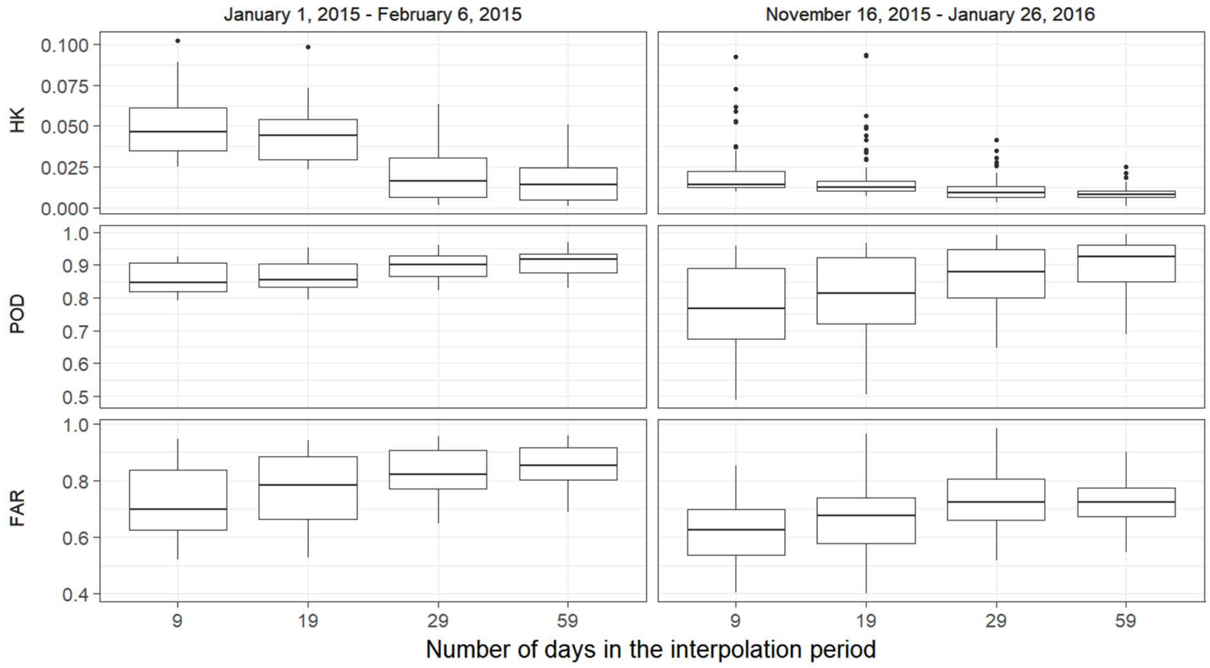


Figure 2-4. Compare the FAR, POD, and HK between different interpolation periods. (Left) The period from January 1, 2015 to February 6, 2015. (Right) The period from November 16, 2015 to January 26, 2016

When we increased the number of days, there was a noticeable gap in FAR and POD between the 19-day calculation unit and the 29-day calculation unit in the first period (January 1, 2015 - February 6, 2015). The FAR third quartile of the 29-day calculation unit was 0.026, which was nearly the 19-day calculation unit FAR first quartile of 0.027; in contrast, the median increased by 0.05 from 0.85 to 0.9 of 19-day POD and 29-day POD, respectively. In general, when using the 29-day interpolation period, the VI algorithm had much better results than using shorter interpolation periods. Furthermore, the 29-day interpolation period had results close to the 59-day interpolation period, which required doubled the computational time.

All in all, in considering the number of consecutive days as a calculation unit, 29 consecutive days seemed to capture most of the dynamical changes and exposure of snow cover, and hence produced satisfactory results when applied to the VI method. Moreover, this time frame is also productive for computational efficiency. From now on, the 29-consecutive-day unit is used as a calculation unit for all the VI implementations.

2.3.1.5. Assessing the accuracy of different days in the interpolation period

The MODIS-SCA products have a lag time of around 2 days with respect to real time. Despite that, the products are very useful in near real-time snow monitoring (Hall, 2002). Naturally, one might raise the question: How accurately can the VI algorithm provide clear-cloud snow images in near real-time? To answer this question, we made an experiment to compare the accuracy of different days in the interpolation period.

We used 15 days from the middle day to the last day of the interpolation period for comparison. The results of the above study period were validated using snow records from around 950 stations over CONUS. Three categorical verification scores were used for the above experiment.

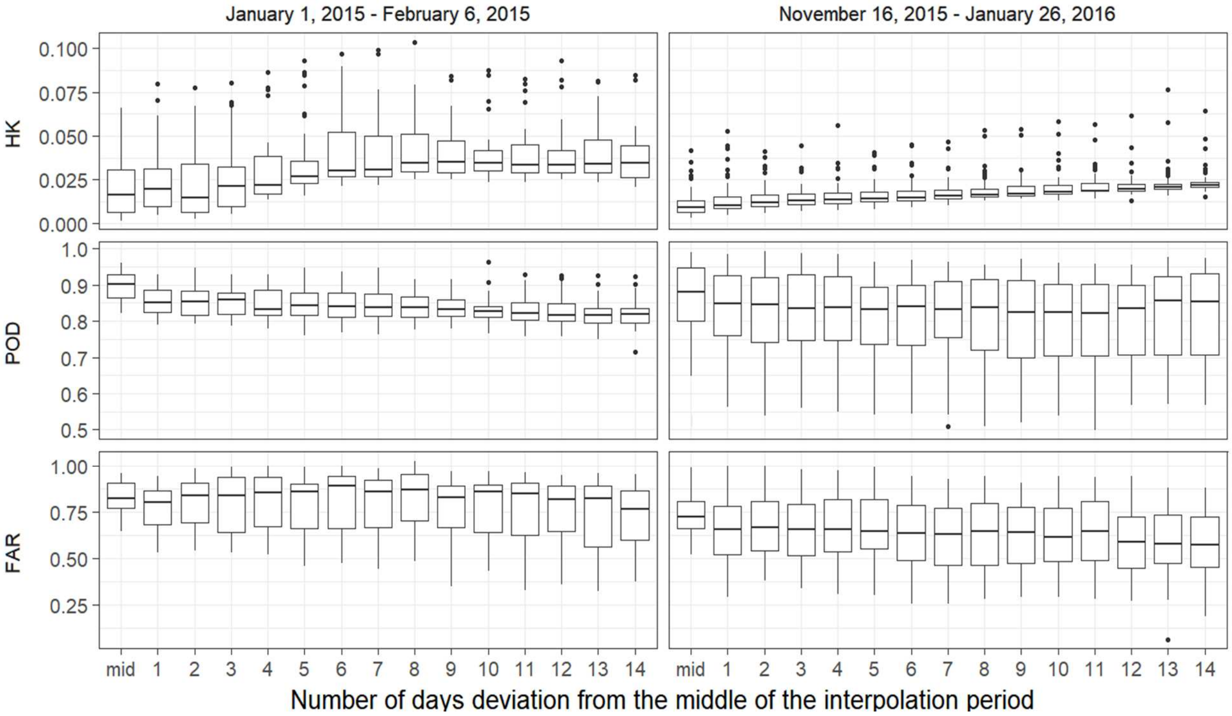


Figure 2-5. Compare the FAR, POD, and HK for different days in the interpolation periods. (Left) The period from January 1, 2015 to February 6, 2015. (Right) The period from November 16, 2015 to January 26, 2016

The experiment clearly demonstrated that middle days had the best results, and the accuracy decreased when approaching the end of the interpolation periods as the snow cover information became more scarce. Yet, both scores from the last days were still considered as adequate as the middle days output from VI algorithm using the 19-day interpolation period. For example, the last day of the period had medians of 0.035, 0.86, and 0.68 with FAR, POD, and HK, respectively, in comparison with 0.04, 0.86, and 0.79 from the middle days of VI algorithm using the 19-day interpolation period. We suggest to use the last days for near real-time snow monitoring; later on, when there is enough information of snow cover boundaries, these days will be updated by re-running the VI method and treating them as middle days of the interpolation periods.

2.4. Dataset creation

The cloud-free dataset was developed by applying the mitigating filters and the VI algorithm introduced in the previous subsections. Regarding the use of VI algorithm, wide ranges of interpolation periods were examined to obtain the optimum results in terms of accuracy; a period of 30 days has been selected. This interpolation period is used in a moving window approach such that the window is centered around the day under consideration in order to efficiently utilize information about snow boundaries. The interpolation period started from February 24, 2000; the whole dataset was created in more than one month, using the High Performance Clusters (HPC) of the University of California, Irvine.

2.5. Validation

2.5.1. MODIS-SCA validation using the SNOTEL network

MODIS-SCA products have been evaluated in numerous studies and proven their accuracy. Nonetheless, due to cloud obscuration, most of the research only selected days with little or no clouds. This may lead to data interruption, especially during snow accumulation periods, when clouds appear more frequently in snow areas. In this study, we decided to count all the numbers of snow stations from November 2016 to the end of February 2017 which fall into five categories: (1) Hit stations, (2) Miss stations, (3) False stations, (4) Correct Negative stations, and (5) Cloud-hindered stations. We took consecutive days in a relatively long period to not only show the dynamic change of cloud and snow, but more importantly to validate the MODIS-SCA accuracy.

Figure 2-6. Number of SNOTEL stations in five categories: Hit, Miss, False, Correct Negative, and Cloud-hindered with respect to the merged MODIS images.

Figure 2-6 demonstrates that, in the first two weeks of winter, the merged MODIS images failed to capture around 30 to 40 stations that showed snow. This is due to a fairly large area (0.05^{o}-by-0.05^{o}) covered by one pixel of MODIS; when snow started accumulating, the snow proportion of that pixel may still be smaller than 50%, which leads to miss detection. Fortunately, after these two weeks the merged MODIS captured the hit stations quite well. Although the number of miss stations throughout the period is low, 2.66% in total number of stations, they located consistantly in complex land surface terrain of the Sierra Nevada Mountains (Xia, et al., 2012). This problem of MODIS-SCA products has been discussed in (Hall, 2002) (Klein & Barnett, 2003) which show that MODIS fractional snow cover using NDSI needs major adjustments for multifarious land cover types areas.

One more thing to notice from the experiment is a very high percentage of cloud-hindered stations (74,608 stations; i.e., around 66%) throughout the period (Table 2-1), which illustrates the limited usage of the products. The last experiment would mention this aspect again in comparing the merged MODIS images with results from the VI algorithm.

Table 2-1. Total number of stations for validation the merged MODIS from November 2016 to February 2017

	Total number of stations					
	Hit	Miss	False	Correct Negative	Cloud Hinder	Total
Merged MODIS	25,869	3,001	579	8,960	74,608	113,017

2.5.2. Landsat validation

2.5.2.1. Region selection

In this study, Landsat 7 ETM+ was used as a baseline to validate the spatial continuity of the snow cover dataset since Landsat has high spatial resolution (30m) and full coverage of the dataset time range (2000 to 2017). To avoid a potential saturation of ETM+ visible bands (Painter, et al., 2009), only Landsat 7 Tier 1 images have been utilized. Data from October to March of each year with a cloud threshold of less than 15% was obtained from the U.S. Geological Survey website (https://landsat.usgs.gov). Landsat 7 Tier 1 product ensures the highest available data quality for time-series processing analysis (LEDAPS; (Masek, et al., 2006)). Tier 1 Landsat data has RMSE no greater than 12 m and can be considered consistent and inter-calibrated across the full collection (https://landsat.usgs.gov).

In order to map snow cover from Landsat images, the SNOWMAP algorithm (Dozier, 1989)(Hall, et al., 1995) was used. It is based on NDSI index which is calculated for Landsat using band 2 and band 5:

$$NDSI = \frac{\text{Band } 2 - \text{Band } 5}{\text{Band } 2 + \text{band } 5} \tag{2-9}$$

When NDSI is greater than or equal to 0.4 and the band 4 reflectance value is greater than 11%, the pixel is classified as snow (Huang, et al., 2011). The resolution of the Landsat snow cover is then up-scaled to 0.05^{o} and re-projected into the geographic coordinate system with the spheroid of WGS84 to match with the MODIS-SCA product resolution and projection. The evaluation was conducted in four regions with different climate conditions, elevations, and land cover, namely, the Seattle region, the Minneapolis region, the Rocky Mountain, and the Sierra Nevada of California. The following paragraphs illustrate the four regions and their characteristics.

A total of 50 samples of Landsat images from March 28, 2000 to February 15, 2017 were used for validation (See Table 2-2). The four selected regions for validation in this study include two high altitude regions, namely the Rocky and Sierra Nevada Mountains. The Rocky Mountains are located between 40.80^{o} and 42.73^{o} N in latitude, and 105.51^{o} and 108.25^{o} W in longitude. This area has a high average altitude of 2500m with a large amount of grasslands. The region's snow regime is "predominantly continental with some pockets of intermountain characteristics" (Painter, et al., 2009). On the other hand, the Sierra Nevada Mountains (37.95^{o} to 39.89^{o} N, 118.84^{o} to 121.45^{o} W) is under great influence of maritime snow climates. Elevation ranges from 2m at the foothills to around 2800m on the mountain (in the sample area) with equal portions of needle leaf forest, savannas, and grasslands.

Moreover, two additional regions were selected for validation. Firstly, a region around Seattle located between 46.46^{o} and 48.42^{o} N in altitude, and 123.32^{o} and 123.04^{o} W in longitude was selected. The region is known for its rainy climate with the Cascade Mountain range located on the east side; winters in this region are typically wet with significant snow accumulation in the mountain area. Land cover is mostly evergreen needle leaf forest and mixed forest. Secondly, a region around Minneapolis located between 43.62^{o} and 45.60^{o} N in latitude, and 92.09^{o} and 94.97^{o} W in longitude was selected. It has an average elevation of 330 m and a land use pattern primarily consisting of cropland/natural vegetation mosaic and urban. Snow is the main form of precipitation from November through March with an annual state-wide average of 110 snow-cover days.

2.5.2.2. Results

Table 2-2. Landsat validation for the performance of cloud-free dataset

Region	Landsat no. and path/row	Date	POD	FAR	Initial cloud ratio in MODIS
Seattle	L7: 46/27	03/28/2004	0.886	0.280	0.035
		10/04/2009	0.896	0.130	0.012
		03/24/2014	0.886	0.236	0.133
		03/29/2016	0.882	0.280	0.027
Minneapolis	L7: 27/29	01/10/2001	1.000	0.130	0.063
		02/27/2001	0.999	0.019	0.285
		03/15/2001	0.988	0.005	0.762
		02/17/2003	0.864	0.063	0.288
		12/23/2005	0.996	0.210	0.223
		01/27/2007	0.994	0.270	0.777
		01/14/2008	1.000	0.195	0.103
		01/30/2008	0.982	0.220	0.856
		12/15/2008	0.999	0.245	0.858
		02/01/2009	1.000	0.186	0.001
		01/06/2011	0.993	0.268	0.110
		03/27/2011	0.959	0.208	0.110
		03/06/2015	0.916	0.246	0.140
		12/03/2015	1.000	0.269	0.021
Rocky	L7: 35/31	01/02/2001	0.994	0.111	0.021
		03/07/2001	0.948	0.168	0.179
		01/13/2005	1.000	0.178	0.113
		12/15/2005	0.995	0.175	0.440
		12/18/2006	1.000	0.228	0.250
		01/22/2008	0.998	0.086	0.046
		03/10/2008	0.981	0.192	0.164
		12/23/2008	0.999	0.215	0.358
		03/29/2009	0.950	0.246	0.389
		11/24/2009	0.993	0.215	0.083
		12/10/2009	0.991	0.134	0.052
		02/15/2011	0.964	0.149	0.115
		02/18/2012	0.998	0.095	0.104
		01/03/2013	1.000	0.134	0.063
		12/05/2013	0.995	0.092	0.072
		01/06/2014	1.000	0.222	0.265
		03/30/2015	0.903	0.149	0.011

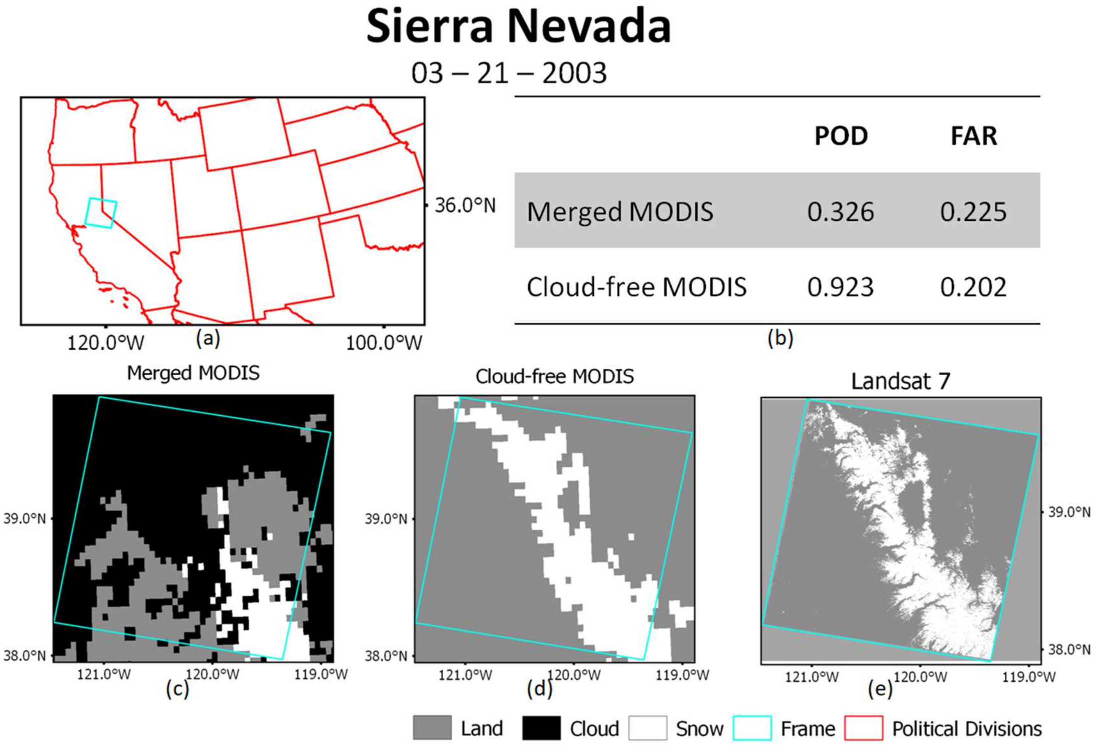
		12/27/2015	1.000	0.120	0.024
		02/15/2017	0.860	0.240	0.169
Sierra Nevada	L7: 43/33	03/28/2000	0.924	0.127	0.183
		02/27/2001	0.904	0.144	0.033
		03/05/2003	0.909	0.204	0.033
		03/21/2003	0.923	0.202	0.652
		03/07/2004	0.935	0.135	0.023
		03/10/2005	0.940	0.162	0.023
		02/15/2008	0.909	0.250	0.039
		03/02/2008	0.902	0.151	0.100
		03/18/2008	0.933	0.181	0.221
		03/05/2009	0.985	0.155	0.176
		03/16/2013	0.928	0.230	0.040
		03/19/2014	0.878	0.224	0.256
		03/24/2016	0.889	0.164	0.464
			0.955	0.179	0.199

We compared the performance of the cloud-free dataset with Landsat using two categorical validation indices, Probability of Detection (POD) and False Alarm Ratio (FAR).

$$POD = \frac{Hit}{Hit + Miss} \tag{2-10}$$

$$FAR = \frac{False}{Hit + False} \tag{2-11}$$

Over the validation scenarios, POD ranged from 0.860 to 1.000 with an average of 0.955. Modest results of POD with a mean of 0.888 came from the Seattle region since this area has a complex topography and dense forests which hindered the original MODIS-SCA product snow detection. Meanwhile, regions in high elevations or frequent snow areas showed high POD. Since the VI algorithm retrieved ground states of cloud hindered pixels, it is also important to validate FAR of the cloud-free dataset. Across 50 validation scenes, the dataset yielded a reasonable average FAR of 0.179 with the highest FAR of 0.28 for 2 days, March 28, 2004 and March 29, 2016, in the Seattle region. The dataset's modest performance in the Seattle region is justified by the rapidly varying topography and dense needle leaf forests in this region. These characteristics impose difficulties in mapping snow for this region from satellites (Rittger, et al., 2013)(Hall & Riggs, 2007)(Liu, et al., 2004)(Liu, et al., 2008).



	POD	FAR
Merged MODIS	0.326	0.225
Cloud-free MODIS	0.923	0.202

Figure 2-7. A validation example over the Sierra Nevada region on March 21, 2003. First row, left to right: a base map with state lines in red and a cyan rectangle to indicate the validation region, a table showing Probability of Detection (POD) and False Alarm Ratio (FAR) results for the Merged MODIS image (second row, left) and the Cloud-free MODIS image (second row, center), validated with the Landsat7 Image (second row, right).

In Table 2-2, we also computed the percentage of clouds in the combined MODIS images from Terra and Aqua for each region to demonstrate the effect of the VI algorithm. In general, since we only selected Landsat images with little or no cloud percentage, the corresponding merged MODIS images are likely to contain less clouds (Table 2-2). However, days when cloud cover polluted heavily the merged MODIS images (e.g., January 30 and December 15, 2008 in Minneapolis or March 21, 2003 in Sierra Nevada), the VI algorithm effectively recovered snow boundaries to correlate well with the Landsat snow-cover maps. Figure 2-7 shows an example of the snow boundary recovered from cloud cover using VI to match with the Landsat image.

2.5.3. Bootstrap testing

In order to evaluate the accuracy of the mitigated filters and the VI algorithm, a cross-validation method is used. The validation process consists of three steps. First, a record of synthetic cloud-covered images was developed. This was performed by selecting combined MODIS snow images over CONUS that satisfies two conditions, namely 20% or less cloud cover area and 4% or more snow cover area. Subsequently, these images were overlaid by cloud cover extracted from MODIS images with a cloud pollution rate higher than 80% to create the synthetic record. Second, each filter and the VI algorithm was applied sequentially to remove clouds from images in the synthetic record. Third, the resulting image after each step was evaluated using the previously selected low cloud cover MODIS images. Table 2-3 shows the evaluation results for each filter and the VI algorithm for each day averaged across different cloud cover scenarios.

Table 2-3. Cross validation for each filter

Days	Filters	Percentage of Cloud Elimination	POD	FAR
Jan 24, 2006	Short-term temporal filter	0.071	0.985	0.016
	Elevation filter	0.021	0.983	0.017
	Spatial filter	0.105	0.981	0.017
	Long-term temporal filter	0.465	0.970	0.025
	VI algorithm	0.339	0.826	0.028
Jan 26, 2007	Short-term temporal filter	0.101	0.983	0.012
	Elevation filter	0.021	0.981	0.013
	Spatial filter	0.104	0.980	0.013
	Long-term temporal filter	0.397	0.981	0.013
	VI algorithm	0.376	0.953	0.017
Jan 31, 2009	Short-term temporal filter	0.135	0.984	0.013
	Elevation filter	0.025	0.981	0.014
	Spatial filter	0.080	0.980	0.014
	Long-term temporal filter	0.447	0.979	0.014
	VI algorithm	0.313	0.944	0.016
Mar 18, 2009	Short-term temporal filter	0.317	0.964	0.031
	Elevation filter	0.049	0.956	0.031
	Spatial filter	0.017	0.954	0.031
	Long-term temporal filter	0.530	0.949	0.031
	VI algorithm	0.087	0.867	0.035
Jan 18, 2013	Short-term temporal filter	0.234	0.988	0.010
	Elevation filter	0.037	0.988	0.010
	Spatial filter	0.050	0.986	0.011
	Long-term temporal filter	0.422	0.978	0.016
	VI algorithm	0.256	0.932	0.025

Mar 13, 2014	Short-term temporal filter	0.066	0.986	0.021
	Elevation filter	0.012	0.982	0.022
	Spatial filter	0.116	0.980	0.022
	Long-term temporal filter	0.453	0.965	0.049
	VI algorithm	0.353	0.917	0.051
Mar 6, 2015	Short-term temporal filter	0.146	0.986	0.012
	Elevation filter	0.035	0.981	0.013
	Spatial filter	0.090	0.979	0.014
	Long-term temporal filter	0.418	0.976	0.012
	VI algorithm	0.310	0.931	0.018
Mar 7, 2015	Short-term temporal filter	0.451	0.980	0.030
	Elevation filter	0.065	0.975	0.031
	Spatial filter	0.000	0.973	0.031
	Long-term temporal filter	0.377	0.971	0.028
	VI algorithm	0.107	0.933	0.029
Nov 23, 2015	Short-term temporal filter	0.107	0.985	0.013
	Elevation filter	0.020	0.980	0.014
	Spatial filter	0.094	0.979	0.014
	Long-term temporal filter	0.416	0.973	0.056
	VI algorithm	0.364	0.863	0.063

The metrics used for evaluation include POD, FAR and cloud removal ratio. The results shown in Table 2-3 demonstrate that each filter contributes to the cloud removal while maintaining high accuracy (i.e. POD and FAR). After complete removal of clouds, the average accuracy metrics of the images are 0.907 and 0.031 for POD and FAR respectively. Furthermore, it can be seen that the VI algorithm generally is the main factor in cloud removal with an average percentage of 0.278. Despite this high cloud removal ratio, the VI algorithm maintains high accuracy with an average change of -0.075 and +0.014 in POD and FAR respectively.

2.6. Case studies

The dataset produced in this study is useful in hydrological studies due to its adequate resolutions and validated accuracy as discussed in the previous section. In this section, we provide two simple applications of the dataset to serve as an example of the potential usages.

2.6.1. CONUS Snow Cover Extent

The annual average snow cover extent over CONUS, measured in million square kilometers, from both the merged MODIS and cloud-free snow cover datasets was compared. The comparison was performed from 2001 to 2016. Results show that statistics of snow cover using the merged MODIS could be substantially different from the cloud-free maps. As shown in Figure 2-8, the merged MODIS images (blue) show significantly less snow cover extent than the cloud-free dataset (red). The cloud-free dataset maintains an average snow cover of 1.342 million km^2 in the period (2001-2016) compared to 0.462 million km^2 from merged MODIS. This shows that using this dataset has major implications in quantifying the amount of snow for different hydrologic processes.

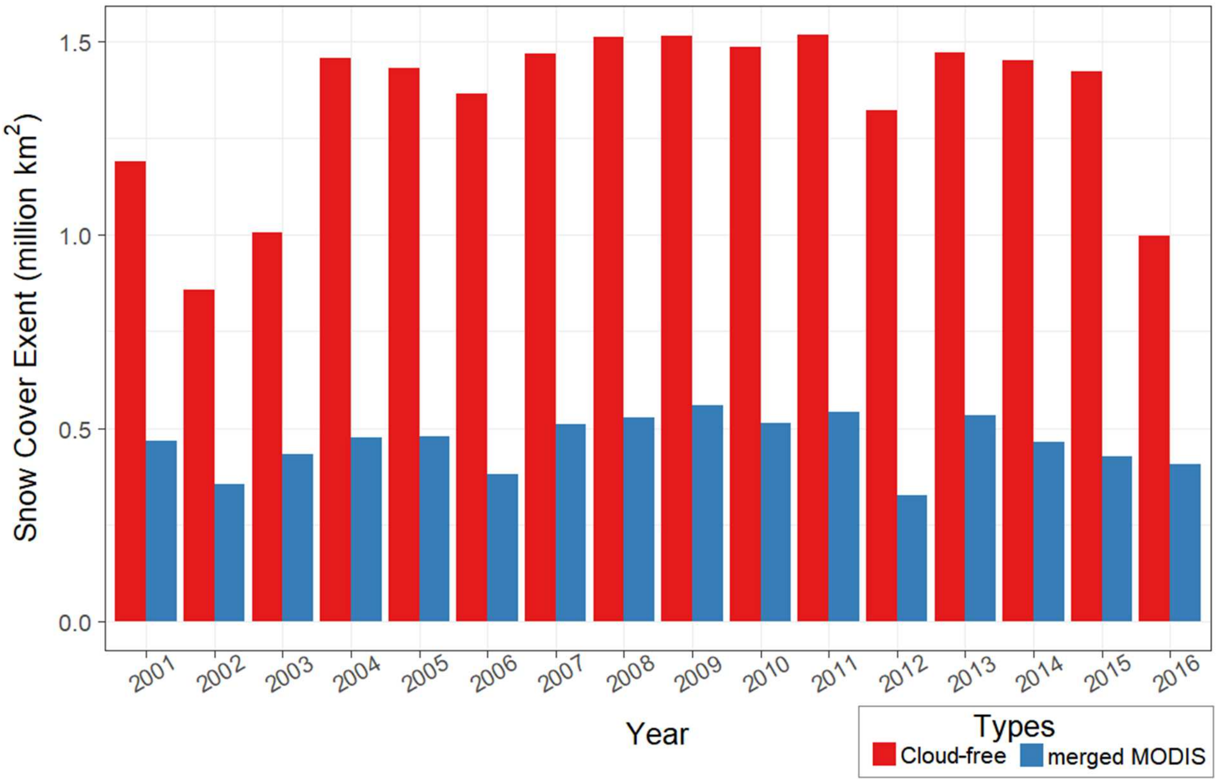


Figure 2-8. Comparison of annual average snow cover extent area over the CONUS computed from merged MODIS and the cloud-free snow cover product. The figure compares the annual average snow cover extent area over CONUS (in million km^2) as computed from (a) the cloud-free snow cover product developed in this study (red bars) and (b) merged MODIS product (blue bars). The time period of the comparison is from 2001 to 2016.

2.6.2. Annual number of snow days

In order to examine the differences in the two datasets beyond the yearly and monthly average of snow extent, we also studied the number of snow days over CONUS. From this perspective, we could compare the snow spatial distribution of the two products. As shown in Figure 2-9, during the period (2001 – 2016), the merged MODIS estimates 100 days as the annual number of snow days in the mountainous area of the Western US which has an altitude range of (910 - 1830 m). However, the cloud-free images estimate the number of snow days in the mountain states and the Sierra Nevada area as 160 to 200 days per year.

In two drought years, 2002 and 2003, there was considerably less snow days in the western mountains and the Midwest states, such as Colorado, Wyoming, and South Dakota, as illustrated by the cloud-free maps. Moreover, the images also show shorter snow seasons from the Northeast to the Northwest of CONUS.

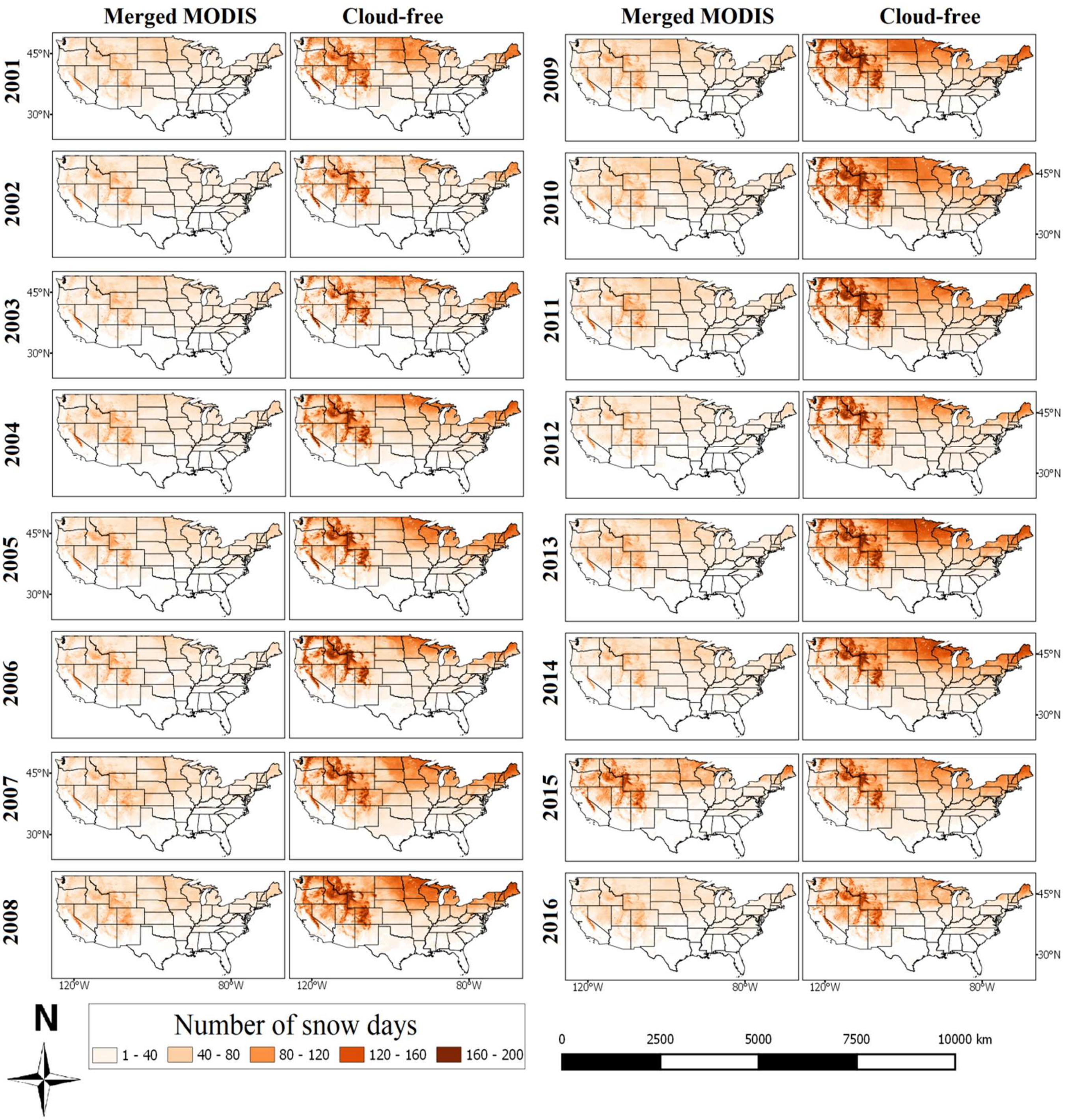


Figure 2-9. Comparison of the annual number of snow days over CONUS computed from merged MODIS and the cloud-free snow cover product. The figure compares the annual number of snow days over CONUS during the period 2001-2016 as computed from (a) Merged MODIS product and (b) the cloud-free snow cover product developed in this study.

The cloud-free dataset provides considerably different estimates regarding both the amount of snow as well as the number of snowy days. This has considerable implications in the results of hydrologic and climate modeling studies.

2.7. Chapter Summary and Conclusion

The VI algorithm introduced by (Xia, et al., 2012) is an automatic and robust approach to remove all of the clouds from the global MODIS snow cover area product (MODIS-SCA). However, the main drawback of their method is computational stability. When applying the original algorithm to larger scales (i.e., over CONUS with respect to the California region or 29 days instead of 5 days for calculation unit), the linear system of the method would likely become ill-conditioned, leading to undesirable "time-varying snow cover boundary" surfaces. This study proposed the MINRES iteration in place of the traditional LU decomposition in determining the 3-d implicit surface to overcome this computational deficiency. Experiments on calibrating both original and modified VI algorithms demonstrate that the latter was more stable, perform more efficiently than the former one. In this study, we fulfilled all experiments using the modified VI algorithm.

Before conducting further experiments, we evaluated the merged MODIS images (i.e., merged products from Terra and Aqua snow cover images of the same day) for their accuracy over the CONUS region using around 950 snow stations from the SNOTEL network since results of the VI algorithm directly inherit from the merged MODIS visible snow cover boundaries. Two conclusions can be made from the evaluation. First, when snow starts accumulating, due to the product spatial resolution (0.05°), if the snow proportion in one merged MODIS pixel is lower than 50%, the pixel will be classified as no snow. This resulted in a fairly high number of missed stations at the beginning of snow seasons. Second, the NDSI algorithm applied for retrieving the MODIS-SCA products required major adjustments in complex land surface areas, as discussed in (Hall, 2002). The problem resulted in systematic errors found in the Mountain states and the Sierra Nevada mountain regions of the merged MODIS images. In general, the average POD of the merged MODIS throughout the study period is 0.807, which is reasonably high, although we have to keep in mind that all the disadvantages of the merged MODIS maps will also affect the resulting accuracy of the VI algorithm.

Regarding the accuracy of the VI algorithm, the collection of constraint points on snow cover boundaries and the selection of number of consecutive days as a calculation unit play crucial roles. We applied the Douglas-Peucker method to choose appropriate representative constraint points of the snow cover boundaries to feed into the VI algorithm. Moreover, we carried out the test to determine an effective interpolation period. Four interpolation periods, namely 9, 19, 29, and 59 consecutive days, were taken into account. Results from different scenarios were validated using the SNOTEL network. In conclusion, the 29-day period has proven its efficiency in both accuracy and computational cost.

From the 3-d implicit surface created by the VI algorithm, results for any day in the period can be deducted. It is logical that the middle day of the period has the most information about snow cover boundaries and hence has the highest accuracy. However, other days in the middle to the end of the period have less information on snow boundaries. Therefore, how accurate are the images for these days? In this study, the experiment shows that the last day of the interpolation period had reasonable results in POD, FAR, and HK. This could be an important application for near real-time snow monitoring.

While the accuracy of the VI algorithm was tested using the bootstrapping method in Section 2.5.3, the accuracy of the whole dataset was validated against snow images derived from Landsat. Overall, the dataset produced in this study is useful in hydrological studies due to its adequate resolutions and validated accuracy as discussed in the previous section.

The future extension of this study involves applying the VI algorithm for near real-time global snow recovery. With the improvement in computational efficiency, the VI algorithm can be implemented in large Climate Zone regions (Kottek, et al., 2006) and then merged into the whole global map. Parallel processing of multiple regions at the same time should be adequate for near real-time global snow recovery.

Chapter 3. Cloud-free MODIS flood maps

3.1. Reviews of flood mapping from space

Flood mapping from space has been used extensively for more than two decades due to its global coverage and accuracy (Smith, 1997)(Frappart, et al., 2006)(Alsdorf, et al., 2007)(Karlsson & Arnberg, 2009)(Stephens, et al., 2012). Low Earth Orbiting satellites often image the entire Earth at a daily or weekly temporal resolution. Moreover, obtaining flood maps from direct satellite observations is relatively straightforward due to the distinguishable reflectance characteristics of water. Among various satellite products, the rapid response NASA MODIS surface reflectance product has proven its effectiveness in flood mapping and modeling(Brakenridge & Anderson, 2006). The product's high resolutions (250 m and daily in spatial and temporal resolutions) enable capturing occurrences of flood events. Various studies have used MODIS to study floods all over the continents: Africa(Wolski, et al., 2017), Asia (Ahamed & Bolten, 2017), Australia (Mohammadi, et al., 2017), Europe (Li, et al., 2015), North America (Zheng, et al., 2017), and South America (Houspanossian, et al., 2018).

However, cloud cover is a major problem which limits the product's usage. Clouds block satellites from capturing images, especially in events such as hurricanes or storms, whose aftermath is flood. Therefore, cloud removal using interpolation algorithms is of utmost importance for accurate flood mapping. Despite the essence of cloud removal, studies attempting to remove clouds from MODIS flood maps while preserving the product's original resolutions are still in an early stage. The only effort that we found to remove clouds from flood maps came from the same group of (Policelli & SlayBack, 2017) when they create composite flood maps of 2,3, and 7 days. While this approach is appropriate in some cases, it increases the product’s latency (Nigro, et al., 2014).

Given an assumption of water bodies persistent characteristic, we used VI to remove clouds from the MODIS flood maps. The resulting cloud-free flood maps and their derived information cloud-free water depth maps, will be continuous, reliable data sources for flood studies. More detail on the implementation of the VI will be discussed in the following sections.

3.2. Study domain and available data

The study domain is the Upper Mississippi River Basin (Figure 3-1) which has an area of 308,810.2 km^2 that ranges between 40^{o}20N and 47^{o}50N in latitude and from 97^{o}20W to 88^{o}12W in longitude. The basin includes 7 rivers (Mississippi River, Minnesota River, St. Croix River, Chippewa River, Wisconsin River, Rock River, and Cedar River), and 11 United States Geological Survey (USGS) stations (Brainerd - USGS 05242300, Henderson - USGS 0532700, St. Paul - USGS 05331000, Hasting - USGS 05331580, Winona - USGS 05378500, McGregor - USGS 05389500, Clinton - USGS 05420500, Joslin - USGS 05446500, Conesville - USGS 05465000, and Keokuk - USGS 05474500).

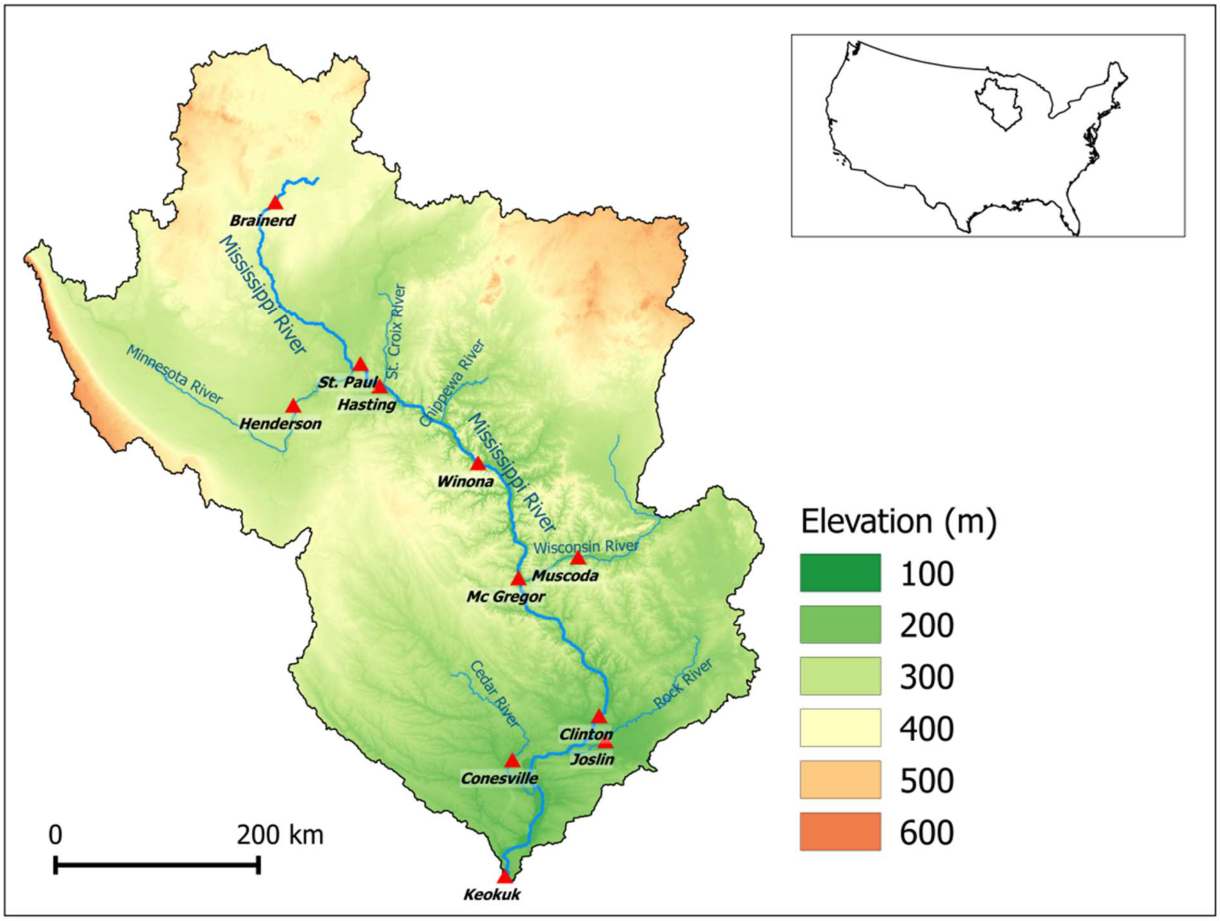


Figure 3-1. The Upper Mississippi River Basin

The Upper Mississippi River Basin is one of major basins in the United States which has a population of 30 million people. According to (UMRBA, 2006), over 60 percent of the basin is cropland or pasture. The general width is large, according to (UMRBA, 2006) "On the upper reaches near the Minnesota-Wisconsin border, the river's floodplain is between 1.5 and 5 kilometers (between 1 and 3 mi) wide. South of St. Louis, Missouri, the alluvial floodplain is approximately 80 kilometers (50 mi) wide.".

NASA MODIS products: MOD09GQ/MYD09GQ version 6 - MODIS Surface Reflectance Daily L2G Global 250m SIN Grid and MOD09GA/MYD09GA version 6 - MODIS Surface Reflectance Daily L2G Global 1 km and 500 m SIN Grid (Vermote, 2015) were used in this research. While the former products provide the surface spectral reflectance of MODIS bands 1 and 2 at 250 m spatial resolution, the latter products provide surface reflectance from MODIS bands 1 through 7 at 500 m and 1 km spatial resolutions. Both surface reflectance products are corrected for atmospheric conditions such as gasses, aerosols, and Rayleigh scattering. Band 1 and 2 of the GQ products were main inputs to produce 250 m resolution MODIS flood maps when band 3 through band 7, azimuth and zenith angles information from GA products were used to detect and remove cloud shadows from resulting flood maps. (For more details refer to sections 3.3.1.1 and 3.3.1.2).

The NASA's MWP uses MOD09GQ as input to produce flood maps. Water pixels are classified based on the reflectance ratio between NIR band and Red band, if the ratio is smaller than 0.7, pixels are classified as water.

In addition, (Policelli & SlayBack, 2017) employ the MOD44W (Carroll, et al., 2017) as the reference water layer, if a water pixel falls outside of the reference water layer, it will be considered as flood pixel. They also used a daily cloud mask from MODIS which has spatial resolution of 1-km to mask out cloud pixels before land/water discrimination.

Landsat 8 satellite images were used to validate MODIS flood maps and inundation maps produced by the model. Landsat 8 satellite captures the entire Earth every 16 days and provides high spatial resolutions (i.e. around 30m), better characterization of land cover images. The employed L1T product has been corrected radiometrically and geometrically using all two spacecraft's sensors, Ground Control Points (CGPs), and DEMs (Landsat 8 Users Handbook). Data from June to September each year (i.e., during flood seasons of the basin) which has less than 15% area covered by clouds was obtained from the U.S. Geological Survey website (https://landsat.usgs.gov).

3.3. Methodology

3.3.1. Classified MODIS flood maps

Flood inundation maps from satellites which are crucial for flood studies are limited by cloud obscuration. Figure 3-2 presents a sequence of steps to develop cloud-free flood maps. These steps are illustrated in the subsequent subsections.

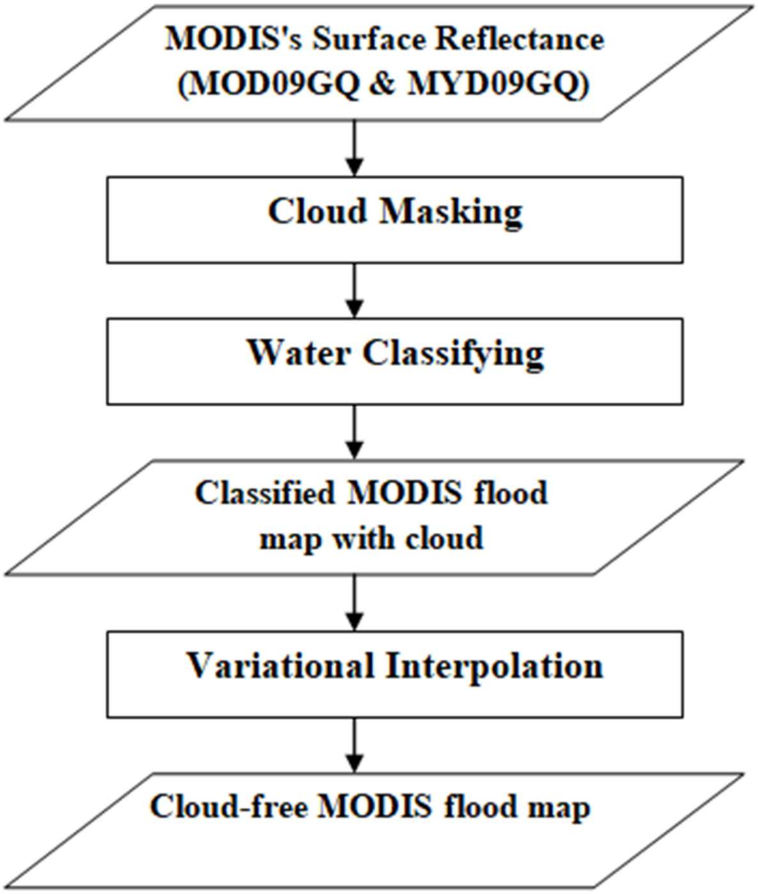


Figure 3-2. Flow chart of the cloud removal algorithm

3.3.1.1. Cloud masking

From the original MODIS reflectance products, we start by masking all the cloud pixels thus keeping only ground pixels in the satellite images. Due to the distinguishable reflectance of clouds at different wave bands (Figure 3-3), we propose a threshold of reflectance between 0.8 and 0.9 in the Red Band (Band 1) of MODIS to mask clouds. This threshold guarantees low false positive rates in identifying cloud pixels.

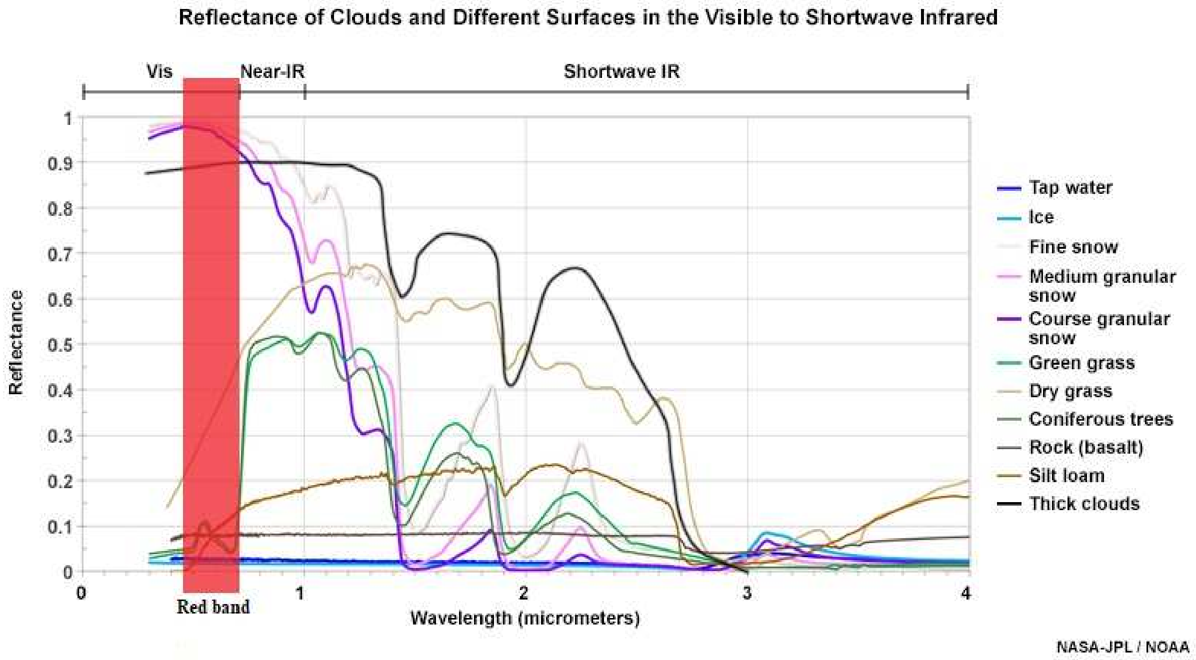


Figure 3-3. Cloud and surface reflectance in the Visible and Shortwave Infrared. (Image originally courtesy of NASA-JPL and NOAA)

Afterwards, the obtained images with non-cloud pixels are used as inputs for the next step of classifying water pixels.

3.3.1.2. Identifying of water pixels

We use Brakenridge's algorithm to detect water based on empirically derived reflectance ratio MODIS (Band 2 / Band 1) threshold(Brakenridge & Anderson, 2006). (Nigro, et al., 2014) reported that the classified MODIS flood maps generated by this algorithm captured 44% of the 53 flood events from various sites in the world in 2013 and 2014. The results demonstrated that the main source of error in the algorithm is that it flags most cloud shadows as water due to their similar reflectance values.

In order to overcome the problem of mistakenly identifying cloud shadows as water, we adopted a scheme that combines geometric and spectral approaches to determine cloud shadow locations (Luo, et al., 2008). According to their approach, from a cloud position on the image (x_{cloud}, y_{cloud}), the nadir projection (x_{nadir}, y_{nadir}) of cloud on the ground is determined based on height of cloud above the surface (h_c), viewing zenith and azimuth angles (θ_v and ϕ_v, clockwise from the true North), and the azimuth angle of the true North from the y axis (γ):

$$\begin{aligned} x_{nadir} &= x_{cloud} + h_c tan\theta_v \ \mathrm{SIN}(\phi_V+\gamma) \\ y_{nadir} &= y_{cloud} + h_c tan\theta_v \ \mathrm{COS}(\phi_V+\gamma) \end{aligned} \tag{3-1}$$

Then the projection of cloud shadow (x_{shadow}, y_{shadow}) is determined by solar zenith and azimuth angles (θ_s and ϕ_s):

$$\begin{aligned} x_{Shadow} &= x_{nadir} + h_c tan\theta_s \ \mathrm{SIN}(\phi_S+\gamma) \\ y_{Shadow} &= y_{nadir} + h_c tan\theta_s \ \mathrm{COS}(\phi_S+\gamma) \end{aligned} \tag{3-2}$$

The classified MODIS flood maps generated from this step are still polluted by cloud pixels that were masked in the first step. Next, we illustrate the final process to completely remove clouds.

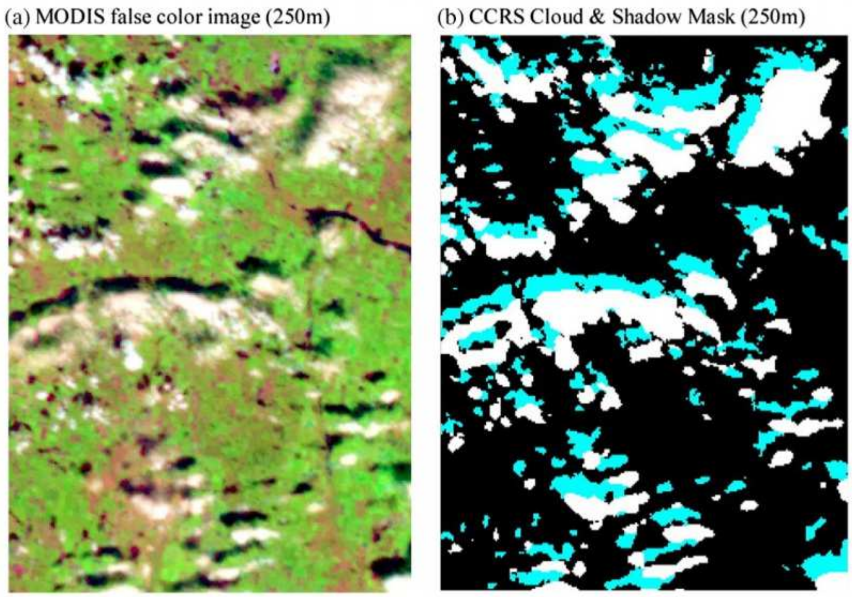


Figure 3-4. (a) MODIS image at 250 m spatial resolution, where some clouds and their shadows can be visually seen. (b) Cloud (white) and cloud shadow (blue) mask derived by the Canada Centre for Remote Sensing (CCRS) scheme. Example taken from (Luo, et al., 2008).

3.3.2. Cloud-free MODIS flood maps

3.3.2.1. Variational Interpolation

In the scope of this study, the VI algorithm is implemented to remove clouds from the classified MODIS flood maps. VI has been effectively used in (Xia, et al., 2012) and (Tran, et al., 2018) to remove cloud from MODIS's SCA product. These studies are based on the hypothesis that the interpolation algorithm works best with smooth three-dimensional implicit surfaces.

$$f(\vec{x})\begin{cases} > 0 & inside\ water\ body \\ = 0 & at\ water\ boundaries \\ < 0 & outside\ water\ body \end{cases} \tag{3-3}$$

Hence, an assumption about the dynamic property of water body boundaries has to be made. Given two reasons: (1) MODIS provides two images of a place in one day and, (2) its Red and Near-Infrared (NIR) bands have high spatial resolutions, we hypothesized that MODIS can capture gradual changes of water bodies. Thus, water bodies represented in three dimensional spaces have smooth implicit surfaces and can be approximated as a linear combination of radial-basis functions established at selected constraint points on the surface (Duchon, 1977), as equation (3-4).

$$f(\vec{x}) = 0 \Longrightarrow \sum_{i=1}^{N} w_i R(\vec{x} - \overrightarrow{x_i}) = 0 \tag{3-4}$$

Here, $\{w_i\}$ is a set of N weights and $\{R(\vec{x} - \overrightarrow{x_i})\}$ is a selected radial-basis function established at N constraints points. As in (Xia, et al., 2012) and (Tran, et al., 2018), we decided to use the thin plate function $R(.) = r^2 \log r$ with $r = \|\vec{x} - \overrightarrow{x_i}\|$ to present the radial-basis function.

With constraint points collected on water body boundaries in discrete time, the weights of those points can be computed by solving the linear system to create the implicit surface.

3.3.3. Determining the flooded area

The obtained images after applying the VI algorithm are completely cloud-free, in other words, each grid either represents water or land. To effectively identify flooded areas, we compare water grids in the map with the static water layers (MOD44W) obtained from NASA. Specifically, if a water grid falls outside the boundary of static water bodies, it is identified as a flooded grid.

3.4. Flood extent validation

In order to map floods from Landsat images, as in Xia et al. (2012), we followed a process outlined in (Brakenridge, et al., 2014). The process was based on Landsat 8 green and near-infrared bands (band 3 and band 5, respectively):

$$NDSI = \frac{\text{BAND } 3 - \text{BAND } 5}{\text{BAND } 3 + \text{BAND } 5} \tag{3-5}$$

When the NDSI is smaller than or equal to -0.04 the pixel is classified as water. By using this combination, the resulting flood products from Landsat avoid misclassifying cloud shadow as water(Brakenridge, et al., 2014). The resolution of the Landsat flood products is then up-scaled into 250-m and re-projected into the geographic coordinate system with the spheroid of WGS84 to match with the other two MODIS products resolution and projection.

For validation, we compared the MWP product with the developed cloud-free flood maps (Cloud-free). The results are shown in Table 3-1

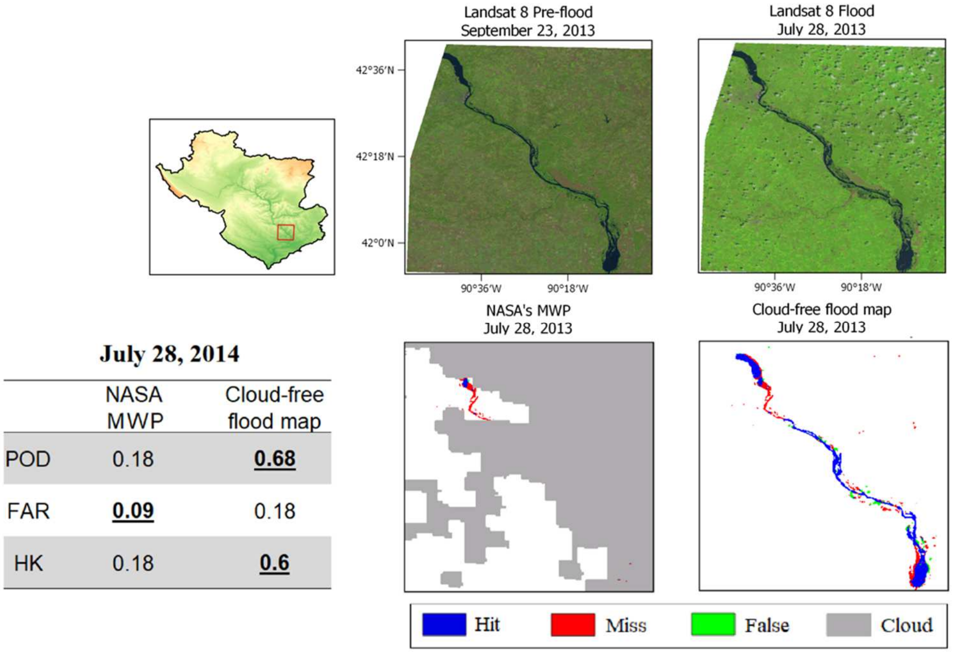
Table 3-1. Categorical validation of NASA MODIS Water Product (MWP) and the cloud-free flood map

Date	***POD***		***FAR***		***HK***		***Cloud percentage***
	MWP	*Cloud-free*	*MWP*	*Cloud-free*	*MWP*	*Cloud-free*	
06-07-2013	0.49	**0.57**	**0.35**	0.49	0.48	**0.51**	70.42
07-25-2013	0.26	**0.52**	0.07	**0.06**	0.26	**0.52**	79.15
08-01-2013	0.24	**0.33**	0.26	**0.07**	0.23	**0.33**	31.37
08-10-2013	0.27	**0.54**	0.30	**0.16**	0.27	**0.53**	60.61
08-26-2013	0.38	**0.43**	**0.13**	0.27	0.38	**0.43**	36.07
07-28-2014	0.18	**0.68**	**0.09**	0.18	0.18	**0.60**	80.15
08-13-2014	0.36	**0.50**	**0.06**	0.21	0.36	**0.50**	17.22
08-29-2014	0.17	**0.38**	**0.08**	0.27	0.17	**0.38**	73.78

Bold values represents better metric performance

Throughout two validation periods of 2013 and 2014, the cloud free flood extent derived in this study has higher POD and HK than the NASA's MWP. While the VI algorithm helped to increase the POD, it also falsely classified some cloud pixel as water which resulted in higher FAR of cloud free flood maps than NASA's MWP in more than half of the test cases. The main reason for this overestimation compared to MWP and explained by higher FAR values is the following. First, it is important to realize that this overestimation is relative to MWP. Second, since in MWP, all cloud grids are masked from the flood inundation map and taking into consideration that clouds will tend to cover most of the flooded area during a storm, the probability that the remaining grids be classified mistakenly as water is relatively low. On the contrary, in our product, there is a higher probability of mistakenly identified water grids. By careful examination of the cloud percentage masked in MWP, one will find that our product relatively overestimates flooded area in days where cloud percentage in MWP is very high.

Figure 3-5 visualizes the flood maps before and after removing clouds on July 28, 2014. When dense clouds prevented satellites from capturing images in NASA's MWP, VI recovered most of the Upper Mississippi river hindered by cloud.



July 28, 2014

	NASA MWP	Cloud-free flood map
POD	0.18	**0.68**
FAR	**0.09**	0.18
HK	0.18	**0.6**

Figure 3-5. Cloud-free MODIS flood map detection in the Upper Mississippi River Basin on July 28, 2014

3.5. Cloud-free MODIS water depth maps

3.5.1. Water depth creation

We adopted the Regression and Elevation-based Flood Information eXtraction (REFIX) method from (Schumann, et al., 2007) to derive water depth maps from the resulting cloud-free MODIS water extent. Here we only give some summary information, and refer the reader to read this article for further details.

The water depth creation procedure is illustrated in Figure 3-6. Firstly, given the essence of high resolution in flood mapping, we upscaled the original MODIS maps from original resolution of 250 m to 30 m in order to match with the SRTM elevation ones.

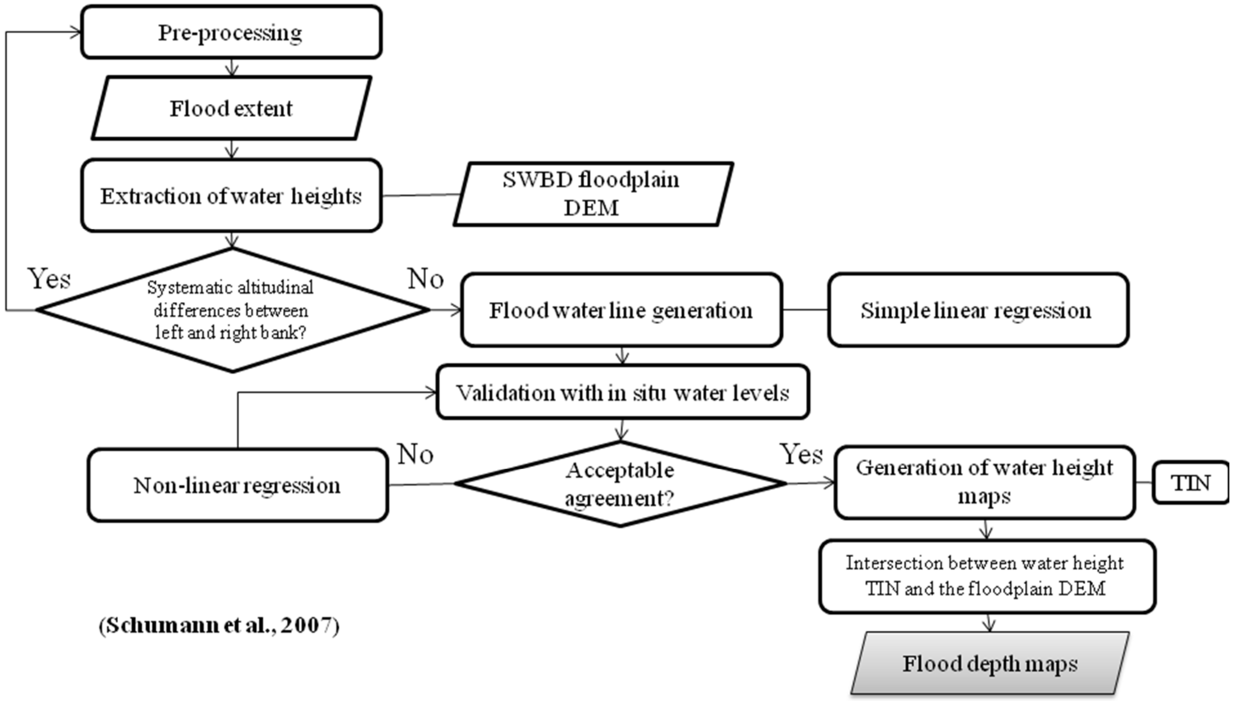


Figure 3-6. Water depth flowchart -- REFIX methodology(Schumann, et al., 2007))

For the REFIX best performance, information of flood boundaries must be accurate (Schumann, et al., 2007). Hence, we also combined the cloud-free MODIS maps with an existing water body data from SRTM project (SWBD; USGS). The cloud-free MODIS maps are main factors in the combined products in detecting flooded areas.

Secondly, the DEM for floodplain is used to extract water heights for both riverbanks at critical sections of the Mississippi river (Schumann, et al., 2007). It is worth noting that, if there is a systematic altitudinal difference between the left and right riverbank, the pre-processing step need to be reviewed. Next, if all seems adequate (Figure 3-7), we will generate a water line for that day based on a simple liner regression model based on a relation between water height and distance downstream. The linear regression is formulated as equation (3-6):

$$H = a.d + b \qquad (3\text{-}6)$$

Where H is an estimated water height (m), a is a slope of the regression line which based on the stream characteristics (m/m), d is a downstream distance (m), and b is an intercept (m).

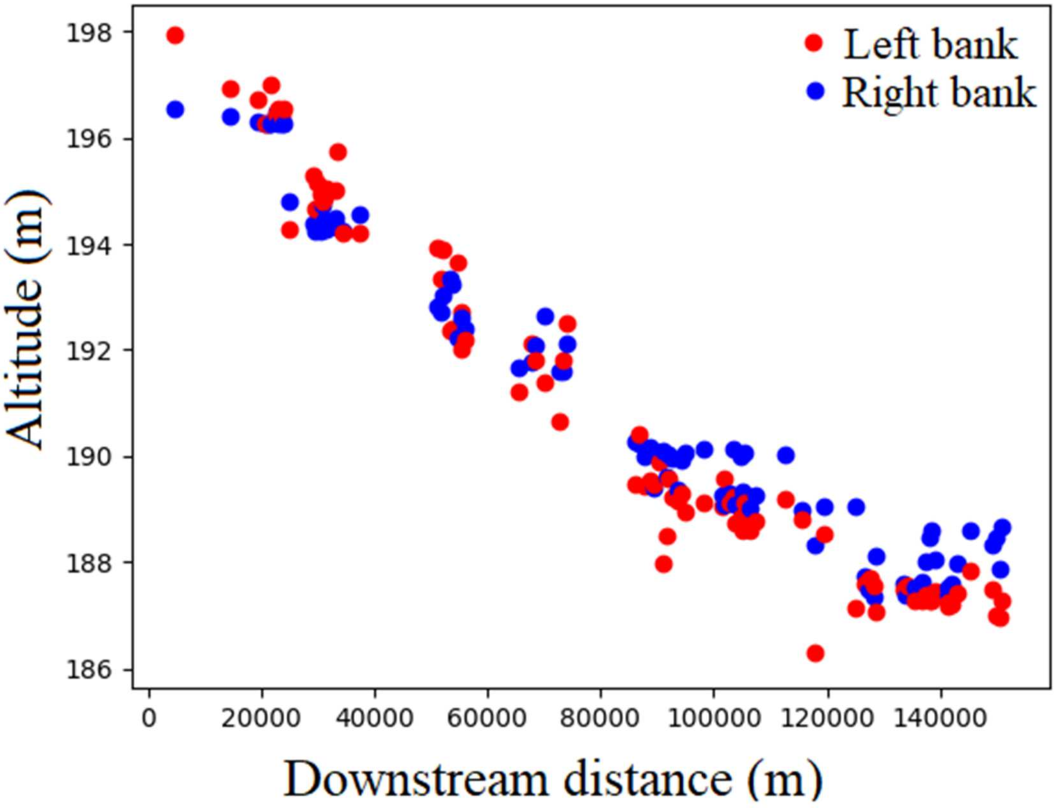


Figure 3-7. Scatter plot of the DEM-derived water levels for both riverbanks for one part of the Mississippi river flows from Winona to McGregor in June 26, 2014

We decided to implement the REFIX method for one part of the Mississippi river from Winona to McGregor for two reasons: (1) although having 11 USGS stations in the basin, only two of them have measurements of the river water level (i.e. Winona and McGregor stations), and (2) through experiments, we found that the stream characteristics for this river part are adequate for deriving an almost linear relationship between water height and downstream distance as equation (5). While the former reason is necessary for validating the estimated water heights, the latter reason is prone more to a practical aspect.

Once the validation results with observed water levels are satisfied, the Triangulated Irregular Network (TIN) mesh will be used to generate the water height map. (Schumann, et al., 2007) stated that: the TIN approach helps both keeping water level horizontal at each cross section and respecting changes in river flow directions. Finally, we subtract the DEM from the water height map to obtain the actual water depth map.

The whole process is automated to ensure creating water depth map for each corresponding cloud-free MODIS flood map. Since we in cooperated in situ water levels validation in creating satisfactory water depth maps, in next step we will only demonstrate the result of the water depth creation process during a flood event from June 26 to July 7, 2014.

3.5.2. Water depth results

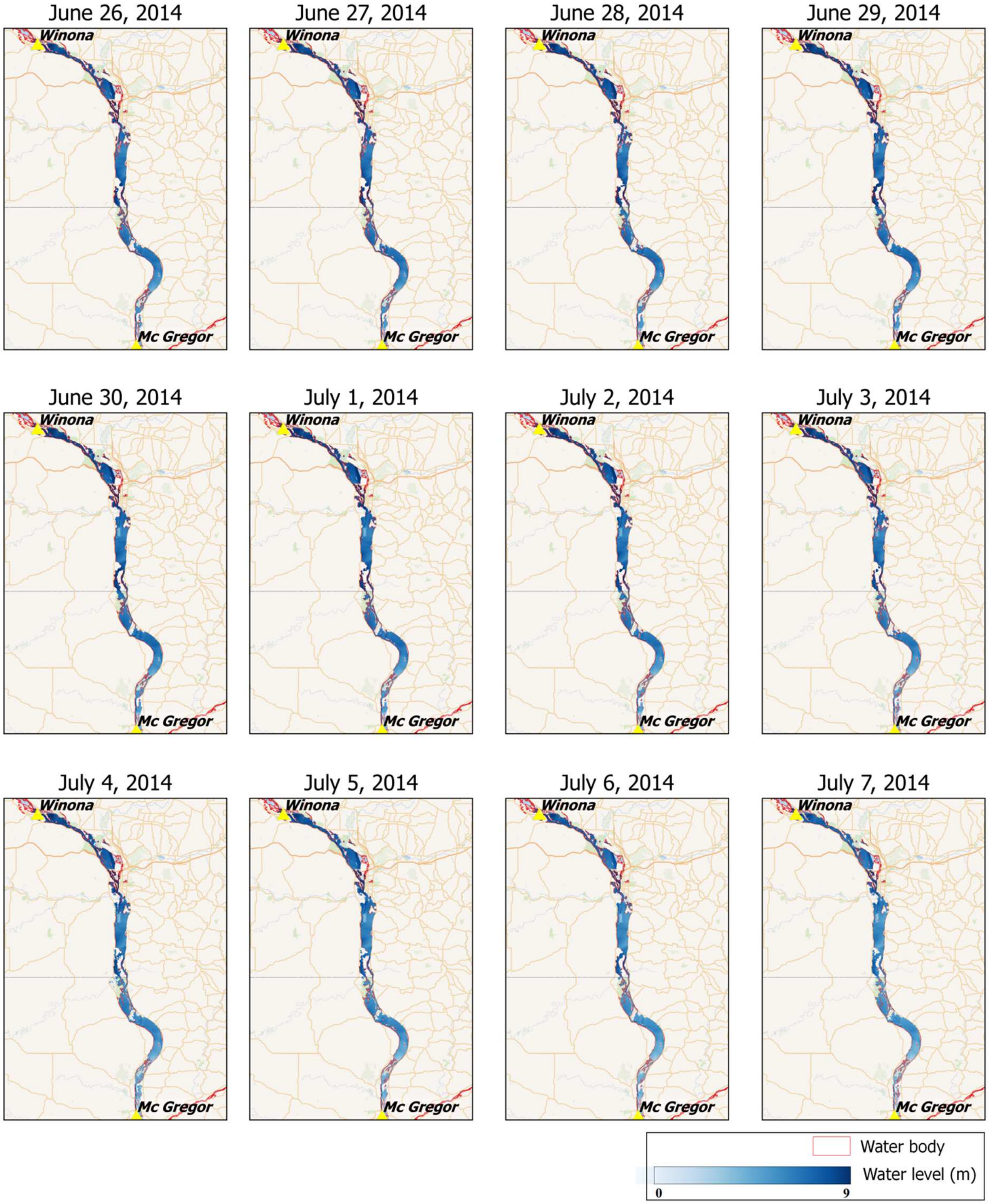

June 26, 2014
June 27, 2014
June 28, 2014
June 29, 2014
June 30, 2014
July 1, 2014
July 2, 2014
July 3, 2014
July 4, 2014
July 5, 2014
July 6, 2014
July 7, 2014

Figure 3-8. Water depth level for one part of the Mississippi river flows from Winona to McGregor from June 26 to July 7, 2014

Figure 3-8 shows the dynamic change of both water extent and water depth over the period of flood. At the start day of the flood period, June 26, 2014, the average water depth for the river part was at the flood stage level of 3.96 m and there was not any flood excess from the water body. Along the period, as the water depth increased, number of places that excess normal water body was also increased. These flooded places were mostly in narrow channels and floodplain areas which lie between two river branches. At the peak day of July 2, 2014, when the water depth at Winona station was 4.72 m (Figure 3-9), there were multiple places that the water extent excess the water body. Moreover, we can clearly see the water depth decreased (the blue color became less) from July 2 to July 7, 2014 in the lower part near McGregor station.

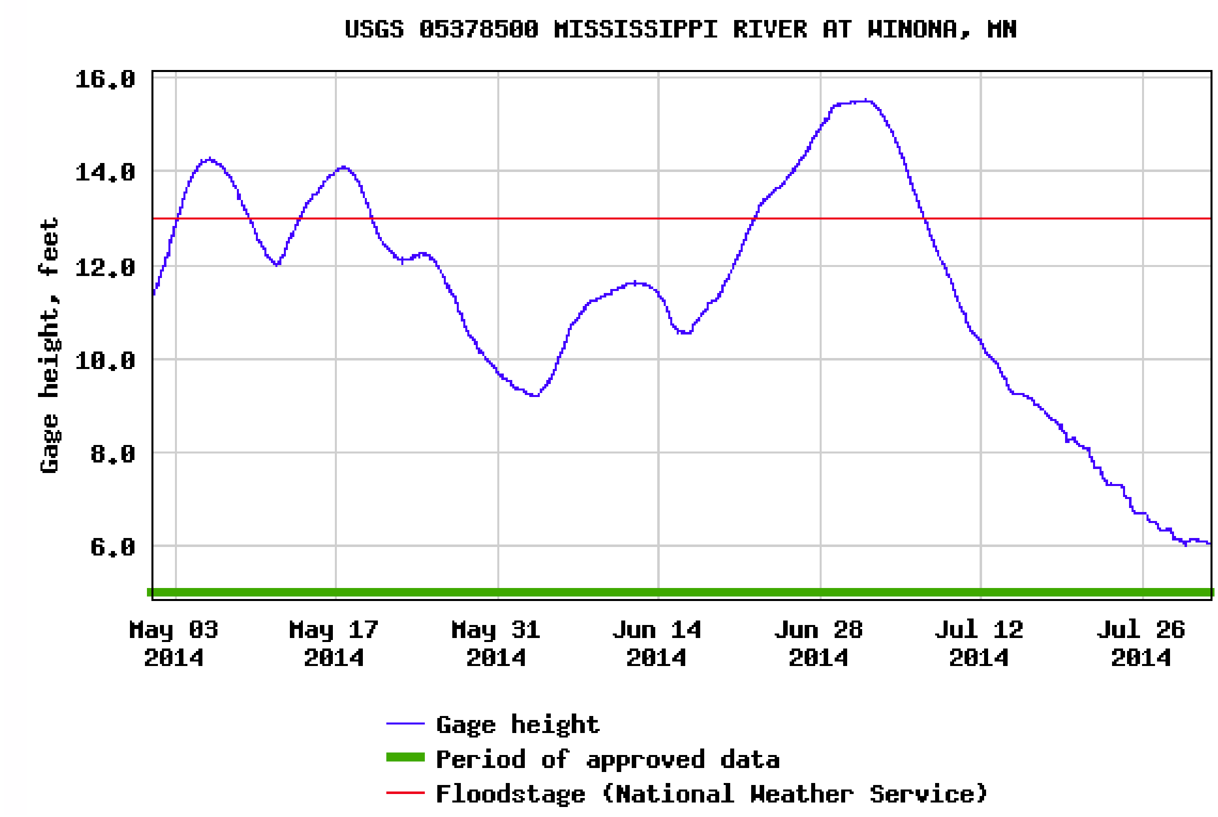


Figure 3-9. Gage height in feet at Winona station from May 1 to July 31, 2014

3.6. Chapter Summary and Conclusion

Through comparison against Landsat 8 observations, we found that the VI algorithm can effectively remove clouds from MODIS flood maps. After comparing with NASA's MWP which also uses the same water detection algorithm, VI removed most of the cloud-hindered parts of the Upper Mississippi River thus providing information about the dynamics of flood events. In the validation period, the recovered MODIS flood maps have on average higher scores of POD and HK, 0.49 and 0.48, respectively, in comparison with ones from NASA's MWP of 0.29 and 0.29, respectively.

While VI correctly estimated the states of most cloud pixels, it also falsely classified some cloud pixels as water which caused the final products often to have higher FAR than the original ones. This problem is under further investigation and might possibly be mitigated by incorporating new geographical information as another dimension of the VI algorithm to improve its interpolation ability. Moreover, during our experiment in some scenarios, if cloud cover occurs over a long period (e.g., more than 5 days), VI cannot gather enough information to reconstruct water bodies accurately. We will propose a quality control screen which filters out parts of rivers or streams where clouds remain there over a certain amount of time. The screen masks these areas as missing data to ensure the overall accuracy. The missing information will then be retrieved by model simulations.

The VI method has already proven its capability in cloud-free snow cover and flood extent recovery. Hence, we can implement the VI in other remote sensing fields such as ocean color or water-mass distribution to break the major barrier, cloud obscuration, which limits the potential of satellite observations.

The REFIX method demonstrated its capability in deriving accurate water depth from water extent in river Alzette, Luxembourg and river Mosel, France (Schumann, et al., 2007). We adopted the method to produce the cloud-free MODIS water depth maps from the cloud-free MODIS water extent maps. The results seem to reflect dynamic changes in river height during a flood event from June 26 to July 7, 2014 for one part of the Mississippi River flows from Winona and McGregor. Since there were only two stations provide water height measurements which were used to decide the function between downstream distance and water height, it was difficult to validate the resulting water depth maps.

Finally, it should be noted that the MODIS water depth maps could be validated using the future Surface Water and Ocean Topography (SWOT) direct observations of water elevation. SWOT is an anticipated, advanced satellite mission which is scheduled to be launched in 2021. SWOT uses a state-of-the-art Ka-band radar interferometer (KaRIn) to provide a set of water body surface elevations with expected high height accuracy ((Rodriguez, 2015); (Biancamaria, et al., 2016)).

Chapter 4. Improving hydrologic modeling using cloud-free flood maps

4.1. Overview of flood modeling studies

While being an important part of the hydrological cycle, floods are also considered as one of the most destructive natural disasters. Hence, flood modeling has emerged since the 1970s within the research community (Teng, et al., 2017). (Teng, et al., 2017) recently provided a thorough review of up-to-date flood inundation models. In this chapter, we only summarize main ideas from their study and refer readers to their paper for more detail.

Hydrodynamic models, including one-dimensional (1D)(DHI, 2003) (Brunner, 2016), two-dimensional (2D)(Sanders, et al., 2010)(Moulinec, et al., 2011)(Nguyen, et al., 2016), and three-dimensional (3D)(Vacondio, et al., 2012)(Prakash, et al., 2014), models "simulate water movement by solving equations derived from applying physical laws to fluid motion with varying degrees of complexity" (Teng, et al., 2017). They are the most widely used tools to simulate detailed flood dynamics thanks to the versatility in adjusting input in order to "investigate the impact of changes in initial conditions, boundary conditions". While 1D models are computationally efficient, (Teng, et al., 2017) pointed out a number of their drawbacks: "the inability to simulate lateral diffusion of the flood wave, the discretization of topography as cross sections rather than as a continuous surface and the subjectivity of cross-section location and orientation". On the other hand, 2D and 3D models are able to simulate inundation with high accuracy, however, they are computationally expensive. Recent advances in 2D model studies focused on improving the computational efficiency by simplification of the two-dimensional shallow water equations (Bates & De Roo, 2000), shared memory multiprocessing programming (Sanders, et al., 2010)(Flamig, et al., 2015), and graphics processing units (GPU) utilization (DHI, 2012).

Furthermore, a main uncertainty source of those hydrodynamic models comes from assumptions of the physical characteristics of the rivers and floodplains (Andreadis, et al., 2007)(Munier, et al., 2015). Two approaches can be taken to reduce the effect of model errors, namely, calibration of hydrological parameters and assimilation of observations. More details about these two approaches will be discussed in Section 4.3.2.

The scope of this study is twofold which includes the use of cloud-free products developed from Chapter 3. First, we used the cloud-free MODIS flood maps to calibrate inundation parameters of a hydrological model in order to produce accurate inundation maps. Second, we used the derived cloud-free water depth to correct the states of a routing model in order to have a better forecast of downstream discharge.

4.2. Study domain and available data

The study domain is the Upper Mississippi River basin (Figure 4-1) as in Chapter 3.

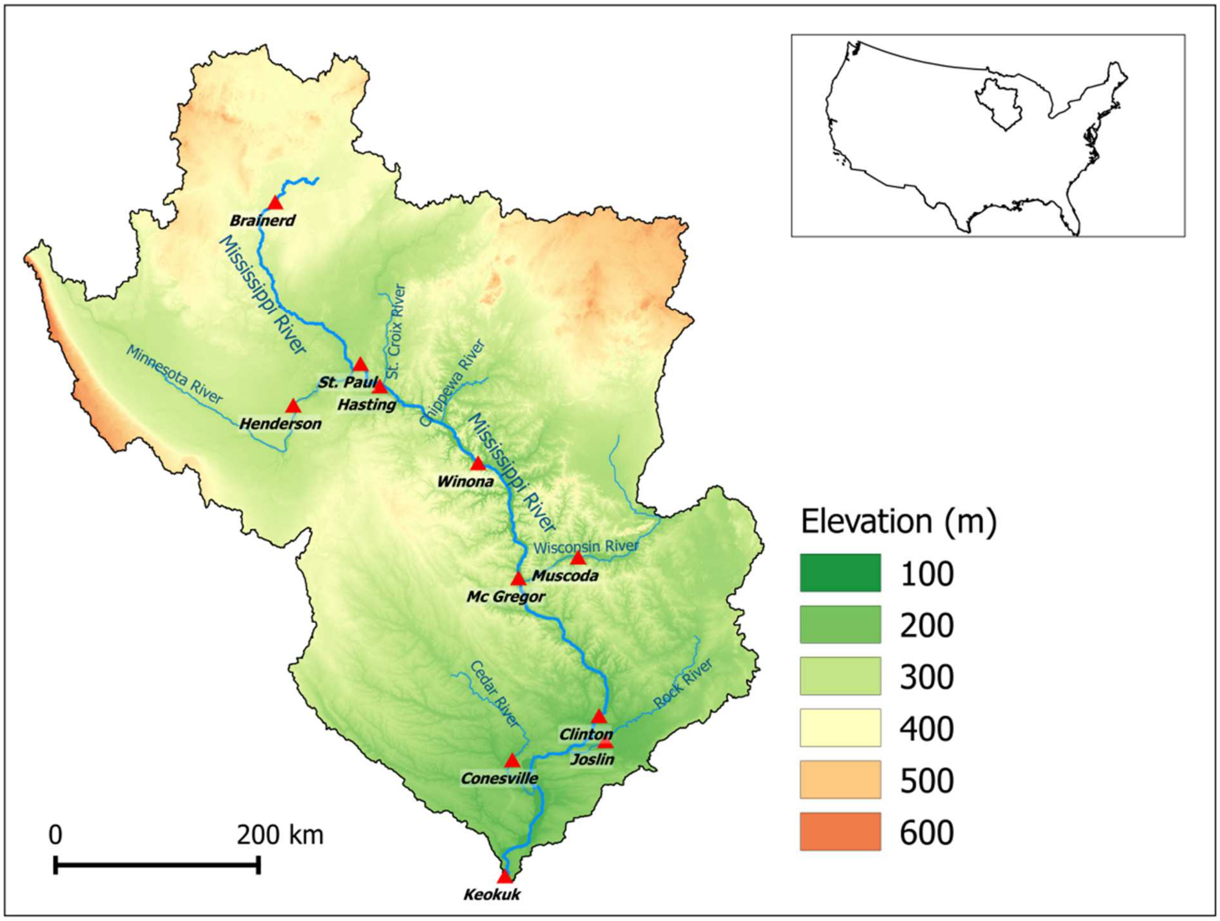


Figure 4-1. The Upper Mississippi River Basin

Streamflow data were available in both hourly and daily time scales, quality-controlled by USGS (USGS 2016). Meteorological data, precipitation and evapotranspiration were from the National Oceanic and Atmospheric Administration (NOAA) and the Famine Early Warning System Network (FEWSNET). NOAA's multi-sensor precipitation data Stage IV (Lin, 2011), produced by integrating data from 12 Contiguous U.S. River Forecast Centers, has 0.04^o spatial resolution and hourly and daily temporal resolutions. Evapotranspiration data, on the other hand, is important for drought and water resource management but not for rainfall-driven flooding events.

Hence, to reduce computing resources, we chose the monthly mean evapotranspiration data from FEWSNET with spatial resolution of 1°, calculated on a spatial basis using the Penman-Monteith equation.

The elevation data used in this chapter is from the Shuttle Radar Topography Mission (SRTM) version 3.0 published in 2013. The product has 30 m spatial resolution with all voids eliminated.

4.3. Methods

4.3.1. Hydrological model framework

The Ensemble Framework for Flash Flood Forecasting (EF5;(Flamig, et al., 2015)) is a C++ based framework for conducting research and operational distributed hydrologic simulations. EF5 consists of multiple components including water balance, river routing and inundation models that can be jointly implemented to simulate flood processes. The whole framework takes advantage of multiple computing cores and parallel computing strategies for faster simulations.

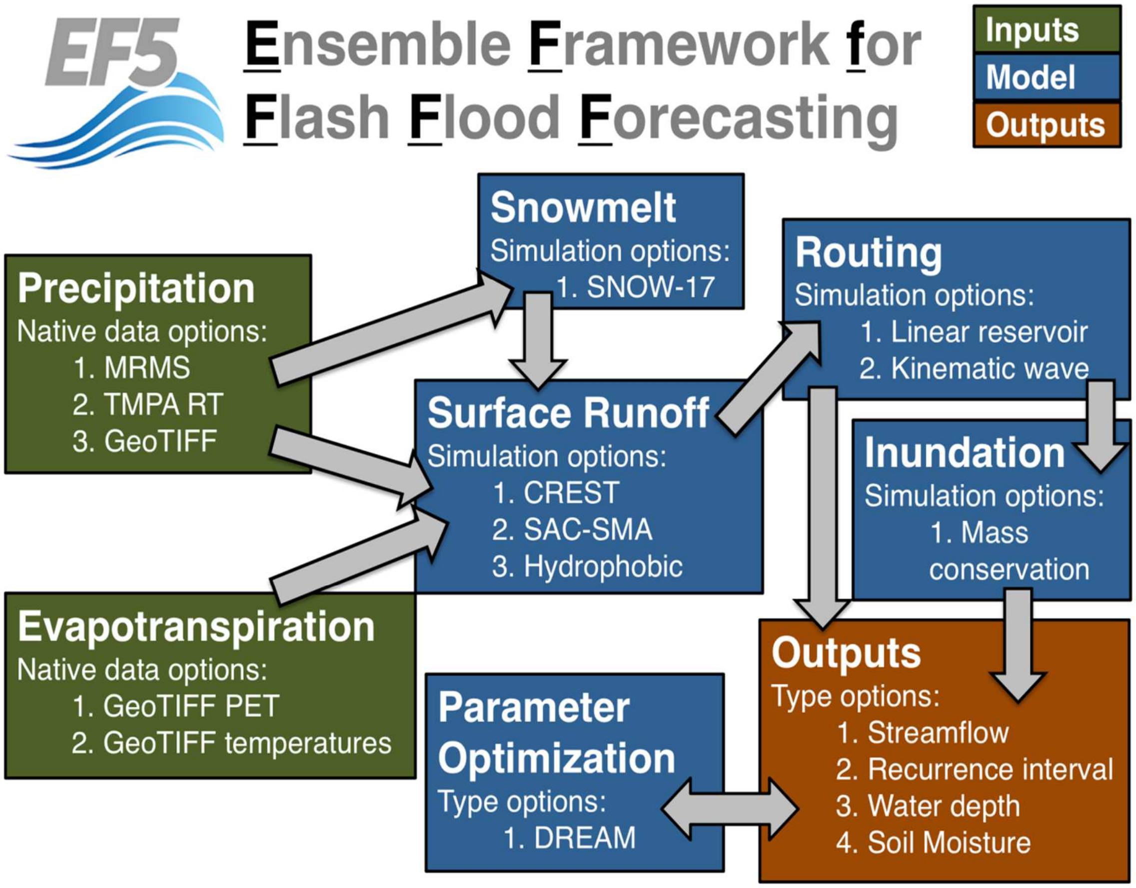


Figure 4-2. EF5 framework. (Image originally courtesy of University of Oklahoma)

Three models of EF5 have been used for this study: The Coupled Routing and Excess STorage (CREST) (Wang, et al., 2011) as a water balance model, the kinematic wave routing (Lighthill & Whitham, 1955) as a routing model and, the simple mass-conserving inundation as an inundation model. They are summarized below.

The CREST is a distributed water balance model developed to simulate the spatial and temporal variation of land surface, and subsurface water fluxes and storage by cell-to-cell simulations. CREST's distinguishing characteristics include: (1) distributed rainfall–runoff generation and cell-to-cell routing; (2) coupled runoff generation and routing via three feedback mechanisms; and (3) representation of sub-grid cell variability of soil moisture storage capacity and sub-grid cell routing (via linear reservoirs).

The kinematic wave routing is used in "many models of open channel flow routing of runoff hydrographs" (Hromadka & DeVries, 1988) because of its simplicity and accuracy. Given solving the momentum and continuity equations for the 1-D full shallow water equations (4-1) and (4-2):

$$\underbrace{\frac{\partial Q_x}{\partial t}}_{local\ acceleration} + \underbrace{\frac{\partial}{\partial t}\left(\frac{Q_x^2}{A}\right)}_{convective\ acceleration} + \underbrace{gA\frac{\partial (h+z)}{\partial x}}_{water\ slope} + \underbrace{\frac{gn^2Q_x^2}{R^{4/3}A}}_{friction\ slope} = \left(\frac{Q}{\alpha}\right)^{\frac{1}{\beta}} \qquad (4\text{-}1)$$

$$\frac{\partial A}{\partial x} + \frac{\partial Q_x}{\partial x} = 0 \qquad (4\text{-}2)$$

where Q_x is volumetric flow rate in the x Cartesian direction, A is the cross sectional area of flow, h the water depth, z the bed elevation, g gravity, n the Manning’s coefficient of friction, R the hydraulic radius, t time and x the distance in the x Cartesian direction. The kinematic wave routing only includes friction slope and water slope, and water slope only includes bed gradient (dz/dx). The water slope neglects local and convective accelerations and free surface gradient (dh/dx).

The last model, mass-conserving inundation is a simple model that computes inundation based on stream flows and cross-sectional area and has been used widely in many studies ((Bates & De Roo, 2000); (Horritt & Bates, 2002); (Bates, et al., 2003); (Jian, et al., 2009)).

Computing cross-sectional area in mass-conserving inundation model is defined as follows:

$$A = \left(\frac{Q}{\alpha}\right)^{\frac{1}{\beta}} \tag{4-3}$$

Where A is the cross-sectional area (m^2), Q is the discharge (m^3/s), α and β are parameters governing routing and overland flow.

Parameters of the three models are described in Table 4-1.

Table 4-1. EF5 parameters

Model	***Parameter***	***Unit***	***Description***
CREST	WM	mm	Maximum soil water capacity
	B	--	Exponent of the variable infiltration curve (VIC)
	IM	--	Impervious area ratio
	KE	--	Multiplier to convert between input Potential Evapotranspiration (PET) and local actual Evapotranspiration (ET)
	FC	mm/h	Soil saturated hydraulic conductivity
	IWU	%	Initial value of soil water
Kinematic	TH	--	Threshold for how many cells must drain into a cell for it to become part of a river in the model
	UNDER	--	Interflow flow speed multiplier

	LEAKI	%	Amount of water leaking out of the interflow reservoir at each time step
	ISU	mm	Initial value of the interflow reservoir
	ALPHA	--	Multiplier in the equation $\mathrm{Q} = \alpha \mathrm{A}^{\beta}$ governs routing
	BETA	--	Exponent in the equation $\mathrm{Q} = \alpha \mathrm{A}^{\beta}$ governs routing
	ALPHA0	--	Multiplier in the equation $\mathrm{Q} = \alpha \mathrm{A}^{0.6}$ governs overland flow
Mass conservation	ALPHA	--	Multiplier in the equation $\mathrm{A} = \left(\frac{\mathrm{Q}}{\alpha}\right)^{\frac{1}{\beta}}$ used for computing cross-sectional area
	BETA	--	Exponent in the equation $\mathrm{A} = \left(\frac{\mathrm{Q}}{\alpha}\right)^{\frac{1}{\beta}}$ used for computing cross-sectional area

4.3.2. Overview of error correction methods

4.3.2.1. Parameter optimization/ Model calibration

The Shuffled Complex Evolution algorithm - University of Arizona (SCE-UA) (Duan, et al., 1992) has been used widely in water resources management ((Sorooshian, et al., 1993)(Yapo, et al., 1996)(Madsen, 2000)(Toth, et al., 2000)(Eckhardt & Arnold, 2001)(Liong & Atiquzzaman, 2004)(Ajami, et al., 2004)(Lin, et al., 2006)(Barati, et al., 2014)(Yang, et al., 2015)), as well as other fields of study, such as pyrolysis modeling

((Ding, et al., 2016)(Hasalova, et al., 2016)) and artificial intelligence (Yang, et al., 2017). (Naeini, et al., 2018) summarized the SCE-UA as:

The SCE-UA framework employs Nelder-Mead simplex (Nelder & Mead, 1965) technique along with the concept of controlled random search (Price, 1987), clustering (Rinnooy Kan & Timmer, 1987), competitive evolution (Hollan, 1975) and complex shuffling (Duan, et al., 1993) to offer a global optimization strategy. By employing these techniques, the SCE-UA algorithm provides a robust optimization framework and has shown numerically to be competitive and efficient comparing to other algorithms for calibrating rainfall-runoff models ((Gan & Biftu, 1996)(Wagener, et al., 2004)(Wang, et al., 2009)(Beven, 2012)).

4.3.2.2. Data assimilation

The Kalman Filter (Kalman, 1960) and the Ensemble Kalman Filter (EnKF) (Evensen, 1994) are summarized in numerous studies, including (Biancamaria, et al., 2011) and (Munier, et al., 2015). Hence, we refer the reader to the original papers for further details, (Munier, et al., 2015) summarized those two methods as:

The Kalman Filter (Kalman, 1960) is a sequential data assimilation scheme widely used in hydrological sciences. At each time step, new observations are combined with the model outputs derived from the simulated state (forecast) to compute an update state (analysis). The latter is obtained by optimally accounting for observation and model errors. Here the system state includes water levels, while observations are the cloud-free MODIS water levels. The original Kalman filter, based on linear systems and Gaussian errors, has been extended to nonlinear models. One of the most popular extension is the EnKF described by (Evensen, 1994). It is based on a Monte Carlo approach to approximate the error covariance matrix, which usually is not known:

$$P_e = \overline{(A - \bar{A})(A - \bar{A})^T} \tag{4-4}$$

where A is the forecast state ensemble, P_e the ensemble covariance matrix, and the superscript T the transposition operator. Each column of matrix A represents a specific ensemble member state.

We implemented the square root analysis algorithm developed by (Evensen, 2004) in order to avoid measurements perturbations. With the same notation as above, the analysis step is given by:

$$A^o = A + P_e H^T (H P_e H^T + R)^{-1} (D - HA) \tag{4-5}$$

where A^o is the analysis state ensemble, D the observation matrix (columns are identical and represent the observation vector), H the observation operator, and R the measurement covariance matrix. Note that R is always assumed diagonal, i.e., each observation error is assumed independent.

4.3.3. Calibration scheme

4.3.3.1. Surface runoff and routing parameters calibration using stream flow gauges

Calibration of surface runoff and routing models is necessary to ensure the accuracy of the flood inundation model. The year 2012 has been selected as the calibration period with calibration cascading from upstream to downstream stations. We used NSE as the objective function in the SCE-UA (Duan, et al., 1992) algorithm. The calibration scheme is described in Figure 4-3.

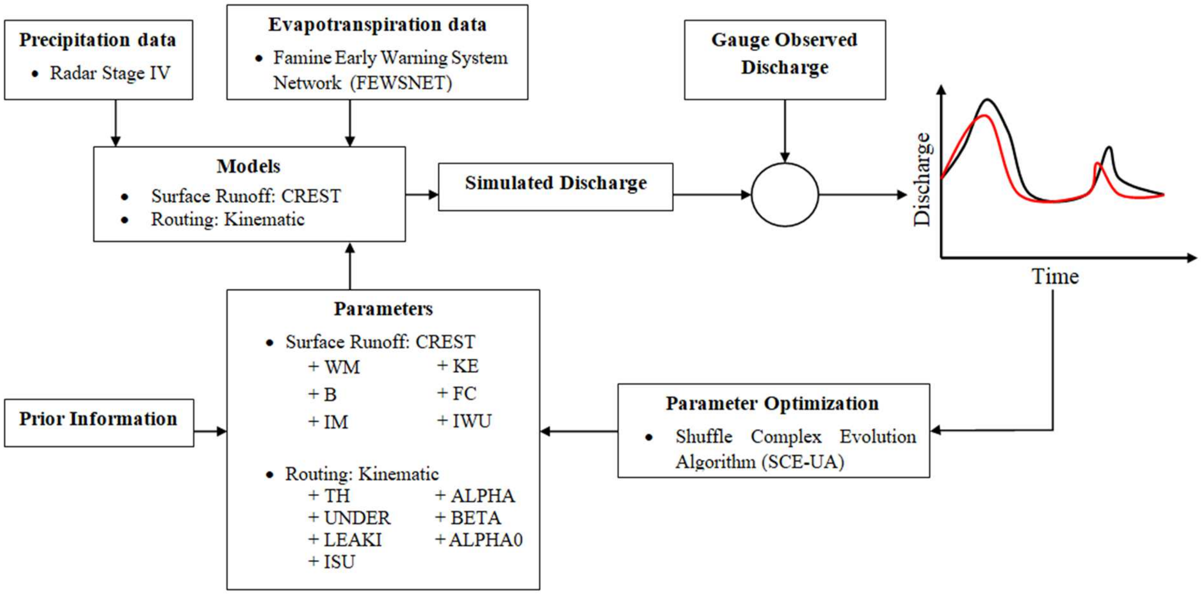


Figure 4-3. Surface runoff and routing parameters calibration

4.3.3.2. Inundation parameters calibration using cloud-free flood maps

In a four-month period from June to September 2012, we calibrated inundation parameters using the cloud-free flood maps as benchmarks. Inundation maps simulated by the model are classified into water/no water maps. One thing to note here, since the inundation model is solely based on the altimetry data, parameters which govern overland flows are not uniformly distributed (i.e. each floodplain region is assigned by specific parameters). Model simulations tend to overestimate inundation areas. Thus, when we validated these maps with the observed benchmarks, the FAR scores were frequently high which indicates a large number of false pixels. Hence, the objective of this calibration process is to minimize the FAR of the output water depth maps.

With regard to the calibration of inundation models, spatial distribution of stations plays an important role since the total flood map is constrained by the flooding extent of each event. Therefore, we took into account all parameters of the basin stations in the optimization scheme so that the "observed" cloud free flood maps can help to reduce simulated overflows.

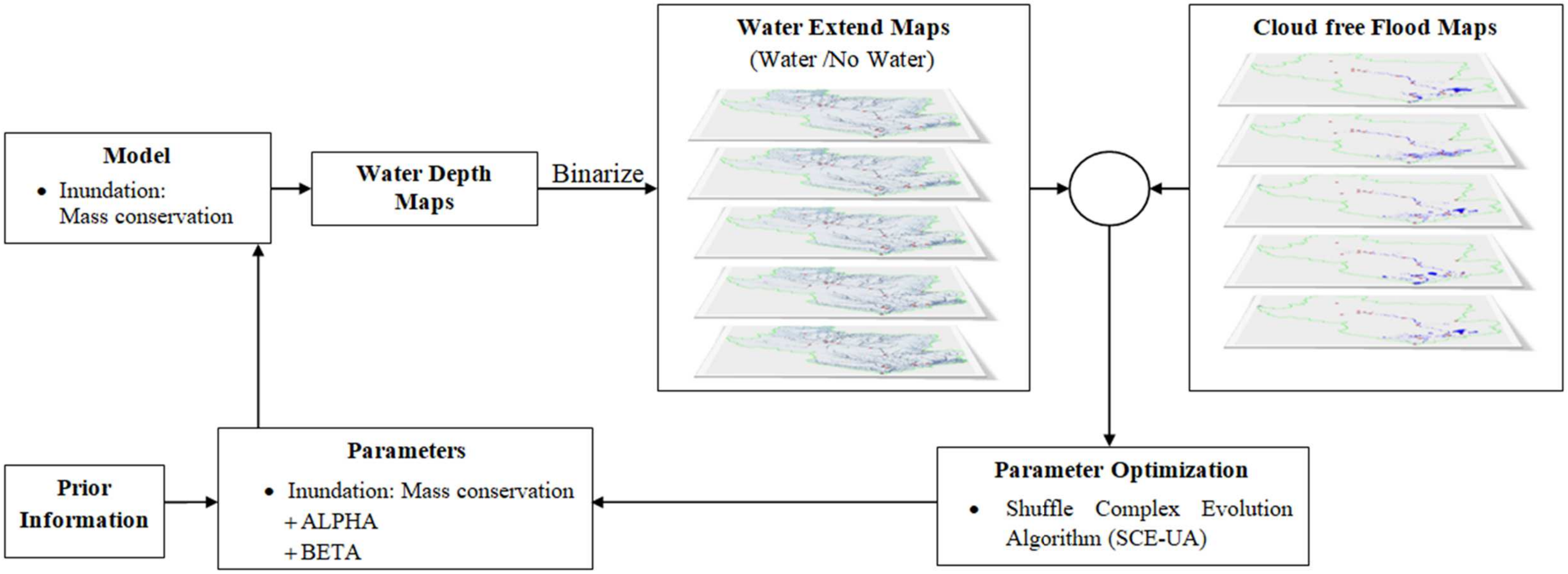


Figure 4-4. Inundation parameters calibration

4.3.4. Assimilation scheme

We applied the assimilation scheme to the kinematic wave routing model. At daily time steps, the cloud-free MODIS water depth maps were used to update the river height using EnKF (Figure 4-5).

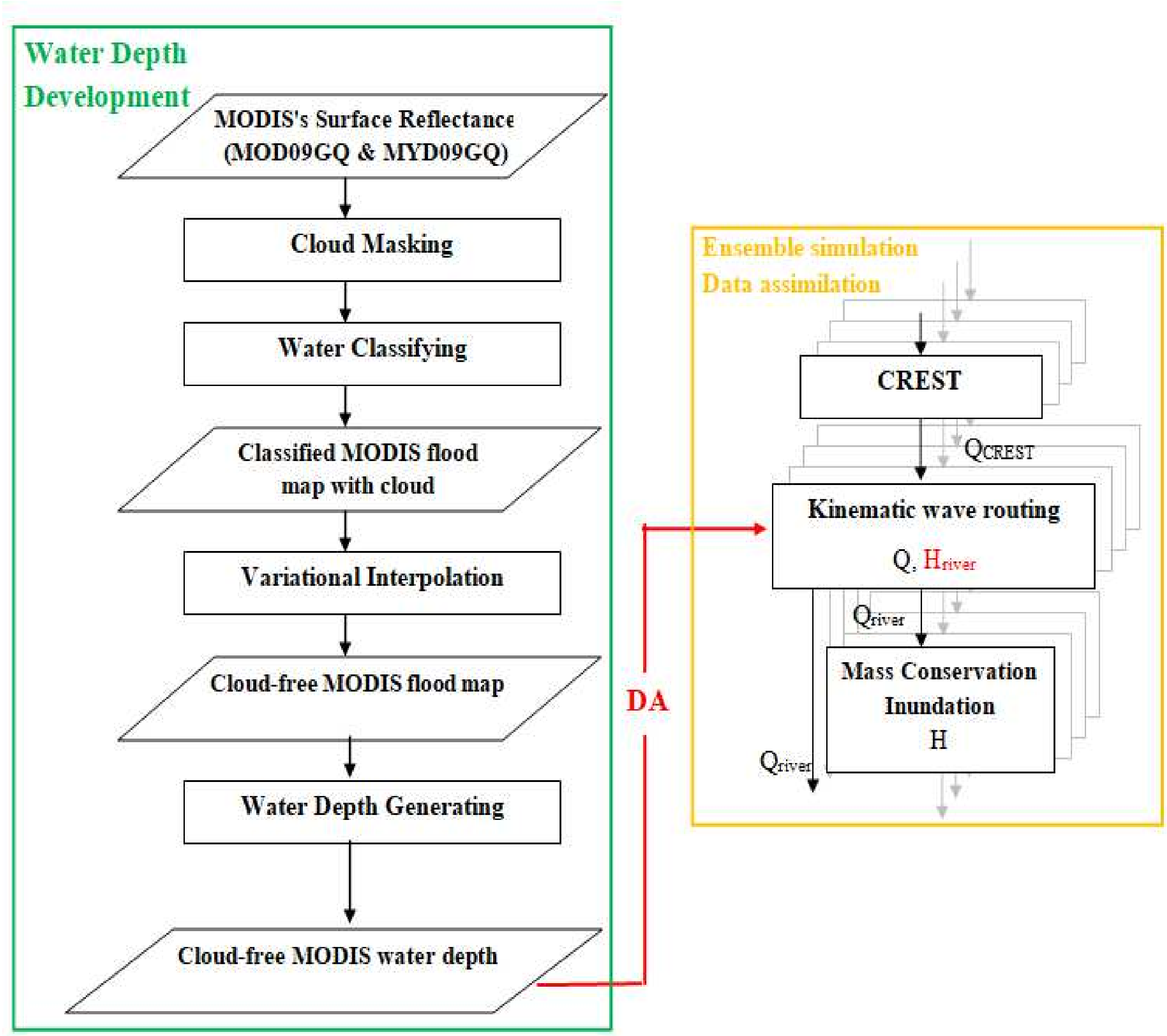


Figure 4-5. Data assimilation framework

4.3.4.1. Ensemble Member Generation

(Munier, et al., 2015) summarized the ensemble member generation step as:

The ensemble of model states is intended to represent model uncertainties, including uncertainties in the meteorological forcings, the model structure, and the parameters. As in (Andreadis, et al., 2007); (Biancamaria, et al., 2011); (Munier, et al., 2015), only the first category is considered here. To generate the ensemble, the forcings of the hydrological model (precipitation, mean and range temperature and wind) were corrupted using the methodology initially developed by (Auclair, et al., 2003) which consists of perturbing the most statistically significant modes. Each forcing field is first decomposed into Empirical Orthogonal Functions (EOFs), then the first modes that explain 95% of the variance are multiplied by a white noise with a 0.2 standard deviation before the signal is reconstructed. The main advantage of the method is that the corrupted fields have coherent spatial and temporal patterns. To limit computational costs, we used 20 ensemble members, which means that 20 corrupted fields of each model forcing have been generated. The example of the spatially averaged precipitation of the 20 members is shown in Figure 4-6.

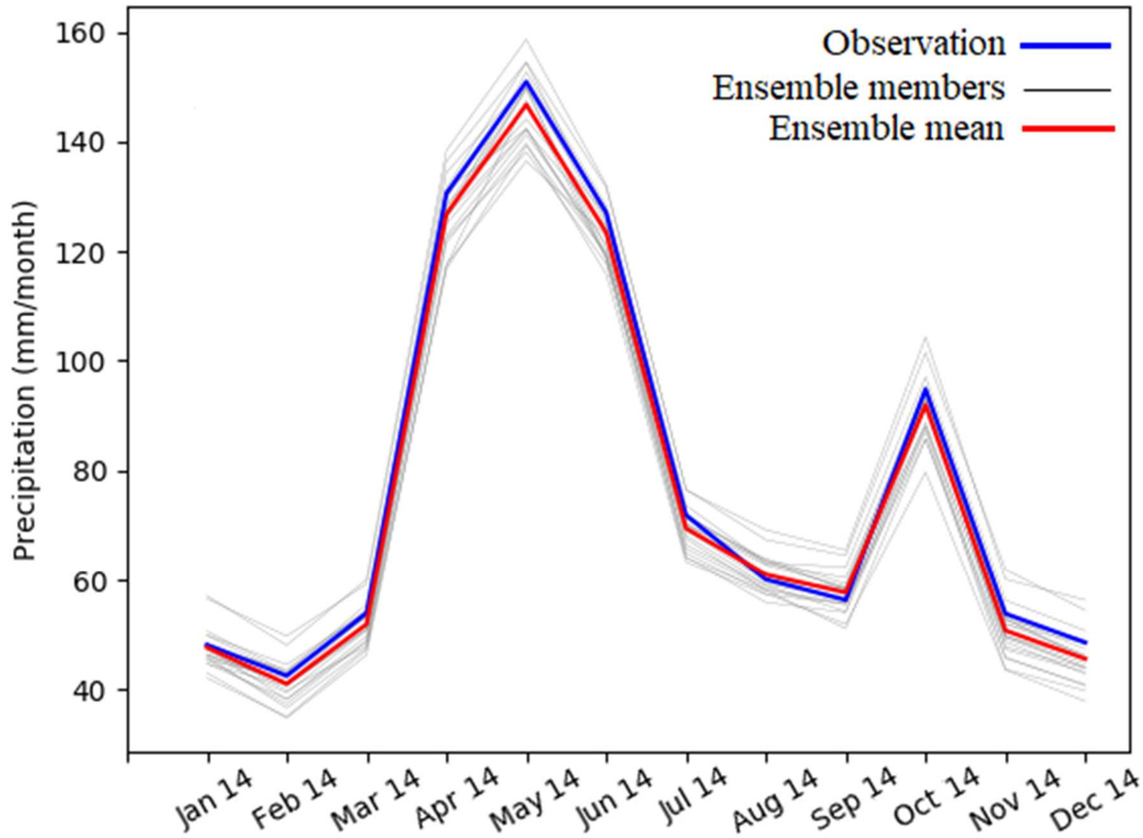


Figure 4-6. Spatial average of precipitation for the 20 members and comparison with the original measurements.

4.3.4.2. Ensemble Kalman Filter (EnKF)

More details on updating the water levels in the kinematic wave routing model using EnKF can be found from (Biancamaria, et al., 2011)(Munier, et al., 2015). Although we did not use the smoother since the water depth was available at every time step, we adopted the scheme from (Biancamaria, et al., 2011) and (Munier, et al., 2015) for the updating/analysis step as from Section 4.3.2.2.

4.4. Results

4.4.1. Model calibration

4.4.1.1. Surface runoff and routing parameters calibration

After calibration of runoff and routing parameters, the EF5 was run for the period of (2013-2014). Simulated discharges from 9 stations were validated with observed values (**Error! Reference source not found.**).

For six stations along the Upper Mississippi River (Brainerd, St. Paul, Hasting, Winona, Clinton, and Keokuk) from upstream to downstream, the correlation coefficient and NSE increased as the discharge increased (six graphs on the left of **Error! Reference source not found.**). For the other three stations in tributary rivers, the average correlation coefficient and NSE are 0.80 and 0.54 respectively (three graphs on the right of **Error! Reference source not found.**).

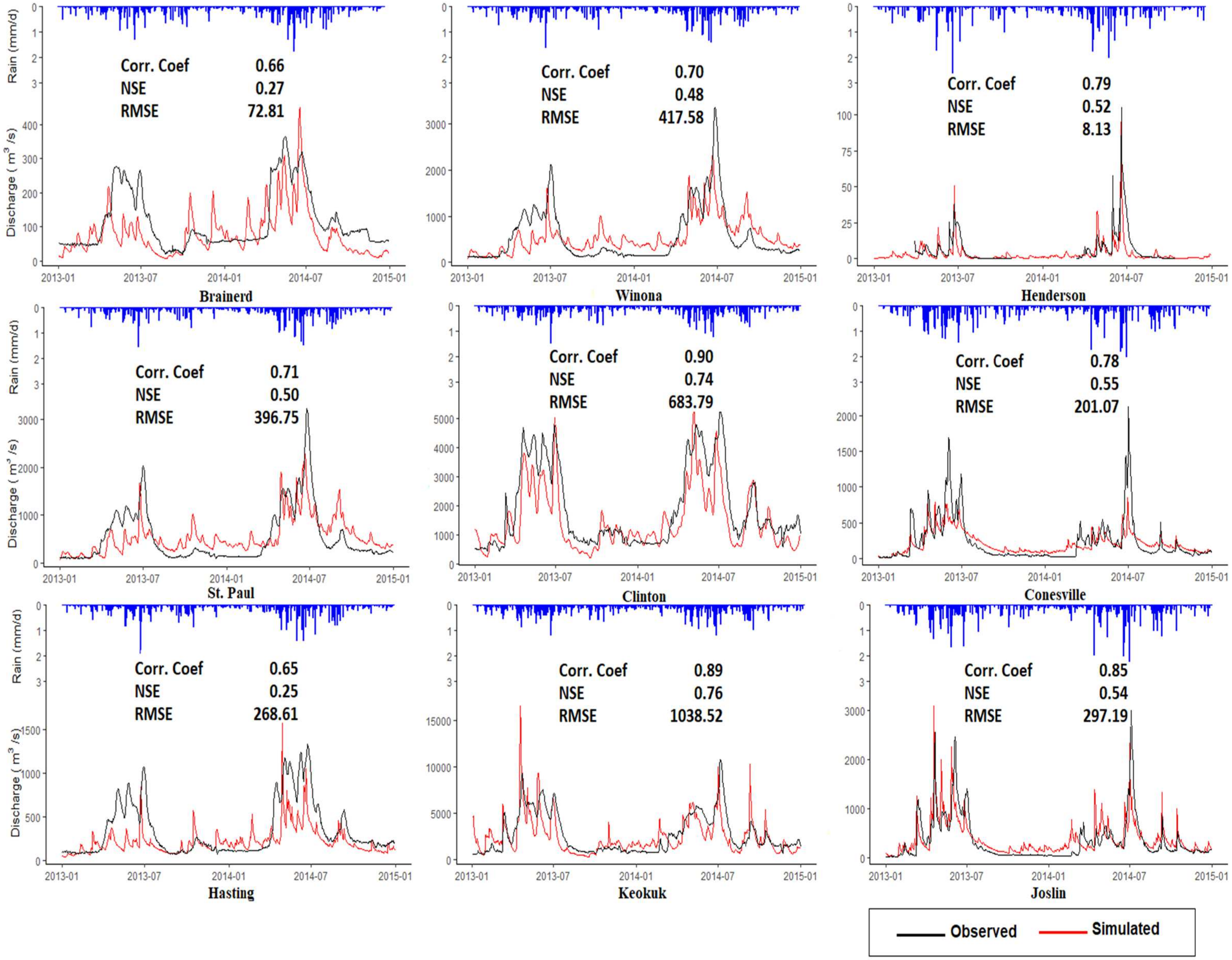


Figure 4-7. Discharge validation for stations in the Upper Mississippi River Basin in 2013 and 2014

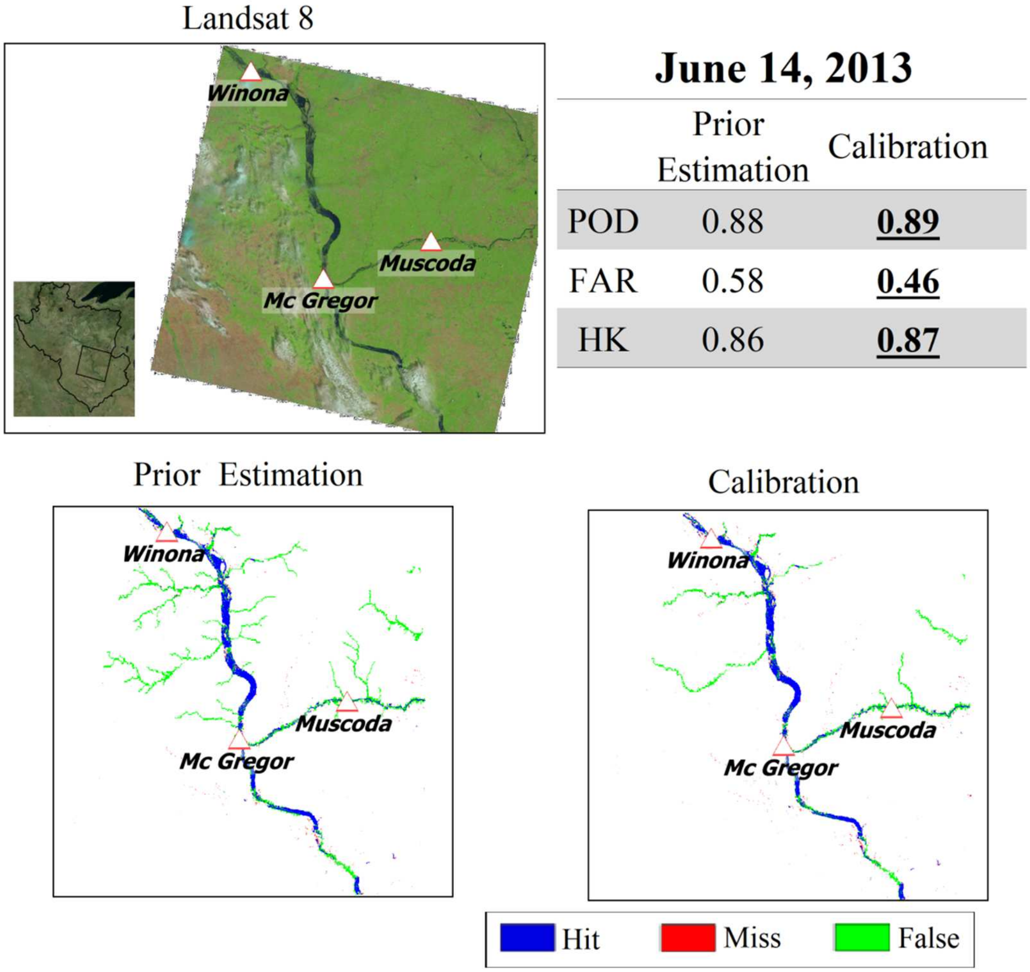
4.4.1.2. Inundation parameters calibration

After calibration using the cloud free flood maps, the resulting inundation maps have much lower FAR than the ones simulated with prior distribution (i.e. less overestimate). From Table 4-2, the post calibration maps have an average FAR of 0.49 in comparison with 0.65 from maps in the other scenario. Also, the inundation maps from the calibrated model detected more water pixels than the inundation maps from model with parameters chosen from prior distributions which resulted in higher POD and FAR.

Table 4-2. Categorical validation of inundation maps before and after calibration

Date	***POD***		***FAR***		***HK***	
	Before	*After*	*Before*	*After*	*Before*	*After*
06-14-2013	0.88	**0.89**	0.58	**0.46**	0.86	**0.87**
06-30-2013	0.69	**0.92**	0.74	**0.63**	0.56	**0.90**
07-16-2013	0.63	**0.83**	0.40	**0.32**	0.57	**0.81**
08-01-2013	0.59	**0.77**	0.48	**0.38**	0.54	**0.75**
08-10-2013	0.71	**0.91**	0.67	**0.57**	0.69	**0.89**
08-26-2013	0.70	**0.84**	0.63	**0.55**	0.67	**0.81**
06-26-2014	0.63	**0.82**	0.86	**0.31**	0.59	**0.78**
07-28-2014	0.68	**0.87**	0.68	**0.52**	0.65	**0.85**
08-13-2014	0.71	**0.89**	0.66	**0.56**	0.68	**0.86**
08-29-2014	0.71	**0.93**	0.81	**0.61**	0.66	**0.89**

*Bold values represents better metric performance



June 14, 2013

	Prior Estimation	Calibration
POD	0.88	**0.89**
FAR	0.58	**0.46**
HK	0.86	**0.87**

Figure 4-8. Categorical validation for inundation maps before and after calibration in June 14, 2013

Figure 4-8 illustrates how cloud free flood maps help simulate inundation maps accurately. On June 14, 2013, inundation parameters in three stations, namely, Winona, Muscoda, and McGregor, after calibration, reduced the FAR from 0.58 to 0.46.

4.4.2. Data assimilation

We analyzed the effects of the cloud-free MODIS data assimilation on the downstream discharge of the basin. We compared simulations by (1) open loop (OL; i.e.

no data assimilation) and (2) EnKF. A six-month simulation was performed (from January 1 to July 10, 2014), results are shown in Figure 4-9.

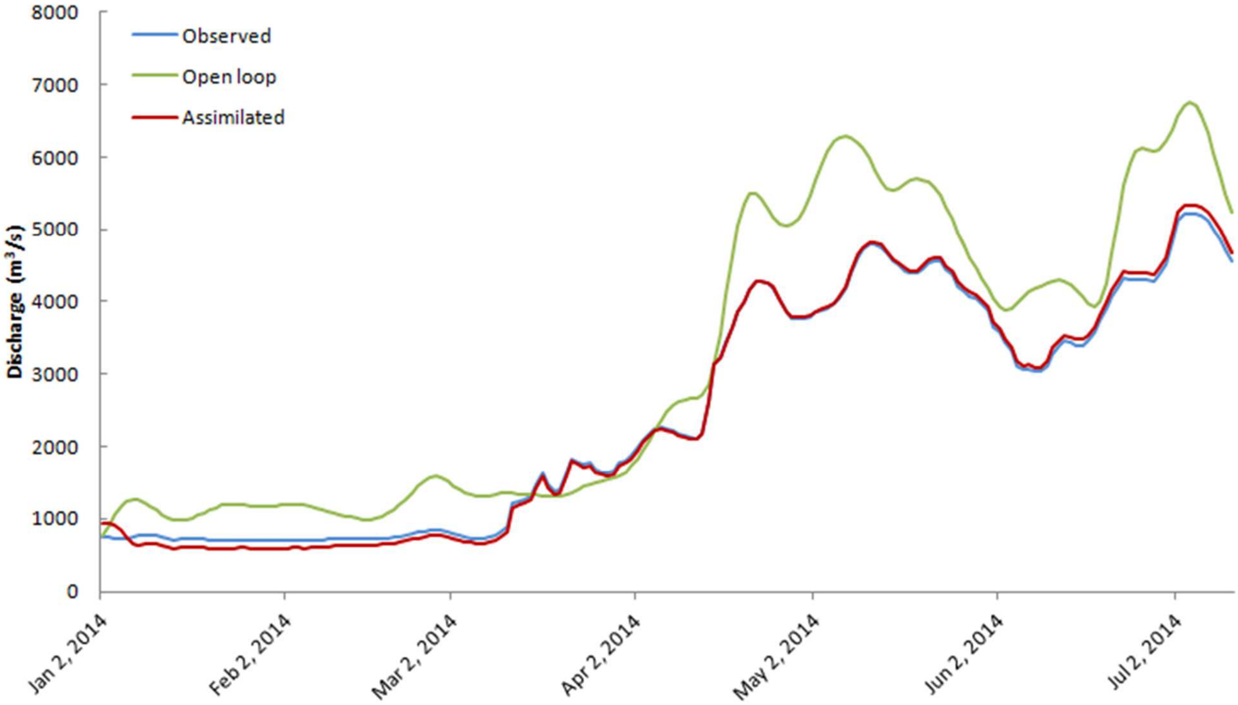


Figure 4-9. Discharge in downstream station

The OL results show that meteorological forcing corruption led to the ensemble mean overestimation of 66.76 % with respect to the observed flow. The application of the EnKF efficiently corrects the water levels and consequently the discharge. Since the cloud-free MODIS is available at daily time steps, which matches with the simulation time step, no smoother needed to propagate the error covariance through time.

Quantitatively, the NSE of the ensemble mean were 0.698, 0.998 for the OL and the EnKF respectively, showing the high efficiency of the EnKF algorithm. The mean ensemble standard deviation were 568.98, 76.48 m^3/s for the two simulations, respectively.

One thing to note is that the assimilation was more persistent during the high-flow period than the low-flow period.

4.5. Chapter Summary and Discussion

With regard to calibrating hydrological parameters, the overall results over the validation periods are satisfactory. Stream flow validations for 2013 and 2014 yield average scores of 0.77, 0.51, and 376.05 for correlation coefficient, NSE, and RMSE respectively in nine stations. More importantly, calibrated flood inundation maps match more closely with images retrieved from Landsat 8 than ones generated from models using prior estimated parameters. From Table 4-2, the average FAR decreased sharply from 0.65 to 0.49 before and after calibration while the average POD increased from 0.69 to 0.87 after calibration. These sufficient results and the fact that MODIS is a global product encourage the use of cloud free flood maps to calibrate hydrological models in remote and ungauged regions.

Nevertheless, an important aspect about the computational efficiency during the flood inundation calibration process should be highlighted. Since the calibration of flood inundation parameters was done for all stations in the basin at the same time, we acknowledge that this process required a lot of computational power. We will use actual numbers of one calibration iteration to illustrate this idea. We used SCE-UA optimization algorithm which was proven as accurate and robust in calibrating hydrological parameters. We calibrated using 11 stations and each station had two parameters, hence there were 22 parameters to be tuned. According to the SCE-UA, then the number of points in each simplex is as follow:

$$N_{point}^{(1)} = 2N_{params} + 1 \tag{4-6}$$

Where $N_{point}^{(1)}$ is a number of points in the first simplex, N_{params} is the number of parameters that need to be optimized. Given $N_{params} = 22$, $N_{point}^{(1)} = 2 * 22 + 1 = 45$ points. Furthermore, if we use a typical number of 5 simplexes, the total number of points will be 225 points. For each iteration of the algorithm, if we ran the model to simulate inundation maps for four summer months in 2012, the computation time could take up to 3 hours. With 100 iterations, if running was done using normal for loops, the calibration process could take two weeks to finish. Therefore, we are parallelizing the SCE-UA to be able to work with a larger number of iterations in order to find the global optimum.

Furthermore, we demonstrated the positive effect of cloud-free MODIS water depth assimilation on simulating and forecasting downstream discharge. We compared the simulation performance with and without data assimilation, and showed that the assimilation of cloud-free MODIS water depth into the hydrodynamic model greatly improved the results.

As in (Munier, et al., 2015), one of the main limitations of this study is the assumption as to the nature errors from input forcings, observations (the cloud-free MODIS water depth), and the model. Here errors in meteorological forcings were accounted for in the ensemble member's generation step, and cloud-free MODIS water depth errors are assumed to be zero-mean Gaussian with a 15 cm standard deviation. Other errors related to observations are neglected. Studies to analyze the effects of such errors are important for the simulation and forecast results. They are scheduled to be done in the future.

To conclude, the purpose of this chapter is to introduce two of the many potential applications of cloud-free MODIS flood products. Furthermore, hydrologic modeling will also help the general public to get a full picture of flood forecasts.

Chapter 5. Conclusions and Future directions

5.1. Summary of findings

The main thrust behind this dissertation is to improve the usefulness of satellite observations in hydrological study. For either monitoring snow covers or water bodies using satellites, the major problem is cloud obstruction. (Xia, et al., 2012) successfully applied the VI method from image processing field to remove clouds from MODIS-SCA products. We improved their algorithm in computational stability in order to apply VI on a much larger scale both in time and space for both MODIS snow and flood products. By making an assumption of the persistent characteristic of water bodies, we created the cloud-free MODIS flood maps. The results are promising.

Furthermore, these space-based observations will also help improve hydrologic modeling by reducing model uncertainties. Also, due to some unresolved problems related to remote sensing such as: cloud persistence and cloud shadow mixing, hydrological models will also help us to get a full picture of flood mapping and forecasting. We used the resulting cloud-free MODIS flood products to reduce model uncertainties through model calibration and data assimilation. During the validation periods, dynamic outputs of the hydrologic model matched closely with observations.

Following is a summary of the key findings addressed in this dissertation.

5.1.1. Cloud-free snow cover maps over CONUS

The second chapter presents a cloud-free snow cover dataset with a daily temporal resolution and 0.05° spatial resolution from March 2000 to February 2017 over the contiguous United States (CONUS). The dataset was developed by completely removing clouds from the original NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) Snow Cover Area product (MOD10C1) through a series of spatiotemporal filters followed by the Variational Interpolation (VI) algorithm; the filters and VI algorithm were evaluated using bootstrapping. The dataset was validated over the period with the Landsat 7 ETM+ snow cover maps in the Seattle, Minneapolis, Rocky Mountains, and Sierra Nevada regions. The resulting cloud-free snow cover accurately captured dynamic changes of snow throughout the period in terms of Probability of Detection (POD) and False Alarm Ratio (FAR) with average values of 0.955 and 0.179 for POD and FAR, respectively. The dataset provides continuous inputs of snow cover area for hydrologic studies for almost two decades. The VI algorithm can be applied in other regions given that a proper validation can be performed.

5.1.2. Cloud-free MODIS flood products

Flood mapping from satellites has major advantages, primarily, global land coverage and direct observation. However, cloud pollution in satellite images limits their practical usability. Cloud removal techniques are popular for snow cover satellite products due to snow persistent characteristic. Among all approaches, the Variational

Interpolation (VI) algorithm, a method that constructs three dimensional space-time surfaces from continuous and adjacent satellite images, proved its robustness and effectiveness in removing clouds from the NASA Moderate Resolution Imaging Spectroradiometer (MODIS) Snow Cover Area (SCA) products. In Chapter 3, we implemented the VI method to remove clouds from the MODIS flood images while preserving the product's original resolutions. The resulting cloud-free flood maps showed an improvement of nearly 70% in average Probability of Detection (POD) (from 0.29 to 0.49) when validated with flood images from Landsat 8. Furthermore, the cloud-free flood maps have been used to derive the cloud-free water depth by using the REFIX method. The cloud-free water depth and the cloud-free flood extent will be used to improve the hydrologic model performance in the next chapter.

5.1.3. Improving hydrologic modeling using cloud-free MODIS flood maps

Hydrologic models are crucial for hydrological studies and water resources management. However, model errors are inevitable due to assumptions of the nature physical characteristics. In order to reduce the effects of errors, two approaches can be taken. The first one is to calibrate parameters of the hydrologic model using observations. The second is to assimilate observations so as to correct the effects of model errors (Munier, et al., 2015). Model calibration and data assimilation using gauge observations have been employed for decades due to the availability of gauge observations. On the other hand, satellite observations and their derived products just emerged from the last decade and immediately proved their importance in improving hydrologic modeling. Many studies have used existing/expected satellite observations to improve the performance of hydrologic models ((Matgen, et al., 2007)(Andreadis, et al., 2007)(Di Baldassarre, et al., 2009)(Schumann, et al., 2009)(Biancamaria, et al., 2011)(Munier, et al., 2015)). In Chapter 4, we utilized the usefulness of the cloud-free MODIS flood products to improve hydrologic modeling. The cloud-free flood maps have been used to calibrate a hydrologic model; the resulting simulated inundation maps had an average POD of 0.87 over the validation period. Moreover, the cloud-free water level maps have been used for data assimilation, hence, greatly improve flood modeling. During the validation period, from January 1 to July 10, 2014, the NSE and the standard deviation were 0.998 and 76.48 m^3/s respectively.

5.2. Future Extensions

This dissertation focused on (1) the developments of cloud-free snow cover maps and cloud-free flood maps, and (2) the improvement of hydrologic modeling when using the resulted cloud-free flood maps to calibrate and assimilate.

Needless to say that further extensions and improvements can be made and following are some of the recommended areas for future studies.

5.2.1. Global cloud-free snow cover maps

The future extension of this study involves applying the VI algorithm for near real-time global snow recovery. With an improvement in computational efficiency, the VI algorithm can be implemented in large Climate Zone regions (Kottek, et al., 2006) and then merged into the whole global map. Parallel processing of multiple regions at the same time should be adequate for near real-time global snow recovery.

5.2.2. Integration with SWOT data product

The soon-to-come Surface Water and Ocean Topography (SWOT) direct observations of water elevation is an advanced satellite mission which is scheduled to be launched in 2021. SWOT uses state-of-the-art Ka-band radar interferometer (KaRIn) to provide a set of water body surface elevations with expected high height accuracy (Rodriguez, 2015)(Biancamaria, et al., 2016). Hence, we can greatly extend and improve the MODIS flood maps using SWOT.

We can, first, use the SWOT data to validate the cloud-free MODIS water depth maps, and then, second, combine those sources to create a timely and accurate water level product.

5.2.3. Near real-time cloud-free global flood map

We plan to build an operational real-time cloud-free global flood mapping using available satellite data resources, currently the Terra and Aqua satellites. The system will be fully automated, ingesting new MODIS images from NASA's ftp website and delivering products on completion via a website. Our goal is that the daily products will be available within 8 hours of Aqua overpass (~8:00 PM local time).

The expected latency will result from product available (3-6 hours), data processing and quality control (2-5 hours), and product display on website (5-10 minutes).

5.2.4. An automated framework for flood forecasting

The assimilation framework proposed in Chapter 4 shown great potential in predicting floods. We plan to apply this framework in other data-scarce regions to utilize more from satellite images. Along with section 5.2.3, whenever a cloud-free MODIS water depth map is available, it will be automatically used to update water levels in the hydrological model.

REFERENCES

Ahamed, A. & Bolten, J. D., 2017. A MODIS-based automated flood monitoring system for southeast asia. *International Journal of Applied Earth Observation and Geoinformation,* Volume 61, pp. 104-117.

Ajami, N. K., Gupta, H., Wagener, T. & Sorooshian, S., 2004. Calibration of a semi-distributed hydrologic model for streamflow estimation along a river system. *Journal of Hydrology,* 298(1-4), pp. 112-135.

Alsdorf, D., Rodriguez, E. & Lettenmaier, D., 2007. Measuring surface water from space. *Reviews of Geophysics,* 45(2).

Andreadis, K. M., Clark, E. A., Lettenmaier, D. P. & Alsdorf, D. E., 2007. Prospects for river discharge and depth estimation through assimilation of swath-altimetry into a raster-based hydrodynamics model. *Geophysical Research Letters ,* 34(10).

Andreadis, K. M. & Lettenmaier, D. P., 2006. Assimilating remotely sensed snow observations into a macroscale hydrology model. *Advances in Water Resources,* 29(6), pp. 872-886.

Auclair, F., Marsaleix, P. & De Mey, P., 2003. Space-time structure and dynamics of the forecast error in a coastal circulation model of the Gulf of Lions. *Dynamics of Atmospheres and Oceans,* 36(4), pp. 309-346.

Barati, R., Neyshabouri, S. & Ahmadi, G., 2014. Sphere drag revisited using Shuffled Complex Evolution algorithm. In: A. J. Schleiss, G. De Cesare, M. J. Franca & M. Pfister, eds. *River Flow 2014.* London: Taylor & Francis Group, pp. 345-353.

Barrett, A. P., 2003. *National operational hydrologic remote sensing center snow data assimilation system (SNODAS) products (digital media) at NSIDC.* [Online] Available at: https://nsidc.org/sites/nsidc.org/files/files/nsidc_special_report_11.pdf [Accessed March 2018].

Bates, P. D. & De Roo, A. J., 2000. A simple raster-based model for flood inundation simulation. *Journal of Hydrology,* 236(1-2), pp. 54-77.

Bates, P. D., Marks, J. & Horritt, M., 2003. Optimal use of high-resolution topographic data in floodinundation models. *Hydrological Processes,* 17(3), pp. 537-557.

Begnudelli, L., Sanders, B. F. & Bradford, S. F., 2008. Adaptive Godunov-Based Model for Flood Simulation. *Journal of Hydraulic Engineering,* 134(6).

Beven, K., 2012. *Rainfall-runoff Modelling: the Primer.* 2nd ed. s.l.:John Wiley & Sons.

Biancamaria, S. et al., 2011. Assimilation of virtual wide swath altimetry to improve Arctic river modeling. *Remote Sensing of Environment,* 115(2), pp. 373-381.

Biancamaria, S., Lettenmaier, D. P. & Pavelsky, T. M., 2016. The SWOT Mission and Its Capabilities for Land Hydrology. *Surveys in Geophysics,* 37(2), pp. 307-337.

Brakenridge, R. & Anderson, E., 2006. MODIS-BASED FLOOD DETECTION, MAPPING AND MEASUREMENT: THE POTENTIAL FOR OPERATIONAL HYDROLOGICAL APPLICATIONS. In: *Transboundary Floods: Reducing Risks Through Flood Management.* Dordrecht: Springer.

Brakenridge, R. et al., 2014. *Rapid Response Mapping of the 2014 Flooding in Paraguay and Southern Brazil, DFO Event # 4150, Area B, Current Conditions and Maximum Flood Extent.* [Online] Available at: http://floodobservatory.colorado.edu/Version3/2014Paraguay4150b.html [Accessed 2018].

Brunner, G. W., 2016. *HEC-RES River Analysis System - User's Manual Version 5.0. US Army Corps of Engineers.* [Online] Available at: http://www.hec.usace.army.mil/software/hec-ras/documentation/HEC-RAS%205.0%20Reference%20Manual.pdf
[Accessed 2018].

Choi, S.-C., 2006. *Iterative Methods for Singular Linear Equations and Least-Squares Problems.* s.l.:PhD thesis.

Cline, D. W. & Carroll, T. R., 1999. Inference of snow cover beneath obscuring clouds using optical remote sensing and a distributed snow energy and mass balance model. *Journal of Geophysical Research: Atmospheres,* 104(D16), pp. 19631-19644.

Crane, R. G. & Anderson, M. R., 1984. Satellite discrimination of snow/cloud surfaces. *International Journal of Remote Sensing,* 5(1), pp. 213-223.

De Maupertuis, P., 1744. *Accord de différentes lois de la nature qui avaient jusqu'ici paru incompatibles.* Paris: Memoires de l'Academie des Sciences de Paris.

DHI, 2003. *MIKE 11-A Modelling System for Rivers and Channels - User Guide.* [Online]
Available at: http://manuals.mikepoweredbydhi.help/2017/Water_Resources/MIKE_11_Short_Introduction-Tutorial.pdf
[Accessed 2018].

DHI, 2012. *MIKE 21-2D Modelling of Coast and Sea.* [Online]
Available at: https://www.mikepoweredbydhi.com/-/media/shared%20content/mike%20by%20dhi/flyers%20and%20pdf/software%20flyers/coast%20and%20sea/mike21-mikepoweredbydhi-productflyer-uk.pdf
[Accessed 2018].

Di Baldassarre, G., Schumman, G. & Bates, P. D., 2009. A technique for the calibration of hydraulic models using uncertain satellite observations of flood extent. *Journal of Hydrology,* 367(3-4), pp. 276-282.

Ding, Y. et al., 2016. Estimation of beech pyrolysis kinetic parameters by Shuffled Complex Evolution. *Bioresource Technology,* Volume 200, pp. 658-665.

Dong, C. & Menzel, L., 2016. Producing cloud-free MODIS snow cover products with conditional probability interpolation and meteorological data. *Remote Sensing of Environment,* 186(1), pp. 439-451.

Douglas, D. & Peucker, T., 1973. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica: The International Journal for Geographic Information and Geovisualization,* 10(2), pp. 112-122.

Dozier, J., 1989. Spectral signature of alpine snow cover from the landsat thematic mapper. *Remote Sensing of Environment,* Volume 28, pp. 9-22.

Dozier, J., Painter, T. H., Rittger, K. & Frew, J. E., 2008. Time–space continuity of daily maps of fractional snow cover. *Advances in Water Resources,* 31(11), pp. 1515-1526.

Duan, Q., Gupta, V. K. & Sorooshian, S., 1993. A Shuffled Complex Evolution Approach for Effective and Efficient Global Minimization. *Journal of Optimization Theory and Applications,* 76(3), pp. 501-521.

Duan, Q., Sorooshian, S. & Gupta, V., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research,* 28(4), pp. 1015-1031.

Duchon, J., 1977. Splines minimizing rotation-invariant semi-norms in Sobolev spaces. In: W. Schempp & K. Zeller, eds. *Constructive Theory of Functions of Several Variables. Lecture Notes in Mathematics.* Berlin: Springer, pp. 85-100.

Eckhardt, K. & Arnold, J., 2001. Automatic calibration of a distributed catchment model. *Journal of Hydrology,* 251(1-2), pp. 103-109.

Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research,* pp. 10143-10162.

Evensen, G., 2004. Sampling strategies and square root analysis schemes for the EnKF. *Ocean Dynamics,* 54(6), pp. 539-560.

Flamig, Z., Vergara, H. J., Clark, R. & Hong, Y., 2015. *Ensemble Framework For Flash Flood Forecasting (EF5).* [Online] Available at: http://ef5.ou.edu/

Frappart, F. et al., 2006. Water volume change in the lower Mekong from satellite altimetry and imagery data. *Geophysical Journal International,* pp. 570-584.

Gafurov, A. & Bardossy, A., 2009. Cloud removal methodology from MODIS snow cover product. *Hydrology and Earth System Sciences,* Volume 13, pp. 1361-1373.

Gan, T. Y. & Biftu, G. F., 1996. Automatic Calibration of Conceptual Rainfall-Runoff Models: Optimization Algorithms, Catchment Conditions, and Model Structure. *Water Resources Research,* 32(12), pp. 3513-3524.

Gauss, C. F., 1829. Ueber ein allgemeines Grundgesetz der Mechanik. *Journal für die reine und angewandte Mathematik,* Volume 4, pp. 232-235.

Gomes, A. et al., 2009. *Implicit Curves and Surfaces: Mathematics, Data Structures and Algorithms.* London: Springer-Verlag.

Hall, D., 2002. MODIS snow-cover products. *Remote Sensing of Environment,* 83(1-2), pp. 181-194.

Hall, D. K. & Riggs, G. A., 2007. Accuracy assessment of the MODIS snow products. *Hydrological Processes,* Volume 21, pp. 1534-1547.

Hall, D. K. & Riggs, G. A., 2016. *MODIS/Terra Snow Cover Daily L3 Global 0.05Deg CMG, Version 6.* Boulder, Colorado USA: NASA National Snow and Ice Data Center Distributed Active Archive Center.

Hall, D. K. et al., 2002. MODIS snow-cover products. *Remote Sensing of Environment,* 83(1-2), pp. 181-194.

Hall, D. K., Salomonson, V. V. & Riggs, G. A., 1995. Development of Methods for Mapping Global Snow Cover Using Moderate Resolution Imaging Spectroradiometer Data. *Remote Sensing of Environment,* 54(2), pp. 127-140.

Hall, D., Riggs, G., Foster, J. & Kumar, S., 2010. Development and evaluation of a cloud-gap-filled. *Remote Sensing of Environment,* 114(3), pp. 496-503.

Harer, S., Bernhardt, M., Siebers, M. & Schulz, K., 2018. On the need for a time- and location-dependent estimation of the NDSI threshold value for reducing existing uncertainties in snow cover maps at different scales. *The Cryosphere,* 12(5), pp. 1629-1642.

Hasalova, L., Ira, J. & Jahoda, M., 2016. Practical observations on the use of Shuffled Complex Evolution (SCE) algorithm for kinetic parameters estimation in pyrolysis modeling. *Fire Safety Journal,* Volume 80, pp. 71-82.

Hollan, J. H., 1975. *Adaptation in natural and artificial systems.* Ann Arbor: University of Michigan Press.

Horritt, M. S. & Bates, P. D., 2002. Evaluation of 1D and 2D numerical models for predicting river flood inundation. *Journal of Hydrology,* 268(1-4), pp. 87-99.

Houspanossian, J. et al., 2018. Long-lasting floods buffer the thermal regime of the Pampas. *Theoretical and Applied Climatology,* 131(1-2), pp. 111-120.

Hromadka, T. V. & DeVries, J. J., 1988. Kinematic Wave Routing and Computational Error. *Journal of Hydraulic Engineering,* 114(2).

Huang, X., Liang, T., Zhang, X. & Guo, Z., 2011. Validation of MODIS snow cover products using Landsat and ground measurements during the 2001–2005 snow seasons over northern Xinjiang, China. *International Journal of Remote Sensing,* 32(1), pp. 133-152.

Jian, C., Hill, A. A. & Urbano, L. D., 2009. A GIS-based model for urban flood inundation. *Journal of Hydrology,* 373(1-2), pp. 184-192.

Kalman, R. E., 1960. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering,* 82(1), pp. 35-45.

Karlsson, J. M. & Arnberg, W., 2009. Quality analysis of SRTM and HYDRO1K: a case study of floodinundation in Mozambique. *International Journal of Remote Sensing,* pp. 267-285.

Klein, A. G. & Barnett, A. C., 2003. Validation of daily MODIS snow cover maps of the Upper Rio Grande River Basin for the 2000–2001 snow year. *Remote Sensing of Environment,* 86(2), pp. 162-176.

Kottek, M. et al., 2006. World Map of the Köppen-Geiger climate classification. *Meteorologische Zeitschrift,* 15(3), pp. 259-263.

Krajci, P., Holko, L. & Parajka, J., 2016. Variability of snow line elevation, snow cover area and depletion in the main Slovak basins in winters 2001-2014. *Journal of Hydrology and Hydromechanics,* 64(1), pp. 12-22.

Krajewski, W. et al., 2017. Real-time flood forecasting and information system for the state of Iowa. *Bulletin of the American Meteorological Society,* 98(3).

Lanczos, C., 1970. *The Variational Principles of Mechanics.* New York: Dover Publications.

Lighthill, M. J. & Whitham, G. B., 1955. On kinematic waves II. A theory of traffic flow on long crowded roads. *PROCEEDINGS OF THE ROYAL SOCIETY OF LONDON,* 229(1178).

Li, H., Li, X. & Xiao, P., 2016. Impact of Sensor Zenith Angle on MOD10A1 Data Reliability and Modification of Snow Cover Data for the Tarim River Basin. *Remote Sensing,* 8(9).

Li, L. et al., 2015. Evaluation of MODIS Spectral Indices for Monitoring Hydrological Dynamics of a Small, Seasonally-Flooded Wetland in Southern Spain. *Wetlands,* 35(5), pp. 851-864.

Lin, J.-Y., Cheng, C.-T. & Chau, K.-W., 2006. Using support vector machines for long-term discharge prediction. *Hydrological Sciences Journal,* 51(4), pp. 599-612.

Lin, Y., 2011. *GCIP/EOP Surface: Precipitation NCEP/EMC 4KM Gridded Data (GRIB) Stage IV Data. Version 1.0.* [Online] Available at: https://doi.org/10.5065/D6PG1QDD [Accessed 13 April 2018].

Liong, S.-Y. & Atiquzzaman, M., 2004. Optimal design of water distribution network using shuffled complex evolution. *Journal of The Institution of Engineers, Singapore,* 44(1).

Liu, Y. et al., 2004. Nighttime polar cloud detection with MODIS. *Remote sensing of environment,* 92(2), pp. 181-194.

Liu, Y., Key, J. R. & Wang, X., 2008. The influence of changes in cloud cover on recent surface temperature trends in the Arctic. *Journal of Climate,* 21(4), pp. 705-715.

Luo, Y., Trishchenko, A. P. & Khlopenkov, K. V., 2008. Developing clear-sky, cloud and cloud shadow mask for producing clear-sky composites at 250-meter spatial resolution for the seven MODIS land bands over Canada and North America. *Remote Sensing of Environment,* 112(12), pp. 4167-4185.

Madsen, H. O., 2000. Automatic calibration of a conceptual rainfall–runoff model using multiple objectives. *Journal of Hydrology,* 235(3-4), pp. 276-288.

Malmros, J. K. et al., 2018. Snow cover and snow-albedo changes in the central Andes of Chile and Argentina from daily MODIS observations (2000-2016). *Remote Sensing of Environment,* Volume 209, pp. 240-252.

Marcil, G.-K., Trudel, M. & Leconte, R., 2016. Using Remotely Sensed MODIS Snow Product for the Management of Reservoirs in a Mountainous Canadian Watershed. *Water Resources Management,* 30(8), pp. 2735-2747.

Masek, J. G. et al., 2006. A Landsat Surface Reflectance Dataset for North America, 1990–2000. *IEEE GEOSCIENCE AND REMOTE SENSING LETTERS,* 3(1), pp. 68-72.

Matgen, P. et al., 2007. Integration of SAR-derived river inundation areas, high-precision topographic data and a river flow model toward near real-time flood management. *International Journal of Applied Earth Observation and Geoinformation,* 9(3), pp. 247-263.

Maurer, E. P., Rhoads, J. D., Dubayah, R. O. & Lettenmaier, D. P., 2003. Evaluation of the snow-covered area data product from MODIS. *Hydrological Processes,* 17(59-71), pp. 59-71.

Micheletty, P. D., Kinoshita, A. M. & Hogue, T. S., 2014. Application of MODIS snow cover products: wildfire impacts on snow and melt in the Sierra Nevada. *Hydrology and Earth System Sciences,* 18(11), pp. 4601-4615.

Mohammadi, A., Costelloe, J. F. & Ryu, D., 2017. Application of time series of remotely sensed normalized difference water, vegetation and moisture indices in characterizing flood dynamics of large-scale arid zone floodplains. *Remote Sensing of Environment,* Volume 190, pp. 70-82.

Molotch, N., S, F., R, B. & S, H., 2004. Estimating the distribution of snow water equivalent and snow extent beneath cloud-cover in the Salt-Verde River basin, Arizona. *Hydrological Processes,* 18(9), pp. 1595-1611.

Moulinec, C. et al., 2011. TELEMAC: An efficient hydrodynamics suite for massively parallel architectures. *Computers & Fluids,* 51(1), pp. 30-34.

Munier, S. et al., 2015. SWOT data assimilation for operational reservoir management on the upper Niger River Basin. *Water Resources Research,* 51(1), pp. 554-575.

Naeini, M. R. et al., 2018. Shuffled Complex-Self Adaptive Hybrid EvoLution (SC-SAHEL) optimization framework. *Environmental Modelling & Software,* Volume 104, pp. 215-235.

Nelder, J. & Mead, R., 1965. A simplex method for function minimization. *The Computer Journal,* 7(4), pp. 308-313.

Nguyen, P. et al., 2015. Flood Forecasting and Inundation Mapping Using HiResFlood-UCI and Near-Real-Time Satellite Precipitation Data: The 2008 Iowa Flood. *Journal of Hydrometeorology,* 16(3).

Nguyen, P. et al., 2016. A high resolution coupled hydrologic–hydraulic model (HiResFlood-UCI) for flash flood modeling. *Journal of Hydrology,* pp. 401-420.

Nigro, J., Slayback, D., Policelli, F. & Brakenridge, R. G., 2014. *NASA/ DFO MODIS Near Real-Time (NRT) Global Flood Mapping Product Evaluation of Flood and Permanent Water Detection.* [Online] Available at: https://floodmap.modaps.eosdis.nasa.gov//documents/NASAGlobalNRTEvaluationSummary_v4.pdf
[Accessed 2 11 2018].

Paige, C. C., Parlett, B. N. & Van der Vorst, H. A., 1995. Approximate solutions and eigenvalue bounds from Krylov subspaces. *Numerical Linear Algebra with Applications,* 2(2), pp. 115-133.

Paige, C. C. & Saunders, M. A., 1975. Solution of Sparse Indefinite Systems of Linear Equations. *SIAM Journal on Numerical Analysis,* 12(4), pp. 617-629.

Painter, T. H. et al., 2009. Retrieval of subpixel snow covered area, grain size, and albedo from MODIS. *Remote Sensing of Environment,* Volume 113, pp. 868-879.

Parajka, J. & Bloschl, G., 2006. Validation of MODIS snow cover images over Austria. *Hydrology and Earth System Sciences,* Volume 10, pp. 679-689.

Parajka, J. et al., 2010. A regional snow-line method for estimating snow cover from MODIS during cloud cover. *Journal of Hydrology,* 381(3-4), pp. 203-212.

Perez, T., Mattar, C. & Fuster, R., 2018. Decrease in Snow Cover over the Aysén River. *Water,* 10(5).

Policelli, F. & SlayBack, D., 2017. *MODIS Near Real-Time Global Flood Mapping Project.* [Online]
Available at: https://floodmap.gsfc.nasa.gov

Prakash, M., Rothauge, K. & Cleary, P. W., 2014. Modelling the impact of dam failure scenarios on flood inundation using SPH. *Applied Mathematical Modelling,* 38(23), pp. 5515-5534.

Price, W. L., 1987. Global optimization algorithms for a CAD workstation. *Journal of Optimization Theory and Applications,* 55(1), pp. 133-146.

Riggs, G. A. & Hall, D. K., 2015. *MODIS Snow Products Collection 6 User Guide.* [Online]
Available at: https://nsidc.org/sites/nsidc.org/files/files/MODIS-snow-user-guide-C6.pdf
[Accessed March 2018].

Riggs, G. A., Hall, D. K. & Roman, M. O., 2017. Overview of NASA's MODIS and Visible Infrared Imaging Radiometer Suite (VIIRS) snow-cover Earth System Data Records. *Earth System Science Data,* Volume 9, pp. 765-777.

Rinnooy Kan, A. H. G. & Timmer, G. T., 1987. Stochastic global optimization methods part I: Clustering methods. *Mathematical Programming,* 39(1), pp. 27-56.

Rittger, K., Painter, H. T. & Dozier, J., 2013. Assessment of methods for mapping snow cover from MODIS. *Advances in Water Resources,* Volume 51, pp. 367-380.

Rodell, M. & Houser, P. R., 2004. Updating a Land Surface Model with MODIS-Derived Snow Cover. *Journal of Hydrometeorology,* Volume 5, pp. 1064-1075.

Rodriguez, E., 2015. *swot.jpl.nasa.gov.* [Online] Available at: https://swot.jpl.nasa.gov/files/swot/SRD_021215.pdf [Accessed 12 January 2018].

Sanders, B., Schubert, J. E. & Detwiler, R. L., 2010. ParBreZo: A parallel, unstructured grid, Godunov-type, shallow-water code for high-resolution flood inundation modeling at the regional scale. *Advances in Water Resources,* 33(12), pp. 1456-1467.

Schneider, D. & Molotch, N. P., 2016. Real-time estimation of snow water equivalent in the Upper Colorado River Basin using MODIS-based SWE Reconstructions and SNOTEL data. *Water Resources Research,* 52(10), pp. 7892-7910.

Schumann, G. et al., 2009. Progress in integration of remote sensing–derived flood extent and stage data and hydraulic models. *Reviews of Geophysics,* 47(4).

Schumann, G. et al., 2007. High-Resolution 3-D Flood Information From Radar Imagery for Flood Hazard Management. *IEEE Transactions on Geoscience and Remote Sensing,* 45(6), pp. 1715-1725.

Schwarzenberg-Czerny, A., 1995. On matrix factorization and efficient least squares solution. *Astronomy and Astrophysics Supplement,* 110(450).

Smith, L., 1997. Satellite remote sensing of river inundation area, stage, and discharge: a review. *Hydrological Processes,* pp. 1427-1439.

Sorooshian, S., Duan, Q. & Gupta, V. K., 1993. Calibration of rainfall-runoff models: Application of global optimization to the Sacramento Soil Moisture Accounting Model. *Water Resources Research,* 29(4), pp. 1185-1194.

Stephens, E., Bates, P., Freer, J. & Mason, D., 2012. The impact of uncertainty in satellite data on the assessment of flood inundation models. *Journal of Hydrology,* pp. 162-173.

Tang, Z. et al., 2017. Spatiotemporal Variation of Snow Cover in Tianshan Mountains, Central Asia, Based on Cloud-Free MODIS Fractional Snow Cover Product, 2001-2015. *Remote Sensing,* 9(10).

Tekeli, A. E., Sonmez, I. & Erdi, E., 2016. Snow-covered area determination based on satellite-derived probabilistic snow cover maps. *Arabian Journal of Geosciences,* 9(3).

Teng, J. et al., 2017. Flood inundation modelling: A review of methods, recent advances and uncertainty analysis. *Environmental Modelling & Software,* Volume 90, pp. 201-216.

Toth, E., Brath, A. & Montanari, A., 2000. Comparison of short-term rainfall prediction models for real-time flood forecasting. *Journal of Hydrology,* 239(1-4), pp. 132-147.

Tran, H. et al., 2018. A cloud-free MODIS snow cover dataset for the contiguous United States from 2000 to 2017. *Scientific Data, Nature.*

Turk, G. & O'Brien, J. F., 1999. *Shape Transformation Using Variational Implicit Functions.* s.l., ACM Press/Addison-Wesley Publishing.

Vacondio, R., Rogers, B., Stansby, P. K. & Mignosa, P., 2012. SPH Modeling of Shallow Flow with Open Boundaries for Practical Flood Simulation. *Journal of Hydraulic Engineering,* 138(6).

Verbyla, D. et al., 2017. Remote Sensing of 2000–2016 Alpine Spring Snowline Elevation in Dall Sheep Mountain Ranges of Alaska and Western Canada. *Remote Sensing,* 9(11).

Vermote, E., 2015. *MOD09GQ MODIS/Terra Surface Reflectance Daily L2G Global 250m SIN Grid V006.* [Online] Available at: https://doi.org/10.5067/modis/mod09gq.006

Wagener, T., Wheater, H. S. & Gupta, H. V., 2004. *Rainfall-runoff modelling in gauged and ungauged catchments.* s.l.:World Scientific.

Wang, J. et al., 2011. The coupled routing and excess storage (CREST) distributed hydrological model. *Hydrological Sciences Journal ,* pp. 84-98.

Wang, X., Xie, H., Liang, T. & Huang, X., 2009. Comparison and validation of MODIS standard and new combination of Terra and Aqua snow cover products in northern Xinjiang, China. *Hydrological Processes,* 23(3), pp. 419-429.

Wang, Y.-C., Yu, P.-S. & Yang, T.-C., 2009. Comparison of genetic algorithms and shuffled complex evolution approach for calibrating distributed rainfall–runoff model. *Hydrological Processes,* 24(8), pp. 1015-1026.

Wolski, P., Murray-Hudson, M., Thito, K. & Cassidy, L., 2017. Keeping it simple: Monitoring flood extent in large data-poor wetlands using MODIS SWIR data. *International Journal of Applied Earth Observation and Geoinformation,* Volume 57, pp. 224-234.

Wunderle, S., Gross, T. & Husler, F., 2016. Snow Extent Variability in Lesotho Derived from MODIS Data (2000–2014). *Remote Sensing,* 8(6).

Xia, Q., Gao, X., Chu, W. & Sorooshian, S., 2012. Estimation of daily cloud-free, snow-covered areas from MODIS based on variational interpolation. *Water Resources Research,* 48(9).

Xia, Q., Gao, X., Chu, W. & Sorooshian, S., 2012. Estimation of daily cloud-free, snow-covered areas from MODIS based on variational interpolation. *Water Resources Research,* 48(9).

Yang, T. et al., 2017. An enhanced artificial neural network with a shuffled complex evolutionary global optimization with principal component analysis. *Information Sciences,* Volume 418-419, pp. 302-316.

Yang, T., Gao, X., Sellars, S. L. & Sorooshian, S., 2015. Improving the multi-objective evolutionary optimization algorithm for hydropower reservoir operations in the California Oroville–Thermalito complex. *Environmental Modelling & Software,* Volume 69, pp. 262-279.

Yapo, P. O., Gupta, H. V. & Sorooshian, S., 1996. Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data. *Journal of Hydrology,* 181(1-4), pp. 23-48.

Yi, S., Wang, Q., Chang, L. & Sun, W., 2016. Changes in Mountain Glaciers, Lake Levels, and Snow Coverage in the Tianshan Monitored by GRACE, ICESat, Altimetry, and MODIS. *Remote Sensing,* 8(10).

Zhang, Y. et al., 2017. Spatial and Temporal Variation Analysis of Snow Cover Using MODIS over Qinghai-Tibetan Plateau during 2003-2014. *Journal of the Indian Society of Remote Sensing,* 45(5).

Zheng, W., Sun, D. & Li, S., 2017. Mapping coastal floods induced by hurricane storm surge using ATMS data. *International Journal of Remote Sensing,* 38(23), pp. 6846-6864.