Spring 2020

HUMAN AND CLIMATE IMPACTS ON FLOODING VIA REMOTE SENSING, BIG DATA ANALYTICS, AND MODELING

Eunsang Cho
University of New Hampshire, Durham

Recommended Citation

Cho, Eunsang, "HUMAN AND CLIMATE IMPACTS ON FLOODING VIA REMOTE SENSING, BIG DATA ANALYTICS, AND MODELING" (2020). *Doctoral Dissertations*. 2500.
https://scholars.unh.edu/dissertation/2500

HUMAN AND CLIMATE IMPACTS ON FLOODING VIA REMOTE SENSING, BIG DATA ANALYTICS, AND MODELING

BY

EUNSANG CHO

B.E., Civil and Environmental Engineering, Hanyang University, 2010

M.S., Water Resources and Environmental Engineering, Hanyang University, 2014

DISSERTATION

Submitted to the University of New Hampshire
in Partial Fulfillment of
Requirements for the Degree of

Doctor of Philosophy
in
Civil Engineering

May 2020

This thesis/dissertation has been examined and approved in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Civil Engineering by:

Dr. Jennifer M. Jacobs
Professor, Civil and Environmental Engineering & Earth Systems Research Center, Institute for the Study of Earth, Oceans, and Space, UNH
Thesis/Dissertation Director

Dr. Jo Sias
Professor, Civil and Environmental Engineering, UNH

Dr. Anne Lightbody
Associate Professor, Earth Sciences, UNH

Dr. Ernst Linder
Professor, Mathematics & Statistics, UNH

Dr. Carrie M. Vuyovich
Physical Research Scientist, Hydrologic Science Laboratory
NASA Goddard Space Flight Center

On May 2020

Original approval signatures are on file with the University of New Hampshire Graduate School.

ACKNOWLEDGMENTS

I am lucky to have Jennifer Jacobs as my advisor. My Ph.D. would not have been made possible without Dr. Jacobs' guidance and encouragement. I cannot say how much I have received unparalleled support from her on all aspects. I may not recognize all her thoughtful consideration, but I've felt that she took care of not only me but also my whole family throughout the UNH life. To me, Dr. Jacobs is a perfect advisor, an amazing person, a life-long mentor.

I extend thanks to my committee members (Drs. Sias, Lightbody, Linder, and Vuyovich) who taught and guided me to complete this dissertation with valuable comments and supports. And my hearty thanks to all my friends, labmates, and collaborators for their help and support on all aspects, Samuel Tuttle (Mount Holyoke College), Simon Kraatz (UMass Amherst), Ronny Schroeder (Embry-Riddle Aeronautical Univ.), Adam Hunsaker, Mahsa Moradi, Jang-Geun Choi, Daniel Macadam (UNH), Xinhua Jia (NDSU), Carrie Olheiser (NOAA OWP), Pedro Restrepo, Mike DeWeese, Brian Connelly (NOAA NCRFC), Mike Cosh (USDA), Sujay Kumar, David Mocko (NASA GSFC), Jill Deines (Stanford Univ.), Rachel McCrary (NCAR), and Hyunglok Kim (UVA). I would also like to show my appreciation to Dr. Minha Choi (SKKU), my M.S. advisor, who guided me to enter the hydrology world.

I am grateful to UNH, NASA, and Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI) for financial support through prestigious fellowships and assistantships. This work was generously supported by the NASA Water Resources Applied Sciences Program (NNX15AC47G), the UNH College of Engineering & Physical Sciences (CEPS) Fellowship 2015-2016, the UNH Summer Teaching Assistant Fellowship 2018, the UNH Dissertation Year Fellowship, 2019-2020, and the Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI) Pathfinder Fellowship 2020.

Last but certainly not least, it's such a blessing to have my family along with me on this journey. I can't express how much I am deeply grateful to my wife, Sera, for love and support. I'd love to thank my happy son, Robin, for his smile that cheers me up. Finally, I am truly grateful to my God, Jesus Chirst, for giving me unfailing love, wisdom, and courage at every step throughout my life.

TABLE OF CONTENTS

TABLE OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

TABLE OF TABLES

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

ABSTRACT

HUMAN AND CLIMATE IMPACTS ON FLOODING VIA REMOTE SENSING, BIG DATA ANALYTICS, AND MODELING

by

Eunsang Cho

Department of Civil and Environmental Engineering, University of New Hampshire, May 2020

Over the last 20 years, the amount of streamflow has greatly increased and spring snowmelt floods have occurred more frequently in the north-central U.S. In the Red River of the North Basin (RRB) overlying portions of North Dakota and Minnesota, six of the 13 major floods over the past 100 years have occurred since the late 1990s. Based on numerous previous studies as well as senior flood forecasters' experiences, recent hydrological changes related to human modifications [e.g. artificial subsurface drainage (SSD) expansion] and climate change are potential causes of notable forecasting failures over the past decade. My dissertation focuses on the operational and scientific gaps in current forecasting models and observational data and provides insights and value to both the practitioner and the research community. First, the current flood forecasting model needs both the location and installation timing of SSD and SSD physics. SSD maps were developed using satellite "big" data and a machine learning technique. Next, using the maps with a land surface model, the impacts of SSD expansion on regional hydrological changes were quantified. In combination with model physics, the inherent uncertainty in the airborne gamma snow survey observations hinders the accurate flood forecasting model. The operational airborne gamma snow water equivalent (SWE) measurements were improved by updating antecedent surface moisture conditions using satellite observations on soil moisture. From a long-term perspective, flood forecasters and state governments need knowledge of historical changes in snowpack and snowmelt to help flood management and to develop strategies to adapt to climate changes. However, historical snowmelt trends have not been quantified in the north-central U.S. due to the limited historical snow data. To overcome this, the current available historical long-term SWE products were evaluated across diverse regions and conditions. Using the most reliable SWE product, a trend analysis quantified the magnitude of change extreme snowpack and melt events over the past 36 years. Collectively, this body of research demonstrates that *human and climate impacts, as well as limited and noisy data, cause uncertainties in flood prediction in the great plains, but integrated approaches using remote sensing, big data analytics, and modeling can quantify the hydrological changes and reduce the uncertainties*. This dissertation improves the practice of flood forecasting in Red River of the North Basin and advances research in hydrology and snow science.

CHAPTER 1

1.1 Introduction

During the last 20 years, the amount of streamflow has greatly increased, and floods have occurred more frequently in the north-central United States. In the Red River of the North Basin (RRB) bordering eastern North Dakota and western Minnesota, six of the 13 major spring snowmelt floods over the past 100 years have occurred since the late 1990s (Rannie, 2015; Rasmussen, 2016).

The National Oceanic and Atmospheric Administration (NOAA) National Weather Service's (NWS) network of River Forecast Centers (RFCs) are responsible for providing river flood forecasts and warnings to protect people and assets. The NWS flood forecasting model estimates the amount of runoff generated from a precipitation or snowmelt event, computes how the water will move downstream through soil storage, and then predicts the flow of water at a given location throughout the forecast period over the U.S. However, accurate flood forecasting has been challenged in the north-central U.S. For example, North Central River Forecasting Center (NCRFC) overestimated a peak flow by 70% of the observed 2013 flow at Fargo, North Dakota (Tuttle et al., 2017; Restrepo, 2014; personal communications with Mike DeWeese, NOAA NCRFC).

There are several potential causes of forecasting failures according to regional flood experts and forecasters. There are limited physics in the NWS flood forecasting system. Even though the current NWS flood forecasting model was launched around the early 2000s, core sub-models (e.g., SAC-SMA and SNOW-17) originally developed over 40 years ago are used with simplistic physics (Burnash et al., 1973; Anderson, 1976). Additionally, while recent land characteristics have been drastically changed by the human modifications in the north-central U.S. (e.g. subsurface drainage expansion), the related physics is not included in the current NWS flood forecasting model. Furthermore, there were very limited reliable SSD records, resulting in a limited understanding of hydrological impacts of the SSD expansion in the north-central U.S. A lack of accurate observations of the water equivalent of the snowpack (snow water equivalent; SWE) is one of the major challenges that hampers accurate flood predictions. The region has relatively sparse observational data with SWE estimates being provided by a few station observations. Even though the operational airborne gamma snow survey collects areal mean SWE data and provides the information to the RFCs, the airborne gamma SWE products have uncertainties due to operational limitations (Carroll, 2001). Lastly, there is limited knowledge of historical snowpack trends for flood forecasters and stakeholders in the north-central U.S. It is known that the north-central U.S. has experienced the largest changes in precipitation and temperature in the U.S. (Peterson et al., 2013). However, historical changes in snowpack and snowmelt are still elusive due to the limited snow records in the north-central U.S. (Cooper et al., 2016; Mote et al., 2018; Pierce et al., 2008). Thus, a better understanding of historical snowpack and snowmelt changes is needed to help flood management and to develop strategies to adapt to climate changes.

My dissertation seeks to tackle a series of challenges needed to overcome operational and scientific gaps for enhanced flood predictions. The dissertation statement is that "*Human and climate impacts, as well as limited and noisy data, cause uncertainties in flood prediction in the great plains, but integrated approaches using remote sensing, big data analytics, and modeling can quantify the hydrological changes and reduce the uncertainties*". In the next 'background' section, the detailed information about current and future flood forecasting models and land characteristics in the north-central U.S. is provided. Subsection 2.1 describes the physics of the current NOAA flood forecasting model with snow and soil moisture sub-models. Section 2.2 introduces the next generation flood forecasting model, National Water Model (NWM), with fundamental physics. Section 2.3 describes regional characteristics in the north-central U.S. with the recent expansion of artificial drainage systems. Section 2.3 provides observations to help flood prediction.

1.2 Background

1.2.1 Operational Flood Forecasting Model

The NOAA NWS network of RFCs is responsible for providing river flood forecasts and warnings to protect people and assets. Among 13 RFCs over the entire U.S., the North Central RFC models river flow in three major basins, Great Lakes, Hudson Bay, and Mississippi river, and issues forecasts at 426 forecast points for 1173 subwatersheds in the north-central U.S.

The NWS flood forecasting model, to be known as the Community Hydrologic Prediction System (CHPS), is a subwatershed-basis spatial scale model and includes the Sacramento soil moisture accounting (SAC-SMA; Burnash et al., 1973; Burnash, 1995) and the SNOW-17 model (Anderson, 1973). SNOW-17 is a conceptual Snow Accumulation and Depletion model that uses air temperature as a proxy for the energy exchange at the snow-air interface with precipitation as the only other input variable. Due to the scarcity of meteorological observations back in the 1970s when the model was developed, the model only uses those two inputs, precipitation and temperature. SNOW-17 is used operationally throughout the RFCs with generally good results but it has difficulties in regions with sparse ground-based observations. Even with accurate temperature and precipitation inputs, estimating snow accumulation is challenging because the form of precipitation is not uniquely a function of air temperature. Local spatial variations in temperature can make distinguishing rainfall from snowfall difficult. Snow metamorphism and melt are also difficult to capture using only temperature and precipitation. For example, sublimation can significantly reduce SWE in regions with blowing snow.

The SAC-SMA model is a spatially lumped continuous soil moisture accounting model that uses heuristic equations governing the transfer of soil moisture from the surface to the soils and the runoff. Using rainfall and snowfall variables calculated from SNOW-17, the model estimates key hydrologic states and processes including upper and lower zone soil moisture, interflow, evapotranspiration, and percolation from a basin. The model is likely suitable for large drainage basins, but not effective for small catchment/watersheds. Because the SAC-SMA runs on a subwatershed-basis, with one value for each model parameter in each subwatershed, the model does not have the ability to represent local variability within a subwatershed such as the recent land-use changes by human activities (e.g. subsurface drainage expansion).

Given the limitation of the forecasting models, flood forecasters augment these models with observational data and their knowledge and experience of flood forecasting. Flood forecast modelers can make manual modifications of the model inputs, parameters, or outputs for the regions where the models are not performing well to improve model performance. Thus, the flood forecasts heavily rely on modeler experiences in regional RFCs. The observational hydrological variables data (e.g. soil moisture and SWE) are ground-based station measurements from the U.S. Army Corp Engineer (USACE) and Soil Climate Analysis Network (SCAN) and airborne gamma data from snow survey program operated by NOAA’s National Operational Hydrologic Remote Sensing Center (NOHRSC). Since there are very sparse ground observations in the north-central U.S. (Tuttle et al., 2017), the modelers primarily consider the airborne gamma radiation data to check the model state variables. If model forecasts are different from observed river flows at major gages beyond typical error ranges, they compare and modify state variables using ground- and airborne measurements. Their modifications are based on +10 years of model experiences (in case of senior levels) with further information from local weather forecast offices (e.g. climatology and localized weather and/or land surface characteristics). The model simulations are typically updated with observation data using the direct insertion assimilation method and regional averaging. The updates had made good forecasts, but recently, the manual modifications did not work (e.g. huge overestimate in 2013 snowmelt flood).

1.3.2 National Water Model: Next-generation Flood Forecasting Standard

Given the limitations of the CHPS model with recent forecast failures, in August 2016, NOAA launched a new integrated hydrologic modeling framework ‘National Water Model (NWM)’ under the new NOAA Water Initiative effort. NOAA intends to replace the RFC CHPS flood forecast models with the National Water Model in the near future. The NWM provides streamflow forecasts over the continental U.S. and represents a major advance over the CHPS platform. The core of the NWM system is the National Center for Atmospheric Research (NCAR)-supported community Weather Research and Forecasting Hydrologic model (WRF-Hydro). It ingests forcing data from a variety of sources including radar-gauge observed precipitation data, and Numerical Weather Prediction forecast data. WRF-Hydro is configured to use the Noah land surface model with multi-parameterization options (Noah-MP) to simulate land surface hydrological processes. Separate water routing modules perform diffusive wave surface routing and saturated subsurface flow routing on a 250 m grid. The NWM runs uncoupled analyses (simulations of current conditions) with look-back periods ranging from 28 hours to 3 hours. These analyses are used to provide initial conditions for the model's forecast runs. Short-range forecasts are executed hourly over the CONUS. The model configurations provide streamflow for 2.7 million river reaches and other hydrologic variables on 1 km and 250 m grids.

Noah-MP is an updated version of the Noah LSM including enhanced hydrological and biophysical processes such as an interactive vegetation canopy (Dickinson et al., 1998), an unconfined aquifer for a dynamic water table and groundwater storage (Ek et al., 2003; Niu et al., 2007), a simple TOPMODEL (TOPography based hydrological MODEL)-based runoff production (Niu et al., 2005), and a multilayer snowpack (Yang and Niu, 2003). However, the current Noah-MP LSM, as well as NWM, does not consider human activities (e.g. subsurface drainage system) (Cohen et al., 2018). In the north-central U.S., the absence of drainage systems in the Noah-MP may generate uncertainties in hydrological processes and flood forecasts (Maidment, 2017).

1.3.3 North-Central U.S.

The North-Central U.S. covers three major basins, Great Lakes, Hudson Bay, and Mississippi river. This work targets the 100,000+km^2 RRB in the Hudson Bay drainage area. The RRB drains parts of western Minnesota, eastern North Dakota, and northeastern South Dakota (Figure 1). The Red River flows north from its headwaters in Wahpeton, North Dakota to Winnipeg, Manitoba, Canada. The Red River in the U.S. is approximately 635 km long and with a drainage area of 101,500 km^2 (Rannie 2016). Current land use in areas is predominantly cropland. The RRB has extremely flat terrain (average slope of 8 cm/km) and poorly permeable clay soil (Tuttle et al., 2017). Thus, the RRB region is very vulnerable to flood events.

Approximately 85% of the annual peak river flows over the last century resulted from the spring snowmelt (Rasmussen 2016). Despite the lower average snowpack compared to snowpacks in the western U.S. (Brasnett 1999), most snowmelt floods in the RRB persist on from weeks to months. For example, the Red River at Grand Forks, ND exceeded the flood stage for 46 days during the 1997 snowmelt flood (Todhunter 2001). The inundatedarea could extend far from the mainstream (e.g., up to 100 km of the floodwater width; Schwert, 2003). The snowmelt floods have damaged private property and infrastructure in this region. The 1997 spring flood caused more than $5 billion of damage in Fargo and Grand Forks, North Dakota, and other communities (Todhunter, 2001).

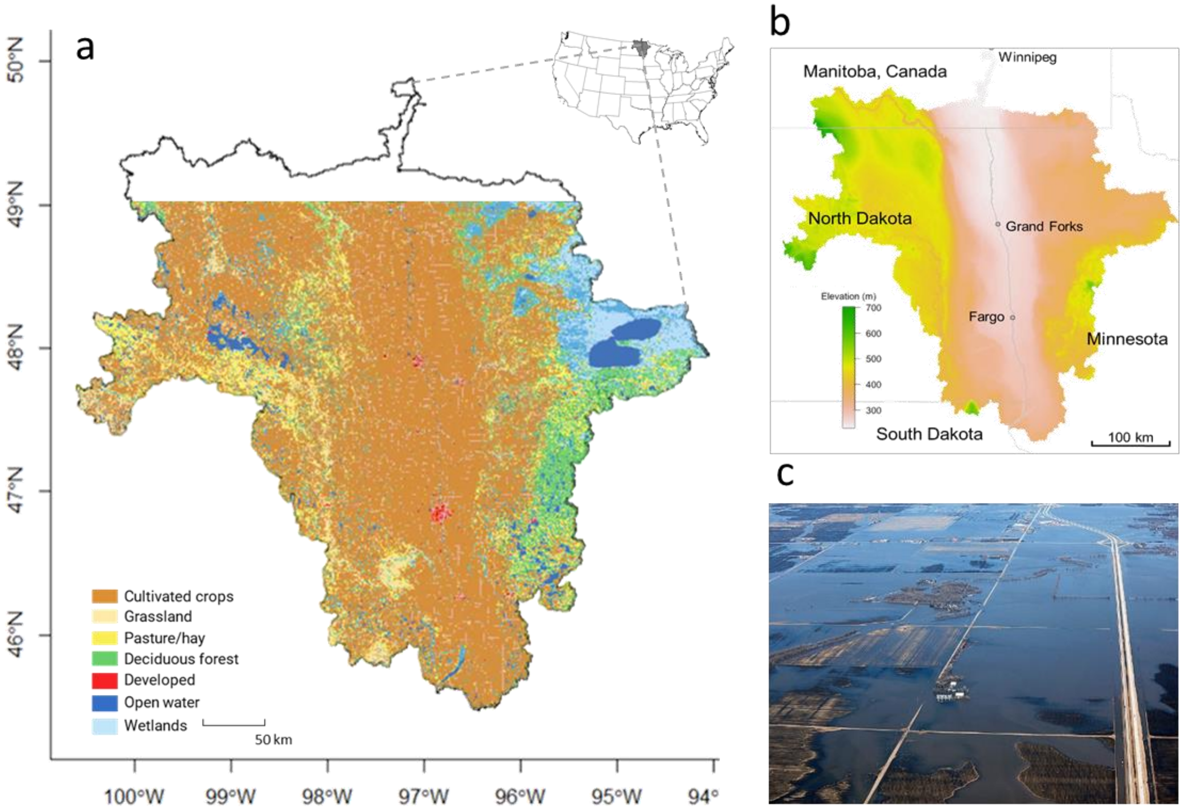


Figure 1. (a) Land cover and (b) Elevation maps of Red River of the North Basin with (c) the spring snowmelt flood, Fargo, North Dakota (April 2011; Courtesy of NOAA NWS)

The north-central U.S., including the RRB, is understudied for snow hydrology and snowmelt-runoff processes as compared to other snow-dominant regions (e.g., western U.S., Sierra Nevada mountains). Ground-based observations of hydrological variables are extremely sparse compared to other regions. In the entire RRB area, frequent, high-quality in-situ measurements of SWE and soil moisture are available from less than 10 and 20 locations, respectively. This scarcity of SWE and soil moisture makes flood forecasting difficult in this region.

Since the 1990s, the north-central regions have experienced remarkable increases in river flow and spring floods. The RRB is one of only two watersheds in the U.S. that is experiencing a demonstrated increase in flood frequency (Hirsch and Ryberg, 2012). The regional hydrological changes are potentially related to the strong manifestation of climate change. The RRB has the largest increase in annual average precipitation and temperature (Peterson et al., 2013). In North Dakota, the annual average temperature for present-day (1986-2016) increased by 1.69°F, relative to the average for the first half of the last century (1901-1960). With a changing climate, artificial subsurface drainage (SSD; so-called 'tile drainage') systems have heavily expanded from 2000 until present (Finocchiaro, 2014; 2016). For example, the total of drainage permits in the Bois de Sioux Watershed, Minnesota were 4.7, 1254, and 2508 km in 1999, 2009, and 2010, respectively (Dollinger et al., 2013). The SSD expansion is considered as a potential issue regarding regional water balance change in the north-central U.S.

1.2.4 Observations

In the north-central U.S., observational hydrological state variables (e.g. soil moisture and SWE) are obtained from ground-based station measurements from the U.S. Army Corp

Engineer (USACE) and Soil Climate Analysis Network (SCAN) and airborne gamma data from snow survey program operated by NOAA's National Operational Hydrologic Remote Sensing Center (NOHRSC). Additional snow observational data consists of snow survey data from local and federal agencies and cooperative network observations. However, because there are few ground station observations and additional snow survey data are not always available in the north-central U.S. (Tuttle et al., 2017), the flood modelers rely heavily on airborne gamma snow program to use reliable SWE observations. Since 1979, the National Oceanic and Atmospheric Administration's (NOAA) Office of Water Prediction's (OWP) airborne gamma radiation snow survey program has provided real-time reliable SWE observations to regional NWS RFCs. The snow survey collects areal mean SWE data over a network of 2,400 flight lines covering 25 states and 7 Canadian provinces (Carroll, 2001). The airborne gamma technique uses the attenuation of the gamma-ray signal by water in the snowpack (any phase) to measure SWE for each flight line. The mean areal gamma SWE value is based on the difference between gamma radiation measurements over bare ground and snow-covered ground.

The accuracy of airborne gamma SWE estimates has been evaluated using numerous ground-based snow observations from snow courses and field campaigns on designated flight lines (Carroll & Schaake, 1983; Carroll & Vose, 1984; Goodison et al., 1984; Glynn et al., 1988; Peck et al., 1971). These studies provided the impetus to develop an airborne gamma SWE program, which has been successfully used for operational flood forecasting over the last 40 years (Carroll, 2001). Currently, the airborne gamma SWE observations support the near-real-time, high spatial resolution (1 km^2 gridded) national snow products (Barrett, 2003).

One drawback of the current gamma snow survey program is an assumption of constant soil moisture conditions from the late fall to winter. Currently, this is a recognized challenge in the program. Flight lines are typically flown once over bare soil in fall to obtain soil moisture and then flown over the snow-covered ground in the winter because the gamma radiation counts include the effect of both SWE and soil moisture. The difference between the fall and winter gamma radiation observations are used to calculate SWE. However, rainfall events or drying after the fall survey, as well as drainage from snow freeze/thaw cycles during the winter, can alter soil moisture. In that case, the program manager should decide to use either the fall survey soil moisture or the default soil moisture (35% of soil moisture) but both may generate uncertainties in gamma SWE. Thus the airborne gamma SWE measurements need to be modified by updating the gamma fall SM estimates immediately before winter onset for enhanced spring flood prediction.

1.3 Objective

The objective of this work is to improve the capability of flood forecasting by overcoming operational and scientific gaps in the current and future flood models for both the practitioners and the research community. A flow chart summarizing current challenges with relevant research tasks is provided in **Figure 2**. This research was conducted in five parts. First, the current flood forecasting model needs both the location and installation timing of SSD and SSD physics. In Chapter 2, SSD maps are developed using satellite “big” data and a machine learning technique. Next, using the maps with a LSM, the impacts of SSD expansion on regional hydrological changes are quantified, described in Chapter 3. With improved model physics, accurate snowpack observations can be directly related to the flood forecasting capacity because most flood events are driven by melting snow in the north-central U.S. However, the inherent uncertainty in the airborne gamma snow survey estimates of SWE hinders accurate flood forecasting model. The third step, described in Chapter 4, is to improve the airborne gamma SWE measurements by updating antecedent soil moisture conditions using satellite observations on soil moisture. From a long-term perspective, flood forecasters and state governments need knowledge of historical changes in snowpack and snowmelt to help flood management and to develop strategies to adapt to climate changes. However, historical snowmelt trends have not been quantified in the north-central U.S. due to the limited historical snow data. To overcome this, the fourth step, described in Chapter 5, is to evaluate the current available historical long-term SWE products across diverse regions and conditions. Using the most reliable SWE product found in Chapter 5, a trend analysis of extreme snowpack and melt events is presented in Chapter 6. Chapter 7 summarizes the main findings and contributions of this work and describes future direction.

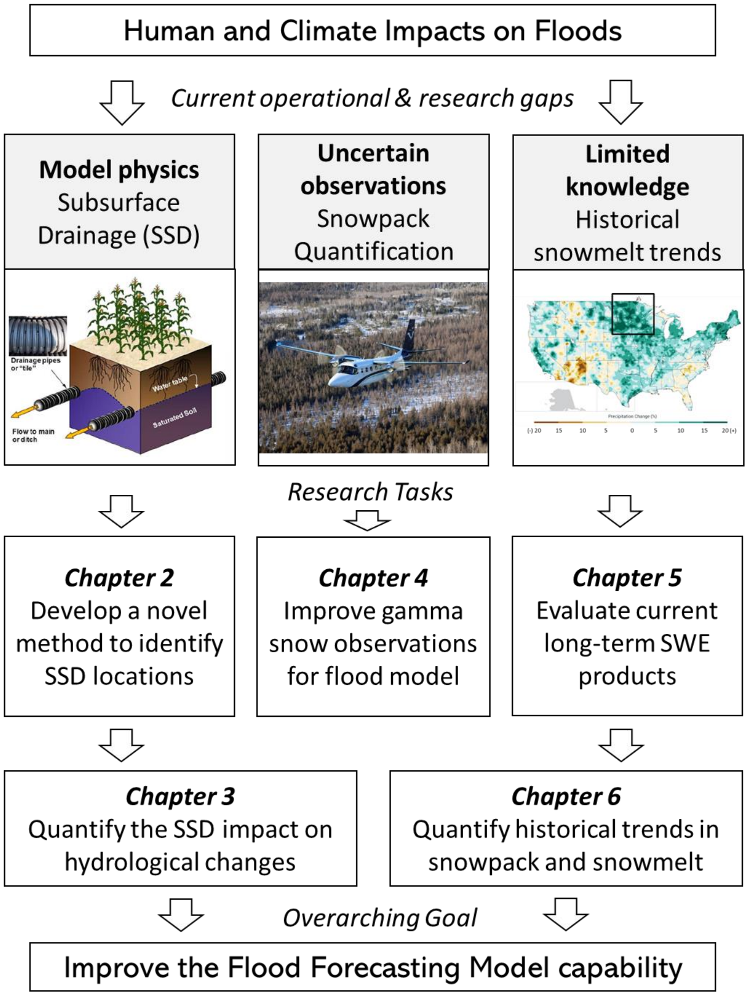


Figure 2. Flow chart of this dissertation with current challenges, research tasks, and overarching goal

CHAPTER 2

Identifying Subsurface Drainage using Satellite Big Data and Machine Learning via Google Earth Engine[1]

2.1 Introduction

In the north central U.S., the amount of streamflow has greatly increased and floods have occurred more frequently during the last 20 years. In the Red River of the North Basin (RRB), six of the 13 major floods over the past 100 years have occurred since the late 1990s (Rannie, 2015; Todhunter, 2001; Tuttle et al., 2017). Numerous studies have been conducted to determine the major causes for the hydrologic changes in the north central U.S. (Foufoula-Georgiou et al., 2015; Frans et al., 2013; Raymond et al., 2008; Schilling et al., 2010). Potential factors include changes in climate, land use and land cover (LULC), including agricultural subsurface drainage installation. Subsurface drainage (SSD) expansion in agricultural landscapes resulting in an increase in cultivated areas is a key cause of regional water balance change (Rogger et al., 2017; Schottler et al., 2014). In the past two decades, SSD systems have exponentially expanded over the agricultural areas in the north central U.S. In the RRB, SSD areas have dramatically increased from 2000 to the present (e.g. in North Dakota, 1.26, 114, and 892 km^2 for 2002, 2008, and 2016, respectively) (Finocchiaro, 2014; 2016; Dollinger et al., 2013).

SD systems are used to remove excess surface water and to lower water tables through subsurface pipe networks installed below the ground surface. When the drainage pipes are

---

[1] Cho, E., J.M. Jacobs, X. Jia, S. Kraatz (2019) Identifying Subsurface Drainage using Satellite Big Data and Machine Learning via Google Earth Engine, *Water Resources Research*, 55(10), 8028-8045, https://doi.org/10.1029/2019WR024892

installed at a certain depth and spacing, the pressure head at the pipes is approximately the atmospheric pressure and the pressure distributions in soil profile horizons change to an equilibrium profile. Thus, the original water tables in the undrained condition are lowered to the equivalent depth of the drainage systems, especially during spring and fall. The effective infiltration rates would be changed by drainage installations due to the altered hydraulic gradient of the upper soil layer above drained pipes (Rodgers et al., 2003; Shokri and Bardsley, 2015; Youngs, 1975).

SSD impacts on runoff, soil moisture dynamics, and evapotranspiration have been studied at a range of spatiotemporal scales (Eastman et al., 2010; Frans et al., 2013; Kelly et al., 2017; King et al., 2014; Kladivko et al., 2004; Lenhart et al., 2011; Rahman et al., 2014; Randall et al., 2003; Schottler et al., 2014; Williams et al., 2015). At a field scale, Kladivko et al. (2004) showed that SSD-induced water yields were 8 to 26% of annual rainfall in southeastern Indiana, depending on the year and the drain spacing. Eastman et al. (2010) found that the subsurface-drained field discharged four times more water than the naturally drained field for their clay loam sites. At a watershed scale, King et al. (2014) reported that about 21% of annual precipitation and 47% of total watershed discharge was generated from SSD in central Ohio. Williams et al. (2015) concluded that SSD discharge contributed 56% of the annual watershed discharge in the same Ohio watershed. At a larger scale, Frans et al. (2013) showed that SSD increased annual streamflow up to 40% locally in the Upper Mississippi River basin. Schottler et al. (2014) compared a change in water yield between two historical periods (1940-1974 and 1975-2009) in watershed scale. They found that SSD expansion is likely the major driver of increased streamflow in 21 Minnesota agricultural watersheds. Kelly et al. (2017) also concluded that the extensive SSD systems in agricultural basins have contributed to the increase in river flow at the large basin scale.

Despite the increased water yield, it is possible that SSD could mitigate downstream flooding by allowing surface runoff to infiltrate and be released at a slower rate. As recently as 2013, the National Oceanic and Atmospheric Administration's (NOAA) National Weather Service (NWS) North Central River Forecast Center (NCRFC) predicted a peak flow that exceeded the observed by 70% in the RRB (Tuttle et al., 2017). Because the current flood forecasting system does not consider SSD information, it is still an open question as to the dominant processes that are affected by SSD in the region. However, it has been observed that as SSD systems have expanded, operational flood forecasting has become more difficult due to limited information about spatial and temporal SSD expansion (personal communications with Mike DeWeese and Pedro Restrepo, NOAA NCRFC).

Due to the paucity of SSD data, the results of the previous studies also had considerable uncertainties. Schottler et al. (2014) indicated that the unexplained portion of evapotranspiration change in the long-term water balance approach is due to SSD change, but did not have the supporting SSD data. While Kelly et al. (2017) had county-level drainage data for five census years to assess SSD impact on runoff patterns, they noted inconsistencies and errors of the census data with concerns about limited SSD records in the U.S.

Most previous studies that have attempted to map SSD systems focused on delineating subsurface drained lines (or areas) at a field or catchment scales and used Geographical Information System (GIS)-based analyses and/or aerial image processing techniques (Naz et al., 2009; Naz and Bowling, 2008; Sugg, 2007; Tetzlaff et al., 2009a & b; Zhang et al., 2014). The 1992 National Resource Inventory (NRI) dataset provided potential extents of subsurface drains in continental United States (Wieczorek, 2004). The NRI dataset was created with GIS and database management tools using collections at more than 800,000 sample sites over the U.S. Sugg (2007) estimated SSD percentage for each county based on the GIS-based soil drainage class. They compared their results with the NRI drainage map and developed a SSD map at the county scale. Sui (2007) also used a GIS-based analysis of land cover, soil, and slope datasets to classify the SSD areas for cropland in Indiana where the soils are poorly drained, and the slope is less than 2%. However, the SSD mapping studies could not validate their results due to the lack of SSD data (Naz et al., 2009; Sugg, 2007). Infrared aerial photographs have been used to map subsurface drain lines and to delineate wet and drained areas in a field (Verma et al., 1996). Soils over subsurface drained areas have higher reflectance in the infrared spectrum because these areas tend to dry faster than the soil at other regions. Previous studies found that the best time to take imagery to be used for SSD delineation is within 3 days after a 25 mm or greater rainfall event (Varner et al., 2002; Northcott et al., 2000). A combination of high resolution (1-m) color (or black and white) infrared aerial images with land cover, soil, and topography data provided a map of individual drainage lines in west-central Indiana (Naz et al., 2009). Tlapakova et al. (2015) provided an example of manifestations of SSD systems in color aerial images and suggested best land conditions for the optimal SSD identification. Using an optical camera and unmanned aerial vehicle (UAV) system, Zhang et al. (2014) developed a mosaiced SSD map from infrared color composite imagery. While the aerial imagery approaches allow targeted study of watersheds, they are expensive and may be limited by weather and the availability of resources.

Satellite remote sensing data offers the ability to observe temporal changes in surface conditions due to SSD at large spatial extents. Gökkaya et al. (2017) and Møller et al. (2018)

provide evidence of SSD induced surface changes using Landsat satellite images. However, they had few satellite observations due to limited cloud-free images and data processing requirements. Jacobs et al. (2017) showed that Moderate Resolution Image Spectroradiometer (MODIS) land surface temperature and Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) soil moisture products could detect physical effects of SSD systems on soil thermal-moisture dynamics. In addition to these products, there are many other satellite products that might show the SSD signature. However, traditional analysis techniques, such as image processing techniques and the GIS-based decision tree classification commonly used in previous studies (Gökkaya et al., 2017; Naz and Bowling, 2008; Sugg, 2007), are not well suited to manage and analyze terabyte-size satellite remote sensing datasets for SSD detection. In these cases, machine learning (ML) techniques have demonstrated value (Belgiu and Drăguţ, 2016; McCabe et al., 2017; Møller et al., 2018; Shen, 2018; Tao et al., 2016).

Random Forest Machine Learning (RFML) is a supervised classification algorithm that constructs a multitude of decision trees and predicts class labels, using a random subset of training samples and variables (Breiman, 2001). The RFML has become popular within the remote sensing and hydrology communities due to its accuracy (Belgiu and Drăguţ, 2016; Gomez et al., 2016; Petty and Dhingra, 2018). For land surface and crop type monitoring, the RFML has been shown to produce higher accuracies than other ML techniques such as Maximum Likelihood Classifier, Neural Network, and Support Vector Machine (Gomez et al., 2016; Ma et al., 2017; Ok et al., 2012). Also, it has been widely used in the field of hydrological predictions due to its capacity to determine variable importance, its robustness to data reduction, and that does not over-fit (Petty and Dhingra, 2018; Shortridge et al., 2016; Wang et al., 2015). Compared to other techniques, however, the RFML method has inherent limitations including 1) complexity which makes less straightforward to understand the relationship in the input data and 2) significant and timing-consuming of computational requirements to construct the algorithm.

The Google Earth Engine’s (GEE) cloud computing platform (Gorelick et al., 2017) provides the ability to manage very large satellite and model datasets to analyze them using ML techniques. The GEE is designed to provide access to high-performance computing resources for processing massive geospatial datasets, without technical hurdles (e.g. data download and storage, handling obscure file formats, and managing databases). Because a variety of geospatial datasets including historical and current satellite and aerial imaging systems can be freely accessed and analyzed, the GEE has been widely used in computationally expensive hydrological, agricultural and socio-economic studies (Deines et al., 2017; Ge et al., 2019; Jin et al., 2019; Xie et al., 2019).

Here, we focus on developing SSD maps to improve the capability of flood forecasting in agricultural landscapes across the RRB. The RFML algorithm is used to develop annual SSD maps in the GEE computing platform. We also seek to understand which of the related, globally available vegetation, thermal, moisture, and climate datasets from multi-scale satellites and models can be used to identify SSD areas and with what accuracy. The accuracy of these maps is assessed using SSD permit records in the Bois de Sioux Watershed (BdSW) in Minnesota and the North Dakota portion of the RRB region (ND-RRB).

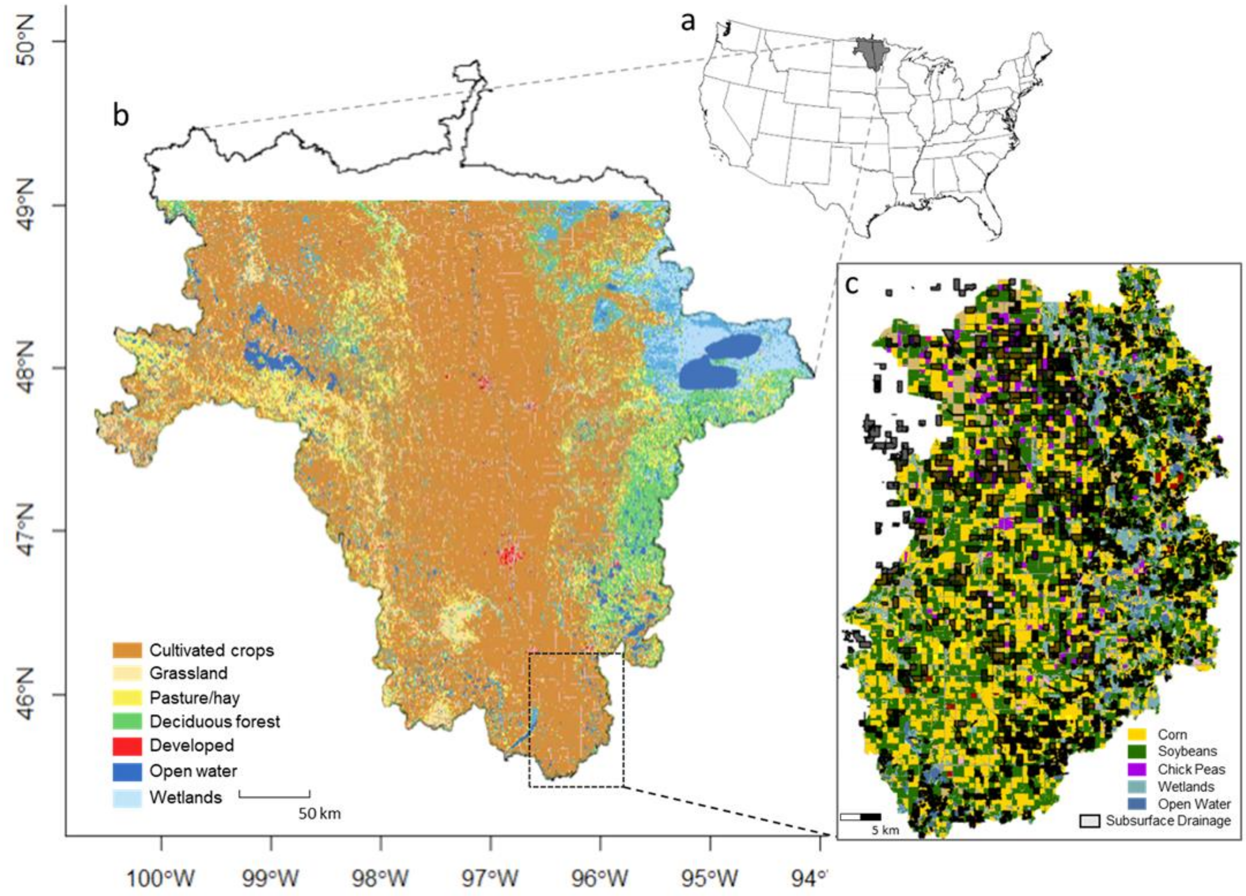


Figure 1. Study area location and land cover map. (a) Red River of the North Basin; (b) Land cover classification from USGS National Land Cover Database 2011; and (c) Cropland Data Layers (CDL) with tile drained area in 2017 noted in Bois de Sioux Watershed (BdSW).

2.2 Study Area

The Red River of the North Basin overlies portions of North Dakota, South Dakota, and Minnesota (Figure 1). Its main stem marks the border between North Dakota and Minnesota. The river flows north from Wahpeton, ND to the U.S.–Canada border, and then through Winnipeg, Manitoba, Canada. The basin drainage area is about 112,200 km^2, with 885 km long from U.S. Geological Survey (USGS) Watershed Boundary Dataset (HUC04). Along the distance of the main stem, it drops only 72 m, for an average gradient of 0.08 m/km. In the RRB, agricultural SSD systems have increasingly used to drain fields since the late 1990s due to the region's flat topography and low-permeability soils. The NOAA flood forecasters and water resource experts in RRB identified the rapid increase in the SSD systems as a likely culprit for the changed river flows and floods because SSD alters direct runoff, soil moisture, and evaporation seasonally (Rijal et al., 2012; Schottler et al., 2014).

2.3 Method

2.3.1 Datasets

Working in the GEE cloud computing platform, the datasets from multi-source satellite and model assimilation products were used (total 1.4 terabytes). Table 1 lists the 36 seasonal and annual layers that were generated including 16 vegetation layers (top 16 lines at the table), 8 soil-climate variable layers (next 8 lines), and 12 thermal-moisture layers (12 lines from the bottom) for 2009, 2011, 2014, and 2017. All 36 input layers were disaggregated to 30 m pixel resolution. The datasets generally fit into three categories: vegetation, thermal-moisture, and climate-land variables. The four years were selected based on land surface conditions (dry/wet) particularly in spring based on spring mean precipitation and soil moisture. 2009 and 2011 were selected because they have experienced spring snowmelt floods over the RRB. Even though years 2014 and 2017 were not as wet as 2009 and 2011, they were included to examine whether the RFML method can capture continuous SSD expansion over the years that have occurred in RRB. The magnitude of year-over-year hydrologic variability for 2009 through 2017 is shown via a hydrograph at Fargo, North Dakota (USGS: 0505400), which a major streamflow gage in the RRB (Figure S1).

Table 1. Summary of variables used in RFML including time period, resolution, and data source. All input variables were accessed through the GEE's data archive, except for the three 30-m soil property datasets from POLARIS (available at www.polaris.earth; Chaney et al., 2016, 2019),

which were manually uploaded to the GEE for RFML classification. 16 vegetation layers appear in the top 16 rows (EVI, GI, NDVI, and NDWI), 12 thermal-moisture layers follow the vegetation layers (SM, LST, STR1 & 2), and 8 soil-climate variable layers are the remaining 8 rows in the table (Preci, Aridity, Cropland, and three soil properties).

Variable (Full Name)	**Short Name**	**Time Period**	**Res (m^2)**	**Source**
Spring mean EVI	EVI_spr_mean	1 May – 30 Jun	30	Landsat
Spring mean GI	GI_spr_mean	1 May – 30 Jun	30	Landsat
Spring mean NDVI	NDVI_spr_mean	1 May – 30 Jun	30	Landsat
Spring mean NDWI	NDWI_spr_mean	1 May – 30 Jun	30	Landsat
Spring range in EVI	EVI_spr_range	1 May – 30 Jun	30	Landsat
Spring range in GI	GI_spr_range	1 May – 30 Jun	30	Landsat
Spring range in NDVI	NDVI_spr_range	1 May – 30 Jun	30	Landsat
Spring range in NDWI	NDWI_spr_range	1 May – 30 Jun	30	Landsat
Summer mean EVI	EVI_sum_mean	1 Jul – 30 Sep	30	Landsat
Summer mean GI	GI_sum_mean	1 Jul – 30 Sep	30	Landsat
Summer mean NDVI	NDVI_sum_mean	1 Jul – 30 Sep	30	Landsat
Summer mean NDWI	NDWI_sum_mean	1 Jul – 30 Sep	30	Landsat
Summer range in EVI	EVI_sum_range	1 Jul – 30 Sep	30	Landsat
Summer range in GI	GI_sum_range	1 Jul – 30 Sep	30	Landsat
Summer range in NDVI	NDVI_sum_range	1 Jul – 30 Sep	30	Landsat
Summer range in NDWI	NDWI_sum_range	1 Jul – 30 Sep	30	Landsat
Spring mean soil moisture	SM_spr_mean	1 May – 30 Jun	25,000	SMOS (NASA-USDA)
Spring range soil moisture	SM_spr_range	1 May – 30 Jun	25,000	SMOS (NASA-USDA)
Spring mean LST	LST_spr_mean	1 May – 30 Jun	1,000	Terra MODIS
Spring range LST	LST_spr_range	1 May – 30 Jun	1,000	Terra MODIS
Spring mean STR 1	STR1_spr_mean	1 May – 30 Jun	30	Landsat
Spring mean STR 2	STR2_spr_mean	1 May – 30 Jun	30	Landsat
Spring range STR 1	STR1_spr_range	1 May – 30 Jun	30	Landsat
Spring range STR 2	STR2_spr_range	1 May – 30 Jun	30	Landsat
Summer mean STR 1	STR1_sum_mean	1 Jul – 30 Sep	30	Landsat
Summer mean STR 2	STR2_sum_mean	1 Jul – 30 Sep	30	Landsat
Summer range STR 1	STR1_sum_range	1 Jul – 30 Sep	30	Landsat
Summer range STR 2	STR2_sum_range	1 Jul – 30 Sep	30	Landsat
Growing season precipitation	Preci_grow	1 May – 30 Sep	4,000	GRIDMET
Early season precipitation	Preci_early	1 Dec – 30 Apr	4,000	GRIDMET
Annual precipitation	Preci_ann	1 Dec – 30 Sep	4,000	GRIDMET
Aridity	Aridity	1 May – 30 Sep	4,000	GRIDMET

Cropland Data Layers	Cropland	NA	30	USDA NASS
Clay percentage	clay_perc	NA	30	POLARIS
Available soil water content	awc	NA	30	POLARIS
Saturated hydraulic conductivity	ksat	NA	30	POLARIS

Seasonal mean and range (max – min) composites of four vegetation indices were produced using spectral reflectance products from Landsat 7 Enhanced Thematic Mapper Plus (ETM+) and Landsat 8 Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) (30 m resolution): (1) the normalized difference vegetation index (NDVI); (2) the normalized difference Water index (NDWI) which is highly sensitive to vegetation water content (Jackson, 2004); (3) the enhanced vegetation index (EVI) which is an improved vegetation index with de-coupling of the background signal of canopy (Huete et al., 2002); and (4) the greenness index (GI) that is sensitive to the irrigation signal (Deines et al., 2017). The vegetation indices were divided seasonally for the spring (April – June) and summer (July – September) periods to include vegetation growth and their seasonal changes into the RFML model. The detailed variable equations are included in Appendix (Text S1).

For thermal-moisture variables, two shortwave infrared transformed reflectances (STR) from Landsat 7 and 8 were used, which have a linear relationship with soil moisture content (Sadeghi et al., 2015). Land surface temperature (LST) from MODIS (1 km resolution) and surface soil moisture from Soil Moisture Ocean Salinity (SMOS) satellite (25 km resolution) were also used, but the soil moisture data were only available from 2010 (Kerr et al., 2010).

Climate-land variables can improve classification accuracy by refining wet versus dry year patterns and including crop type and soil property effects. Total precipitation for the growing (May to Oct) and non-growing (Dec in the previous year to Apr) seasons, and aridity (precipitation scaled by reference evapotranspiration, May to Aug) were assembled from the University of Idaho Gridded Surface Meteorological Dataset (4 km resolution) (Abatzoglou, 2013). Annual crop types from Cropland Data Layers (CDL) were provided by the USDA National Agricultural Statistics Service (NASS). Three soil property maps, available water content, saturated hydraulic conductivity, and clay percent of the soils at 0-5 cm, from PLARIS database (30 m spatial resolution) (Chaney et al., 2016, 2019) were also used in the RFML classification. Land cover and slope information were used to make the non-SSD area (e.g. non-agricultural and high slope area). We identified low gradient cultivated crop areas (slope < 2%) using the USGS National Land Cover Dataset (NLCD) and the USGS National Elevation Dataset (Naz et al., 2009). The input products with coarse resolutions (e.g. 1, 4, or 25 km grid) were disaggregated/resampled to the finer resolution (30-m grid) using a nearest neighbor resampling by default in the GEE (https://developers.google.com/earth-engine/resample).

2.3.2 Subsurface drainage permit records for training and validation data

Two separate SSD permit records were used to develop training points and to validate the RFML maps, assuming the permit records are ground “truth” SSD measurements: (1) a sub-basin SSD records obtained from the BdSW district in Minnesota (http://www.bdswd.com) and (2) the USGS records obtained from the North Dakota State Water Commission (Finocchiaro, 2016). The BdSW SSD permit records were collected from 1999 to the present over the BdSW region in Minnesota (Figure 1c). The annual SSD records contain locations of subsurface permit lines and the request and approved dates as GIS shape files as well as engineering design specifications. SSD installation is estimated to occur within three months of permit approval. Because the BdSW SSD record is a line shape file, the SSD lines were buffered to provide an effective extent. A 30 m buffer (15 m buffer on either side of the line) was used based on typical SSD separation and guidance from the region's agricultural engineers (Naz et al., 2009). The USGS SSD records (https://www.sciencebase.gov) were issued by the ND State Water Commission and collected by the USGS over the North Dakota from 1993 to 2016 (Finocchiaro, 2016). The USGS SSD records provide polygon outlines of the permit areas and approval dates.

Previous studies used the US Census of Agriculture drainage data (USDA National Agricultural Statistics Service, 2014; Kelly et al., 2017; Krapu et al., 2018). The Census data are extremely limited because the five available census years only provide a single SSD value for each county and year in several US states, do not include areas less than 2 km^2 (Kelly et al., 2017). In contrast to previously used Census SSD data, the BdSW and USGS SSD permit records provide greatly improved information (e.g. exact SSD locations and approval dates) and are well-suited for developing training and validation data.

That said, the BdSW and USGS SSD records are not perfect. Errors in the records may occur if farmers did not install the system or if they were installed them later than originally planned. The permit records also depend on an institution's policy. The North Dakota SSD permit policy was changed in 2011, likely resulting in uncertainties about the SSD permit records (North Dakota Century Code; https://www.legis.nd.gov/cencode/t61.html). After 2011, farmers in SSD no longer needed to obtain a permit to install SSD systems if the SSD footprint is less than 0.32 km^2 (80 acres). Thus, in small fields, SSD is underreported.

The RFML uses the satellite products to identify changes in surface vegetation and soil water state that result from SSD. However, even within a single field, SSD effects depend on the soils, slope, and vegetation as well as the distance from the SSD. The satellite product's spatial resolution (30 m) is relatively fine compared to a field scale and captures within field variations of SSD effects. Additionally, farmers install SSD systems over their fields with different SSD intensities (e.g. depth and spacing) depending on field-specific soils, crop type, and cost (Blann et al., 2009). Thus, matching satellite detected effects of SSD to permitted SSD locations is somewhat problematic. Neither the USGS polygon outlines of fields with SSD nor the static 30-m buffered SSD lines provided for the BdSW SSD, areas can be expected to perfectly capture the portion of the field that is affected hydrologically by SSD as resolved by 30 m satellite observations.

In this study, the annual accumulated SSD permit records were used to classify SSD and undrained (UD) areas. The low-slope cropland areas (slope < 2%) without the SSD permit areas were defined as the UD areas. Pixels were then randomly selected from the buffered SSD and UD areas using a random sample generator in the R package. For the BdSW, the total number of sample pixels is 2164, 2150, 4710, and 4746 for 2009, 2011, 2014, and 2017, respectively. For the ND-RRB, training sample pixels were directly selected from the accumulated SSD and UD areas for each year. There were total 9016, 8880, 8766, and 8754 sample pixels for 2009, 2011, 2014, and 2017, respectively. For each region and year, half of the sample pixels were randomly selected as training pixels and the remaining 50% were used to validate the model outputs.

2.3.3 Random Forest machine learning (RFML) classification

RFML is an ensemble ML classification method comprised of a collection of tree-structured classifiers (Breiman, 2001). The major principle behind ensemble learning methods is that a group of weak classifiers (or learner) can be joined to form a strong classifier. In ensemble learning, two well-known methods are boosting (Shapire et al., 1998) and bootstrap aggregation (or “bagging”; Breiman, 1996) of classification trees. Compared to a single classification tree, the bagging method is used to reduce the variance of the tree. The method creates several subsets of bootstrapped samples from original training dataset chosen randomly with replacement. Each collection of subset samples is used to independently train a classification tree. In the end, an ensemble of all different trees (models) is constructed. A simple majority vote is taken for prediction which is more robust than a single classification tree. However, bagging method as an ensemble learning often do not work because classification trees in bagging are developed independently by using all variables. The bagging method is allowed to look through all variables to choose the best split-point (specific variable and its value) at each node in each tree. If there exists one very strong variable for prediction, most or all of the bagged trees use the strong predictor in the top split. In this case, most bagged trees look very similar and their predictions also highly correlated. This means that the results from the highly correlated trees does not accomplish a substantial reduction in variance over a single tree.

To overcome the limitation, the RFML is an improved extension over the bagging which applies randomness to the procedure when taking a subset of variables rather than using all variables to grow trees. In other word, while in decision tree each node is split using the best among all variables, in a random forest each node is split by the best among the subset of variables (Liaw and Wiener, 2002). For example, the first tree in a random forest is constructed using a few variables, not all 36 variables, and the other trees can be developed by using different numbers of variables until each node is ‘pure’. The development procedure in RFML model and classification processes are illustrated in Figure 2.

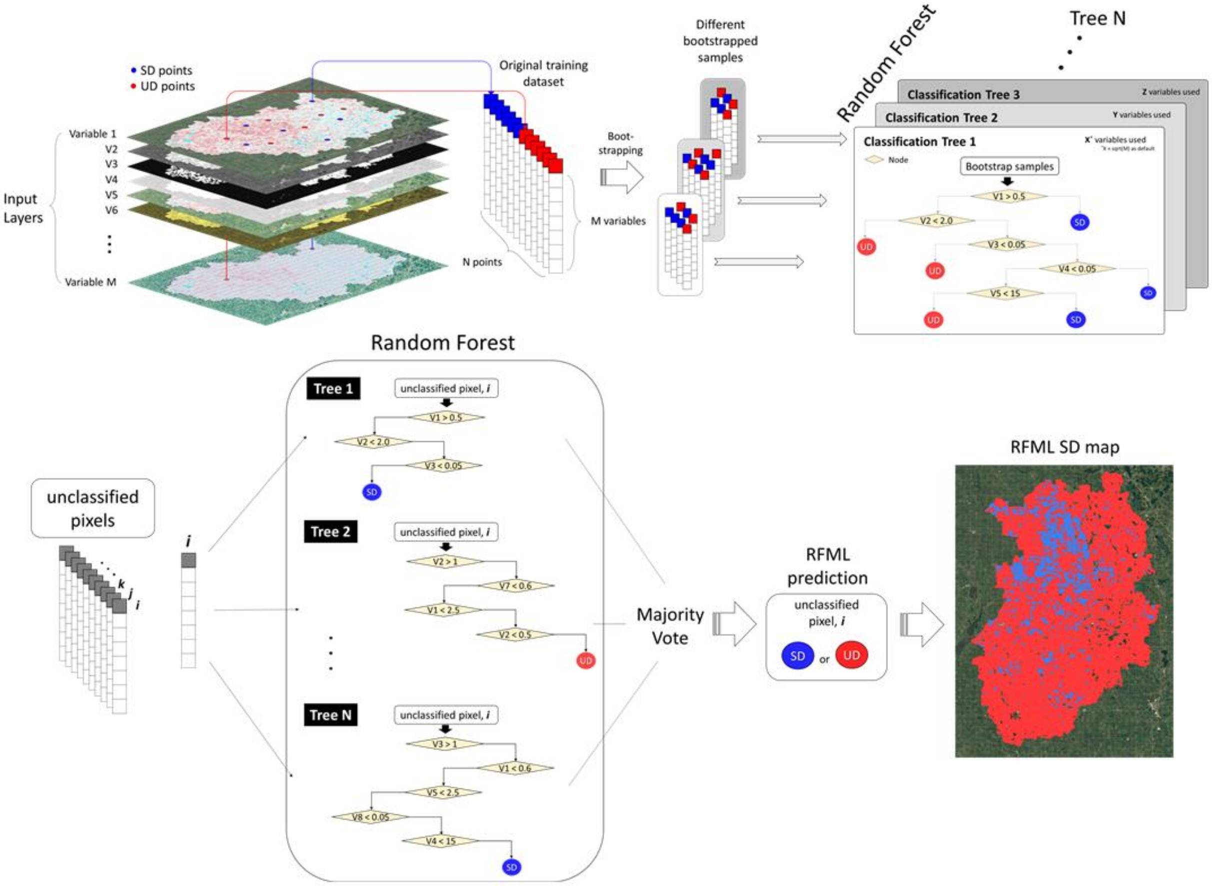


Figure 2. Scheme of construction of the Random Forest Machine Learning (RFML) model using training data and classification processes using the RFML model for classifying subsurface drainage (SSD) / undrained (UD) areas

With the growth of satellite ‘big’ data in hydrology, the RFML was widely used for tasks such as streamflow prediction and flood risk assessment, which have been notoriously difficult with traditional approaches (Belgiu and Drăguţ, 2016; Ma et al., 2017; Petty and Dhingra, 2018; Wang et al., 2015). In this study, to determine if the RFML SSD outputs are affected by spatial scale (basin versus watershed), we developed and ran the RFML model using the same input variables for two regions with different scales, separately.

For each of the training pixels, values were extracted from the 36 input layers to train the RFML algorithm. The full training dataset was used to train RF classifiers for each year in the

GEE. An RF classifier was performed with 300 trees. We applied the annual classifier to the corresponding year, 2009, 2011, 2014, or 2017. After the initial classification, a 3 x 3 majority filter was applied to remove isolated SSD pixels which were sparsely scattered on maps, because SSD systems are usually installed in fields (e.g. a few hundred meters). In RFML, the outcome of implicit feature relevance for each variable is visualized by the Gini index (Breiman, 2001). A Gini index analysis was conducted using R (Liaw and Wiener, 2002) because the GEE does not provide relative importance metrics. The mean decrease in Gini index is a measure of how each variable contributes to the RFML classification.

The Gini index, $i(\tau)$, at each node (τ) within a tree (T) of the RFML is defined as:

$$i(\tau) = 1 - \sum_{j=1}^{n} {P_j}^2 \quad \text{within the tree T} \tag{1}$$

where P_j is the fraction of the N_j samples from class j out of the total of N samples at node τ in T. For a binary class j = {SSD, UD} like the current study, the Gini index is calculated by

$$i(\tau) = 1 - {P_{SD}}^2 - {P_{UD}}^2 \tag{2}$$

The decrease in Gini index, $\Delta i(\tau)$, that results from splitting the samples to two sub-nodes τ_{SD} and τ_{UD} (with respective sample fractions $P_{SD} = \frac{N_{\tau_{SD}}}{N_\tau}$ and $P_{UD} = \frac{N_{\tau_{UD}}}{N_\tau}$) by threshold t_θ on variable θ is defined as

$$\Delta i_\theta(\tau) = i(\tau) - P_{SD} \cdot i(\tau_{SD}) - P_{UD} \cdot i(\tau_{UD}) \tag{3}$$

Mean decrease in Gini index for a variable θ is the average of a variable's total decrease in node impurity over all trees N_T in the forest, weighted by the proportion of samples for all nodes τ where variable θ is used.

$$Mean\ decrease\ in\ Gini\ index(\theta) = \frac{1}{N_T}\sum^{T}\sum^{\tau} pr\,(\tau)\cdot \Delta i_{\theta}(\tau, T) \tag{4}$$

where $pr(\tau)$ is the proportion N_τ/N of samples reaching τ.

2.3.4 Accuracy assessment (Validation)

The BdSW and USGS SSD permit records were used separately to perform an accuracy assessment based on a pixel-level confusion matrix and subwatershed- and subbasin-level statistics. For the BdSW, a pixel-by-pixel comparison was conducted. The number of correct and incorrect predictions was summarized as a confusion matrix using the validation pixels, 1082, 1075, 2355, and 2373 pixels for years 2009, 2011, 2014 and 2017, respectively. For the subwatershed-level accuracy assessment within BdSW, the RFML SSD area and the SSD permit area were aggregated for each of the 34 subwatersheds after masking all training pixels. For the larger scale analysis, a pixel-level comparison was conducted in the same way with the BdSW analysis, but using a larger numbers of validation pixels, 4508, 4440, 4383, and 4377 pixels for years 2009, 2011, 2014 and 2017, respectively. For the subbasin-level accuracy assessment, RFML SSD areas and the USGS SSD permit data were aggregated and compared using the 48 NOAA river forecasting subbasins.

2.4 Results and Discussion

2.4.1 Classification Performance

The RFML classifier, using a combination of satellite-based vegetation, thermal, and soil moisture products, along with soil property and climate variables, prodSuced annual SSD maps for BdSW and ND-RRB in 2009, 2011, 2014 and 2017. Using 2,240 SD and 4,630 UD validation pixels, the pixel-level evaluation at BdSW had an overall accuracy of 77% (True positive: 1,018 SSD pixels and True negative: 4,262 UD pixels) for the four years with accuracies ranging 72% to 84% for individual years (Table 2). For undrained pixels, the RFML model was 92% accurate with a range of 88% to 98%. SSD pixels had relatively lower accuracies with 45% total accuracy. In the BdSW, there is good qualitative agreement between the SSD expansion maps, SSD permit areas and RFML maps (Figure 3a). The RFML model results indicate that SSD extent is 2.5, 3.4, 11.2, and 16.1% of total BdSW area for 2009, 2011, 2014, and 2017, respectively. These values are quantitatively similar to the extent found using the SSD permit records, 1.9, 3.2, 10.3, and 14.3%, from 2009, 2011, 2014 and 2017, respectively. The RFML SSD extents are slightly greater than those determined from permit data, by 0.2 – 1.8%.

Table 2. Point-based accuracy assessment for the four wet years (2009, 2011, 2014, and 2017) between RFML predicted values and SSD-permit based data in the BdSW and the ND-RRB.

Year	Class	BdSW			RRB		
		RFML SSD	RFML UD	Overall Accuracy	RFML SSD	RFML UD	Overall Accuracy
2009	SSD	19.8%	4.3%	79.0% (850/1,076)	59.4%	1.0%	90.7% (4,170/4,596)
	UD	80.2%	95.7%		40.6%	99.0%	
2011	SSD	35.9%	2.4%	83.9% (894/1,066)	40.3%	1.3%	86.6% (3,909/4,512)
	UD	64.1%	97.6%		59.7%	98.7%	
2014	SSD	51.9%	8.7%	77.3% (1,820/2,355)	26.5%	2.1%	82.8% (3,693/4,461)
	UD	48.1%	91.3%		73.5%	97.9%	

2017	SSD	48.7%	12.4%	72.3% (1,716/2,373)	19.8%	2.0%	81.6% (3,632/4,453)
	UD	51.3%	87.6%		80.2%	98.0%	
Overall Accuracy		45.4% (1,018/ 2,240)	92.1% (4,262/ 4,630)	**76.9%** (5,280/ 6,870)	39.9% (1,380/ 3,460)	98.4% (14,024/ 14,247)	**87.0%** (15,204/ 17,708)

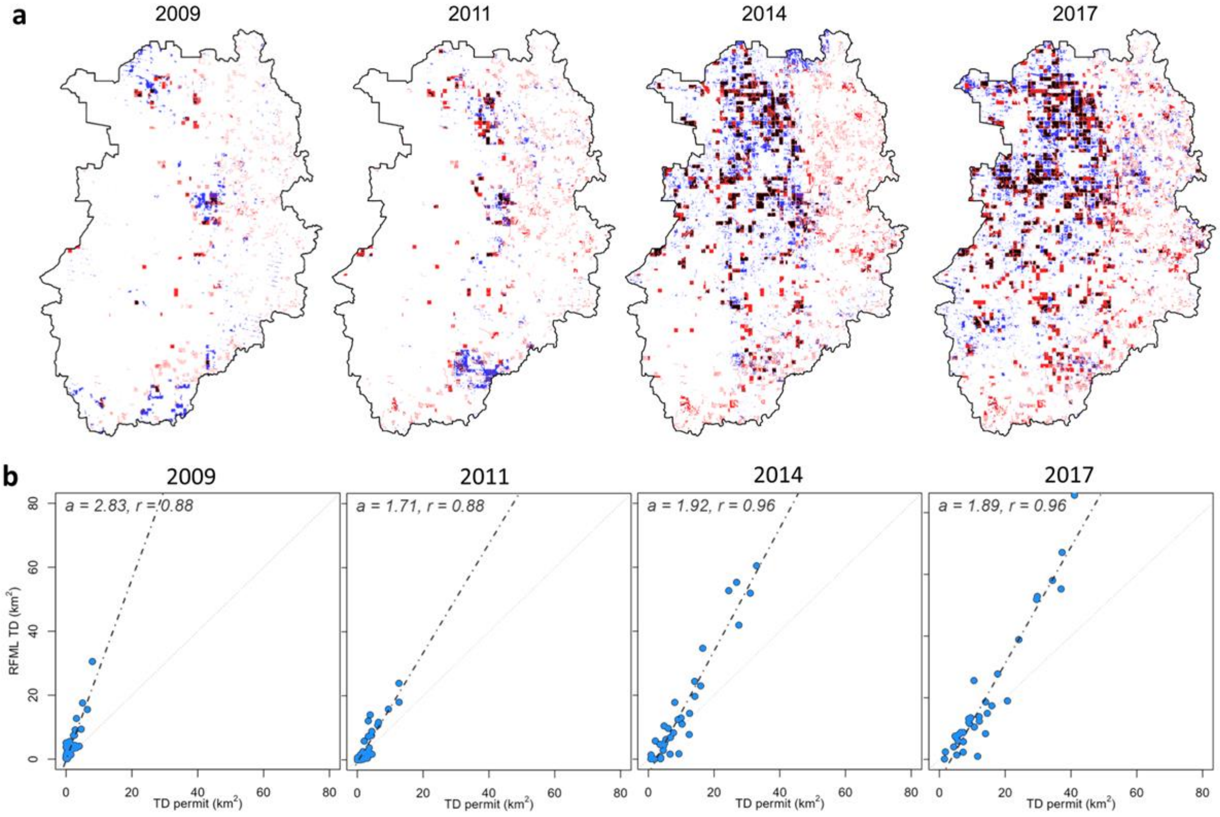


Figure 3. (a) Subsurface drainage expansion in Bois de Sioux watershed, Minnesota in 2009, 2011, 2014, and 2017 from SSD permit records (red color) and predicted SSD areas (blue color) derived by Random Forest machine learning (RFML) classification in the GEE. Black color indicates overlapped SSD areas of the two sources. (b) Subwatershed (HUC12)-level accuracy assessment over BdSW, Minnesota (N=34). Subsurface drained permit area from the BdSW district permit records compared with subsurface drained area from RFML classified maps against a 1:1 line (light dashed). Agreement between the two datasets was assessed with correlation coefficient (r) metrics from simple linear regression (trend line = thick dashed line, a = slope).

Aggregated to the subwatershed-level (HUC12), the RFML SSD estimates showed strong correlation (r = 0.88 – 0.96) with SD permit areas (Figure 3b). However, RFML consistently overestimated subsurface drained areas in each subwatershed in BdSW. The overestimated SSD was also found in other dry years (see Figure S2). A review of individual fields suggests that the

RFML model may be capable of identifying SSD effects even in small areas within a field where SSD systems can exist, but for which there is no permit record (Fields 1 and 2 in Figure 4). The RFML identified numerous small fields as having SSD that were likely not included in the permit record because permits are not required when a field is smaller than 0.32 km^2. Additionally, the RFML detected the extent of the installed SSD effect appears to frequently extend well beyond the 30 m buffer recommended in earlier literature and expert guidance (All fields in Figure 4).

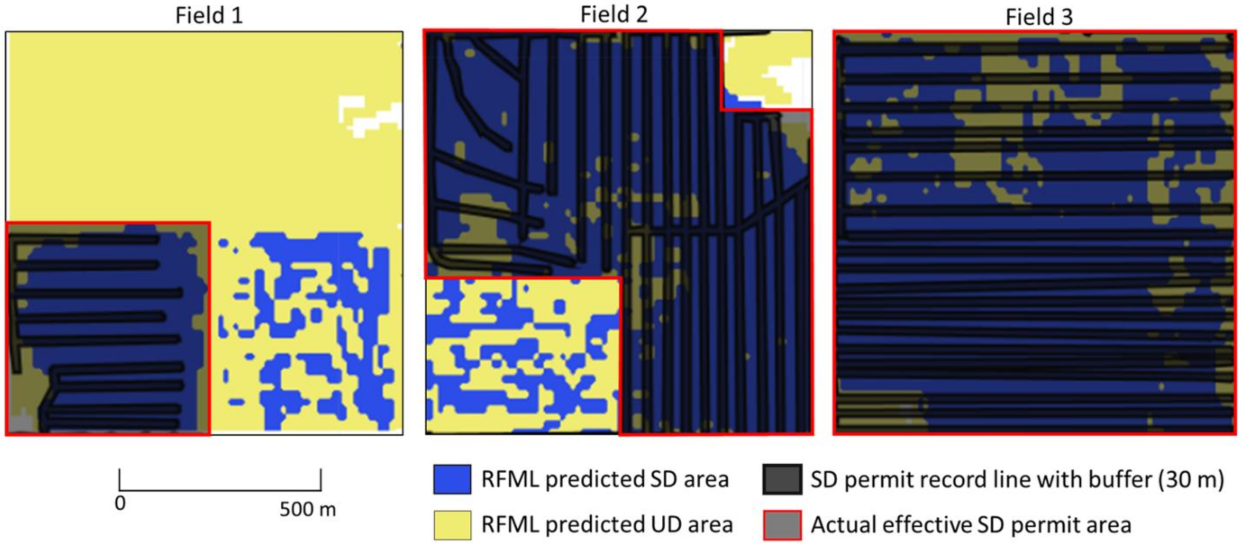


Figure 4. Examples of fields showing areal difference between subsurface drainage (SSD) permit area using buffer function and actual SSD effective area in Bois de Sioux watershed, Minnesota. These examples indicate that SSD permit buffered areas in this study were underestimated in these fields compared to actual SSD effective areas.

For the ND-RRB region, the RFML model achieved an overall accuracy of 87%. Class specific SSD and UD accuracies ranged from 20 to 59% and 98 to 99% with overall accuracies of 40% and 98%, respectively. In both regions, overall accuracies in the early years (2009 and 2011) are higher than those in recent years (2014 and 2017). SSD systems were originally installed at those sites that needed them most. Therefore, training points developed in early years may retain stronger SSD/UD character in this region. A subbasin-level comparison between the

RFML maps and the USGS SSD permit areas conducted for the NOAA river forecasting subbasins found *r*-values ranged from 0.77 to 0.84 for the four years (Figure 5).

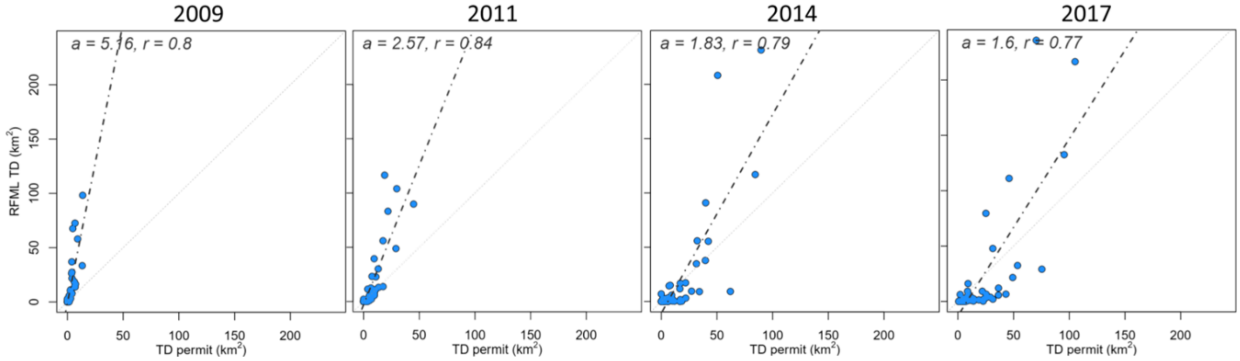


Figure 5. NOAA subbasin-level accuracy assessment over ND-RRB (N=48). NOAA subbasin is hydrological unit to operate the river forecasting system, NOAA RFCs. Subsurface drained permit area from the BdSW district permit records compared with subsurface drained area from RFML classified maps against a 1:1 line (light dashed). Agreement between the two datasets was assessed with correlation coefficient (r) metrics from simple linear regression (trend line = thick dashed line, a = slope). Note that the ranges of y-axis are different.

Again, the RFML overestimated the subsurface drained area, especially in the few subbasins which have dense SSD areas. There are very few SSD areas in the northern part of the RRB. SSD areas are concentrated in the southern part of the RRB (Figure 6a). In North Dakota, the 2017 predicted SSD map near Sheyenne National Grassland showed good spatial agreement with the SSD permit map (Figure 6b and c). However, the RFML maps appear to underestimate SSD areas in Minnesota areas compared to previous findings (Kelly et al., 2017; Nakagaki & Wieczorek, 2016). This indicates that additional training points in Minnesota are required to develop more accurate RFML models.

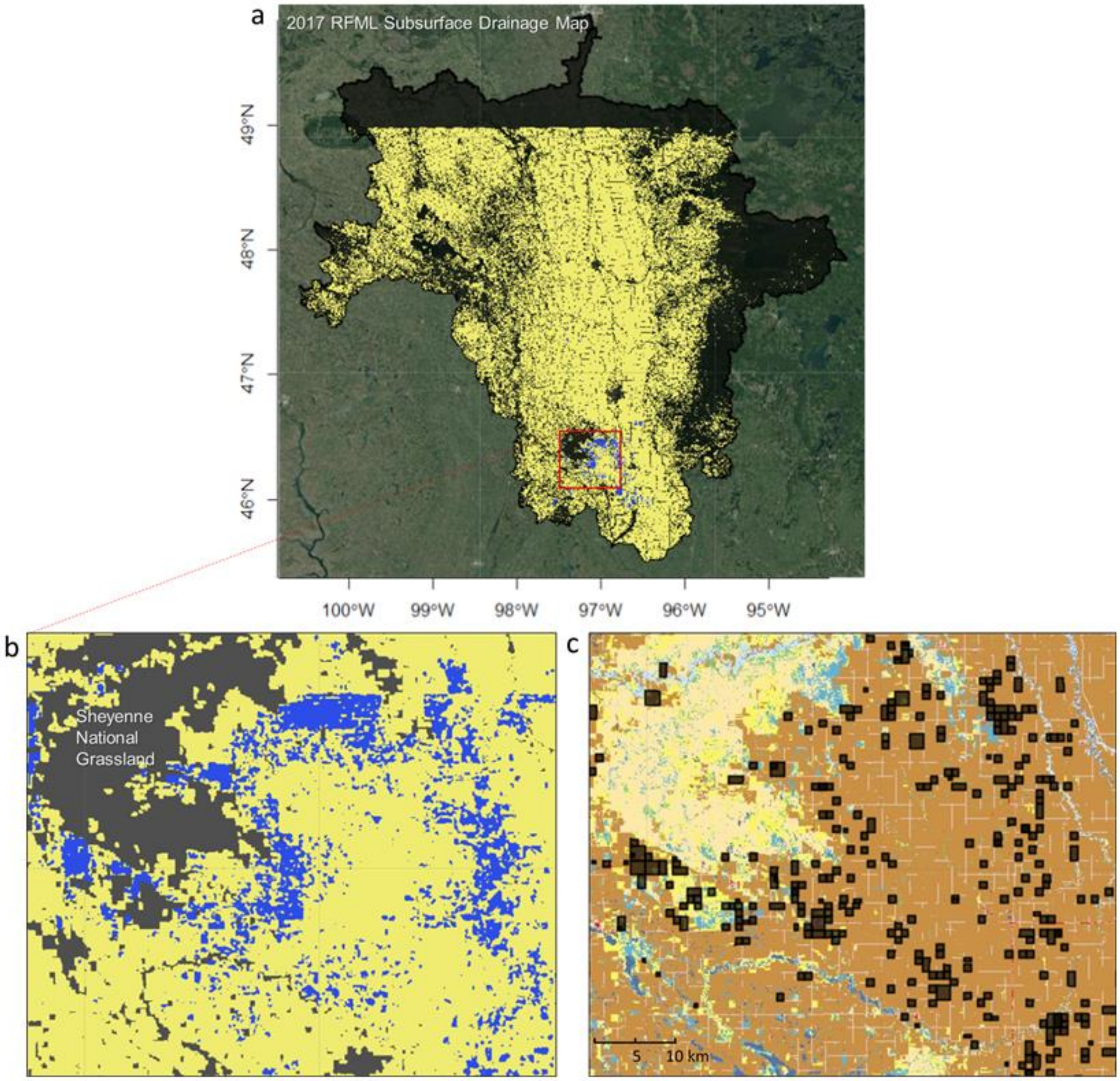


Figure 6. (a) Subsurface drainage map from RFML over the RRB in 2017. (b) A close-up map near Sheyenne National Grassland in North Dakota. Blue colors indicate predicted SSD areas. Yellow colors indicate undrained area. (c) USGS SSD permit records overlaying the NLCD 2011 with same legends in Figure 1 (Finocchiaro, 2016)

2.4.2 Variable Importance

The mean decrease in Gini index was used to determine the relative contribution of each of the 36 input variables for the SSD classification. Larger mean decreases in Gini index indicate variables that play a greater role in partitioning the data into the SSD/UD classification. Soil properties (available water content, awc, clay percentage, clay_perc, and saturated hydraulic conductivity, ksat, in this study) ranked the highest for both regions (Figure 7). Climate variables, precipitation, and aridity also were important, especially for the larger scales. For both regions, LST contributed strongly to the classification. Soil moisture showed minimal importance even though subsurface drains are intended to enhance drainage. This may be due to the coarse resolution (25 km) from the SMOS satellite observations. The importance of spring thermal and wetness variables (e.g. LST and STR2) is noted. These indices warrant further study for use in SSD/UD classification in other agricultural regions. Interestingly, no vegetation-related variables were in the top ten. NDWI scored relatively high among the four vegetation indices, indicating only water-related vegetation variables may enhance accuracy in this region.

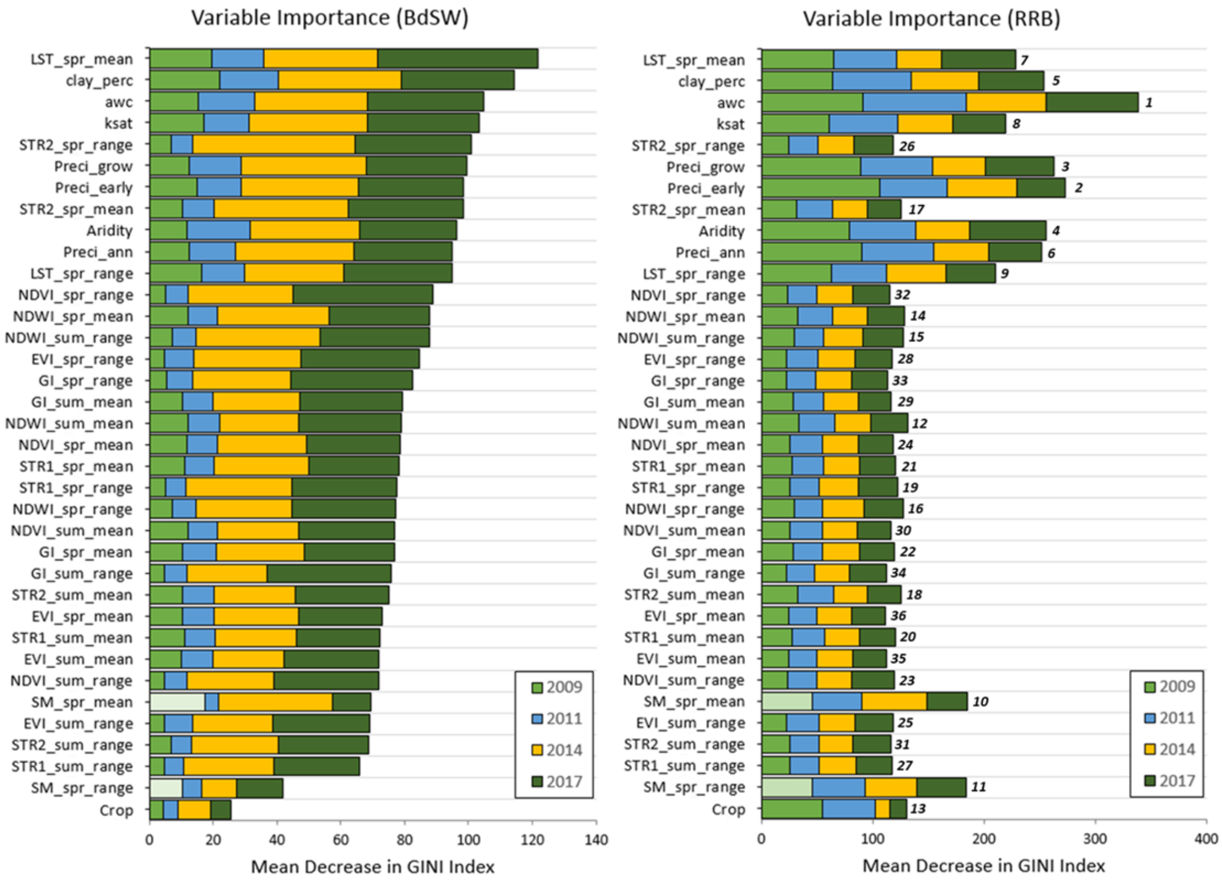


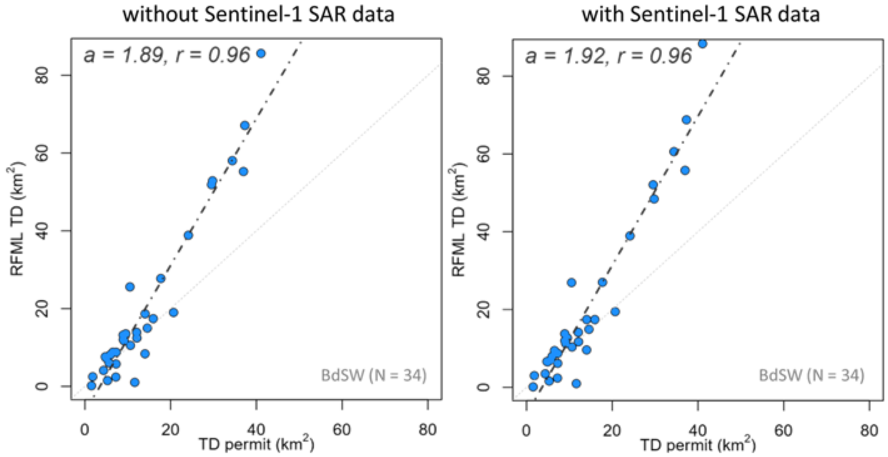
Figure 7. Variable importance in the RFML classification for two regions with different spatial scale (a) BdSW and (b) ND-RRB. For BdSW, variables with their short names were arranged from largest (top) to smallest (bottom) of the accumulated mean decrease in Gini index. Variables in RRB was arranged in same order to those of BdSW. The numbers at the edge of the bar indicate the ranks of each variable. Due to the absence of SMOS soil moisture in 2009, we calculated mean decreases in Gini index of the spring soil moisture mean and range by averaging the other three years' values. Their full names were given in Table 1.

It is possible that the accuracies in the RFML SSD map are improved with new relevant data as an input variable. To test this, Sentinel-1 Synthetic Aperture Radar (SAR) Ground Range Detected C-band backscatter data (VV polarization, ImageCollection ID: COPERNICUS/S1_GRD in GEE) was included in current RFML model as additional input variables (two spring mean and range layers) in 2017. In BdSW, the RFML SSD map with the Sentinel-1 SAR information shows slightly better accuracies than the original SSD without Sentinel-1 SAR (Table 3). The point-based accuracies in RFML SSD and UD predictions were improved by 0.3% (from 48.7% to 49%) and 0.9% (from 87.6% to 88.5%), respectively (the overall accuracy from 72.3% to 73.0%). In the subwatershed-level assessment, the two SSD maps with/without Sentinel-1 SAR have the same correlations (r = 0.96) with similar slopes (Figure 8). However, in the ND-RRB, there is no clear improvement in SSD map accuracies based on the both point-based and subbasin-level assessments. Given that the Sentinel-1 SAR backscattering signal is directly related to surface soil moisture, we expect any improvements of the SSD prediction map by Sentinel-1 data would be much clearer in a wet year. This also suggests that the current RFML SSD model can be steadily improved by including (or replacing) new SSD-related variable information. The Sentinel-1 SAR and RFML SSD maps were provided in Appendix (Figure S3). (Note: subwatershed-level accuracy assessments over the BdSW using the 10 most important variables only are provided in Figure S4).

Table 3. Comparison of RFML SSD maps between with and without Sentinel-1 Synthetic Aperture Radar (SAR) C-band backscatter data based on point-based accuracy assessments in 2017

Year	**Class**	**without Sentinel-1**			**with Sentinel-1**		
		RFML SSD	RFML UD	Overall Accuracy	RFML SSD	RFML UD	Overall Accuracy
BdSW	SSD	48.7%	12.4%	**72.3%** (1716/2373)	49.0%	11.5%	**73.0%** (1732/2373)
	UD	51.3%	87.6%		51.0%	88.5%	
ND-RRB	SSD	19.8%	2.0%	**81.6%** (3,632/4,453)	19.6%	1.8%	**81.7%** (3,639/4,453)
	UD	80.2%	98.0%		80.4%	98.2%	

(a) BdSW



(b) ND-RRB

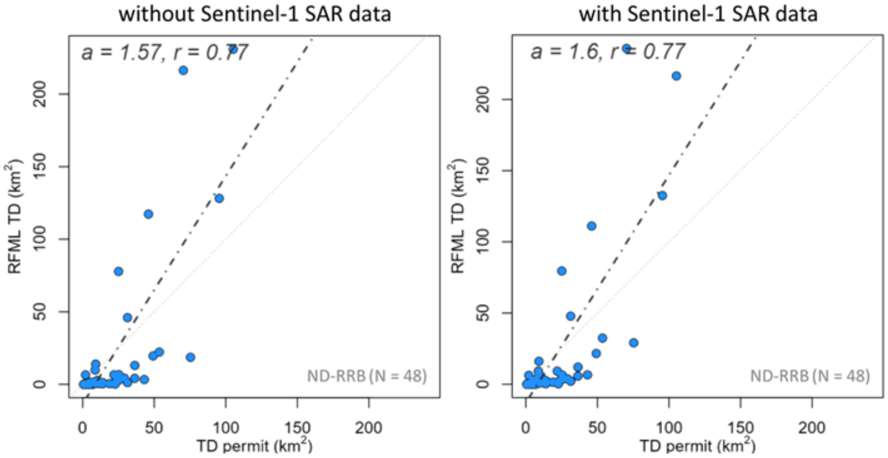


Figure 8. Comparison of RFML SSD maps between with and without Sentinel-1 Synthetic Aperture Radar (SAR) C-band backscatter data based on (a) Subwatershed (HUC12)-level accuracy assessment over BdSW, Minnesota (N=34) and (b) NOAA subbasin-level accuracy assessment over ND-RRB (N=48).

2.4.3 Comparison with Recent Studies

In the RRB, Kelly et al. (2017) reported that the 2012 SSD area was 1,340 km^2, 2.0% of the entire basin area, using the county-level agricultural census drainage data (USDA National Agricultural Statistics Service, 2014). This is larger than our predicted SSD areas (916 km^2) for 2011. There are two potential reasons for the difference. They defined the "RRB region" as being upstream of Grand Forks, North Dakota in United State (67,005 km^2), which is the southern part of our RRB. We used the entire RRB region except for the area in Canada (101,500 km^2) where the northern part of the RRB is almost entirely undrained. There is also a year gap between our results in 2011 and SSD estimates in 2012 from Kelly et al (2017). The USGS SSD permit records for the RRB region in North Dakota indicated an increase of 114 km^2 between the two years. There is likely a similar increase in Minnesota (Dollinger et al., 2013).

Most previous studies were conducted at a smaller scale (e.g. field or catchment scale) than the current study and used stepwise GIS-based analyses and aerial image processing techniques (Naz and Bowling, 2008; Naz et al., 2009; Tetzlaff et al., 2009a; Zhang et al., 2014). They showed spatial agreement with overall accuracies of 78% (Tetzlaff et al., 2009a) and 85% (Naz and Bowling, 2008), which are similar to the performance of the current study (76 – 86%). Zhang et al. (2014) and Naz and Bowling (2008) partially explained the causes of discrepancies in SSD estimates within fields in the current study (e.g. Figure 4). In Canadian subsurface drained fields, Zhang et al. (2014) used unmanned aerial vehicle (UAV)-based NDVI and found within field NDVI differences due to SSD line locations. Naz and Bowling (2008) also found that within-field soil variability can lead to SSD misclassification. Satellite data were also used by Møller et al. (2018) to identify subsurface drained areas in a 43,000 km^2 region in Denmark using an ensemble of ML models with similar input variables to the current study. Møller et al.

(2018) is the sole previous study applying ML methods to SSD detection. However, they only used one month (March 2014) of Landsat 8 imagery. Their final ensemble contained 36 unique models that predicted SSD areas with an accuracy of 76.5%. The results from our current study have better accuracies of 76.9 to 87.0%. This suggests that using an ensemble of multi-source satellite data including seasonal thermal, reflectance, and vegetation input variables could improve results. They also found soil property (e.g. clay content) to be the most important variable, followed by precipitation. This corresponds with our finding that available water content of the soil is the most important variable. Clay percentage ranked in the top five in the RRB region. Climate variables are important at larger scales (Møller et al., 2018; Tetzlaff et al., 2009b). Additionally, we found that LST is the most important variable at a relatively small scale. This seems reasonable considering that drainage systems have significant impacts on surface heat flux and land surface water dynamics. Jacobs et al. (2017) found that spring LST, obtained by subtracting past mean values (2002-2006) from recent values (2013-2017), has significant relationships ($r^2 = 0.85$ and 0.83, respectively) with the SSD density based on a subwatershed-level analysis.

Previous studies also identified uncertainties. Tetzlaff et al. (2009b) noted the difficulty of acquiring aerial images at the right time associated with rainfall events and vegetation growth for a large area. Sugg's (2007) GIS analysis based on soil drainage class and land cover in the Midwest U.S. overestimated total SSD in Minnesota by 3,643 km^2 compared to the 1992 National Resource Inventory (NRI) including inflated estimates of SSD for the RRB region. Their GIS method identified large areas in northwest Minnesota as SSD areas because they are poorly drained soils and cultivated lands. However, the actual SSD installations result from not only geophysical characteristics, but also from socio-economic demand for drainage (Blann et al., 2009). Care must be taken to differentiate between models that predict potential areas requiring SSD systems based on drainage properties versus those that are able to discern areas where SSD has been installed.

Belgiu and Drăguţ (2016) found that the RFML method can handle multi-source satellite data dimensionality and multicollinearity with fast processing and insensitivity to overfitting. However, it tends to be sensitive to training samples (Colditz, 2015), which correspond with our finding in the process of this analysis. We found that the outputs from the RFML method of the current study were sensitive to the proportion of SSD/UD training samples in several trials (not shown). The proportional allocation of SSD/UD training samples to each class based on SSD permit records achieved the best results because the UD class has much larger areas and requires more training samples than the SSD class that occupies limited areas. Further investigations are needed to better understand sample proportioning for RFML. Furthermore, studies are needed that compare the performance among multiple ML methods.

2.5 Conclusion and Future Perspectives

Subsurface drainage systems were mapped at 30 m resolution by leveraging a ML technique and multi-source “big” data sets from operational satellites, Landsat-based vegetation indices (NDVI, EVI, NDWI, and GI) and STR, MODIS LST, and SMOS soil moisture, along with USGS National Land Cover and Slope Datasets, USDA Cropland Data Layer, soil properties from POLARIS, and climate variables from GRIDMET over the RRB region. RFML was conducted in the GEE cloud computing platform, and used SSD permit records from the USGS and the BdSW district for training and validation. The RFML maps showed spatial agreement with SSD permit records and correlated well with HUC12 subwatershed statistics.

The RFML maps appear to be capable of identifying within field variations in SSD effects and capturing the overall SSD expansion over time including for those fields whose acreage was less than that required to be permitted. Soil properties, climate variables, and LST are the strongest predictors of SSD. Predictor variables differed between the two scales, suggesting that SSD models are sensitive to the spatial scale. Using the Sentinel-1 SAR data, we demonstrated the RFML SSD model could be further improved with new relevant data. This ML technique can be readily applied to other regions and future years to provide updated information about SSD expansion to regional water managers and flood forecasters. However, this technique relies on the availability of baseline datasets (e.g., permit records) and more of these datasets may be needed for other regions.

There are future opportunities to further improve the SSD classification (or similar work with demanding LULC detection/classification) using ML algorithms. As a limitation of the current RFML method like other non-deep learning algorithms, the input layers must be developed from raw data with formulas or retrieval algorithms provided by experts for each input data and can be labor-intensive. In this context, deep learning (DL) has substantial potential to overcome this weakness. The DL method, a layered structure of advanced artificial neural network algorithm, allows the automatic extraction of features from raw data by capturing abstract spatial or temporal structures hidden in data (Bengio et al., 2013; Shen, 2018). Also, the use of new remote sensing platforms such as CubeSat and Unmanned Aerial Vehicles can add value for enhanced SSD identification (McCabe et al., 2017; NASA, 2017; NASA CubeSat Launch Initiative, 2018; Planet Team, 2018). For example, more than 130 CubeSats launched by Planet (http://www.planet.com) currently provide daily visible (Red-Green-Blue) and near-infrared imagery with ultra-high resolutions (e.g. 3 meter and 72 centimeter), capturing daily near-global coverage (Planet Team, 2018). This imagery could potentially greatly improve SSD identification with ML or DL methods.

CHAPTER 3

Quantifying Impacts of Subsurface Drainage Expansion on Hydrologic Response in the Red River of the North Basin[2]

3.1 Introduction

During the last 20 years, the amount of streamflow has greatly increased and floods have occurred more frequently in the north-central U.S. In the Red River of the North Basin (RRB), six of the 13 major floods over the past 100 years have occurred since the late 1990s (Rannie, 2015; Rasmussen, 2016). Previous studies have been conducted to determine the major causes for the hydrologic changes (Foufoula-Georgiou et al., 2015; Raymond et al., 2008; Schottler et al., 2014). Potential factors include climate change and agricultural practices including agricultural subsurface drainage (SSD) installation. Artificial SSD system in agricultural landscapes alters surface water dynamics (Blöschl et al., 2007; Schottler et al., 2014). In the past two decades, SSD systems have greatly expanded over the agricultural areas in the north-central U.S. (International Joint Commission, 2000; Sands, 2012). In the Eastern part of the North Dakota and South Dakota, SSD permits have dramatically increased from 2000 until present (**Figure 1**) (Finocchiaro, 2014; 2016). For example, SSD permit growth in the Bois de Sioux Watershed, Minnesota has exponentially increased from 1999 to 2011 (e.g. 1999, 2009, 2010: 2.9, 779.3, 1,558.3 miles of tile drainage, respectively) (Bois de Sioux Watershed district; http://www.bdswd.com/).

---

[2] Cho, E., J.M. Jacobs, S.V. Kumar (2020) Quantifying Impacts of Subsurface Drainage Expansion on Hydrologic Response in the Red River of the North Basin (*in preparation*)

The impacts of artificial SSD system on hydrologic responses have been studied at a range of spatiotemporal scales (Eastman et al., 2010; King et al., 2014; Lenhart et al., 2011; Rahman et al., 2014; Randall et al., 2003; Sands et al., 2008; Williams et al., 2015). At a field scale, Kladivko et al. (2004) showed that SSD -induced water yields were 8 to 26% of annual rainfall in southeastern Indiana, depending on the year and the drain spacing. Eastman et al. (2010) found that the SSD field discharged four times more water than the naturally drained field for their clay loam sites. At a watershed scale (3.89 km^2), King et al. (2014) reported that about 21% of annual precipitation and 47% of total watershed discharge was generated from SSD in central Ohio. Williams et al. (2015) concluded that SSD contributed 56% of the annual watershed discharge in the same Ohio watershed. From a long-term perspective, Schottler et al. (2014)'s watershed scale comparisons between two historical periods (1940-1974 and 1975-2009) found that artificial drainage is the major driver of increased streamflow in 21 Minnesota agricultural watersheds. They found that climate change and crop conversion explained less than half of the streamflow increases that occurred during the later period. The remainder was highly correlated with the recent SSD expansions in Minnesota. Frans et al. (2013) showed that SSD amplified annual streamflow by up to 40% locally in the Upper Mississippi River basin. However, at the basin scale, there are some disagreements among previous studies about SSD effects on soil water storage. For example, Schottler et al. (2014) argued soil storage change was indistinguishable and reported evaporative losses due to SSD

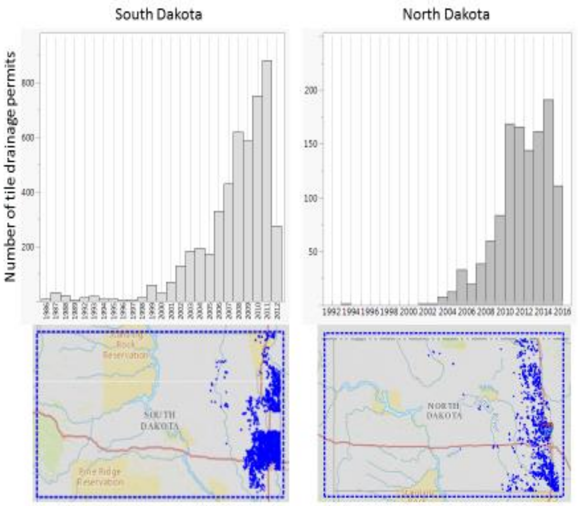


Figure 1. Time series of the number of annual SSD permits and SSD locations by permits in South Dakota and North Dakota

expansion resulted in increased streamflow during the last 70 years in 21 Minnesota watersheds. Kelly et al. (2017) found that soil water storage decreased due to artificial drainage by 30 – 200% and the loss of the storage appeared to have amplified river flow over the last 79 years in Midwest U.S.

To quantify the impact of the SSD expansion on hydrological responses, land surface models (LSMs) can be an effective tool at a large scale. LSMs have evolved in the last few decades with advancing high-performance computing, multiple ground-based networks, remote sensing techniques, and hyper resolution (Baldocchi et al., 2001; McCabe et al., 2017; Wood et al., 2011). Recent studies have investigated the effects of human modifications (e.g. irrigation, deforestation, and urbanization) on hydrological and atmospheric responses using LSMs (Mahmood et al., 2014; Pei et al., 2016; Zhang et al., 2017). One such LSM is the community Noah LSM with multi-parameterization options (Noah-MP) (Yang et al., 2011). Based on the Noah LSM, Noah-MP has added hydrological and biophysical processes such as an interactive vegetation canopy (Dickinson et al., 1998), an unconfined aquifer for a dynamic water table and groundwater storage (Ek et al., 2003; Niu et al., 2007), a simple TOPMODEL (TOPography based hydrological MODEL)-based runoff production (Niu et al., 2005), and a multilayer snowpack (Yang and Niu, 2003). In the recent national hydrologic modeling framework, the National Water Model (NWM), launched by National Oceanic and Atmospheric Administration (NOAA) Office of Water Prediction (https://water.noaa.gov/about/nwm), the Noah-MP LSM is used to simulate land surface processes. The NWM simulates observed and forecasts streamflow over the continental U.S. and is intended to replace the thirteen individual river forecasting center (RFC) flood forecasting models. However, the current Noah-MP LSM, as well as NWM, do not have SSD scheme (Cohen et al., 2018). The absence of SSD information in the Noah-MP can generate uncertainties in hydrological processes and operational flood forecasting in the north-central and Midwest U.S. (Maidment, 2017)

This study focuses on adding an SSD scheme to the Noah-MP LSM within the NASA Land Information System and quantifying SSD impacts on basin-level hydrology. We hypothesize that the SSD will decrease surface and root zone soil moisture above tiles (A lateral effective width of a tile [30 m] is enough to drain a whole field) and amplify surface runoff due to a decreasing soil water storage. This paper is organized as follows. Section 2 details the concept of the SSD system and its impact on water balance from previous studies. Section 3 describes the SSD data and the Noah-MP model with the proposed SSD scheme. Section 4 describes the study area with land characteristics and datasets. Section 5 details the results and compares hydrological variables with and without SSD conditions. Conclusions are drawn in Section 6.

3.2 Background: Principles of subsurface drainage

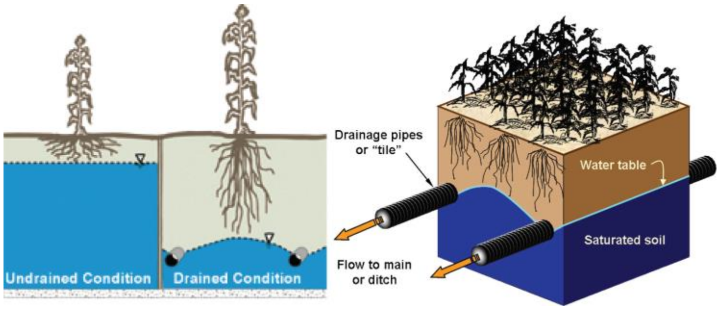


Figure 2. Tile drainage effect to improve root growth of crop in soils (*Blann et al.*, 2009)

In poorly drained agricultural areas, SSD system is indispensable to manage crop water availability to maximize plant growth (**Figure 2**) (Blann et al., 2009; Skaggs et al., 1994). SSD, also known as tile drainage, is used to remove excess surface water and to lower water tables through subsurface pipe networks installed below the ground surface. SSD system plays an essential role in agricultural water management to improve crop productivity worldwide (Blann et al., 2009). The water table changes due to SSD systems can be explained by Darcy’s law (Shokri and Bardsley, 2015). When tiles are installed at a certain depth and spacing, the pressure head at the tiles is approximately the atmospheric pressure and the existing pressure distributions in soil profile horizons change to a new equilibrium. The original water tables in the undrained condition are lowered to the depth of the SSD systems, especially during spring and fall (**Figure 3**). The effective infiltration rates would be changed by SSD installations due to the altered hydraulic gradient of the upper soil layer above tiles (Rodgers et al., 2003; Shokri and Bardsley, 2015; Youngs, 1975). SSD systems have the greatest impact when intense precipitation exceeds the infiltration rates. In general, SSD is most active from April to June after snow melts and before crops mature. The contribution of SSD to streamflow can be significant, especially in smaller watersheds. During mid to late summer, the SSD effects gradually decrease due to increased root water uptake by mature crops. In the fall, the effects increase again after crops are harvested and before soils freeze which is driven by seasonal changes in crop water use. In agricultural regions, groundwater tables generally show strong seasonal variations, higher in spring and late fall and deeper in the middle of summer due to the higher temperature and crop transpiration. SSD plays a role in maintaining relatively constant water tables, even though there are still seasonal changes (Rijal et al., 2012).

SSD also alters evapotranspiration (ET) rates due to a modified water table and soil moisture conditions. Rijal et al. (2012) experimentally evaluated the effect of SSD installation (Note: They used a “controlled” SSD system which differs

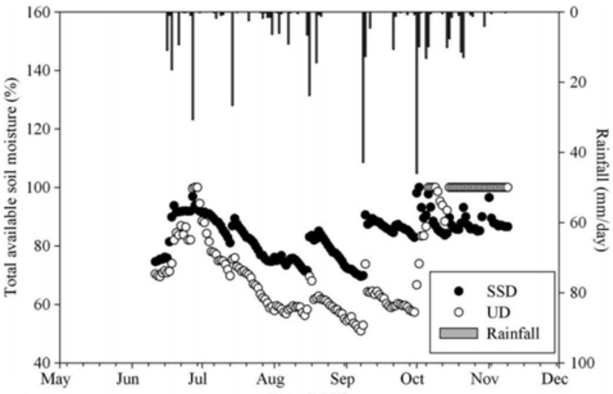


Figure 3. Difference in total soil moisture between subsurface tile drained (SSD) and undrained (UD) fields [*Bowman et al.*, 2015; *Rijal et al.*, 2012]

with “conventional” SSD system - farmers can open/close tile outlets and do subirrigation using a pumping water system through the tile network depending on the soil conditions) for corn and soybean ET in North Dakota. They found that, during the early and very late growth stage, the ET rates in SSD fields were 6% lower than in undrained (UD) fields. The difference was probably due to the removal of the excess soil water and subsequent decreases in evaporation in the surface soils. However, in the summer, the ET rates in SSD fields were 31% and 14% greater than non-tiled areas for corn and soybean, respectively. The higher transpiration rates are thought to be due to a more constant root zone soil moisture resulting from the controlled SSD and deeper roots. For the entire growing season, total ET rates were 16 and 7% higher in the SSD fields than the UD fields for corn and soybean, respectively. Schottler et al. (2014) emphasized the decrease in evaporation in the spring as a major cause of the recent increase in annual stream flows in the Upper Mississippi River. They suggested that, in May and June, the SSD reduces water residence time primarily on the surface storage and drains rapidly to a river. In other words, the increased water yield, which exceeded 50%, resulted from a shorter residence time of the water on the surface and in the root zone soils and subsequent decreases in ET (mainly evaporation) due to SSD expansions.

3.3 Noah-MP

3.3.1 Model Description

Noah-MP was enhanced from the original Noah land surface model through the addition of improved physics and multi-parameterization options (Niu et al., 2011; Yang et al., 2011). The improved physics includes an interactive vegetation canopy, a dynamic groundwater component, and a multilayer snowpack. The multi-parameterization options provide users with multiple choices of parameterizations in leaf dynamics, canopy stomatal resistance, soil moisture factor for stomatal resistance, and runoff and groundwater. Noah-MP delivers better performance compared to earlier versions of Noah LSM in the simulation of runoff, soil moisture, snow processes, and skin temperature across the globe (Cai et al., 2015; Yang et al., 2011). In this study, the Noah-MP version 3.6, running at 0.125° spatial resolution, is used to evaluate the SSD impacts on regional water balance via the framework of NASA's Land Information System (LIS; Kumar et al., 2006). The atmospheric forcing data for runs are from the National Land Data Assimilation System-Phase 2 (NLDAS-2) at hourly temporal resolution and 0.125° spatial resolution (Xia et al., 2012). NLDAS-2 meteorological forcing datasets are downscaled from the North American Regional Reanalysis data, supplemented with ground observational data sets. The forcing inputs include near-surface air temperature, relative humidity, surface pressure, wind, and downward longwave radiation. Parameters and initial condition settings are provided in Table 1. Other parameters and modes used as "default mode" in the NASA LIS system (https://lis.gsfc.nasa.gov/software/lis). Ancillary datasets are used as follows: Moderate Resolution Imaging Spectroradiometer (MODIS)-based International Geosphere-Biosphere Programme (IGBP) land cover map, the National Cooperative Soil Survey and the State Soil Geographic (STATSGO)- Food and Agriculture Organization (FAO) blended soil texture map, Shuttle Radar Topography Mission (SRTM; Native) elevation map, National Centers for Environmental Prediction (NCEP) monthly greenness fraction climatology, and NCEP monthly albedo climatology. Details of the soil water dynamics such as soil infiltration can be found in Niu et al. (2007). The SSD scheme of this study is assumed as a conventional (uncontrolled) SSD system.

Table 1. Summary of parameters, options, and initial conditions used to run the Noah-MP LSM

Parameters & options	Initial conditions	Values

Distance between drains	30 m	Soil temperatures	273 K
Drainage depth from the surface	1 m	Soil moistures	0.2 m^3/m^3
Radius of drain	10 cm	Canopy air temperature	273 K
Depth of bedrock	25 m	Intercepted liquid water	0.0
Dimensionless factor in Kirkham's equation	30	Intercepted ice mass	0.0
Specific yield	0.2	Snow height	25 cm
Vegetation model option	dynamic	Snow water equivalent	5 cm
Canopy stomatal resistance option	Ball-Berry	Depth to water table	2.5 m
Soil moisture factor for stomatal resistance option	Noah	Water storage in aquifer	4900 m^3
Runoff and groundwater option	SIMGM	Stem mass	3.33 g/m^2
Frozen soil permeability option	NY06	Leaf mass	9.0 g/m^2
Snow surface albedo option	CLASS	Mass of fine roots	500.0 g/m^2
Lower boundary of soil temperature	Noah	Snow and soil temperature time scheme	semi-implicit

3.3.2 Subsurface drainage schemes

In this study, the physically-based Hooghoudt and Kirkham SSD equations were adapted from the Modified SWAT model to be incorporated into the Noah-MP LSM (Moriasi et al., 2012; Du et al., 2005). A conventional ("uncontrolled") SSD system was intended to model in the Noah-MP LSM (**Figure 4**). The SSD scheme used in the Modified SWAT model is operated by water table conditions. If water tables are below the surface but above tile location, the drainage flux (q, mm/h) is calculated as

$$q = \frac{8K_e d_e m + 4K_e m^2}{CL^2} \tag{1}$$

where K_e is unsaturated vertical hydraulic conductivity (mm/h), L is the distance between drains (mm) and C is the ratio of the average flux between the drains to the flux midway between the drains (typically assumed to be 1; Moriasi et al., 2012). m is a midpoint of the water tables from tiles (Hooghoudt, 1940; Moody, 1966). *de* is the height of the drain from the impervious layer. If the water table rises to fill the surface (e.g. ponded water), Kirkham's (1957) equation is applied as

$$q = \frac{4\pi K_e(t+b-r)}{gL} \tag{2}$$

where *t* = depth of surface ponding water (m), *b* = depth of drains (m), *r* = radius of drainpipe (m), and *g* = dimensionless factor. Lastly, if the drainage flux is greater than the 20 mm/day, the drainage flux is set equal to the maximum drainage coefficient (DC; 20 mm/day) (Moriasi et al., 2012; Skaggs 1980).

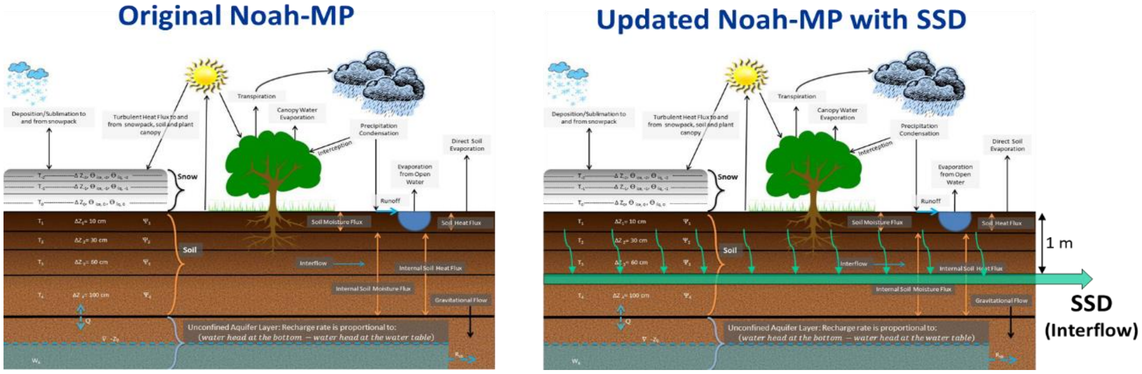


Figure 4. The proposed SSD scheme in the Noah-MP land surface model

If water tables are lower than the tile location, the Hooghoudt's equation (Eq. 1) is applied with adjustment of a parameter ($0 \le m \le 1$) is used as a proportion of available water to soil porosity in soils above tiles as,

$$m = \frac{\theta_t - \theta_{wit}}{\theta_{max}} \tag{3}$$

where Θ_t is the current volumetric soil moisture content (m^3/m^3) at time t in soil columns above tile location. Θ_{wit} is a permanent wilting point (m^3/m^3) and Θ_{max} is saturated soil moisture (m^3/m^3) which is soil porosity of the soils.

A diagram explaining the SSD scheme's process within a pixel is provided in **Figure 5**. The SSD experiment begins by checking if the current pixel has a higher water table than the tile location. If yes, the scheme separates if the water table is higher than the surface (e.g. ponding). For a ponded condition, drainage flux is calculated by Eq. (2). If the water table is between the surface and tile location, drainage flux is calculated by Eq. (1). If the water table is lower than the tile but soil moisture above the tile (3rd layer) is larger than wilting point soil moisture at a pixel, the drainage flux is obtained by Eq. (2) with a new parameter from Eq. (3). Then the calculated drainage flux for each pixel is multiplied by a corresponding SSD fraction as a weighting factor (0-1) from the spatial SSD fraction input.

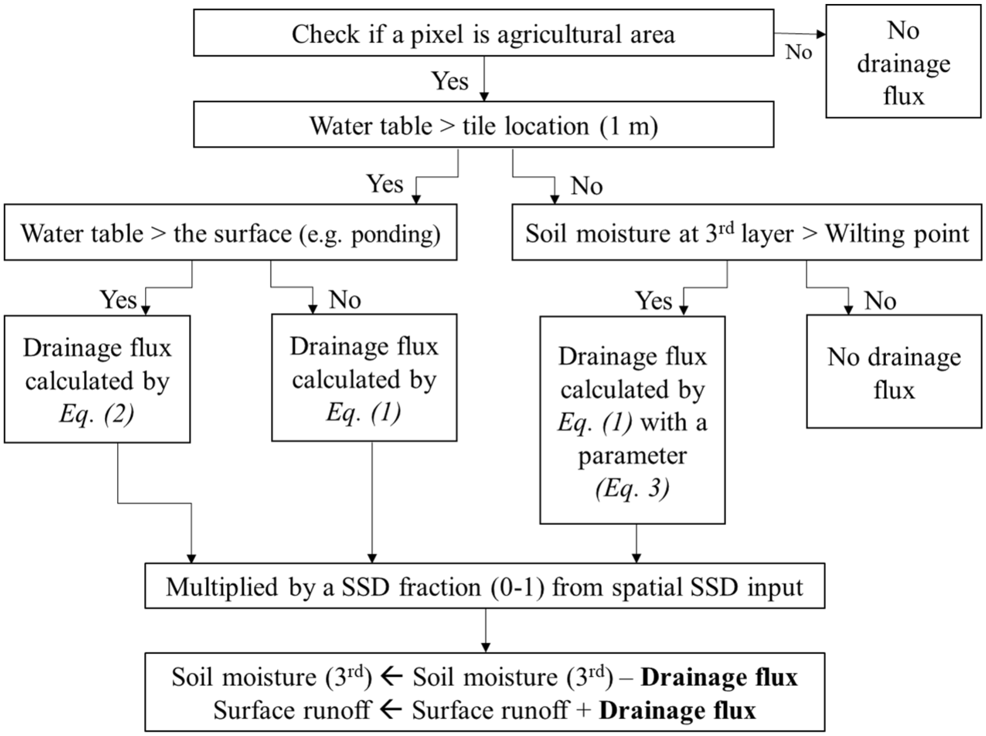


Figure 5. A diagram explaining the SSD scheme's process in the Noah-MP

3.4 Study Area and Data

3.4.1 Red River of the North Basin

The SSD modifications to Noah-MP are demonstrated in the North-Central U.S. For this, the RRB region is an ideal basin to test the SSD impact on regional water balance because the amount of streamflow in the Red River has greatly increased for the last 20 years with the SSD expansion (Rannie, 2015). The RRB drains parts of western Minnesota, eastern North Dakota, and northeastern South Dakota (Figure 6). The Red River flows north from its headwaters in Wahpeton, North Dakota to Winnipeg, Manitoba, Canada. The Red River in the U.S. is approximately 635 km long and with a drainage area of 101,500 km^2 (Rannie 2016; Tuttle et al., 2017). Current land use in areas is predominantly cropland (Homer et al., 2015). The RRB has extremely flat terrain (average slope of 8 cm/km) and poorly permeable clay soil (Stoner et al., 1993). Thus, the RRB region is very vulnerable to flood events. Approximately 85 % of the annual peak river flows over the last century resulted from the spring snowmelt (Rasmussen 2016). Despite the lower average snowpack compared to snowpacks in the western U.S. (Brasnett 1999), most snowmelt floods in the RRB persist on from weeks to months. For example, the Red River at Grand Forks, ND exceeded the flood stage for 46 days during the 1997 snowmelt flood (Todhunter 2001). Inundation area could extend for long distances from the mainstream (e.g., up to 100 km of the floodwater width; Schwert 2003). The snowmelt floods have damaged private property and infrastructure in this region. The 1997 spring flood caused more than $5 billion of damage in Fargo and Grand Forks, North Dakota, and other communities (Todhunter, 2001).

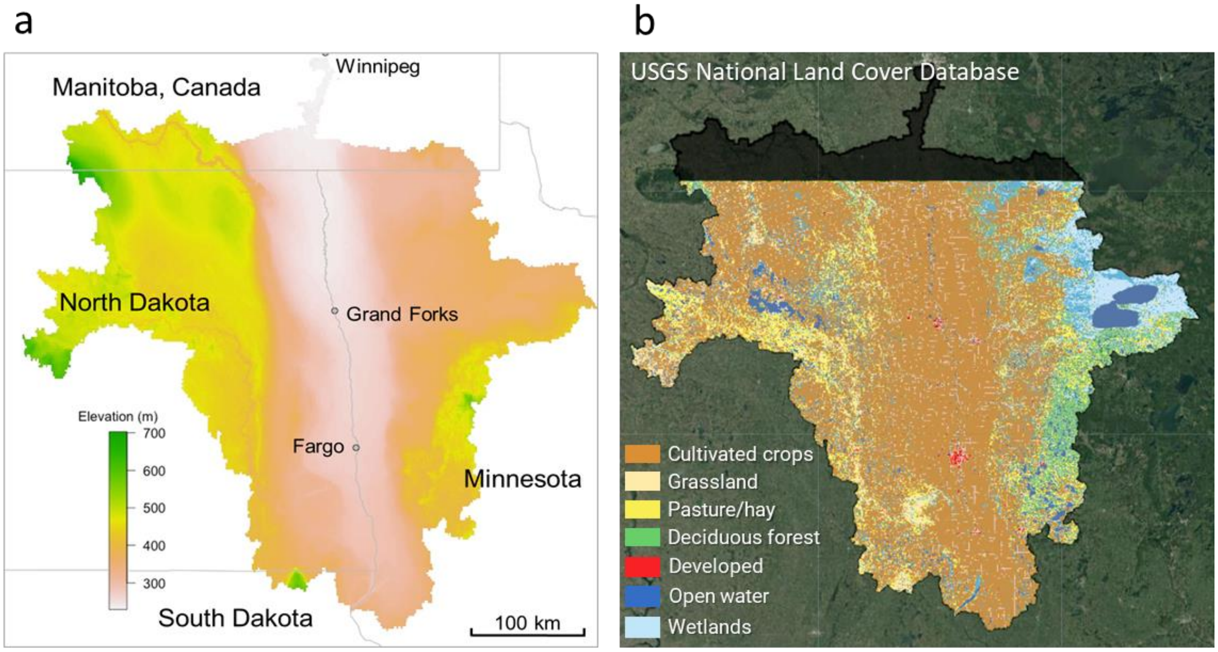


Figure 6. (a) Land cover and (b) elevation maps of Red River of the North Basin

3.4.2 Subsurface drainage map

Recent hydrologic changes with increasing flood events in RRB could be due to the extensive SSD installation [Blöschl et al., 2007; Miller and Frink 1984]. Even though the farmers have expansively installed SSD over the RRB, the spatial distribution of SSD is little documented. To overcome this, Cho et al. (2019) developed the 30-m high-resolution SSD maps over the RRB region by leveraging a Random Forest machine learning (RFML) classification method and multisource “big” data sets from operational satellites, soil properties, and climate variables using the GEE cloud computing platform. The maps were validated with SSD permit records with both point-level (accuracies of 76.9–87.0%) and subwatershed-level statistics (correlation = 0.77–0.96). In this study, the RFML SSD map is used to apply realistic SSD distribution into the Noah-MP. To match the Noah-MP spatial resolution (12.5 km), we upscale the 30-m binary (drained vs. undrained) map into a 12.5 km SSD fraction map (**Figure 7**). The pixel value in the upscaled map indicates a portion of the SSD area within a pixel. The portion is used as a weighting factor (e.g. a fraction of 10% will be 0.1 of a weighting factor) when the SSD flux is calculated. The RFML SSD maps are freely available on Hydroshare at http://www.hydroshare.org/resource/f2f7a9cfbae1451f85b5c0dc3938b9a1.

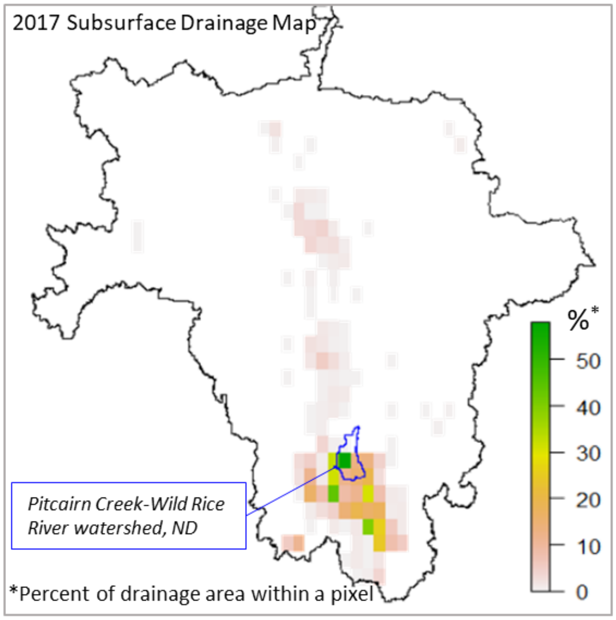


Figure 7. Upscaled Random Forest Machine Learning-based subsurface drainage (RFML SSD) map of Red River of the North Basin [This map was modified from 30-m high resolution SSD map from Cho et al. (2019)]. Pitcairn Creek-Wild Rice River watershed, ND has the largest SSD-installed HUC10 (USGS Hydrologic Unit) watershed in the RRB which is used for vaildation.

3.4.3 Soil Moisture Active Passive (SMAP) soil moisture

In this study, the observation-based assimilated Soil Moisture Active Passive (SMAP) product is used to validate and evaluate the modeled soil moisture's performance from Noah-MP with/without the SSD scheme and the RFML SSD map. The SMAP satellite's L-band radiometer has provided global soil moisture measurements at 2–3 days revisit time (6:00 A.M./P.M. local time) since March 31, 2015 (Chan et al., 2016; Entekhabi et al., 2010). The SMAP soil moisture has been widely used to better understand soil moisture dynamics, land-atmospheric interactions, and hydrological extremes at local and global scales (Cho et al., 2020; Colliander et al., 2017; Ma et al., 2019; Zwieback et al., 2018). To effectively monitor surface and root zone soil moisture conditions for agricultural purposes, the NASA's Goddard Space Flight Center in cooperation with USDA Foreign Agricultural Services and USDA Hydrology and Remote Sensing Laboratory recently developed the NASA-USDA SMAP global soil moisture product at 0.25° x 0.25° spatial resolution. The product is generated by integrating SMAP Level 3 soil moisture observations into the modified two-layer Palmer model using a 1-D Ensemble Kalman Filter data assimilation approach (Bolten et al., 2010; Sazib et al., 2018). The NASA-USDA SMAP soil moisture data is freely available via the Google Earth Engine computing platform (https://explorer.earthengine.google.com/#detail/NASA_USDA%2FHSL%2FSMAP_soil_moisture).

3.4.4 Experimental design

Three scenarios are performed to assess the effect of the agricultural SSD system on hydrological responses (e.g. soil moisture, evapotranspiration (ET), and surface runoff): 1) no SSD (or undrained; UD), 2) SSD over the entire area (hereafter fSSD), and 3) realistic SSD based on the RFML SSD map (hereafter SSD+RFML). The effect of SSD is demonstrated through comparisons between the three scenarios. The domain is the RRB region, gridded at 0.125 resolution, and the experimental run period is a 2017 entire calendar year, and evaluation was conducted from April 1 to October 30, 2017.

3.5 Results and Discussion

3.5.1 Average Basin Impacts between UD and fSSD conditions

Noah-MP simulations with and without the SSD scheme show a difference in surface and deep soil moisture, total ET, and surface runoff for the RRB. The time series of the basin-average soil moisture from April to October in 2017 is shown in **Figure 8a**. For all four soil layers in Noah-MP (0-10 cm, 10-40 cm, 40-100 cm, and 100-200 cm), the SSD soil moisture is drier than the UD soil moisture. In the spring (April to May), there are larger differences in soil moisture between SSD and UD as compared to the summer (June to August) and the early fall (September to October). The differences in soil moisture gradually decrease from spring to summer season, and then increase again in October.

The SSD surface and root zone soil moisture (0 – 10 cm and 10 – 40 cm) decreases faster after rainfall events compared to that of UD. There is little difference in the near-surface soil moisture (0 – 10 cm and 10 – 40 cm) from August to September. Soil moisture at 40 – 100 cm has the greatest difference between SSD and UD conditions. This soil layer is immediately above the SSD system. In the summer, the SSD soil moisture decreases up to 0.13 m^3/m^3. In the deep soil layer below the SSD system (100 – 200 cm), SSD soil moisture is constant around 0.19 m^3/m^3 while UD soil moisture increases slightly after large rainfall events. In total soil columns (0 – 200 cm), the total soil water for the UD run is 441 mm while the SSD run has 363 mm.

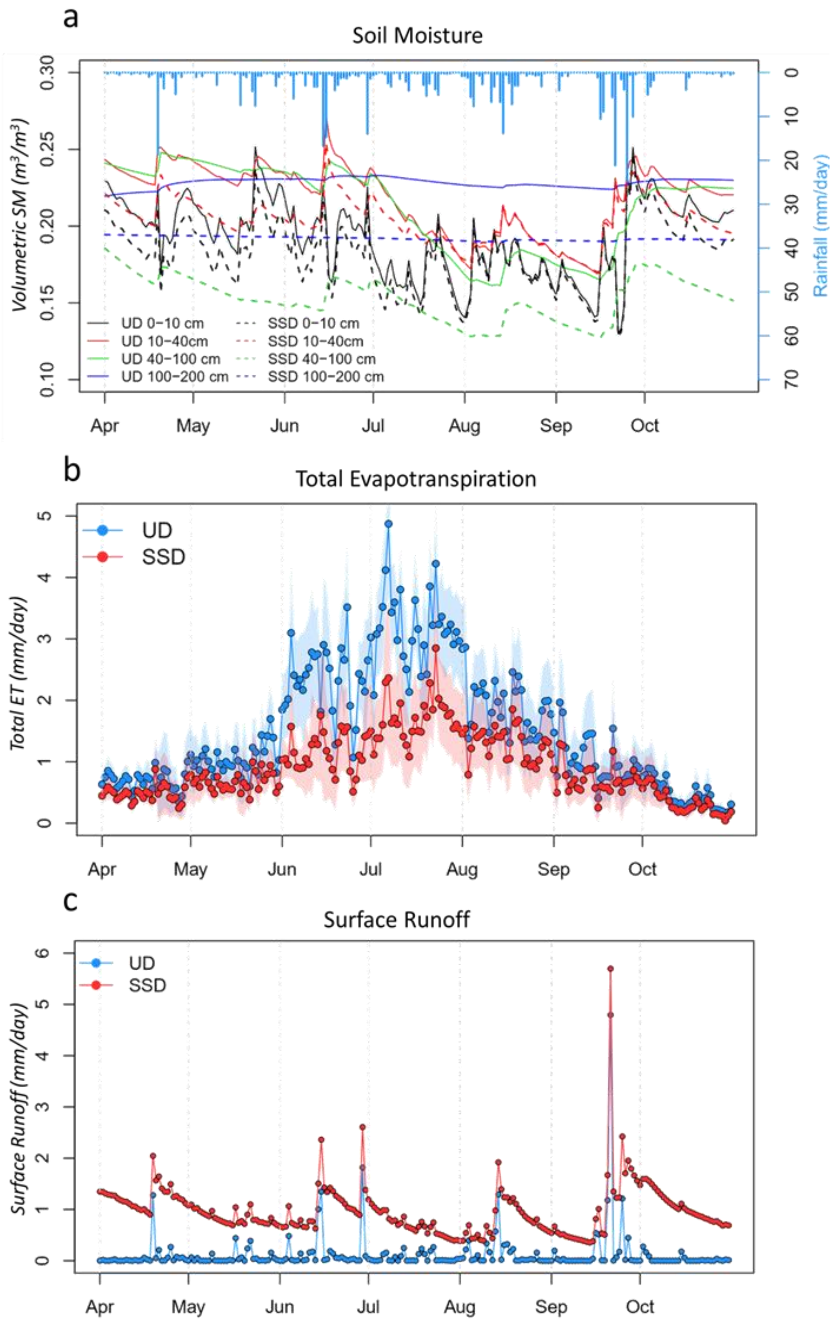


Figure 8. Comparison of basin-average (a) volumetric soil moisture (0-10 cm, 10-40 cm, 40-100 cm, and 100-200 cm), (b) total evapotranspiration, and (c) surface runoff between UD and full SSD condition (The light blue and red colors in the evapotranspiration indicate spatial variations (the average plus or minus one standard deviation) over the RRB

Spatial mean ET rates over the RRB have large differences between fSSD and UD from July to August and small differences in the spring and fall seasons (**Figure 8b**). For fSSD, surface runoff continually appears around 1 mm/day, but only intermittently for UD conditions (**Figure 8c**). During the middle of the summer, the amount of the surface runoff with UD is slightly larger than that of SSD. The total surface runoff to precipitation ratios is 0.57 [surface runoff + SSD (208 mm) / precipitation (365 mm)] which is much larger than the ratio for UD (0.073) (**Table 2**).

Table 2. The amount of water balance components in the study period from April-1 to October-31, 2017 (dSM = change in soil moisture; dGW = change in groundwater storage; Total: ET + surface runoff + subsurface runoff + dSM + dGW)

Experiment	Precipitation (mm)	ET (mm)	Surface runoff (mm)	Subsurface runoff (mm)	dSM (mm)	dGW (mm)	Total (mm)
UD	365	326	27	49	-8	-24	369
fSSD	365	195	208	32	-33	-32	369
SSD+RFML	365	311	46	47	-11	-25	368

3.5.2 Comparison between UD and SSD+RFML conditions

Hydrological changes in actual SSD expansion with the RFML SSD map (SSD+RFML) were evaluated as compared to UD and fSSD conditions (**Figure 7**) from Cho et al. (2019). Compared to the fSSD condition, changes in basin-average soil moisture, ET, and surface runoff in the SSD+RFML condition are relatively minor (**Table 2** and **Figure 9**). SSD soil moisture has similar magnitudes as well as temporal variations to the UD soil moisture. Soil moisture at depth of 40 – 100 cm only decrease by 0.01 m^3/m^3. However, there are notably drier soils for areas with SSD (**Figure 10**). For the entire period, monthly average SSD soil moisture is lower than UD over areas near the border of North Dakota and Minnesota except for soil moisture at the 2nd layer for August and September. As mentioned above, the soil moisture at the third layer from the surface (40 – 100 cm) has the largest differences particularly in the southern part of the RRB region.

At the basin scale, there is little difference in the ET rate in the spring and fall periods (**Figure 9**). In the summer, the ET rates with SSD+RFML condition are slightly lower than the ET with UD condition with basin-average differences of 3.6, 5.7, and 2.5 mm/month in July, August, and September, respectively. The monthly ET difference map (**Figure 11)** shows a localized reduced ET in SSD regions of up to -40 mm/month in the summer season (and up to -13 mm/month in both spring and fall seasons). In the time series of surface runoff, there are increases in basin-mean surface runoff even though the SSD system is in partial areas in the RRB. The surface runoff with SSD+RFML occurs continually around 0.1 mm/day. When there are intense rainfall events (e.g. 17-April and 20-September), the peak runoff with SSD+RFML is higher than UD up to 0.35 mm/day. In April, the basin-average surface runoff increases 3.3 mm/month due to the SSD system which is a relatively large amount compared to summer (e.g. 2.1 and 1.0 mm/month for July and August, respectively). In the monthly surface runoff difference map, there is localized increased surface runoff near the main stem of RRB in most periods (**Figure 11)**.

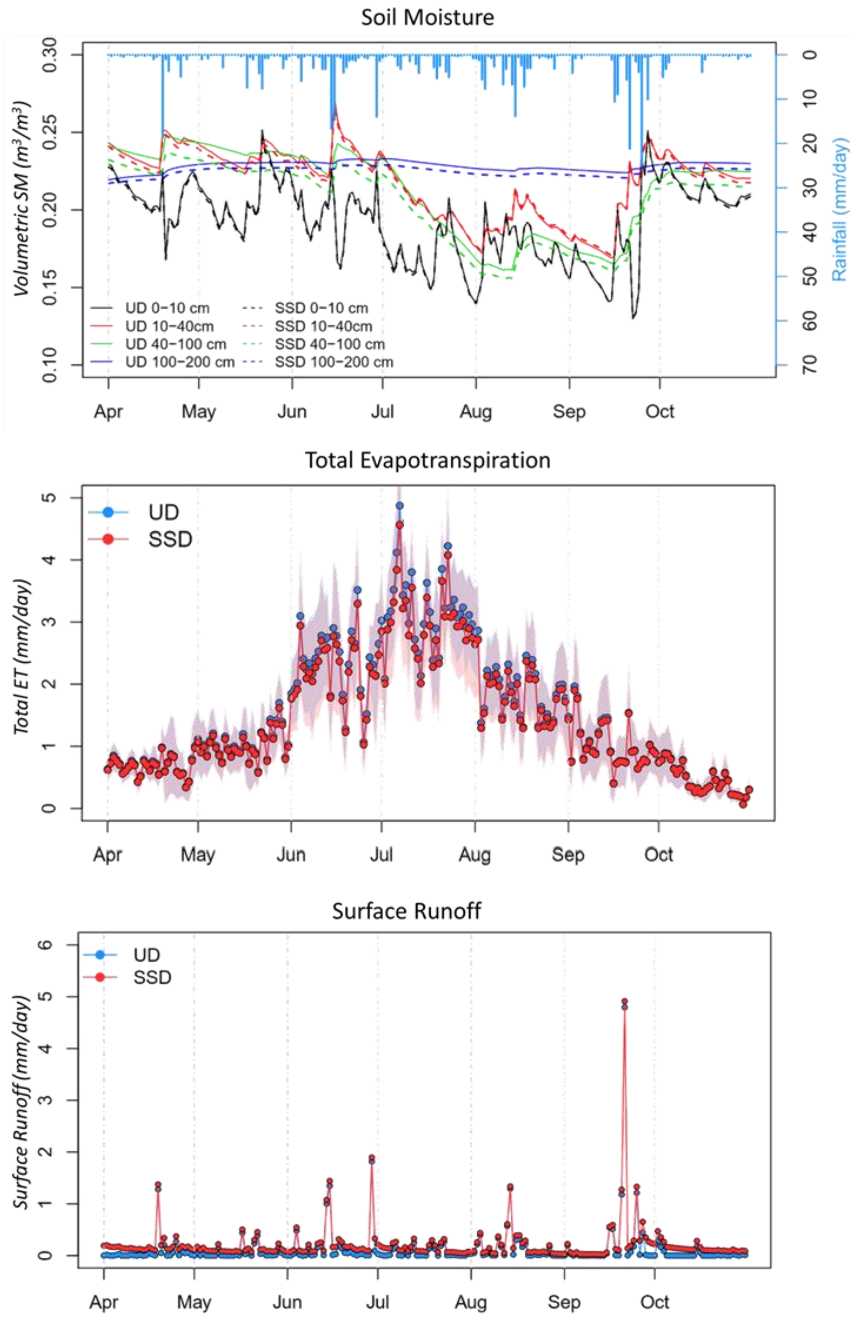


Figure 9. Comparison of basin-average volumetric soil moisture (0-10 cm, 10-40 cm, 40-100 cm, and 100-200 cm), total evapotranspiration, and surface runoff between UD and partial SSD condition based on the RFML SSD map

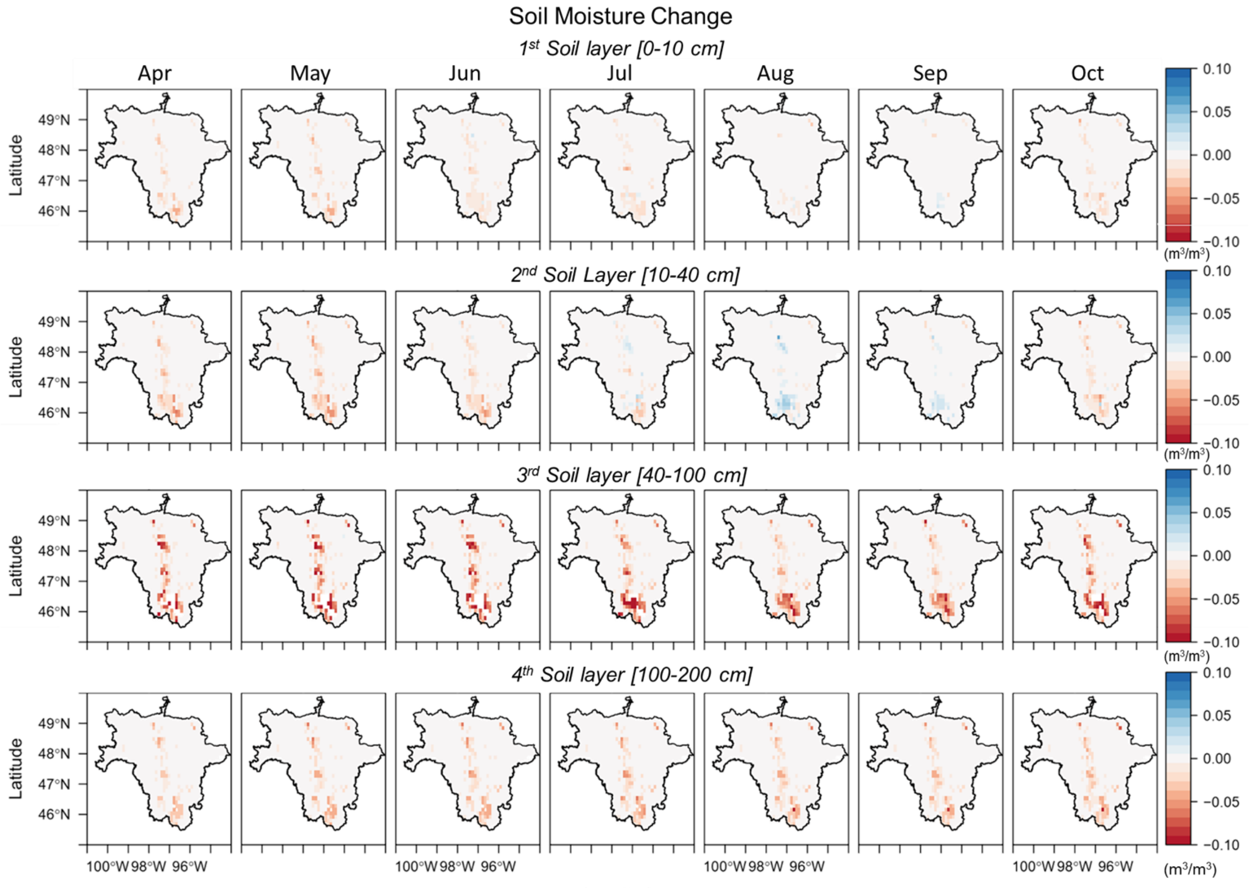


Figure 10. Changes in monthly volumetric soil moisture maps due to SSD expansion (SSD+RFML minus UD)

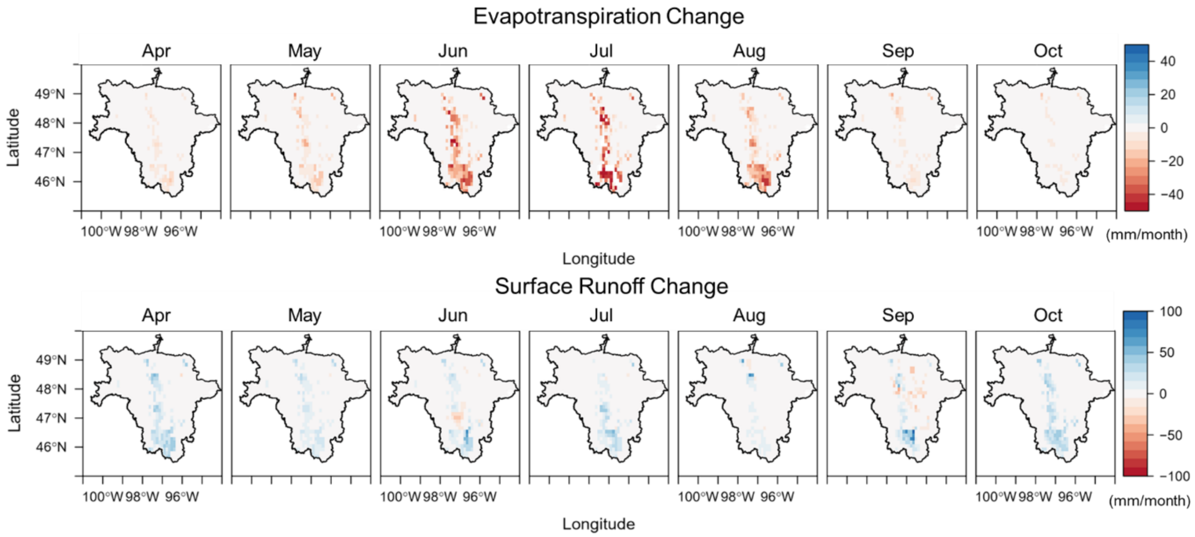


Figure 11. Changes in monthly evapotranspiration and surface runoff maps due to SSD expansion (SSD+RFML minus UD)

3.5.3 Validation with SMAP observations

To validate soil moisture estimates from the Noah-MP with the SSD scheme using the RFML map, SMAP satellite-based surface and root zone soil moisture products were used. The monthly total soil water in the soil column (0 – 200 mm) is compared in **Figure 12**. Temporal changes in SMAP soil moisture are much larger than Noah-MP soil moisture. There are minimal differences between UD and SSD+RFML conditions but soil water near the areas existing SSD expansion slightly decreases. The decreases in soil water near Fargo, ND are captured by SMAP soil moisture. There is relatively lower soil water compared to other areas for the entire period.

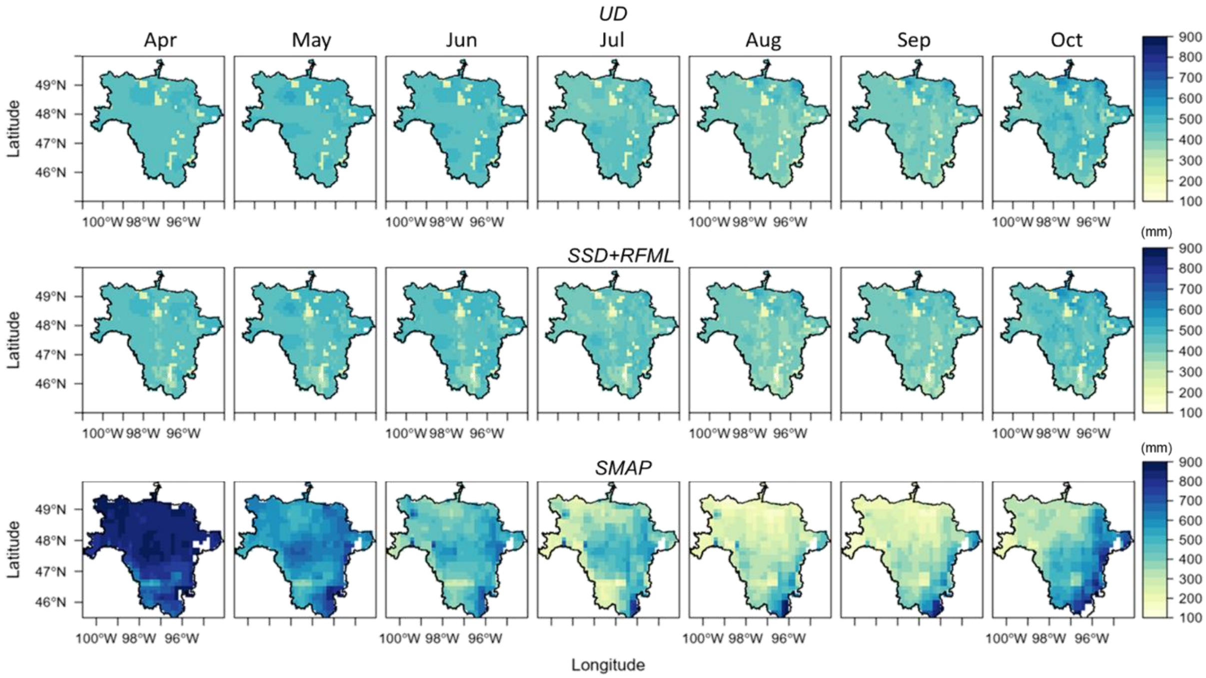


Figure 12. Monthly normalized soil moisture changes for UD, SSD+RFML, and SMAP

To evaluate the soil moisture in a region where the SSD system has expansively installed in the RRB, the watershed-mean Noah-MP soil moisture for Pitcairn Creek-Wild Rice River watershed (624 km^2; **Figure 7**) was evaluated compared to SMAP soil moisture. **Figure 13** shows that the SMAP soil moisture has better agreement with the Noah-MP SSD+RFML soil moisture than the Noah-MP soil moisture without the SSD scheme. For the surface soil moisture, there is a slight improvement in agreement (Pearson linear correlation coefficient (r) = 0.57, unbiased Root Mean Square Difference [ubRMSD] = 0.030, and bias = 0.12 for the SSD+RFML) compared to without SSD scheme (r = 0.56, ubRMSD = 0.032, and bias = 0.13). The Noah-MP soil moisture with the SSD scheme provides better agreement with the SMAP for root zone soil moisture (40-100 cm) near SSD tiles. The Noah-MP soil moisture with SSD+RFML has a higher correlation (r = 0.36) and lower errors (ubRMSD = 0.034 and bias = 0.12) with SMAP than the Noah-MP with UD (r = 0.32, ubRMSD = 0.041, and bias = 0.02).

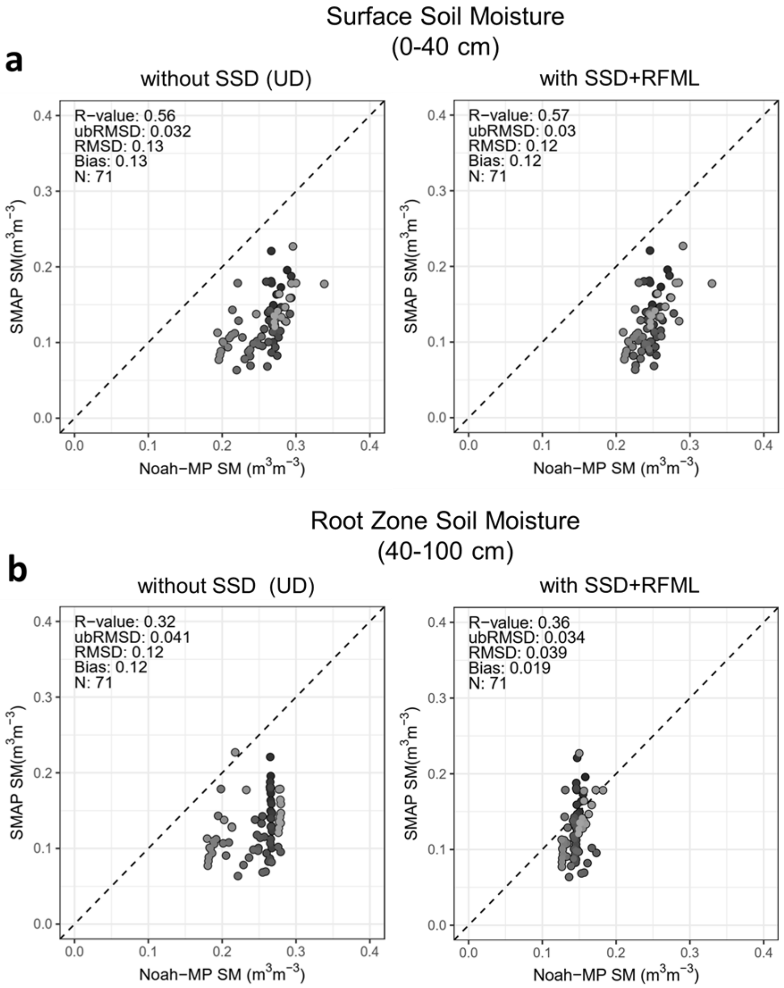


Figure 13. Comparison of surface and root zone soil moisture between UD (without SSD scheme) and SSD+RFML conditions with SMAP satellite-based soil moisture.

3.6 Discussion

3.6.1 Comparison with Recent Studies

In this study, the total soil moisture (0 – 200 cm) and ET from UD to fSSD condition decrease by 78 mm and 131 mm, respectively. This result provides insights to address some disagreements among previous studies about SSD effects in regional water balance in light of the change in soil water storage. Kelly et al. [2017] stated that soil water storage has decreased due to SSD and that the loss of the storage may lead to the amplified river flow. Schottler et al. [2014] argued that soil storage change was indistinguishable and that the evaporative losses due to SSD expansion resulted in increased streamflow. The result from this study indicates that increases in streamflow with SSD are generated by both soil water storage and ET changes. Considering the linkage of soil moisture to ET, a decrease in the soil water storage by the SSD system consequently leads to a decrease in ET (negative feedback; Seneviratne et al., 2010). Overall, the soil moisture decreases and, consequently, decreases in ET cause an amplification of surface runoff, particularly in the spring season.

The surface runoff to precipitation ratios (0.57) in the fSSD condition from this study is similar to previous findings by Eastman et al. (2010). Through field experiments between the naturally drained and SSD fields in southern Quebec, Canada, they found that the total outflow to precipitation ratio at clay loam site was 0.63 [ratio of surface runoff (157 mm) + SSD (358 mm) to precipitation (793 mm)] and 0.88 [ratio of surface runoff (247 mm) + SSD (737 mm) to precipitation (1123 mm)] for 2005 and 2006 water years, respectively. Our results show a relatively lower difference in the surface runoff between UD and fSSD condition. While the total outflow (surface runoff + SSD) from SSD fields was four times higher than the outflow (surface runoff only) from naturally drained fields, the SSD run generates 208 mm of total runoff (surface runoff + SSD) which is about 7.7 times than the UD run (surface runoff only).

3.6.2 Limitations

The SSD scheme developed in this study is based on a conventional SSD system. Further parameterization of other types of the SSD scheme (e.g. controlled SSD system) in the model will be necessary. Under the conventional (or uncontrolled) SSD system, drained water flows directly from the SSD outlet into the surface runoff. However, the conventional SSD system can reduce soil moisture excessively in dry years leading to loss of water and nutrients from fields (Iowa State University, 2017). The conventional SSD systems in the north-central and the Midwest U.S. have been updated to a “controlled” SSD system to overcome the limitations (Ng et al., 2002; Tan et al., 1998). Like conventional SSD, control structures are opened in spring, permitting free drainage to remove excess water and improve aeration in the soils. However, in the summer, control structures are used to store water for crops and some of the control structures supply water “sub-irrigation” through the SSD using a pumping system to achieve the desired soil moisture condition (Drury et al., 2009; Tan et al., 1998). Thus, The SSD scheme introduced here is simplified from a management perspective and does not account for other types of the SSD system that could modify hydrologic responses.

There are inherent limitations in the Noah-MP LSM related to the SSD scheme. As mentioned above, the Noah-MP uses a constant bedrock depth (25 m) over the continental U.S. which is extremely different from the observed depths to bedrock in the RRB. For example, the depths to bedrock in Fargo, North Dakota were found within a depth range of 92 m ranging from 40 m to 133 m (North Dakota Geological Survey, 2011). This may generate uncertainties in quantifying surface-groundwater interaction. Thus, multi-model comparison studies or field data comparison will be required to constrain uncertainties and identify gaps in model physics.

3.7 Conclusion and Future Perspective

For the last two decades, the streamflow in the upper Midwest has markedly increased, and major spring floods have frequently occurred at the same time that climate and agricultural management practices have changed. However, existing operational flood forecasting models do not include SSD. This study investigates the impacts of the recent SSD expansion on regional hydrological changes (soil moisture, ET, and surface runoff) using Noah-MP LSM. The SSD scheme is applied to the Noah-MP using a high-resolution SSD map. We found that the SSD system generates losses of soil water storage. The drying rates of surface soil moisture with SSD are faster compared to that of UD. There are considerable differences in total ET between SSD and UD over the summer season. We found that decreasing soil moisture directly due to SSD installation results in decreasing ET and SSD generates large amounts of surface runoff, particularly in the spring season, indicating that the recent increases in streamflow and spring floods may be partly due to the SSD expansion. In validation, the Noah-MP soil moisture with SSD+RFML has better agreement with SMAP soil moisture as compared to the Noah-MP soil moisture without the SSD scheme. A next step is to run the Noah-MP with a routing model (e.g. HyMAP) to generate streamflow rates in the Red River and to compare them to historical USGS stream gauge observations.

The development of a “controlled” SSD scheme in the Noah-MP should be required with the advanced SSD maps having controlled and uncontrolled SSD spatial information, separately. The physics of the controlled SSD system with sub-irrigation are likely opposite to the conventional SSD in soil moisture-limited regime (e.g. drought). Furthermore, there are no currently available SSD records with controlled vs. conventional types, separately, and the operation of the controlled system is subjective based on field-by-field decisions made by the farm owners. Further investigation is necessary to properly address these challenges.

Lastly, our findings from this study can help to improve the capability of the National Water Model (NWM). The NOAA recently launched the NWM that simulates advanced observed and forecast streamflow over the continental U.S. The NWM uses the Noah-MP LSM to simulate land surface processes without considering the SSD system (https://water.noaa.gov/about/nwm). As the NWM becomes the flood forecasting standard, applying the SSD scheme may be able to lead to better flood forecasts in the drainage-dominant regions.

CHAPTER 4

Improvement of Operational Airborne Gamma Radiation Snow Water Equivalent Estimates using SMAP Soil Moisture[3]

4.1 Introduction

In snowmelt-dominated regions, water resources management and flood predictions rely on accurate snowpack measurements (De Roo et al., 2003; Liu et al., 2012). The most important snowpack measure for streamflow prediction is snow water equivalent (SWE), which is the depth of liquid water that would result if the entire snowpack melted (Bergeron et al., 2016). In the north-central U.S. and southern Canada, accurate flood predictions are needed to help communities prepare for flood events and allocate flood management resources. However, flood prediction is hampered by insufficient information about the magnitude and spatial distribution of SWE and snowmelt across the landscape (Tuttle et al., 2017; Schroeder et al., 2019). In the flood-prone Red River of the North in Minnesota and North Dakota in U.S and Manitoba in Canada (Rannie, 2015; Stadnyk et al., 2016; Todhunter, 2001; Wazney and Clark, 2015), the National Weather Service (NWS) North Central River Forecasting Center (NCRFC) overestimated a peak flow by 70% of the observed 2013 flow in the region. The flood forecasters indicate that uncertainties in SWE spatial distribution as well as antecedent soil moisture estimates were potential causes of the forecasting's failure (personnel communication, Mike DeWeese NOAA NCRFC).

---

[3] Cho, E., J.M. Jacobs, R. Schroeder, S.E. Tuttle, C. Olheiser (2020) Improvement of airborne gamma radiation snow water equivalent measurements using SMAP soil moisture, *Remote Sensing of Environment*, https://doi.org/10.1016/j.rse.2020.111668

Since the late 1970s, satellite passive microwave sensors such as the Scanning Multichannel Microwave Radiometer (SMMR) aboard the NASA Nimbus-7 satellite (temporal coverage: 1978 – 1987), and the Special Sensor Microwave/Imager (SSM/I) and SSMIS aboard the Defense Meteorological Satellite Program (DMSP) series of satellites (F8, F11, F13, and F17: 1987 – current) have provided useful snowpack information globally (Armstrong et al., 1994; Derksen et al., 2005; Foster et al., 2005; Pulliainen and Hallikainen, 2001; Tait, 1998). The Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) aboard the NASA Aqua satellite and AMSR2, a follow-on instrument of AMSR-E onboard the Japan Aerospace Exploration Agency (JAXA) Global Change Observation Mission 1-Water (GCOM-W1) satellite, have successfully provided snow depth and SWE for the past two decades (Dai et al., 2012; Kelly et al., 2003; Kelly, 2009; Cho et al., 2017). SWE from current satellite-based microwave sensors has proven to be a valuable asset for improving snowmelt streamflow predictions at a watershed scale (approximately 47,000 km^2; Vuyovich and Jacobs, 2011). Accurate SWE information at smaller scales remains challenging due to the coarse spatial resolution (e.g. 25 km by 25 km; 625 km^2) of passive microwave satellite observations. In addition, wet snow and variations in snow grain size make the microwave satellite retrieval of SWE difficult (Armstrong et al., 1993; Tuttle et al., 2017; Vuyovich et al., 2017).

Snow observations from airborne platforms can fill the knowledge gap between ground and satellite microwave remote sensing observations of snow (Painter et al., 2016). Airborne gamma-ray spectrometry supports operational snowpack monitoring efforts (Bland et al., 1997; Carroll, 2001; Grasty, 1982; Ishizaki et al., 2016). Since the 1980s, airborne gamma radiation snow surveys conducted by the NOAA's Office of Water Prediction (OWP; and formerly by the National Operational Hydrologic Remote Sensing Center [NOHRSC]) have provided SWE observations to regional NWS RFCs across the U.S. (Carroll, 2001; Mote et al., 2003). The historical 40 years gamma SWE record was proven as reliable long-term reference SWE observations across the U.S. and southern Canada (Cho et al., 2019). The SWE data are also assimilated into NOAA NWS's NOHRSC SNOw Data Assimilation System (SNODAS) (Barrett, 2003; Clow et al., 2012; Hedrick et al., 2015).

Terrestrial gamma-ray emission from radioisotopes in surface soils (~ top 20 cm) is attenuated by water in the liquid or solid form (Carroll, 2001; Peck et al., 1980). The difference between gamma radiation measurements taken in the fall (without snow) and in the winter (with snow) forms the basis of gamma-ray based airborne SWE measurements. Each flight line's footprint is approximately 4.5 – 6 km^2 (15 – 20 km long and about 300 m wide). Flight lines are measured once in the fall (in October or November) and then revisited several times throughout the winter (January to April) to estimate SWE (Carroll, 2001). The operational gamma SWE measurements are considered to be accurate assuming that SM conditions measured during the fall survey remain unchanged prior to winter surveys. However, SM conditions can change due to late-season rainfall events and early-winter snowmelt, which can result in large gamma SWE errors (e.g. NASA SnowEx Science Plan; Durand et al., 2019). Tuttle et al. (2018), for example, found a root mean square difference of 42.7 mm between AMSR-E SWE and airborne gamma SWE in the Northern Great Plains, including parts of the North Dakota, South Dakota, Minnesota, and Iowa, the United States and southern Canadian prairies. They mentioned that a large portion of the error was likely due to the assumption that SM remains constant from fall into winter.

Beginning with the SMMR from 1978 to 1987, satellite active and passive microwave sensors such as AMSR-E (2002 – 2011), ASCAT (Advanced Scatterometer; 2007, 2012, and

2018 – present, from Metop-A, B, and C, respectively) and SMOS (Soil Moisture and Ocean Salinity; 2010 – present) have provided surface SM. Two recent instruments are the AMSR2 (2012 – present) and SMAP (Soil Moisture Active Passive; 2015 - present). The L-band radiometer aboard the National Aeronautics and Space Administration's (NASA) SMAP satellite is well suited for measuring surface SM (Entekhabi et al., 2010). SMAP was launched in January 2015 and provides SM measurements globally every 2-3 days. SMAP SM observations have been used to study soil moisture dynamics (Akbar et al., 2018; Kim et al., 2019; McColl et al., 2017), which are important for hydrological and agricultural applications, such as flood detection (Fournier et al., 2016), irrigation signals (Lawston et al., 2017), and drought monitoring (Mishra et al., 2017), at both global and regional scales. However, satellite microwave-based SM products have well-known limitations for representative depths (~ upper few centimeters) and high uncertainties over dense-vegetated areas (Jackson & Schmugge, 1991; Entekhabi et al. 2014; Chan et al., 2018).

The physics used to estimate SM differ between gamma radiation and satellite microwave sensing. The gamma radiation method uses the difference between the naturally occurring terrestrial gamma radiation flux from wet and dry soils (Carroll, 1981; Jones & Carroll, 1983). The gamma flux from the ground is a function of the water mass and constant radioisotope concentration near the surface. The mass of the moisture regardless of any phase of water affects the attenuation. Increasing SM increases the gamma radiation flux attenuation and decreases the gamma flux at the ground surface. Passive microwave sensors estimate the soil dielectric constant using the observed brightness temperature (Tb) of the land surface (Jackson et al., 1993). Using the estimated dielectric constant, a dielectric mixing model leverages the large difference between the dielectric constants of the soil particles (~4) and water (~80) to obtain the amount of SM with soil texture information. In the single channel algorithm (SCA) in the NASA SMAP standard products, the vertically polarized Tb observations by SMAP L-band are converted to emissivity using ancillary physical temperature (Chan et al., 2018; Dong et al., 2018; O'Neill et al., 2015; updated 2019). The derived emissivity is corrected for surface roughness and vegetation to obtain soil emissivity. The soil emissivity is related to the dielectric properties of the soil and the incidence angle. The Fresnel reflection equation (Ulaby et al., 1986) is then used to determine the dielectric constant.

Land surface model (LSM) provides an alternative source of simulated SM products and have been vetted in weather and climate models as well as hydrological extreme monitoring (e.g. drought and floods) (Koster et al., 2009). The North American Land Data Assimilation System Phase 2 (NLDAS-2) provides simulated SM products for central North America using four land surface models, Noah (Ek et al., 2003; Wood et al., 1997), Mosaic (Koster and Suarez, 1996), Sacramento soil moisture accounting (SAC, Burnash, 1995), and the Variable Infiltration Capacity (VIC, Liang et al., 1994), which have high spatial (12.5 km by 12.5 km) and temporal (hourly) resolution (Xia et al., 2014).

This study seeks to identify which of the aforementioned SM products can improve airborne gamma SWE estimates by updating the ("baseline") fall operational gamma SM estimates to account for changes in SM conditions after baseline gamma flights. This study aims to answer the following four research questions:

1. Are temporal changes in SM from satellite and LSM model products similar to each other after baseline gamma flights?
2. Which satellite and LSM SM products have strong agreement with operational airborne gamma SM?

3. How much does updating the baseline operational gamma SM change gamma SWE estimates?
4. Does the updated gamma SWE improve agreement with independent SWE observations?

4.2 Study Concept

Operational airborne gamma radiation snow surveys rely on the assumption that the SM measured during the fall survey remains constant prior to winter SWE surveys. When SM conditions evolve due to drying, rainfall events, and/or early-winter snowmelt, gamma SWE estimates biases result. Figure 1 shows an example of a SMAP soil moisture time series from the “ND440” flight line footprint, the gamma SM estimate for the flight line, and the daily rainfall and soil temperature data in Mooreton, North Dakota from North Dakota Agricultural Weather Network (NDAWN, https://ndawn.ndsu.nodak.edu) are also shown. The figure illustrates the soil moisture changes after the fall baseline gamma SM survey and their potential influence on the winter gamma SWE estimates. From the 9 November 2016 baseline gamma SM survey, SMAP SM evolves until 1 December 2017 with a net 0.12 m^3/m^3 increase. The gamma SWE estimated on 18 January 2017 using the baseline gamma SM value attributes all the additional gamma radiation attenuation in the winter measurement to SWE rather than accounting for the increase in soil moisture post-baseline survey. If the baseline gamma SM were updated to reflect the fall SM changes, then the operational gamma SWE should be reduced to reflect that portion of the attenuation of gamma radiation due to an increase in SM. Thus, gamma SWE estimates may be improved using an updated gamma SM estimate.

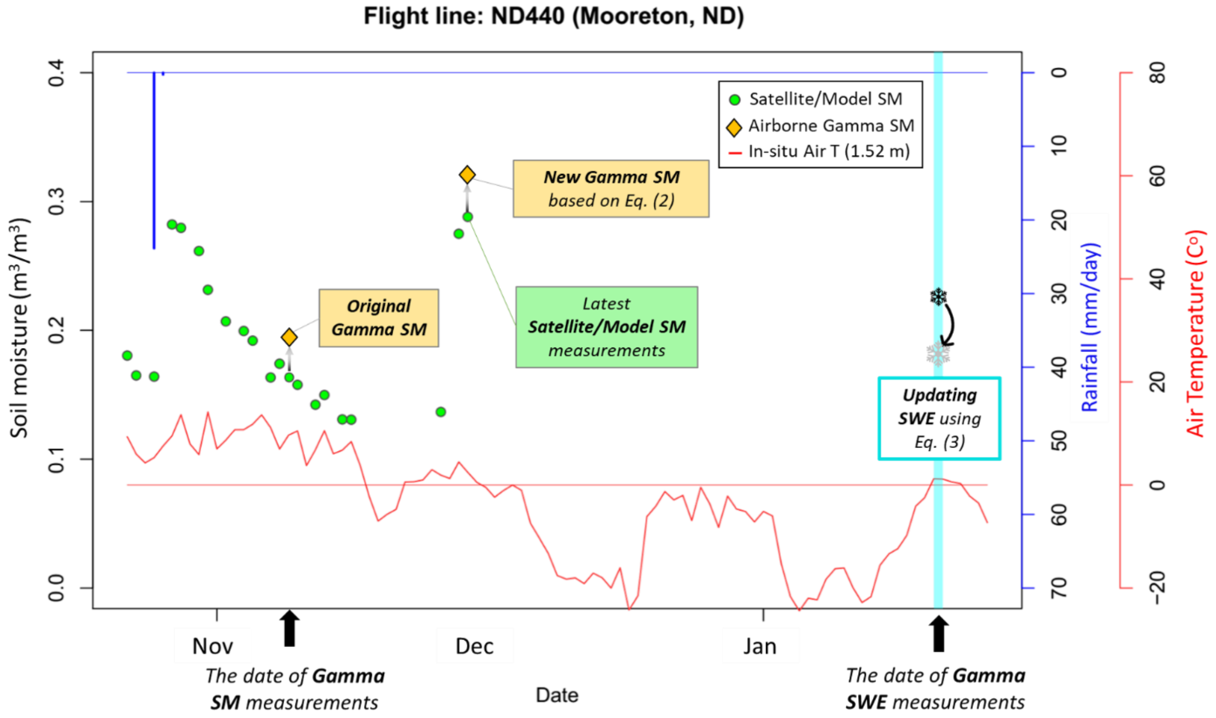


Figure 1. An example time series of satellite/model soil moisture (SMAP enhanced products in this figure) within the given flight line footprint and NOAA operational gamma soil moisture along with daily rainfall and air temperature in 2016 to 2017 from a North Dakota Agricultural Weather Network (NDAWN) station at Mooreton, ND. The ND440 flight line was flown over the Mooreton station. The increase in SMAP soil moisture in December was due to early snowmelt from 26 to 30, November. The errors of the SMAP product (ubRMSE < 0.04 m^3/m^3) meet the mission performance criteria from previous studies (Chen et al., 2018; Colliander et al., 2018).

4.3 Study Area

The study area comprises parts of the north-central and northeast United States and southern Canada (Figure 2), including parts of four RFCs (Missouri Basin RFC (MBRFC), North-Central RFC (NCRFC), North-East RFC (NERFC), and Mid-Atlantic RFC (MARFC)) and two Canadian Provinces including Saskatchewan (SK) and Manitoba (Winnipeg). The RFC boundaries (black lines) were designated by the NOAA NWS Integrated Hydrologic Automated Basin Boundary System to support river flow and flood forecasting throughout the United States. Gamma surveys occur in each regional RFC. The gamma flight lines in Figure 2 were flown from September 2015 to April 2018 (black lines). The flight times range from 9 AM to 6 PM according to weather conditions and operations schedule (https://www.nohrsc.noaa.gov/snowsurvey/photos/). The region is dominated by three land cover types, forest (19%, Deciduous broadleaf forest and Mixed forest), croplands (77%, Croplands and Cropland/Natural vegetation mosaic), and grasslands (4%) from Global Mosaics of the Moderate Resolution Image Spectroradiometer (MODIS) land cover type data (MCD12Q1) using the International Geosphere-Biosphere Programme (IGBP) Land Cover Type Classification (Channan et al., 2014). Airborne gamma surveys in the western U.S. were excluded because most of there SM estimates from 2015 to 2018 used a subjective estimate ('SE') or unknown type ('0') (https://www.nohrsc.noaa.gov/snowsurvey).

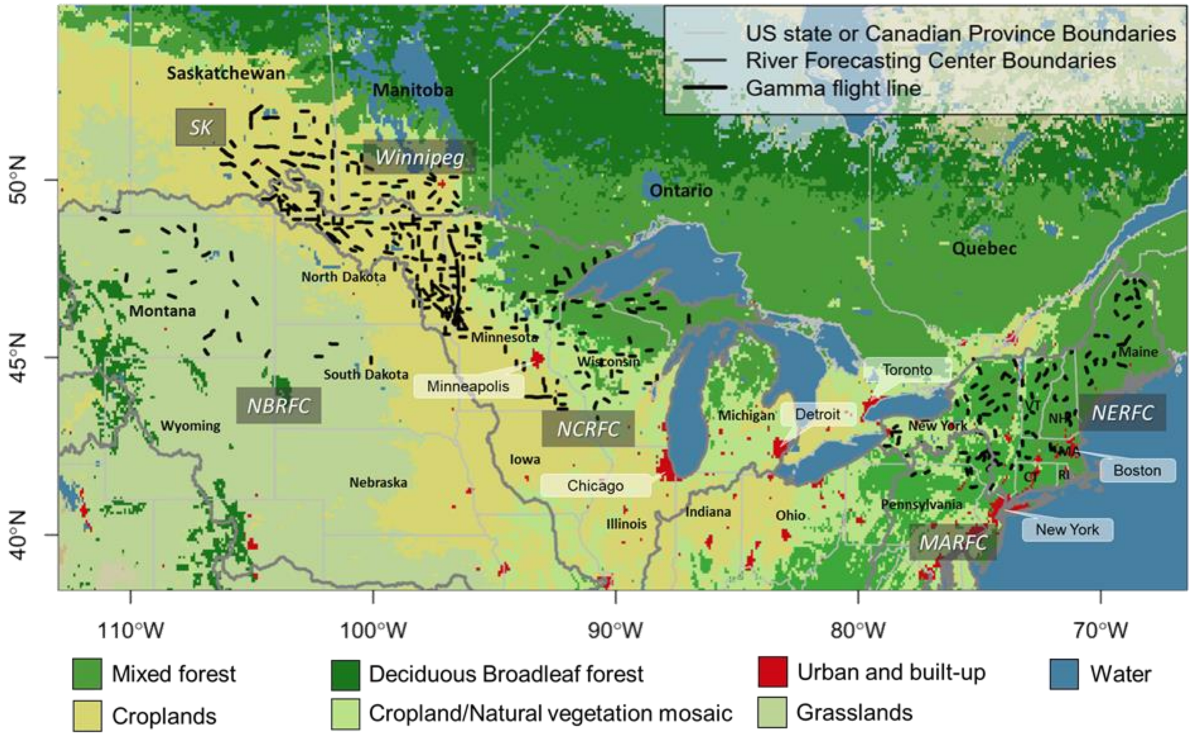


Figure 2. Land cover map of the study area of the north-central and eastern United States and southern Canada with the NOAA airborne gamma flight lines surveyed from 2015 to 2018 (N = 574, blue lines with cyan borders) with River Forecasting Center (RFC) boundaries (black lines) along with U.S. states and Canadian province boundaries (gray lines). The land cover map is

from Global Mosaics of the Moderate Resolution Image Spectroradiometer (MODIS) land cover type product (MCD12Q1).

4.4 Data and Methodology

This study uses a number of SM and SWE products (Table 1). The details of each data product appear in the following sections.

Table 1 Summary of soil moisture and snow water equivalent products including data type, period, footprint/grid size, and source used in this study

Data	Product	Type	Period	Footprint /Grid size	Source
SM & SWE	NOAA gamma	Airborne gamma radiation	2015-2018	5-7 km^2	NOAA
SM	SMAP enhanced	Satellite passive microwave	2015-2017	9 km	NASA
SM	NLDAS-2 Mosaic	Land surface model	2015-2017	12.5 km	NOAA
SM	AMSR2 LPRM	Satellite passive microwave	2015-2017	25 km	NASA
SWE	SSMIS	Satellite passive microwave	2016-2018	25 km	NASA
SWE	GlobSnow	Assimilation	2016-2018	25 km	ESA
SWE	SCAN	In-situ station	2017-2018	point	USDA
SWE	USACE	In-situ field survey	2017-2018	point	USACE

4.4.1 NOAA Airborne gamma survey

The NWS gamma flight line network includes over 2,400 flight lines covering 29 U.S. states and seven Canadian provinces (Carroll, 2001; Peck et al., 1980). Since 1979, the NWS gamma radiation snow survey program has made about 27,000 gamma SWE measurements over North America via the NOHRSC website (http://www.nohrsc.noaa.gov/snowsurvey/). This study uses the 770 airborne SWE observations made from 2015 to 2018 with 413 flight lines in the study area including 648 observations in non-forested areas. A typical flight line is approximately 300 m wide and 16 km long (5 km^2 footprint). The gamma survey SM and SWE observations are areal-average values for each flight line footprint, while satellite and model products used in this study are provided as pixel values.

The airborne gamma radiation technique measures the attenuation of the terrestrial gamma radiation signal due to the intervening water mass (Carroll, 2001; Peck et al.,1971). The gamma flux near the ground surface originates primarily from the ^{40}K, ^{208}Tl, and ^{238}U radioisotopes in the soil. In a typical soil, 91% of the gamma radiation signal is emitted from the top 10 cm of the soil and 96% and 99% from the top 20 cm and 30 cm, respectively (Zotimov, 1968). Airborne gamma fall SM measurements can be made for a given flight line if background terrestrial gamma count rates ($^{40}K_0$, $^{208}Tl_0$, and gross count, GC_0) and coincident background SM (SM_0), and gamma count rates are available. Ground-sampled SM data collected over calibration flight lines are used to determine background SM (Jones and Carroll, 1983). Three independent SM values are calculated using the attenuation of the gamma radiation counts. SM values are calculated using gamma count rates from the ^{40}K window (1.36 - 1.56 MeV), ^{208}Tl (2.41 – 2.81 MeV) window, and GC spectrum (0.41 to 3.0 MeV) by the following equations (Carroll, 1981; 2001)

$$SM\left({}^{40}K_c\right) = \frac{\frac{{}^{40}K_0}{{}^{40}K_c}(100+1.11SM_0)-100}{1.11} \qquad \text{Eq. (1)}$$

$$SM\left({}^{208}Tl_c\right) = \frac{\frac{{}^{208}Tl_0}{{}^{208}Tl_c}(100+1.11SM_0)-100}{1.11} \qquad \text{Eq. (2)}$$

$$SM(GC_c) = \frac{\frac{GC_0}{GC_c}(100+1.11SM_0)-100}{1.11} \qquad \text{Eq. (3)}$$

$$SM_c = 0.346 \cdot SM\left({}^{40}K_c\right) + 0.518 \cdot SM\left({}^{208}Tl_c\right) + 0.136 \cdot SM(GC_c) \qquad \text{Eq. (4)}$$

where $^{40}K_c$, $^{208}Tl_c$, and GC_c are current uncollided gamma count rates in windows ^{40}K, ^{208}Tl, and GC, respectively, and $^{40}K_0$, $^{208}Tl_0$, and GC_0 are background uncollided gamma count rates. A single current SM estimate (SM_c, in units of percent by weight) is calculated by multiplying the three current SM estimates by weighting factors, 0.346, 0.518, and 0.136 for ^{40}K, ^{208}Tl, and GC, respectively (Jones & Carroll, 1983). Only the single, weighted SM (SM_c) is reported as antecedent fall SM which is used in this study. The fall SM survey data are available as Standard Hydrometeorological Exchange Format (SHEF) product through the NWS NOHRSC website (https://www.nohrsc.noaa.gov/snowsurvey/).

The operational gamma SWE measurements are made using the following equations:

$$SWE\left({}^{40}K\right) = \frac{1}{A} \cdot \left[ln\left(\frac{{}^{40}K_b}{{}^{40}K_s}\right) - ln\left(\frac{100+1.11 \cdot SM\left({}^{40}K_s\right)}{100+1.11 \cdot SM\left({}^{40}K_b\right)}\right)\right] \qquad \text{Eq. (5)}$$

$$SWE\left({}^{208}Tl\right) = \frac{1}{A} \cdot \left[ln\left(\frac{{}^{208}Tl_b}{{}^{208}Tl_s}\right) - ln\left(\frac{100+1.11 \cdot SM\left({}^{208}Tl_s\right)}{100+1.11 \cdot SM\left({}^{208}Tl_b\right)}\right)\right] \qquad \text{Eq. (6)}$$

$$SWE(GC) = \frac{1}{A} \cdot \left[ln\left(\frac{GC_b}{GC_s}\right) - ln\left(\frac{100+1.11 \cdot SM(GC_s)}{100+1.11 \cdot SM(GC_b)}\right)\right] \qquad \text{Eq. (7)}$$

$$SWE_{gam_{oper}} = 0.346 \cdot SWE\left({}^{40}K\right) + 0.518 \cdot SWE\left({}^{208}Tl\right) + 0.136 \cdot SWE(GC) \qquad \text{Eq. (8)}$$

where $SM\left({}^{40}K_b\right)$, $SM\left({}^{208}Tl_b\right)$, and $SM(GC_b)$ are SM values by weight (%) over bare ground and $SM\left({}^{40}K_s\right)$, $SM\left({}^{208}Tl_s\right)$, and $SM(GC_s)$ are SM values over snow-cover ground. $^{40}K_b$, $^{208}Tl_b$, and GC_b are uncollided gamma count rates over bare ground and $^{40}K_s$, $^{208}Tl_s$, and GC_s for snow-covered ground. $SWE_{gam_{oper}}$ is the operational gamma radiation SWE estimate (g/cm^2) reported in the SHEF product (Carroll and Schaake Jr, 1983; Carroll, 2001). Based on previous studies, errors of the airborne gamma SM measurement range from - 9.9 to 2.9% of percent bias (Carroll, 1981). Errors of the gamma SWE were about 12.1% over agricultural areas in the Upper Midwest U.S. and 1.3 – 24% over forested areas of the Lake Superior basin, U.S. and Saint John River basin, Canada. (Carroll and Carroll, 1989a; Carroll, 2001; Glynn et al., 1988). Glynn et al. (1988) indicate that the potential sources of errors in gamma SWE estimates include gamma count statistics, navigation, and biomass.

The airborne gamma SM estimate is provided as “percent SM by weight” which is the weight of SM divided by the weight of dry soil multiplied by 100 from approximately the top 20 cm of soil. To compare the gamma SM (by weight, %) to the gridded SM products (volumetric content, m^3/m^3), the units of SM were matched. The percent airborne gamma SM by weight was converted to volumetric SM contents (m^3/m^3) using the constant bulk density (1.295 g/cm^3) based on a dominant soil bulk density in this study area (Dobson et al., 1985). Our results show that using a constant bulk density as compared to individual bulk density for each gamma footprint using the 1-km POLARIS soil datasets (available at www.polaris.earth; Chaney et al., 2016) does not generate additional residual errors in the comparison between gamma SM and other SM products (Figure S1 & S2).

4.4.2 Soil moisture (SM)

4.4.2.1 SMAP enhanced SM

The NASA’s SMAP satellite’s L-band radiometer has provided global SM measurements at 6:00 A.M./P.M. local time at 2–3 days revisit time since March 31, 2015 (Chan et al., 2016; Entekhabi et al., 2010). The SMAP SM product has been validated using ground-based observations and various assimilation products at a global scale (Kim et al., 2018; Colliander et al., 2017; Ma et al., 2019; Zhang et al., 2019; Zwieback et al., 2018).

The SMAP enhanced L3 SM, released in December 2016, is derived from SMAP Level-1C (L1C) interpolated brightness temperatures using Backus-Gilbert optimal interpolation techniques (O'Neill et al., 2018). The SMAP enhanced SM product (9 x 9 km^2) retrieved by the SCA (V-pol) has a finer grid posting relative to the SMAP native products (36 x 36 km^2) although the enhanced footprint's contributing domain is ~ 33km is similar to the native 36 km resolution (Chan et al., 2018). In this study, the SMAP level 3 radiometer global daily EASE-Grid 2.0 (Equal-Area Scalable Earth Grid 2.0) enhanced soil moisture (V002) for the descending overpass (6 A.M.) is used from September 2015 to March 2018. This product (V002) has an improved depth correction for effective soil temperature, which reduced the dry bias in the initial version product (V001) (O'Neill et al., 2018).

4.4.2.2 AMSR2 SM

The AMSR2 passive microwave sensor, a follow-on of the AMSR-E sensor aboard the Aqua satellite, was launched on the GCOM-W1 satellite in May 2012 (Imaoka et al., 2010). The AMSR2 provides daily scans at 1:30 A.M. (descending) / P.M. (ascending) local time with 1–2 days revisit time. There are three widely used AMSR2 surface SM products generated from different algorithms, the LPRM (Land Parameter Retrieval Model) (Owe et al., 2008), the JAXA algorithm (Koike, 2013; Cho et al., 2015) and the SCA (Single Channel Algorithm; Bindlish et al., 2018). The LPRM uses the dual-polarization Tb observations at individual (C or X) bands to retrieve surface SM and vegetation optical depth via a forward radiative transfer model (Owe et al., 2008). This study uses the LPRM AMSR2, Level 3 gridded X-band (10.7 GHz) SM from the ascending overpass, expressed on a regular 1/4° spatial grid (25 km).

4.4.2.3 NLDAS-2 Mosaic SM

The NLDAS-2 is an offline modeling system, running four land surface models [Noah, Mosaic, Sacramento soil moisture accounting (SAC), and the Variable Infiltration Capacity (VIC) model] on a 1/8° spatial grid (12.5 km) over the continental United States (CONUS). NLDAS-2 uses meteorological forcing data (e.g. downward short/longwave radiation, precipitation, 2-m air temperature, 2-m specific humidity, and 10-m wind speed) to run the land surface models to produce water and energy fluxes and state variables (Xia et al., 2012). The NLDAS-2 has SM products from four land surface models (Mosaic, Noah, SAC, and VIC) (Xia et al., 2014). The Mosaic model has three soil layers: 0–10 cm, 10–60 cm, and 60–200 cm (Koster & Suarez, 1996). In this study, the Mosaic 12:00 PM SM at a depth of 0-10 cm is used to represent modeled SM values, because the Mosaic SM had a stronger agreement with the airborne gamma SM than the Noah and VIC SM products from the surface soil layer [0-10 cm] (Figure S3). The SAC SM was not compared because it uses a single soil layer with no surface soil moisture.

In summary, this study uses SMAP and AMSR2 SM products as well as the NLDAS-2 Mosaic SM product. Active microwave satellite (e.g. ASCAT) SM is not included because recent studies found that passive microwave SM (e.g. SMAP/SMOS) products generally have a stronger agreement with in-situ observations or reanalysis SM products than ASCAT SM over our study region (Al-Yaari et al., 2014; Kim et al., 2018).

4.4.3 Snow water equivalent (SWE)

4.4.3.1 SSMIS SWE

The SSMIS sensor onboard the Defense Meteorological Satellite Program (DMSP) F17 platform has provided daily brightness temperature (Tb) measurements with near-complete global coverage from December 2006 to the present. In this study, F17 SSMIS SWE (SWE_{SSMIS}) was estimated using the Chang-type algorithm (Armstrong and Brodzik, 2001; Chang et al., 1987) with modified coefficients developed by Brodzik (2014) as follows:

$$SWE_{SSMIS} = a \cdot Tb_{H,19GHz} - b \cdot Tb_{H,37GHz} - c \qquad \text{Eq. (9)}$$

where a, b, and c are given as 4.807 mm/K, 4.792 mm/K, and 6.036 mm, respectively. $Tb_{H,19GHz}$ and $Tb_{H,37GHz}$ are the brightness temperature at 19 and 37 GHz horizontal polarization, respectively. The DMSP SSM/I-SSMIS Pathfinder Daily EASE-Grid Brightness Temperatures (Version 2) are provided on a 25-km grid on the National Snow & Ice Data Center website (https://nsidc.org/data/nsidc-0032; Armstrong et al., 1994). SSMIS Tb data from the descending overpass (6 A.M.) were used to minimize the potential error by wet snow (Derksen et al., 2000).

4.4.3.2 GlobSnow SWE

The European Space Agency GlobSnow project provides long-term gridded daily SWE maps with 25 km x 25 km spatial resolution from 1979 to current for the Northern Hemisphere, except for glaciers and mountainous regions (Takala et al., 2011). The GlobSnow SWE utilizes an assimilation approach, which combines ground-based synoptic snow depth station data (using constant snow density, 0.24 kg/m^2) with passive microwave satellite measurements via the Helsinki University of Technology (HUT) snow emission model (Takala et al., 2011; Pulliainen, 2006). Ground-based point snow depth measurements are from the World Meteorological Organization weather stations. The GlobSnow SWE has two versions, GlobSnow-2 from 1979 to 2016 (archive_v2.0; http://www.globsnow.info/swe/archive_v2.0/) and GlobSnow-1 from 2011 to current (near-real-time; http://www.globsnow.info/swe/nrt/). The retrieval accuracy is the same between the GlobSnow-1 and 2, but the GlobSnow-2 SWE was improved for northern boreal forest and tundra regions (Luojus et al., 2014). Due to the current study period from 2015 to 2018, the daily GlobSnow-1 SWE was used to evaluate the updated gamma SWE.

4.4.3.3 Ground-based SWE

Compared to the western U.S., there are few SWE stations in the north-central and northeastern U.S. Daily SWE measurements at the Glacial Ridge, Minnesota (ID: 2050; Latitude/Longitude: 47.72°/96.26°; Elevation: 343 m) operated by the United States Department of Agriculture (USDA) Soil Climate Analysis Network (SCAN) were compared to the updated gamma SWE measurements. The SCAN site land cover is "croplands" with a "prairie" snow classification. Two gamma flight lines, MN119 and MN120, are located near the SCAN site with the flight lines' midpoints approximately 9.8 km (northwards) and 29.7 km (southwards), respectively, from the SCAN site. The two flight lines' land cover is also "cropland" and their elevations are about the same (Figure S4). Further details can be found on the SCAN website (https://wcc.sc.egov.usda.gov/nwcc/site?sitenum=2050).

The United States Army Corps of Engineer (USACE) ground-based snow survey data were collected by the USACE St. Paul District to determine snowpack SWE for spring flood risk assessment and water resources management. Their survey measurements sampled the snowpack at representative locations. At each site and date, at least four SWE samples were taken, each approximately 3–4 m apart, using a snow tube (3.81 cm diameter), then averaged to a single mean SWE value. This study uses the weekly USACE SWE observations from 2017 to 2018 at Baldhill, ND (Latitude/Longitude: 47.03°/-98.08°), Orwell, MN (46.22°/-96.18°), and Traverse, MN (45.86°/-96.57°). The gamma flight lines closest to each site with a distance between the midpoint of flight line and the site are ND432 and ND433 (10.6 km and 26.3 km from Baldhill), MN126 and MN129 (24.8 km and 19.2 km from Orwell), and ND441 and MN124 (13.8 km and 22.6 km from Orwell). The detailed gamma flight line locations are provided in Supplementary material (Figure S4).

4.4.4 Methodology

For comparison to the airborne gamma SWE data, the satellite or model pixels overlapped by the given flight line footprint were weighted according to a portion of the footprint within each pixel. Only flight lines having more than 50% of the footprint covered by satellite observations were used in this analysis. For a detailed process with a schematic diagram, please refer to Tuttle et al. (2018).

After one SM product (in this case, the SMAP enhanced SM) was selected based on the statistical agreement (e.g. correlation coefficient and unbiased root mean square difference) with operational baseline gamma SM, a linear regression model that minimizes the sum of squared residuals (ε_i) was developed to relate coincident gamma SM ($SM_{gam,i}$) and the satellite (or model) SM ($SM_{sat,i}$) measurements.

$$SM_{gam,i} = a \cdot SM_{sat,i} + b \pm \varepsilon_i \qquad \text{Eq. (10)}$$

where *i* is flight line number, *a* is the slope and *b* is the y-intercept of the linear regression equation. ε_i is a residual error (m^3/m^3) between operational gamma SM and satellite (or model)

SM for each flight line. Based on the model, new, updated gamma SM estimates were calculated by applying the latest antecedent SM of the chosen product into the linear regression model. It is assumed that the residual, ε_i, is largely generated from differences between the two products' representative areas and land surface characteristics for each flight line. Thus, the residuals are included in the updated gamma SM.

The change in airborne gamma SWE, $\Delta SWE_{gam,i}$, resulting from a change in antecedent SM in the unit of percentage (%) in soil was calculated using Carroll (2001) as follows:

$$\Delta SWE_{gam,i} = \frac{25.4}{A} \cdot \left[\ln \left(\frac{100 + 1.11 \cdot SM_{gam_{oper},i}}{100 + 1.11 \cdot SM_{gam_{upd},i}} \right) \right] \qquad \text{Eq. (11)}$$

where $\Delta SWE_{gam,i}$ is the change in snow water equivalent (mm), *A* is a radiation attenuation coefficient of water which is equal to 0.1482 (Carroll, 2001). 25.4 is used to convert "inches" to "mm" from Equation 3 in Carroll (2001). 1.11 represents the ratio of gamma radiation attenuation in water to air (Carroll, 1981). $SM_{gam_{oper},i}$ is operational gamma SM by weight (%) measured in the fall survey and $SM_{gam_{upd},i}$ is the updated gamma SM by weight (%). A schematic diagram of the methodology is provided in the Supplementary materials (Figure S5). The agreement between airborne gamma survey and satellite/model SM (or SWE) products was quantified by the Pearson's linear correlation coefficient, *R*, the mean bias, *Bias*, the root mean square difference, *RMSD*, and the unbiased RMSD, *ubRMSD*. The equations are available in the Supplementary material (Text S1).

4.5 Results

4.5.1 Change in the soil moisture after baseline gamma flights from satellite and model products

Figure 3 compares the change in NLDAS-2, SMAP, and AMSR2 regional SM maps from the dates of the baseline fall gamma flights until the last observation before freeze onset. As an example, in 2016 most gamma SM flights occurred about 25 October and the latest observation available prior to freezing onset was on 29 November. After the fall gamma flights, SM changes vary by year and location. These changes are typically caused by later rainfall, early-winter snowmelt, and/or freeze/thaw events, suggesting that an adjustment of the baseline gamma SM is necessary for accurate gamma SWE survey.

In 2015, the change in NLDAS-2 and SMAP SM from November 25 to December 12 show similar spatial patterns. Surface soils became wetter in the north-central U.S. and drier in the northeastern U.S. The increases in SMAP SM are greater than NLDAS in Minnesota, North Dakota, and South Dakota where many of the gamma flights occurred. The AMSR2 SM change is remarkably different from NLDAS-2 and SMAP SM. AMSR2 shows drying in Minnesota and most Canadian provinces. In 2016, SM changes clearly differ by data source between 25 October and 29 November. SMAP has a strong drying signal of up to -0.17 m^3/m^3 in north-central and eastern U.S. as well as Saskatchewan and Manitoba, Canada. However, NLDAS-2 and AMSR2 SM in the same regions get wetter by up to 0.12 and 0.25 m^3/m^3, respectively. In the Midwest, AMSR2 shows that SM increases exceed 0.25 m^3/m^3. In 2017, there are clear decreases in NLDAS-2 and SMAP SM from 25 October to 13 December in the Midwest. The drying of SMAP (~0.20 m^3/m^3) is stronger than that of NLDAS-2 (~0.10 m^3/m^3). NLDAS-2 captures modest wetting in Canada, which is not seen by SMAP and AMSR2 SM because these datasets are provided for only limited areas in Canada, due to data masking from soil freeze or snow cover.

In general, SMAP SM changes are spatially similar to NLDAS-2 SM changes but have amplified drying (and wetting). AMSR2 has extreme SM changes considering the normal range of volumetric SM and differs spatially from SMAP and NLDAS-2, which may reflect the much thinner and closer-to-the-surface sensing depth of AMSR2 as compared to SMAP and NLDAS-2's deeper sensing depths.

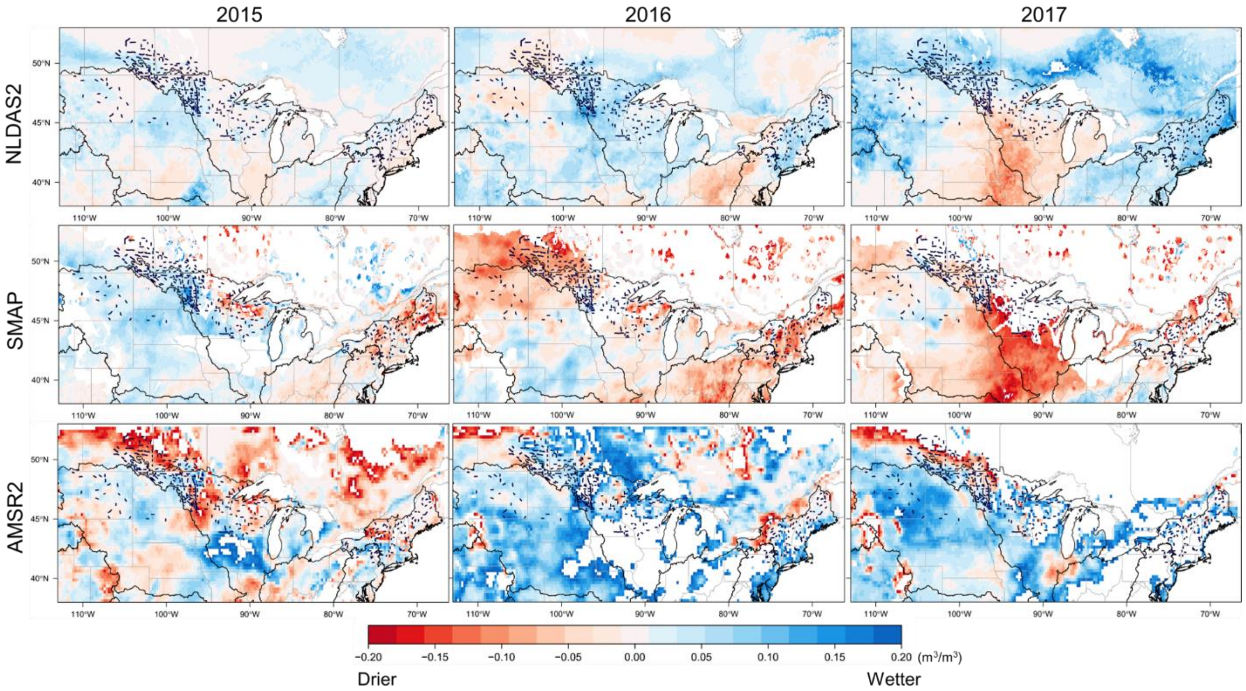


Figure 3. **SM difference maps for NLDAS-2, SMAP, and AMSR2 for the years 2015 to 2017.** SM differences are calculated between the date of the fall baseline gamma flights and the date of the last SM observation prior to freezing onset. A past 5-day composite SM map was used to eliminate spatial gaps.

4.5.2 Airborne gamma SM versus satellite and model SM products

To identify which satellite or model SM product agrees best with gamma SM, the gamma SM data were compared to NLDAS-2, SMAP, and AMSR2 SM products. Because the performance of the microwave SM products typically weakens with increasing vegetation density (Jackson & Schmugge, 1991; Wang et al., 1982; Mladenova et al., 2014), the comparison is conducted with and without forest areas. When forested areas are included, NLDAS-2 SM has better agreement with operational gamma SM than the two satellite SM products (Table 1). There is little difference in agreement between NLDAS-2 mosaic SM and operational gamma SM with/without forest classes (Figure 4a & b). However, the agreement between SMAP and gamma SM clearly differs by a land cover (Figure 4c & d). A majority of the SMAP SM values with a wet bias occur for flights over forests. For the Deciduous broadleaf forest and Mixed forest classes, there are large errors with SMAP SM compared to gamma SM (*Bias*: 0.11 and 0.19 m^3/m^3 and *RMSD:* 0.17 and 0.21 m^3/m^3, respectively). For the AMSR2 comparison, most SM values over forested areas were excluded due to poor data quality before the analysis, but the remaining SM values show a wet bias, similar to SMAP SM, in forested regions (Figure 4e). AMSR2 SM has an extreme wet bias (0.13 m^3/m^3) even in non-forested areas. In non-forested regions, SMAP SM shows very strong agreement with gamma SM as compared to AMSR2 and NLDAS-2 SM (Table 2). The results indicate that SMAP SM values from forested areas (e.g. Deciduous broadleaf forest and Mixed forest) do not agree with the gamma observations and these land uses should be excluded if updating gamma SWE with SMAP SM. A linear regression model between SMAP and operational gamma SM [Eq. (10)] was developed using only the values from non-forested regions for the next step. Comparison between operational gamma SM and SMAP, AMSR2, and NLDAS-2 SM products for forested regions only, are provided in Figure S6.

Table 2 Agreement between NOAA airborne gamma SM and NLDAS-2 Mosaic SM, SMAP enhanced SM, and AMSR2 SM with/without the SM values from forested areas

Data	with forested areas					without forested areas				
	N	*R*	*ubRMSD* (m^3/m^3)	*RMSD* (m^3/m^3)	*Bias* (m^3/m^3)	N	*R*	*ubRMSD* (m^3/m^3)	*RMSD* (m^3/m^3)	*Bias* (m^3/m^3)
NLDAS-2	342	0.53	0.07	0.08	-0.03	277	0.53	0.07	0.08	-0.03
SMAP	342	0.52	0.10	0.10	0.02	277	0.69	0.06	0.06	-0.02
AMSR2	287	0.43	0.08	0.15	0.13	278	0.45	0.07	0.15	0.13

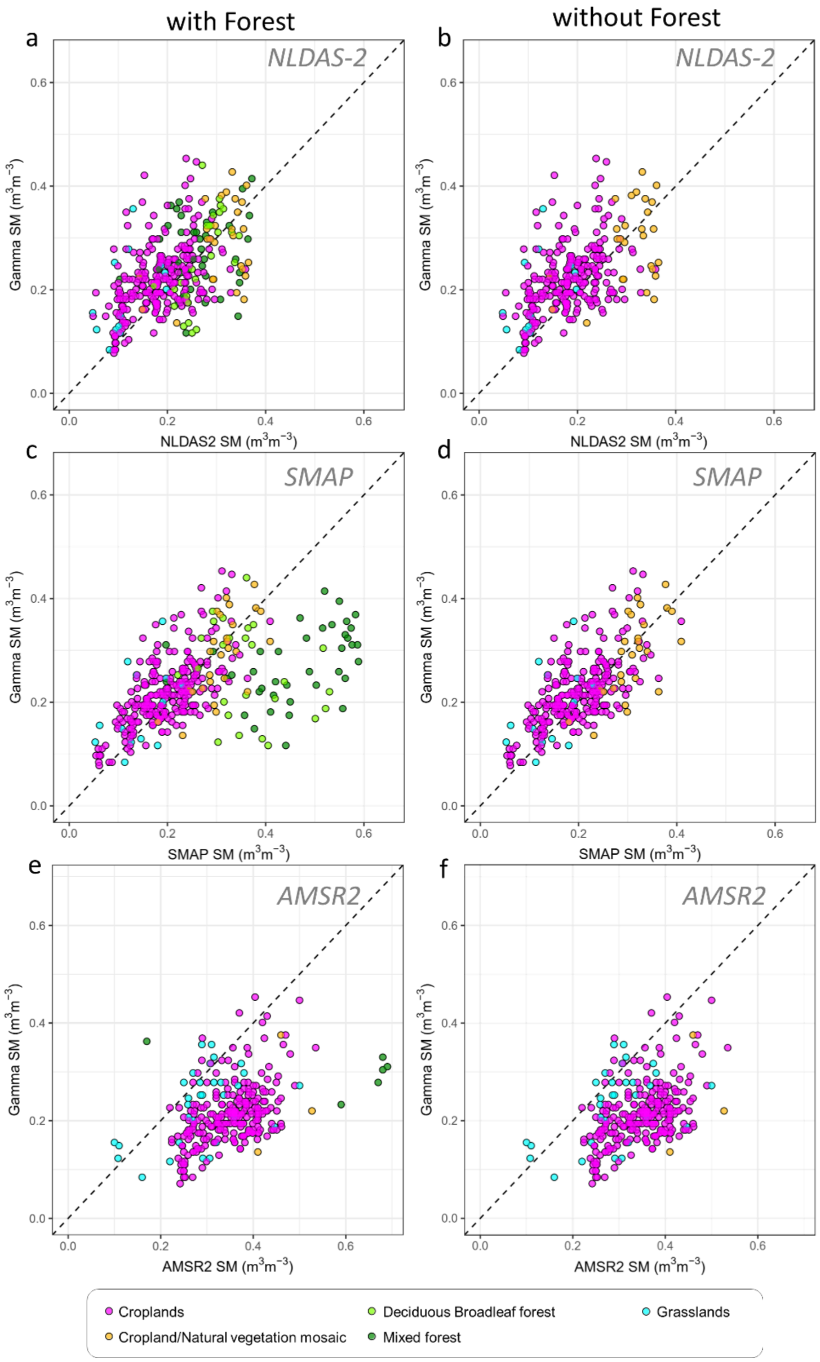


Figure 4. Comparison of NOAA airborne gamma soil moisture with (a, b) Phase 2 of the North American Land Data Assimilation System (NLDAS-2) Mosaic SM, (c, d) Soil Moisture Active Passive (SMAP) Level 3 enhanced soil moisture, and (e, f) Advanced Microwave Scanning Radiometer 2 (AMSR2) SM within the given flight line footprints with/without the SM values from forested areas.

4.5.3 Enhancement of gamma SWE by updating baseline SM

When the operational, baseline gamma SM in non-forested regions from 2015 to 2017 are updated using SMAP SM, the gamma SWE values change. Figure 5a displays SMAP SM changes measured between the date of the fall baseline gamma flights and the date of the last SM observation before freeze-up as well as the corresponding operational and SMAP-updated airborne gamma SM estimates. The SMAP-updated gamma SM were calculated using the linear regression model between airborne gamma and SMAP SM, slope (a) = 0.69 and y-intercept (b) = 0.083 [Eq. (10)]. The slope indicates that SMAP SM is more sensitive than gamma SM. Considering the two methods' different representative soil depths, it is reasonable that SMAP's surface SM would tend to have higher variability than the deeper gamma SM.

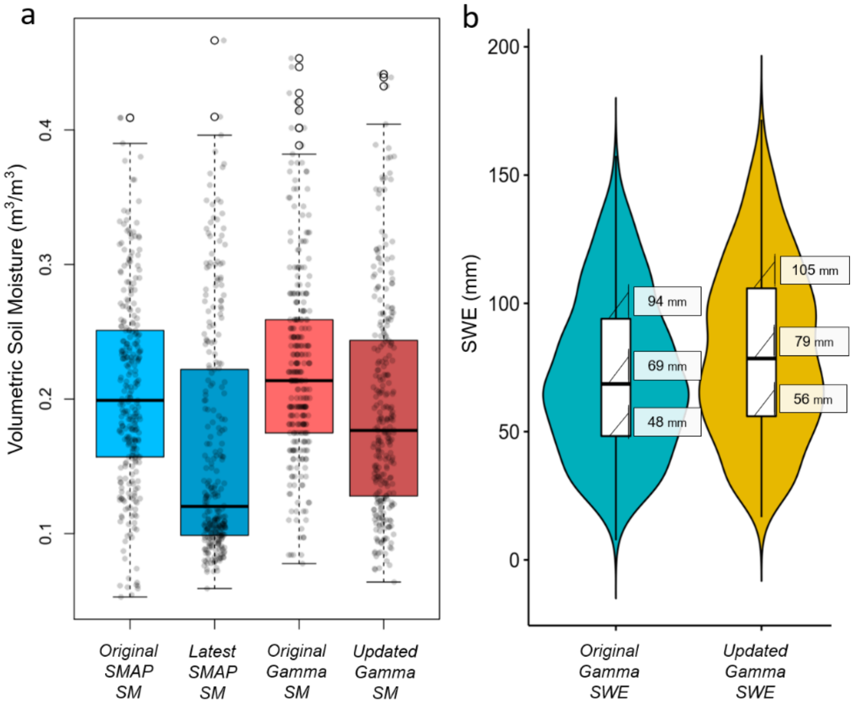


Figure 5. (a) Boxplots of SMAP SM at original (operational) and latest available dates and original and updated gamma SM for entire flight lines in the non-forested region from 2015 to 2017, along with (b) the corresponding original and updated gamma SWE. (a) The small circles are individual SM data (no meaning for a spread in the horizontal direction) and the larger circles are outliers. The bold line within each colored box is median, and the upper and bottom sides of

the box are the upper (75%) and lower (25%) quantiles of the data. (b) The width of the leaf-shape boxplot shows the relative amount of the SWE data at that magnitude.

The SMAP SM immediately before freeze-up (mean: 0.16 m^3/m^3, median: 0.12 m^3/m^3) is typically lower than the SM on the date of the fall baseline gamma flights (mean: 0.21 m^3/m^3, median: 0.20 m^3/m^3), indicating that for this study period most of the region dried in late fall to early winter. Note: a large portion of the gamma SM flights (193 of total 277 flight lines) occurred in fall 2016 when there was an average of 0.05 m^3/m^3 (median: 0.09 m^3/m^3) decrease in SMAP SM. As the SMAP SM differences between the baseline and latest available SM decrease, the gamma SM differences should also decrease following the linear regression model [Eq. (10)]. The SMAP-updated gamma SM is drier by an average of 0.03 m^3/m^3 than the operational baseline gamma SM. The SMAP-updated gamma SM also appears to have a greater interquartile range (IQR; total: 0.12 m^3/m^3) than the operational baseline gamma SM (0.08 m^3/m^3). This indicates that the residual values (ε_i) in the linear regression model comprise a considerable proportion of the variation in SMAP-updated gamma SM.

Using the SMAP-updated SM for each flight footprint, a new, SMAP-updated gamma SWE was calculated using Eq. (10). The original, operational gamma SWE values (mean: 72 mm, median: 69 mm) were adjusted upward by 15% (mean: 82 mm, median: 79 mm) when accounting for the changes in baseline SM (Figure 5b). In summary, decreases in the baseline SM by an average of 0.03 m^3/m^3 (gamma) and 0.05 m^3/m^3 (SMAP) generate average increases in gamma SWE of about 10 mm. Individual gamma SWE estimates have different SM changes due to the variations by year and flight line as presented in Figure 6. 75% of the SM values became drier and the remaining 25% became wetter, but with SM differences ranging from 0.22 to -0.25 m^3/m^3 and gamma SWE changes ranging from -30 to 41 mm.

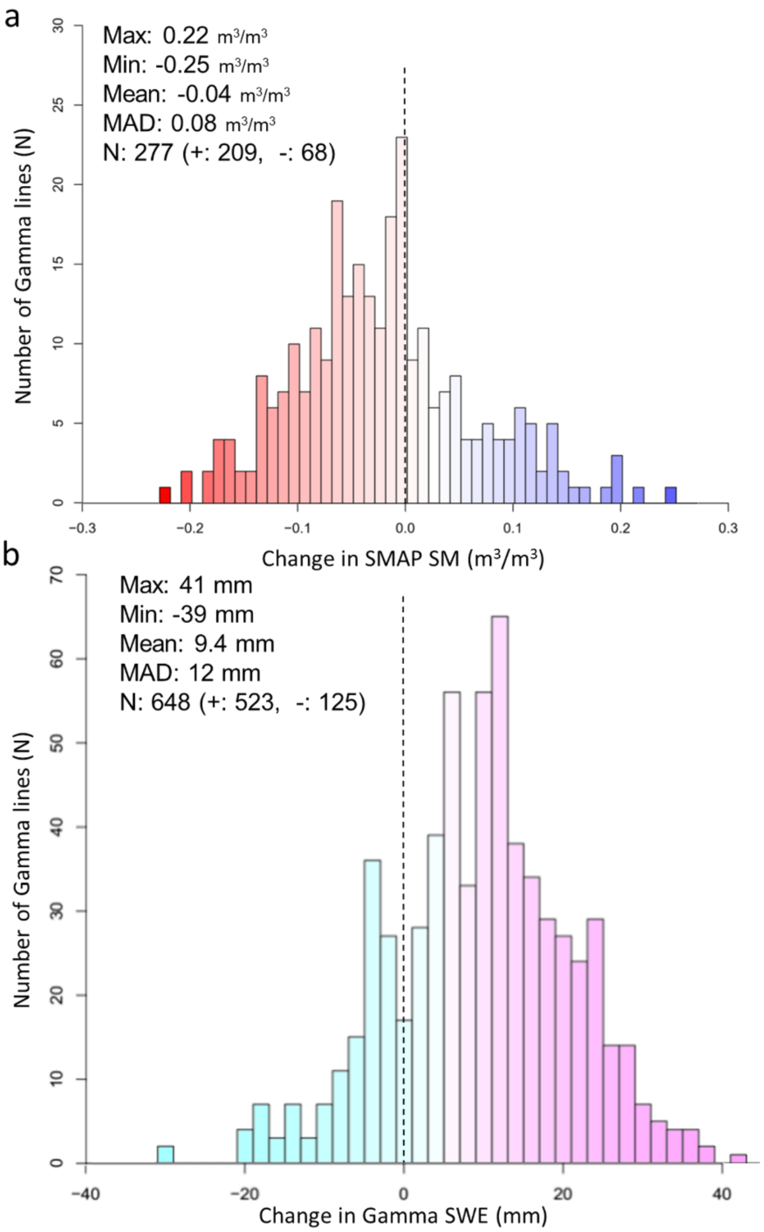


Figure 6. Histogram of (a) changes in SMAP SM and (b) NOAA airborne gamma SWE from the date of the baseline fall gamma flights to the date immediately before winter freeze-up

4.5.4 Evaluation of the updated gamma SWE

To evaluate the SMAP-updated gamma SWE, satellite-based SWE measurements from SSMIS passive microwave sensors were used. Flight lines in forest-dominant regions were excluded because SSMIS underestimates SWE compared to airborne gamma SWE over the forested areas (Figure S7). Figure 7 shows that the SSMIS SWE has better agreement with SMAP-updated gamma SWE than with the operational gamma SWE. When the SSMIS SWE exceeds 125 mm, the SMAP-updated gamma SWE values with high DOY converge toward the 1:1 line. The agreement between the two SWE estimates was improved for each land cover type when gamma SWE was updated with SMAP SM (Figure S8). In grassland, the SSMIS SWE had a higher correlation and lower *ubRMSD* with SMAP-updated gamma SWE as compared to the agreement with the operational SWE. There were also modest improvements in the agreement statistics in croplands, except for *Bias*, which increases from -1.8 to -11 mm.

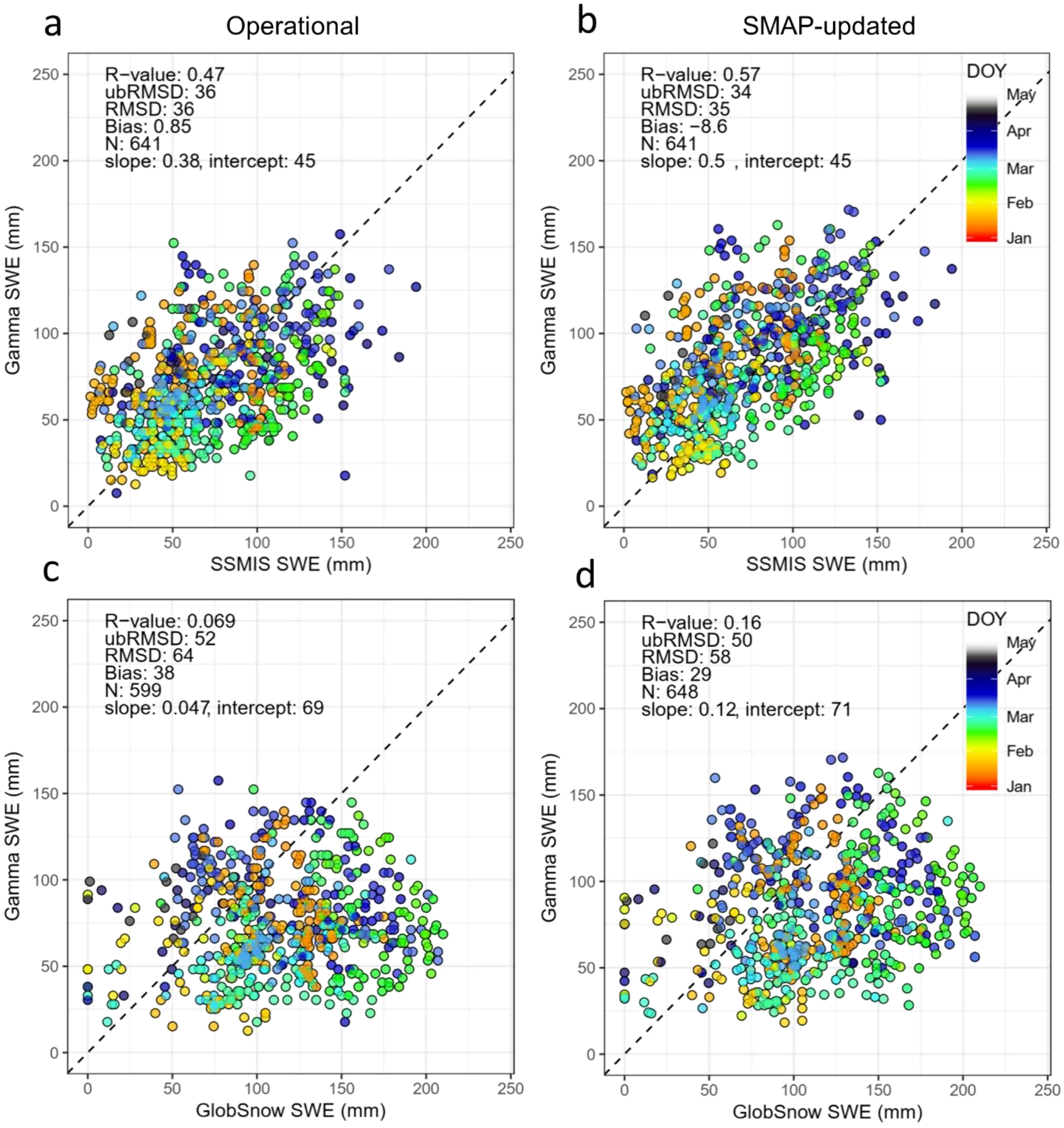


Figure 7. Comparison between operational and SMAP-updated NOAA airborne gamma snow water equivalent with (a, b) satellite-based snow water equivalent from Special Sensor Microwave Imager Sounder (SSMIS) and (c, d) ESA GlobSnow assimilation SWE within the given flight line footprint. The points are colored by day of year (DOY).

To further validate the SMAP-updated gamma SWE, ground-based SWE measurements were obtained from the Glacial Ridge SCAN site snow pillow. Even though there are only five coincident gamma SWE observations, the gamma SWE captures the SWE evolution of the in-situ data for the two years (Figure 8). In 2017, gamma SWE updates were only 3 mm because of the limited changes in the baseline SM. In 2018, the operational gamma SWE values are updated by about 20 mm due to the large decrease in the antecedent SM. The updated gamma SWE shows a higher correlation (R = 0.95 with $p < 0.01$) with in-situ SWE than the operational gamma SWE (R = 0.75 with p = 0.15; Figure 8b). The slope and y-intercept of the updated SWE are also much closer to the 1:1 line. While the operational gamma SWE overestimated SWE by 8 mm in 2017, it underestimated SWE by 12 mm in 2018. The updated gamma SWE biases are consistent for both years.

A final comparison was conducted using the weekly SWE samples from the United States Army Corps of Engineer (USACE) at three sites (Baldhill, ND, Orwell, MN, and Traverse, MN) in the north-central U.S. (see Figure S4). The USACE SWE shows better agreement with the SMAP-updated SWE (R = 0.71 with p = 0.075) than the operational gamma SWE (R = 0.65 with p = 0.12; Figure 9).

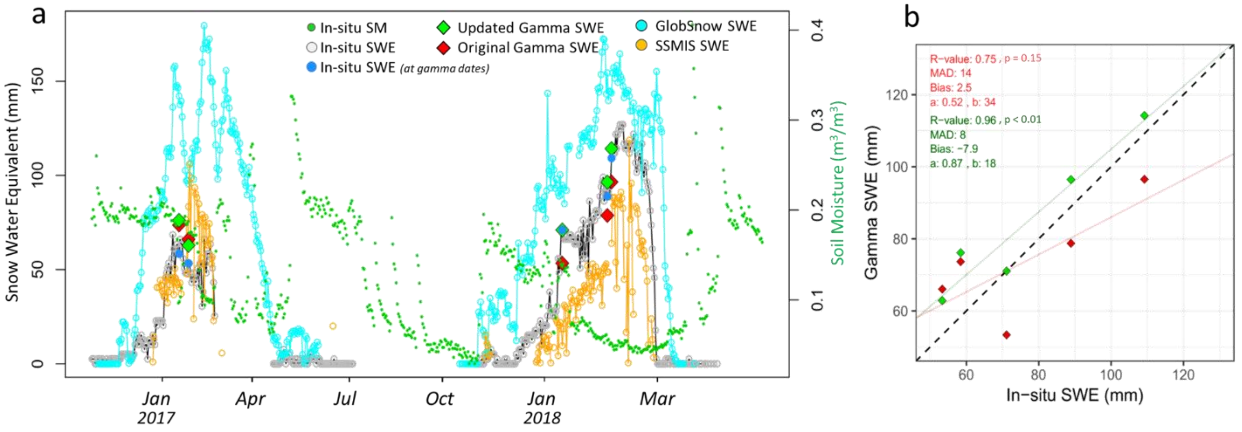


Figure 8. (a) Time series of in-situ SM and SWE measurements with the operational and SMAP-updated gamma SWE at the Glacial Ridge Station, Minnesota (ID: 2050) from the Soil Climate Analysis Network (SCAN) and (b) agreement between the in-situ SWE and the operational and

SMAP-updated gamma SWE. The red points in both plots indicate the operational gamma SWE, while the green points indicate SMAP-updated gamma SWE.

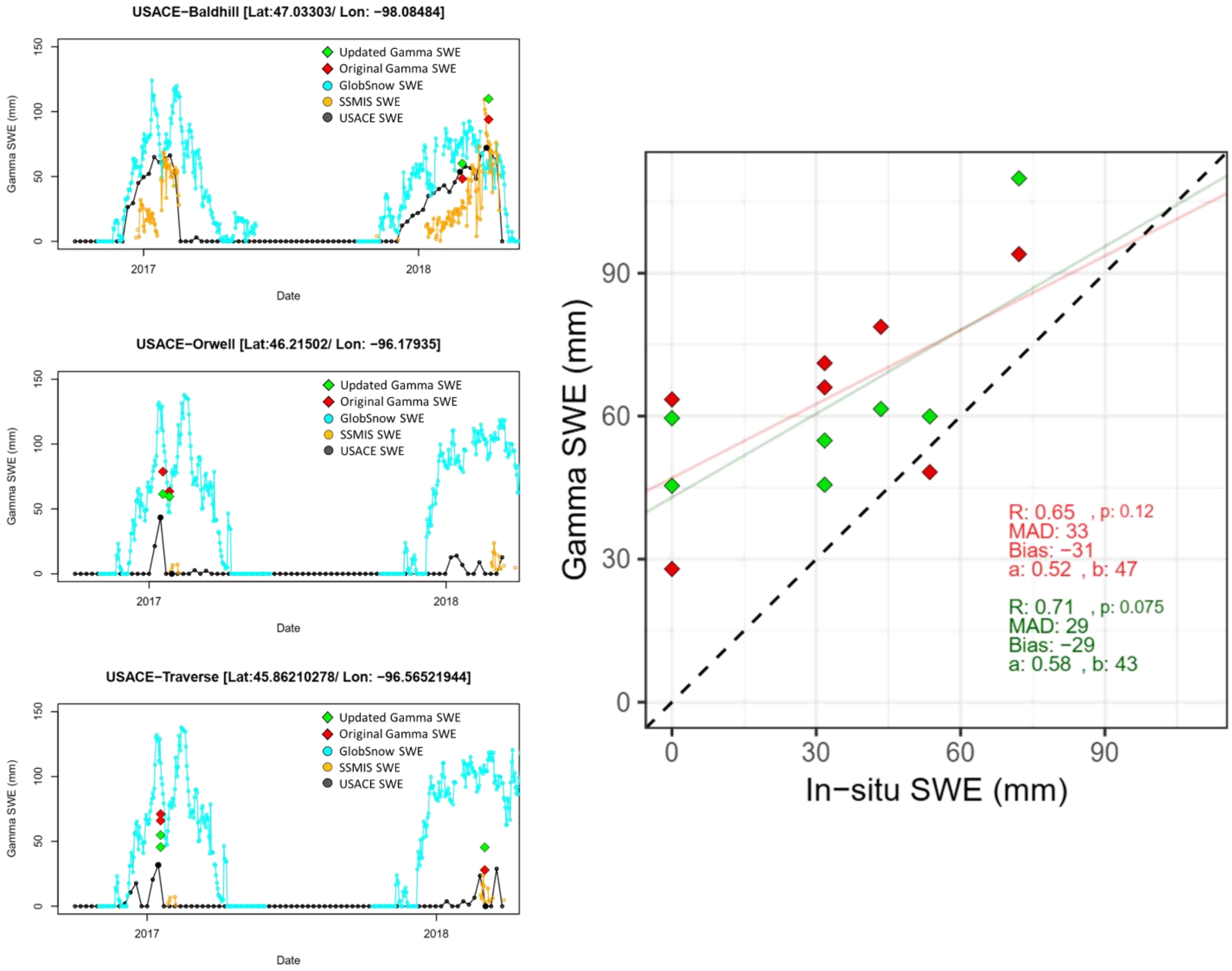


Figure 9. (a) Time series of in-situ SWE measurements with the operational and SMAP-updated gamma SWE at three sites (Baldhill, ND, Orwell, MN, and Traverse, MN) from the United States Army Corps of Engineers (USACE) and (b) agreement between the in-situ USACE SWE and the operational and SMAP-updated gamma SWE. The red points in both plots indicate the operational gamma SWE, while the green points indicate SMAP-updated gamma SWE.

According to antecedent soil moisture conditions and melting processes, a response of peak flood flow to SWE can be varied by annually. A sensitivity analysis of streamflow to the amount of SWE was conducted for recent flood years (2009, 2011, and 2013) in the Buffalo basin, RRB. In 2013, the peak flow can increase by 130% with a 25% (43 mm) increase in annual maximum SWE. Based on the SMAP-based corrected SWE increased up to 41 mm from original SWE (Figure 6), the improved SWE in flood years can result in remarkable improvement in flood prediction.

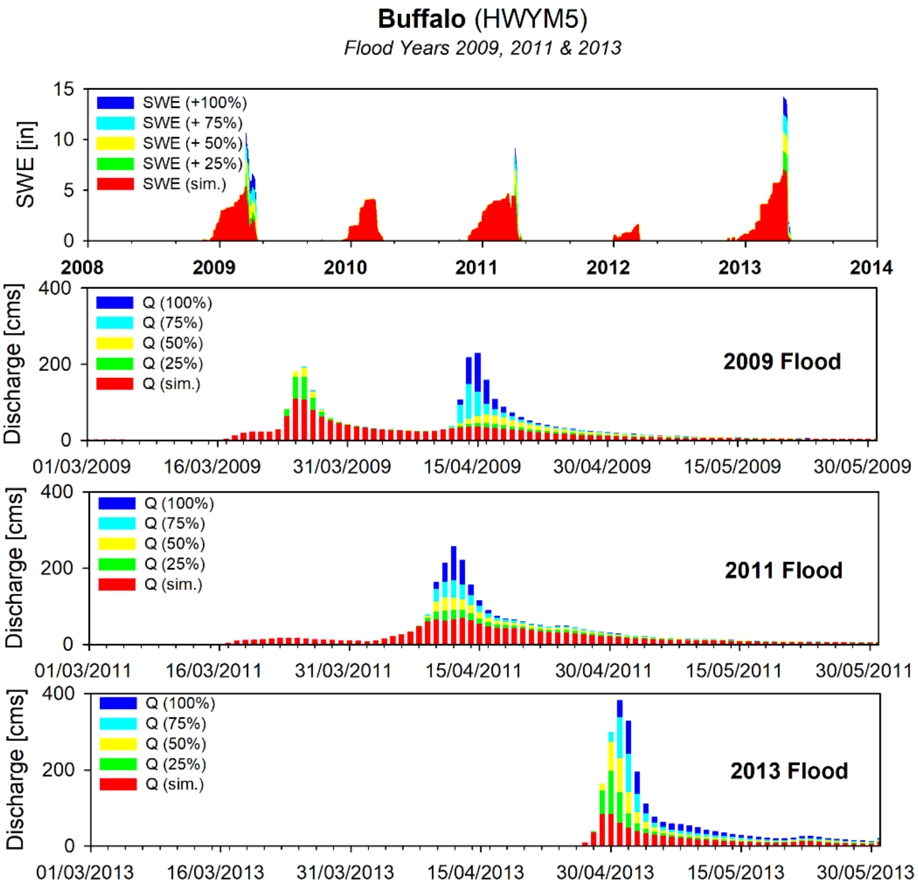


Figure 10. Sensitivity analysis of streamflow to SWE increases in the Buffalo basin, Minnesota in the RRN region for recent flood years (2009, 2011, and 2013) via the NOAA flood forecasting model (Schroeder et al., 2017).

4.6 Discussion

4.6.1 Evaluation of soil moisture

The superior agreement of SMAP products with gamma SM in non-forested areas could be caused by its finer spatial resolution (9 km x 9 km) as compared to AMSR2 (25 km x 25 km) and NLDAS-2 (12.5 km x 12.5 km). Considering that the typical gamma flight line has a 5–7 km^2 footprint, the finer resolution of the SMAP enhanced SM may lead to less spatial heterogeneity error within the pixels (Loew, 2008; Chan et al., 2018). However, Cho et al. (2018) found that the gamma SM also had better agreement with SMAP standard SM (36 km x 36 km; SPL3SMP) than with either the AMSR2 or the NLDAS-2 mosaic SM products. This result is similar to Kim et al.'s (2018) finding that in-situ SM showed better agreement with the SMAP standard SM than with either AMSR2 or Global Land Data Assimilation System (GLDAS) SM products (25 km x 25 km). This suggests that the L-band frequency (1.4 GHz) of the SMAP radiometer might lead to better performance regardless of spatial resolution (Chan et al., 2018). The greater penetration depth of the L-band could be also more representative of the ~20 cm depth of the gamma SM. In dense-forested areas with high vegetation canopy, it is extremely difficult to obtain accurate SM retrievals using the SMAP L-band and AMSR2 X-band frequencies. The AMSR2 X-band SM product over the Deciduous broadleaf forest and Mixed forest regions are typically masked with the data quality flag. In non-forested regions with bare ground or low vegetation canopy, the L-band SM performs better than X-band because the L-band frequency can partly penetrate low vegetation canopy while the higher-frequency X-band experiences greater attenuation (Kim et al., 2018; Jackson & Schmugge, 1991).

In the Deciduous broadleaf forest and Mixed forest classes, the operational gamma SM had a poorer agreement with SMAP SM than with NLDAS-2 SM, which agrees with previous validation studies of passive microwave SM products, including the SMAP radiometer. A known limitation of passive microwave soil moisture retrievals is that dense vegetation canopy over the soil surface reduces the sensitivity of the relationship between emissivity and SM (Jackson & Schmugge, 1991; Wigneron et al., 2003), even though the L-band microwave frequency yields relatively good results under vegetation covers relative to other, higher frequencies because of its higher penetration depth (Vittucci et al., 2016; Entekhabi et al., 2010). Due to the extremely high optical depth of forests, there is little chance of reliably estimating SM conditions. For forest types, Chan et al. (2016) found larger biases and *ubRMSD* between SMAP and in-situ SM measurements at core validation sites (CVS), relative to other land cover types.

While SMAP SM has a wet bias in forest areas, there is no bias between operational gamma SM and NLDAS-2 SM due to land cover. Considering that NLDAS-2 Mosaic SM is estimated based on a physical land surface model (Koster & Suarez, 1996), it is likely that gamma SM is less affected by vegetation effects than passive microwave (SMAP and AMSR2) SM. The airborne gamma radiation technique depends on historical data to establish the relationship between gamma count rates and SM and determine a standardized gamma count rate at 35% gravimetric SM values for each calibration flight line (Carroll, 1980; 2001; Jones and Carroll, 1983). This suggests that the vegetation effect on airborne gamma radiation observations is minimal. Change in vegetation conditions by season are also minor because most gamma SM observations – to estimate antecedent SM prior to soil freezing – are measured in late fall (e.g. October or November) (Carroll, 2001). For these reasons, the gamma SM appears to be reliable in forested regions and has the potential to be used beyond its operational estimates of SWE. However, further investigation is still required to determine how gamma fluxes from the soil are attenuated by vegetation characteristics (e.g. type, height, and density) and how much the attenuation impacts SM estimates (Woods 1965; Schetselaar & Rencz, 1997; Ahl and Bieber, 2010).

Previous studies typically evaluated airborne gamma radiation SM with ground-based SM observations. With an average of 25 samples gravimetric SM measurements per flight line, Carroll (1981) and Jones and Carroll (1983) found airborne gamma SM had strong agreement

(R^2 = 0.87 and 0.84, *RMSD* = 3.2 and 3.9%, respectively). The airborne gamma radiation technique is considered to be a reliable method to estimate areal mean SM measurements.

No previous studies have compared gamma SM observations to satellite-based active and passive microwave or LSM SM, even though there are a series of satellite-based microwave sensors (e.g. SSM/I, AMSR-E/2, ASCAT, SMOS, and SMAP) and numerous evaluation studies since the early 1980s (e.g., Al-Yaari et al., 2014; Babaeian et al., 2019; Mladenova et al., 2014; Kim et al., 2018; Xia et al., 2014). This may be due to the operational mission of the airborne gamma program. As mentioned earlier, the airborne gamma radiation SM data collected by the NOAA NWS's Airborne Gamma Radiation Snow Survey Program is intended primarily to estimate SWE, not SM itself, and to provide the SWE data to the RFCs for use in the snowmelt flood forecasts. In light of the gamma radiation SM performance forests, gamma SM may help estimate SM in forested-dominated regions; one of the current challenges in the SM remote sensing community. As an independent asset, the airborne gamma radiation SM dataset can be utilized to evaluate current and future SM products from various satellites and land surface models to improve hydrological models.

4.6.2 Evaluation of SWE

The SMAP-updated gamma SWE agreement with satellite SWE is better than the previous findings by Tuttle et al. (2018). Tuttle et al. (2018) compared the operational gamma SWE to AMSR-E SWE estimates over the Northern Great Plains from 2002 to 2011. Their correlation coefficient (0.36) and *RMSD* (43 mm) is relatively poor compared to the SMAP-updated gamma SWE results and even the operational SWE. This may be due to different study periods between the two studies (2002-2011 versus 2015-2018). Their statistics could include a few erroneous SWE values during 2009 and 2011 when there were snowmelt floods. The improved agreement of the SMAP-updated SWE with in-situ SWE, satellite microwave SWE, and GlobSnow SWE suggest that a portion of the error in operational gamma SWE caused by antecedent SM can be reduced using this proposed method.

Compared to the operational gamma SWE, the SMAP-updated SWE has better agreement with the limited available datasets including in-situ, satellite-based SSMIS, and GlobSnow assimilated SWE, but positive biases with in-situ and SSMIS SWE (10.4% and 11.8% respectively). Carroll and Schaake Jr (1983) also found that the airborne gamma SWE data tend to overestimate the ground-based data by approximately 10%. This may be due to the airborne gamma radiation method detecting water in all phases, including ground ice, standing water, and superimposed SM in the soil surface (Carroll, 2001), which might not be included in SWE observations from ground samples and snow stations. A snow pillow measures only the mass of the overlaying snowpack (Goodison et al., 1981) and has inherent limitations because the heat exchange between the snow and soil is disrupted, likely causing SWE underestimation (Bland et al., 1997). The current study suggests the method improves gamma SWE estimates but further validation with purposefully designed in-situ SWE measurements is needed.

4.6.3 Limitations

When the linear regression model between operational airborne gamma SM and SMAP SM was developed, the residual errors (ε_i) for each flight line were included in the model, assuming that the errors reflect the physical properties of the land surface within each line footprint (e.g. soil properties, elevation, slope, and inner spatial heterogeneity) (Clark et al., 2011). A residual analysis conducted with land surface characteristics (clay percentage, saturated hydraulic conductivity, and mean elevation and slope) to identify physical properties related to the errors and to assess the appropriateness of the model did not find any statistically significant relationships. Carroll and Carroll (1989b) found that gamma SWE is systematically underestimated when large SWE variability occurs within a flight footprint. Because the gamma technique principles, measuring water mass by attenuation, are the same for SM and SWE, it possible that SM variability could cause gamma SM to be underestimated. High-resolution soil properties and SM-related variables (e.g. land surface temperature / Sentinel-1 SAR backscatter) could be used to understand spatial heterogeneity impacts and to improve the operational gamma SM methodology (Das et al., 2019).

A well-known issue when validating gridded satellite products with in-situ (or different platform) measurements is the difference in spatial scales between the observations and the ability of the sub-grid scale measurements to capture the variability within the satellite footprint (Gruber et al., 2013; Colliander et al., 2017). Tuttle et al. (2018) noted that SWE spatial variability affects the gamma versus satellite SWE comparison because of the different spatial scales for the gamma footprint and the satellite pixel. The different observation scales may contribute to the residual errors in the linear regression model between the gamma and SMAP SM. The gamma SM lines often comprise parts of multiple SMAP pixels. The weighted mean SMAP SM was found for the given flight footprint. However, the weighted mean SM is derived from Tb observations that are partly from outside of the flight line footprint, thus introducing representativeness errors into the linear model. Further studies are required to identify physical characteristics that might be related to the residual errors in the model.

There may be temporal differences between airborne gamma radiation observations and the satellite and model products in this study for SM and SWE comparisons. The gamma flight overpass times range from 9 AM to 6 PM while the sun-synchronous SMAP, AMSR2, and SSMIS sensors have constant local overpass times. Recognizing that SM has diurnal changes (Jackson, 1973), the linear regression model between the operational airborne gamma and SMAP SM could be improved if the measurement time of the gamma flight data were known and the comparison included only those observations where measurement times were similar. It is also possible that this approach would improve if NLDAS-2 SM were used instead of SMAP SM because hourly NLDAS-2 SM values are available (Xia et al., 2015).

The different representative depths among the SM data sources also add error. The passive microwave sensors measure surface SM from the top few centimeters, with a depth that varies with the amount of soil moisture and its distribution (Njoku and Kong, 1977; Escorihuela et al., 2010). The L-band SMAP SM captures approximately the top 5 cm of the soil (O'Neill et al., 2018; McColl et al., 2017) whereas the X-band AMSR2 penetration depth is close to 1 cm (Bindlish et al., 2017) because lower-frequency microwave radiation generally penetrates soil and vegetation canopy more effectively than higher-frequency bands (Jackson & Schmugge, 1991). However, airborne gamma SM is derived from a larger depth range than the penetration depth of any current passive microwave satellite instrument (Carroll, 2001) with 91% of the gamma flux emitted from the upper 10 cm of the soil, and 96% from the upper 20 cm (Zotimov, 1968; Jones and Carroll, 1983). While the different sensors' representative depths are not dissimilar, the modest difference in representative depths could still cause errors, especially during dynamic wetting or drying (e.g., right after rainfall events).

4.7 Conclusion

In this study, a linear regression method was developed to improve operational airborne gamma SWE estimates in non-forested regions by updating the fall baseline SM using the SMAP enhanced SM product. Based on limited comparisons, the SMAP-updated SWE improves agreement with satellite and in-situ SWE observations. This preliminary study identified the need to further test the approach as well as opportunities to improve the approach using higher-resolution/evolving independent products. For example, the Copernicus Sentinel-1 Synthetic Aperture Radar (SAR) provides 1-km C-band backscatter data. Because the SAR backscatter is directly related to surface SM condition, the Sentinel-1 SAR-based information could improve antecedent SM estimates over the gamma flight lines. However, current satellite SM observations offer little value for improving the gamma estimates in forested areas. In densely vegetated regions SM from LSMs, applied using this study's approach, could improve the operational gamma SWE regardless of land cover type. In the United States, snowmelt flood predictions are challenged by limited ground observations and rely heavily on the airborne gamma SWE product which is also used to support the operational SNODAS product. This study shows that the typical SWE corrections are on the order of 10 mm. Thus, the soil moisture corrections would be most important for regions having shallow snowpacks and snowmelt-driven flooding that is highly sensitive to modest SWE differences. Finally, gamma SWE estimates also serve as independent SWE measurements that are useful for evaluating satellite and modeled SWE products. An updated airborne gamma SWE with reduced errors will better support the evaluation of SWE products from current and future satellite missions and regional/global land surface or climate models.

CHAPTER 5

The value of long-term (40 years) airborne gamma radiation SWE record for evaluating three observation-based gridded SWE datasets by seasonal snow and land cover classifications[4]

5.1 Introduction

Snow impacts human activity across the U.S. as a source of water, hydropower and potentially flooding. In the western U.S., snow supplies 70% of the annual water supply valued at more than $348 billion per year (Adams et al., 2004). In the north-central and northeastern U.S., snow meltwater is a dominant driver of severe spring flooding (Wazney et al., 2015; Stadnyk et al., 2016; https://www.weather.gov/dvn/summary_SpringFlooding_2019). Accurate, timely estimates of the snowpack are required over the U.S. to help monitor and manage seasonal snow and melt. The most hydrologically relevant measure of the snowpack is snow water equivalent (SWE), which describes the amount of water stored in the snow.

In addition to real-time snow observations, a long-term record of SWE is important for identifying climate variability and trends and for developing a climatology of the snowpack. As changes in seasonal snow have recently accelerated across the U.S. in the last few decades (Ashfaq et al., 2013; Georgakakos et al., 2014), reliable long-term SWE measurements are further needed for effective water management and flood risk assessments (Zeng et al., 2018).

---

[4] Cho, E., J.M. Jacobs, C. Vuyovich (2020) The value of long-term (40 years) airborne gamma radiation SWE record for evaluating three observation-based gridded SWE datasets by seasonal snow and land cover classifications, *Water Resources Research*, 56(1), https://doi.org/10.1029/2019WR025813

Point-based long term SWE records from snow station networks (e.g. NRCS Snowpack Telemetry (SNOTEL)) have provided high-quality measurements. The time series of SWE observations have been used to evaluate and validate trends and seasonal variabilities in snowpack in numerous previous studies in the western U.S. (Cooper et al., 2016; Mote et al., 2018; Pierce et al., 2008). However, point measurements do not necessarily represent the snowpack distribution especially in areas with spatially heterogeneous terrain (Molotch & Bales, 2005). In order to overcome the limitations, observation-based gridded SWE products have been developed using satellite remote sensing and/or in-situ snow station networks with assimilation techniques.

A series of passive microwave satellite sensors have provided a potential source of spatially distributed SWE information. For more than 30 years, the Special Sensor Microwave Imager (SSM/I) and SSM Imager/Sounder (SSMIS) aboard the Defense Meteorological Satellite Program (DMSP) series of satellites (1987 – current) have provided long term SWE information at a global scale (Derksen et al., 2005; Foster et al., 2005; Tait, 1998). Passive microwave SWE retrieval algorithms typically use empirical relationships between SWE or snow depth and the difference between the brightness temperatures at two different passive microwave frequencies: a low frequency, 18–19 GHz, and a higher frequency, typically around 37 GHz. However, known sources of error hamper the operational use in many regions over the U.S. Carroll et al. (1989) mentioned major difficulties that tend to be inherent in the satellite-based passive microwave SWE products. First, the passive microwave algorithms tend not to work under a deep snowpack (greater than approximately 200 mm SWE), which is called “saturation effect” because the higher frequency microwave signal is no longer detectable (Dong et al., 2005; Vuyovich et al. 2014). Second, the general tendency is for algorithms to not estimate SWE sufficiently well in forested or heavily vegetated regions (Foster et al., 2005; Vander Jagt et al., 2013). Third, even in flat areas with sparse vegetation, the passive microwave signal is highly sensitive to even small amounts of liquid water in the snowpack (Kang et al., 2014; Walker and Goodison, 1993). Consequently, the SWE values are unreliable under wet snow conditions.

To directly assimilate satellite-based passive microwave emission, Pulliainen (2006) developed a Bayesian-based assimilation technique that weighs the passive microwave satellite data and the interpolated ground-based snow depth observations using a semi-empirical radiative transfer model. This assimilation technique was integrated into the GlobSnow project supported by the European Space Agency (ESA) to produce a long-term SWE dataset for the Northern Hemisphere (Takala et al., 2011). Because ground-based snow depth was used to generate interpolated effective grain size data, and simulate Tb and SWE, the accuracy of the SWE is expected to be better than those of typical stand-alone channel differencing algorithms from passive microwave satellite sensors (e.g. Chang et al., 1987; Kelly, 2009). In fact, previous studies found that the GlobSnow SWE has better performance with point ground SWE measurements as compared to the empirical SWE algorithms (Hancock et al., 2013; Mudryk et al., 2015; Larue et al., 2017). However, the GlobSnow-2 SWE still has large uncertainties with RMSE of 94 mm (36%) in wet and deep snow conditions and forested regions, probably due to the inherent sources of error in passive microwave signal (Larue et al., 2017). Hancock et al. (2013) using GlobSnow v1.0 also reported that occasional abrupt changes were found in the product.

The University of Arizona (UA) recently developed a 4-km gridded long-term SWE dataset from 1982 to 2017 (hereafter UA SWE) over the conterminous U.S. (Broxton et al. 2016a) by combining high-quality point SWE measurements with finer spatial resolution gridded precipitation and temperature datasets. The observation-based UA SWE product is produced by interpolating thousands of ground-based measurements of SWE and snow depth from SNOTEL (Serreze et al., 1999) and the NWS Cooperative Observer Program (COOP) network sites and 4-km gridded PRISM precipitation and temperature data (Daly et al., 2008). The quality and reliability of the UA SWE data have been demonstrated compared to reanalysis and land assimilation products (Broxton et al., 2016b) and independent observation-based products (Dawson et al., 2018). Dawson et al. (2018) reported that the passive microwave AMSR-E SWE product and two satellite-merged SWE products (GlobSnow-2 and Canadian Sea Ice and Snow Evolution Network, CanSISE) have large differences from the UA SWE data (Mean Absolute Difference, MAD: 46% to 59%), especially in forested regions. They also found that there was a good agreement in basin-averaged SWE between UA and Airborne Snow Observatory (ASO) products (32 flight measurements) (Painter et al., 2016) in the upper Tuolumne basin, California’s Sierra Nevada (correlation: 0.98 and MAD: 51.5 mm [30%]). However, since the test with ASO SWE had a limited number of flights over a sparsely forested region (< 30% of tree cover fraction; Dawson et al., 2018), there is still a need to further evaluate the accuracy of the UA SWE over more heavily vegetated regions.

Due to the lack of reliable independent SWE records, an evaluation of the currently available long-term SWE products at a continental scale has been limited. The airborne gamma snow survey operated by the National Oceanic and Atmospheric Administration’s (NOAA) National Operational Hydrologic Remote Sensing Center (NOHRSC) has substantial potential to evaluate the gridded SWE products over the conterminous U.S. The NOAA airborne gamma snow survey was designed to help hydrologists and flood forecasters in the National Weather Service offices, regional river forecasting centers, and other U.S. and Canadian federal agencies improve operational spring flood predictions and water supply outlooks (Peck et al., 1980). Since 1979, the snow survey has collected areal mean SWE data over a network of 2,400 flight lines covering 25 states and 7 Canadian provinces (Carroll, 2001). The airborne gamma technique uses the attenuation of the gamma-ray signal by water in the snowpack (any phase) to measure SWE for each flight line. The mean areal gamma SWE value is based on the difference between gamma radiation measurements over bare ground and snow-covered ground.

The accuracy of airborne gamma SWE estimates has been evaluated using numerous ground-based snow observations from snow courses and field campaigns on designated flight lines (Carroll & Schaake, 1983; Carroll & Vose, 1984; Goodison et al., 1984; Glynn et al., 1988; Peck et al., 1971). In prairie regions, with mean ground-based SWE between 20 – 150 mm, the root mean square error (RMSE) of airborne gamma SWE is less than 10 mm (4 – 10%) based on a few hundred samples within the flight lines (Carroll & Schaake, 1983; Carroll., 1983). In densely forested regions, Carroll & Vose (1984) found that the airborne gamma SWE had a low bias and an RMSE of 23 mm as compared to the mean ground-based SWE. In that study of the Lake Superior and Saint John basins in the U.S, SWE measurements ranged from 20 to 480 mm based on approximately 200 snow depth and 20 snow density measurements distributed along the length of each flight line (total 72 lines). These studies provided the impetus to develop an airborne gamma SWE program, which has been successfully used for operational flood forecasting during the last 40 years (Carroll, 2001). Currently, the airborne gamma SWE observations, as well as ground-based and satellite snow covered area observations over the U.S., support the NOHRSC near-real-time, high spatial resolution (1 km^2 gridded) Snow Data Assimilation System (SNODAS) products (Barrett, 2003).

An evaluation of the currently available long-term SWE products is important to differentiate among environments having a strong agreement or large differences among products. The 40-year airborne gamma SWE record is ideal to evaluate the gridded observation-based SWE products according to seasonal snow and land cover characteristics. Even though the gamma SWE record has limited spatial and temporal coverages compared to the coarser satellite and assimilation products, the record provides accurate and reliable SWE in a wide range of snow and land characteristics, especially in forested regions where SWE estimation remains challenging (Cho et al., 2018). The wide range encompasses a broader sample of snow conditions over a longer time period than is collected using any other aerial observations (e.g. airborne lidar) and provides spatially-integrated observations over a footprint closer to satellite resolution than point observations.

This study focuses on providing a comprehensive examination of three currently available observation-based gridded SWE datasets, spaceborne passive microwave SSM/I-SSMIS (hereafter SSMI/S) SWE, GlobSnow-2 SWE, and UA SWE, using the airborne gamma SWE record from 1982 to 2017 over the conterminous U.S. We hypothesize that SSMI/S, using a standard Chang-type microwave retrieval algorithm, and GlobSnow-2 products will show similar patterns as compared to the airborne gamma SWE estimates, i.e. good agreement in regions with little vegetation and relatively homogeneous terrain (e.g. cropland and grassland land cover types and “prairie” of seasonal snow class) and lower agreement in regions with dense vegetation and heterogeneous terrain (e.g. forested-type land cover types and “warm forest” and “maritime” of seasonal snow classes). We also hypothesize that UA SWE will compare favorably to the gamma SWE even in forested or heavily vegetated regions due to the lower impact of forest fraction on both products (Cho et al., 2018; Dawson et al., 2018). Finally, we hypothesize that known gamma weaknesses under heterogeneous conditions will be revealed as larger differences between gamma SWE and the gridded products. We evaluate the three long-term daily SWE estimates against the historical airborne gamma SWE record by land cover type, seasonal snow cover classification, and the degree of the tree fraction and land heterogeneity across the U.S.

This paper is organized as follows. Section 2 describes the study area with land cover types, seasonal snow classes, and tree fractions for a categorized evaluation and sensitivity test of the SWE products. Section 3 describes the three gridded SWE datasets and airborne gamma SWE. The methodologies, which include a resampling method and the calculation of agreement statistics, are described in section 4. Section 5 details the results of spatial comparison of three products (Section 5.1), their differences by seasonal snow and land cover classes (Section 5.2), and the effect of tree fraction and topographical characteristics (Section 5.3). Section 6 offers discussions about the similarities, differences, and new findings in our results with respect to previous studies, and potential limitations of gamma SWE estimates. Conclusion and future perspectives are drawn in section 7.

5.2 Study Area

The study area comprises the conterminous U.S. where all SWE data used in this study are available (Figure 1), including parts of eight NOAA river forecasting centers (RFCs), North-Central (NC), Missouri Basin (MB), Ohio (OH), North-East (NE), Mid-Atlantic (MA), Colorado Basin RFC (CB), Northwest (NW), and California Nevada (CN). The RFC boundaries were designated by the NOAA NWS Advanced Hydrological Prediction Service to manage regional river flow and support flood forecasting over the U.S. The airborne gamma radiation survey provides fall soil moisture and winter SWE measurements to each regional RFC. The NOAA NOHRSC gamma radiation survey network is comprised of over 2,400 flight lines over the U.S. including Alaska and southern Canada. Among them, 1,812 gamma flight lines over the conterminous U.S. were used in this study. The study region is dominated by seven types of the International Geosphere-Biosphere Programme (IGBP) land cover type, evergreen needleleaf forest, deciduous broadleaf forest, mixed forest, croplands, cropland/natural vegetation mosaic, grasslands, and woody savannas (Channan et al., 2014; Figure 1a). The study area is also classified by six seasonal snow classes, tundra, taiga, maritime, ephemeral, prairie, and warm forest by the Sturm's seasonal snow cover classification (Sturm et al., 2010; Figure 1b). Due to a very limited number of gamma SWE observations, the woody savannas land cover type and the ephemeral snow classes were excluded in this study. The annual VCF from the NASA Making Earth System Data Records for Use in Research Environments (MEaSUREs) was used to estimate fractional tree cover at each gamma line (Hansen & Song, 2018; Figure 1c).

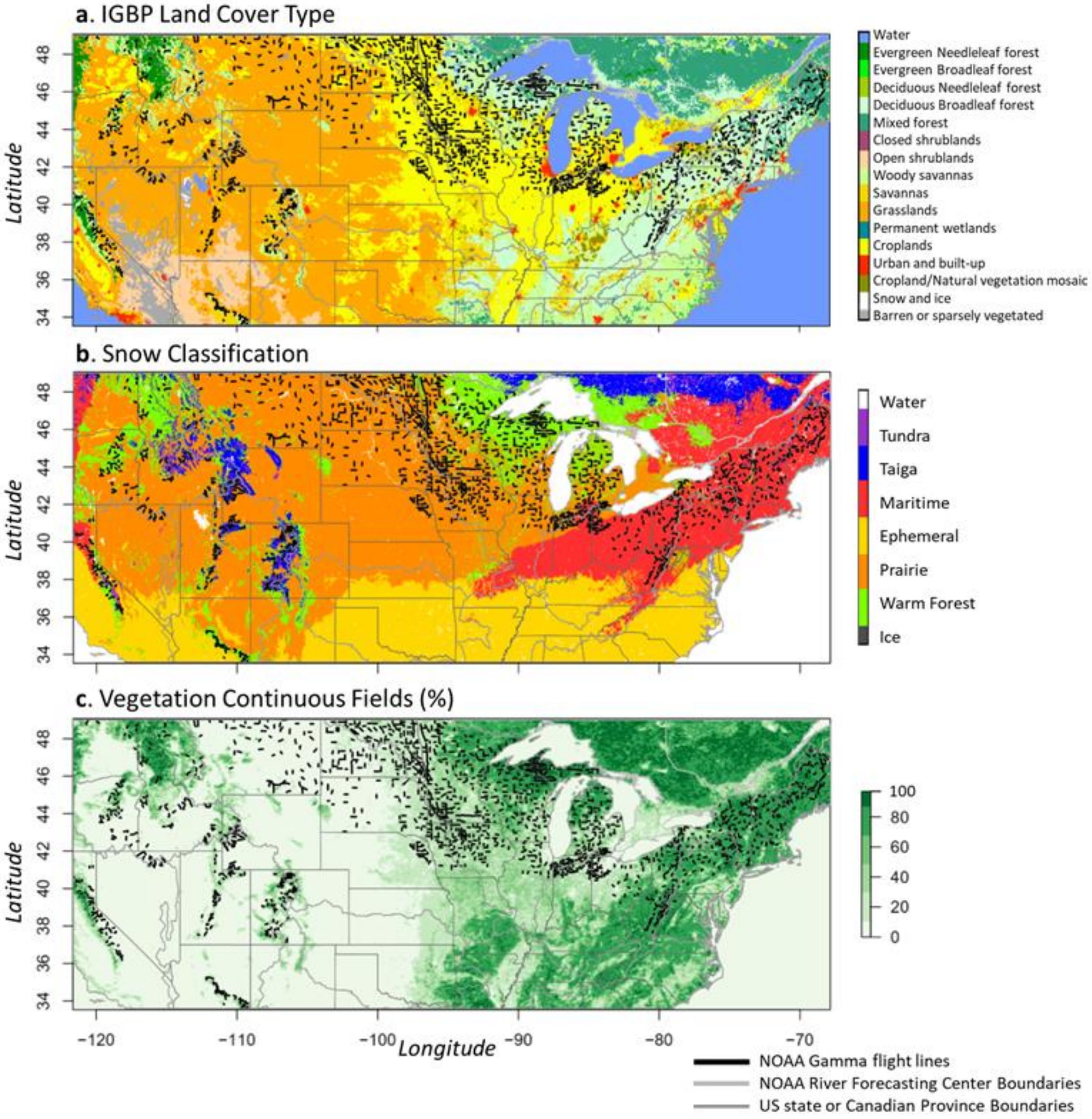


Figure 1. (a) IGBP land cover type, (b) Sturm's seasonal snow classification, and (c) Vegetation Continuous Field maps of the study area over the conterminous United States with NOAA gamma flight lines (N =1,812)

5.3 Data

5.3.1 Airborne gamma radiation SWE

The airborne SWE estimates are obtained using a gamma radiation detector onboard a low-flying aircraft (at an altitude of 150 m above the ground). This instrument measures the natural terrestrial gamma radiation emitted from trace elements of ^{40}K, ^{238}U, and ^{232}Th radioisotopes in the upper 20 cm of soil. In the airborne gamma technique, the attenuation of the gamma-ray signal by water mass in the snowpack (any phase) over a flight line is used to estimate SWE directly. The airborne gamma SWE value is estimated using the difference between the rates of gamma radiation particles over bare ground (attenuation by soil moisture only) and snow-covered ground (attenuation by soil moisture and snowpack).

The operational gamma SWE data are calculated using the following equations:

$$SWE(^{40}K) = \frac{1}{A} \cdot \left[\ln\left(\frac{^{40}K_b}{^{40}K_s}\right) - \ln\left(\frac{100+1.11 \cdot \mathrm{SM}(^{40}K_s)}{100+1.11 \cdot \mathrm{SM}(^{40}K_b)}\right)\right] \qquad \text{Eq. (1)}$$

$$SWE(^{208}Tl) = \frac{1}{A} \cdot \left[\ln\left(\frac{^{208}Tl_b}{^{208}Tl_s}\right) - \ln\left(\frac{100+1.11 \cdot SM(^{208}Tl_s)}{100+1.11 \cdot SM(^{208}Tl_b)}\right)\right] \qquad \text{Eq. (2)}$$

$$SWE(GC) = \frac{1}{A} \cdot \left[\ln\left(\frac{GC_b}{GC_s}\right) - \ln\left(\frac{100+1.11 \cdot SM(GC_s)}{100+1.11 \cdot SM(GC_b)}\right)\right] \qquad \text{Eq. (3)}$$

$$Gamma\ SWE = \ 0.346 \cdot SWE(^{40}K) + 0.518 \cdot SWE(^{208}Tl) + 0.136 \cdot SWE(GC) \qquad \text{Eq. (4)}$$

where ^{40}K, ^{208}Tl, and the total count windows (GC) are uncollided gamma count rates over bare ground ($^{40}K_b$, $^{208}Tl_b$, and GC_b) and snow-covered ground ($^{40}K_s$, $^{208}Tl_s$, and GC_s). SM($^{40}K_b$), SM($^{208}Tl_b$), and SM(GC_0) are gravimetric soil moisture contents over bare ground and SM($^{40}K_s$), SM($^{208}Tl_s$), and SM(GC_s) are gravimetric soil moisture contents over snow-cover ground detected by ^{40}K, ^{208}Tl, and the total count windows (GC), respectively. A is a radiation attenuation coefficient in water, 0.1482. $Gamma\ SWE$ is a single average gamma radiation SWE (g/cm^2) value for the entire flight line reported via the NOHRSC website (http://www.nohrsc.noaa.gov/snowsurvey/).

Since 1979, the operational NWS gamma radiation snow survey has provided about 27,000 gamma SWE measurements over the entire U.S. and southern Canadian provinces via the NOHRSC website (http://www.nohrsc.noaa.gov/snowsurvey/). A typical flight line covers approximately 5 km^2 with a swath 300 m wide and 16 km long. The gamma SWE observations are areal-average values for each flight line footprint. In this study, 20,738 airborne gamma SWE observations covering 1,812 flight lines flown from January 1982 to May 2017 within the conterminous U.S. are used (**Table 1**). Flight lines in which the majority type/class does not exceed 50% of the gamma footprint were considered "*unclassified*" and were excluded from each analysis by the land cover type and snow classification.

Table 1. Overview of the number of the NOAA airborne gamma radiation flight lines and SWE observations by land cover types and snow classes (Note: The gamma SWE values in the *woody savannas*, *ephemeral*, and *unclassified* are excluded in this study).

Land cover types	**Flight lines**	**Total**	**1982-1989**	**1990-1999**	**2000-2009**	**2010-2017**	**Snow cover type**	**Flight lines**	**Total**	**1982-1989**	**1990-1999**	**2000-2009**	**2010-2017**
Evergreen Needleleaf forest	192	1,735	283	716	549	187	*Tundra*	99	1,346	94	631	523	98
Deciduous Broadleaf forest	165	2,609	143	516	1,290	660	*Taiga*	72	925	62	392	391	80
Mixed forest	220	3,961	470	894	1,342	1,255	*Maritime*	391	5,487	485	1,299	2,365	1,338
Grasslands	368	2,640	267	1,002	1,122	249	*Prairie*	1,079	10,704	2,486	3,302	2,222	2,694
Croplands	610	7,140	1,931	2,174	1,038	1,997	*Warm forest*	69	644	159	156	157	172
Cropland/ Natural veg.	*245*	2,582	407	873	823	479	*Ephemeral*	3	5	1	4	0	0

Woody savannas	12	71	48	21	0	2	*unclassified*	99	1,627	262	412	506	447
Total	1,812	20,738	3,549	6,169	6,164	4,829	***Total***	1,812	20,738	3,549	6,169	6,164	4,829

5.3.2 Spaceborne Passive Microwave SSM/I and SSMIS SWE

The series of SSM/I and SSMIS sensors aboard the DMSP series of satellites operated by the U.S. Department of Defense provides daily brightness temperatures at approximately 6 A.M. (descending) / P.M. (ascending) local time with global coverage from July 1987 to the present. The microwave frequencies at 19, 22, 37 (SSM/I and SSMIS), 85 (SSM/I only), and 91 GHz (SSMIS only) are sampled in both horizontal and vertical polarizations, except the 22 GHz (vertical only). In this study, SSMI/S SWE was estimated using a Chang-type algorithm (Chang et al., 1987) with original coefficients for F8-F13 (Armstrong and Brodzik, 2001) and modified coefficients for F17 developed by Brodzik (2014) as follows:

$$SWE_{SSM/I} = 4.77 \cdot (Tb_{H,19GHz} - Tb_{H,37GHz} - 5) \qquad \text{for F8, 11, \& 13 SSM/I} \qquad \text{Eq. (5)}$$

$$SWE_{SSMIS} = a \cdot Tb_{H,19GHz} - b \cdot Tb_{H,37GHz} - c \qquad \text{for F17 SSMIS} \qquad \text{Eq. (6)}$$

where a, b, and c are given as 4.807 mm/K, 4.792 mm/K, and 6.036 mm, respectively. $Tb_{H,19GHz}$ and $Tb_{H,37GHz}$ are brightness temperatures at 19 and 37 GHz horizontal polarization, respectively. F8, 11, 13, and 17 are the DMSP platform ID. The Tb at the descending overpass (6 A.M.) was used to minimize error by wet snow (Derksen et al., 2000). The DMSP SSM/I-SSMIS Pathfinder Daily EASE-Grid Brightness Temperatures (Version 2) from July 1987 to May 2017 were used in this study, and freely available at the National Snow & Ice Data Center website (https://nsidc.org/data/NSIDC-0032; Armstrong et al., 1994).

5.3.3 GlobSnow-2 SWE

The GlobSnow project, funded by the European Space Agency (ESA), provides gridded daily SWE maps with 25 km spatial resolution from 1979 to 2016 (GlobSnow-2, archive_v2.0) for the Northern Hemisphere, except for glaciers and mountainous regions. The GlobSnow SWE utilizes an observation-based data assimilation approach combining ground-based synoptic snow depth station data with passive microwave satellite measurements (Takala et al., 2011). Ground-based point measurements of snow depth are taken from the European Centre for Medium-Range Weather Forecasts (ECWMF) WMO weather stations. The final product uses daily Tb at 19 and 37 GHz in vertical polarization from the series of passive microwave radiometers (SMMR 1979–1987, SSM/I 1987–2009, and SSMIS 2010–2014) (GlobSnow-2 Final Report; Luojus et al., 2014).

The basis of the GlobSnow SWE processing system is presented by Pulliainen (2006) and Takala et al. (2011), though a brief description of the processes for SWE retrievals is given here. Snow depth (SD) maps are produced using ordinary kriging interpolation technique for synoptic weather station SD observations. The gridded SD is used as input to simulated Tb using the HUT snow emission model. The model describes Tb as a function of a single-layer snowpack (depth, snow density, and effective grain size) and forest canopy. The model is fit to satellite observed Tb values by optimizing effective snow grain sizes with constant snow density (0.24 kg/m^2) at the locations where the weather station snow depth values are available. A map of spatially continuous effective snow grain size is developed by the weather station-based snow grain size estimates using a kriging interpolation technique. The SD and effective snow grain size maps are used to initiate the HUT model as inputs and generate gridded Tb simulations. The Tb simulations are then assimilated with satellite Tb observations by using adaptive weights on the satellite-observed Tb according to their variances. A final SWE is estimated using the assimilated Tb with the maps of the effective grain size and land cover information (Takala et al., 2011).

Compared to the previous v1.0 and v1.3 SWE products, there are enhancements to the GlobSnow-2 SWE (archive_v2.0) product including the improved quantification of data uncertainty characteristics, homogenization of the multiple-year snow depth measurements from synoptic weather stations, and re-processing of the long-term SWE datasets. While the retrieval accuracy of the GlobSnow-2 SWE has remained the same for most regions, it showed slight improvements for northern boreal forest and tundra regions in Canadian reference data (Pulliainen et al., 2014). It should be noted that the inter-sensor systematic bias is not corrected in this product (Takala et al., 2011) whereas previous studies have shown inter-sensor biases in the Tb between the SMMR, SSM/I, and SSMI/S sensors (Derksen et al., 2003; Royer & Poirier, 2010; André et al., 2015; Cho et al., 2017). In this study, the daily GlobSnow-2 SWE was used from January 1982 to December 2016 which was obtained from http://www.globsnow.info/swe/archive_v2.0/.

5.3.4 UA SWE

The UA SWE is an observation-based 4-km gridded SWE dataset recently developed by combining the SNOTEL SWE and NWS COOP snow depth measurements with the gridded PRISM precipitation and temperature data over the conterminous U.S. (Zeng et al., 2018). A new interpolation technique was used to produce the gridded SWE dataset based on the ratio of observed SWE over net accumulated snowfall (accumulated snowfall minus cumulative snow ablation), rather than SWE itself (Broxton et al., 2016a). For this, daily precipitation data was divided into daily snowfall vs. rainfall using a daily air temperature (2-m) threshold and then accumulated snow was calculated as the sum of the daily snowfalls. Daily cumulative snow ablation at each station (or grid cell) is taken from a relationship between the SNOTEL-based cumulative snow ablation and cumulative degree days above 0°C established by the entire SNOTEL network (Figure 2b in Broxton et al., 2016a). The UA SWE product also uses a newly developed snow density parameterization, which was called "SNODEN" (Dawson et al., 2017). This parameterization attempts to include physical processes (e.g. temperature-based aging, overburden snowpack, liquid water from snowmelt) based on the SNOTEL SWE and air temperature at 2 m by seasonal snow cover classes. The parameterization was used to convert COOP snow depth data into SWE estimates and compute gridded snow depth from the gridded SWE data. In this study, the daily UA SWE is used from January 1982 to December 2017. The product is available from the NASA National Snow and Ice Data Center (NSIDC) website (https://nsidc.org/data/nsidc-0719).

5.3.5 Land cover type, snow classification, tree cover fraction, and topographic heterogeneity

In this study, the IGBP land cover type, the Sturm et al. snow classification, and VCF data were used to evaluate the long-term SWE estimates with the airborne gamma SWE considering land cover and snow characteristics as well as the tree cover fraction across the U.S. The Terra and Aqua combined Moderate Resolution Imaging Spectroradiometer (MODIS) land cover data (MCD12Q1; Version 6) provides global land cover types at yearly intervals. Among the six different classification schemes in the MCD12Q1 data, the IGBP land cover classification (Type 1) provides 17 classes to meet the needs of the IGBP core science projects (Loveland & Belward, 1997; Channan et al., 2014). The criteria were used for dividing the classes (e.g. leaf longevity [evergreen vs. deciduous] and leaf type [broad vs. needle]). The Sturm et al. seasonal snow cover classification is defined by a unique ensemble of stratigraphic and textural characteristics of snow covers (Sturm et al., 1995; 2010). The snow classification is primarily grouped by the effects of climate (temperature, precipitation, and wind) on the snow cover properties (e.g. snow textures, layers, and lateral variability).

The annual NASA MEaSUREs VCF product (VCF5KYR; Version 1) provides global fractional vegetation cover including three layers, percent tree cover, percent non-tree vegetation, and percent bare ground, at 0.05 degree spatial grid from 1982 to 2016 (Hansen and Song, 2018). In this study, the percent tree cover is used to conduct a sensitivity analysis of tree cover on the gridded SWE products. Because the fractional tree cover could have changed during the last 40 years, the annual VCF value was obtained for each gamma SWE year. The elevation data (0.0083-degree grid) used in this analysis were aggregated from the Shuttle Radar Topography Mission (SRTM) 90 m resolution elevation data. The slope and elevation range maps (0.0083-degree grid; approximately 1 x 1 km grid) were obtained using the “terrain” function in the “raster” R-package, which is computed using the elevation data according to Wilson et al. (2007). The elevation range value is defined as the difference between the maximum and the minimum aggregated elevation value of a cell and its eight surrounding cells. Three topographic values (slope, elevation range, and elevation) are computed for each gamma footprint by areal-weighted average.

5.4 Methodology

The gamma SWE observations were compared to the three gridded 4 km or 25 km SWE products. The original 4 km UA SWE product was used without upscaling to 25 km grid. The gamma flight lines frequently overlapped more than one SWE pixel. The area-weighted average SWE within the effective footprint of each gamma flight line was calculated. For each flight line, an effective polygonal measurement footprint was determined using a buffer function “gBuffer” in the “rgeos” R package around the given flight line, with a fixed diameter of 330 m (Carroll, 2001; Tuttle et al., 2018). On any day when gamma SWE measurements were collected, the gridded SWE data on the same day within the given gamma line footprint were averaged by weighing according to the portion of the footprint contained within each SWE pixel. If the gridded SWE data were not available for a portion of the flight line footprint, the weighted SWE value was calculated using the area of the footprint with available data. However, gamma flight lines with gridded SWE data covering less than 50% of the area of flight footprint were excluded from the analysis. For further details, refer to Figure 2 in Tuttle et al. (2018). The same method was applied to the tree cover fraction and three topographic characteristic data sets (slope, elevation range, and elevation). The seasonal snow cover and land cover classes for each line were selected using the “majority” function, instead of areal-weighted mean values. The classes covering more than 50% of the footprint area were used to classify the footprints in this analysis.

For evaluation, the agreements of the three gridded SWE products with airborne gamma radiation SWE were quantified by the Pearson’s linear correlation coefficient, *R*, mean absolute difference, *MAD*, the proportion of *MAD* to the mean gamma SWE ($\overline{SWE_{gam}}$), *MAD%*, and the root mean square difference, *RMSD*, as follows:

$$R = \frac{\mathrm{cov}(SWE_{grid}, SWE_{gam})}{std(SWE_{grid}) \cdot std(SWE_{gam})} \quad \text{Eq. (7)}$$

$$MAD = \frac{1}{n}\sum_{t=1}^{n}\left|SWE_{grid,i} - SWE_{gam,i}\right| \quad \text{Eq. (8)}$$

$$MAD\% = \frac{(\frac{1}{n}\sum_{t=1}^{n}|SWE_{grid,i} - SWE_{gam,i}|)}{\overline{SWE_{gam}}} \times 100 \qquad \text{Eq. (9)}$$

$$RMSD = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(SWE_{grid,i} - SWE_{gam,i})^2} \qquad \text{Eq. (10)}$$

SWE_{grid} and SWE_{gam} refer to one of the three gridded SWE and the airborne gamma SWE values. cov(•) and std(•) refer to the covariance and standard deviation values of the given product, respectively.

5.5 Results

5.5.1 Comparison of three SWE products with airborne gamma SWE

Agreement statistics (R-value and MAD, %) were calculated for each gamma flight line having 6 or more pairs of SWE values (Figures 2 & 3). For an example flight line, WY122, in Wyoming, the 32 pairs of historical gamma SWE and corresponding UA SWE have an R-value = 0.88 (see Figure S3). SSMI/S and gamma SWE generally have a poor correlation with some regional differences (Figure 2a). In the north-central U.S., the agreement is better than other regions (for example, mean correlation and MAD of North Dakota are 0.52 and 38%, respectively). However, there are poorer agreements over northern Michigan and northeastern U.S. In the western U.S., the correlation is also extremely low, particularly in Colorado. The MAD spatial patterns are more readily seen (Figure 3a). MAD values exceed 75% near Lake Michigan and from Pennsylvania to Maine and are extremely high in the western U.S.

The GlobSnow-2 generally agrees better with the gamma SWE than the SSMI/S result in western Minnesota, Iowa, Wisconsin, Michigan, and the northeastern U.S. where there are densely forested areas (Figures 2b and 3b). However, in some regions, the agreement between

GlobSnow-2 SWE and gamma SWE is weaker than that of SSMI/S SWE. For example, the GlobSnow-2 SWE has lower R-values (Mean: 0.36) and higher MADs (Mean: 43%) than SSMI/S SWE in North Dakota. While there are a limited number of comparisons in the western U.S., because GlobSnow-2 SWE is masked in mountainous areas (Takala et al., 2011), there is no marked difference between GlobSnow-2 and SSMI/S.

UA SWE strongly agrees with gamma SWE across all regions (Figures 2c & 3c). In the north-central U.S., there are very high correlation coefficients (mean: 0.78) with small MADs (24 mm, 31%). In the northeastern U.S., UA SWE (mean R-value: 0.71 and MAD: 35mm, 44%) also has better agreement with gamma SWE than the SSMI/S and GlobSnow-2 SWE (mean R-value: 0.23 and 0.65, and MAD: 93 mm, 83% and 54 mm, 65%, respectively). While the two satellite-based SWE products had very poor agreement with gamma SWE in the western U.S., the UA SWE agrees better, particularly in the Rocky Mountain regions including Wyoming, Arizona, Colorado, and Utah (mean R-value: 0.07, 0.29, and 0.70 and MAD: 127, 121, and 71 mm (80, 78, and 59%) for SSMI/S, GlobSnow-2, and UA SWE, respectively).

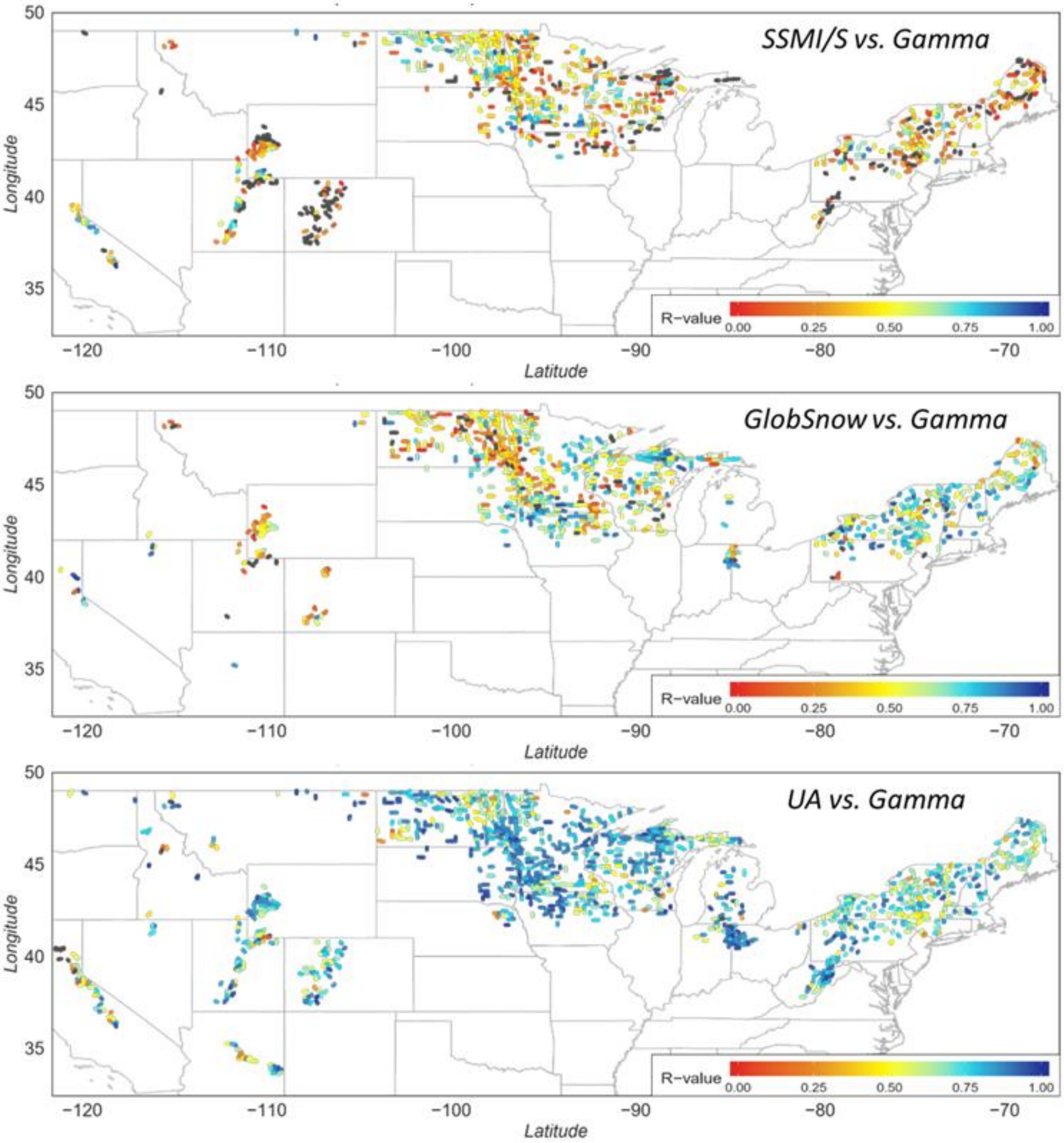


Figure 2. Correlation (R-value) maps of between daily SSMI/S, GlobSnow-2, and UA snow water equivalent with daily NOAA airborne gamma radiation snow water equivalent for each gamma flight line from 1982 to 2017 (Black color represents that the R-value is a negative value).

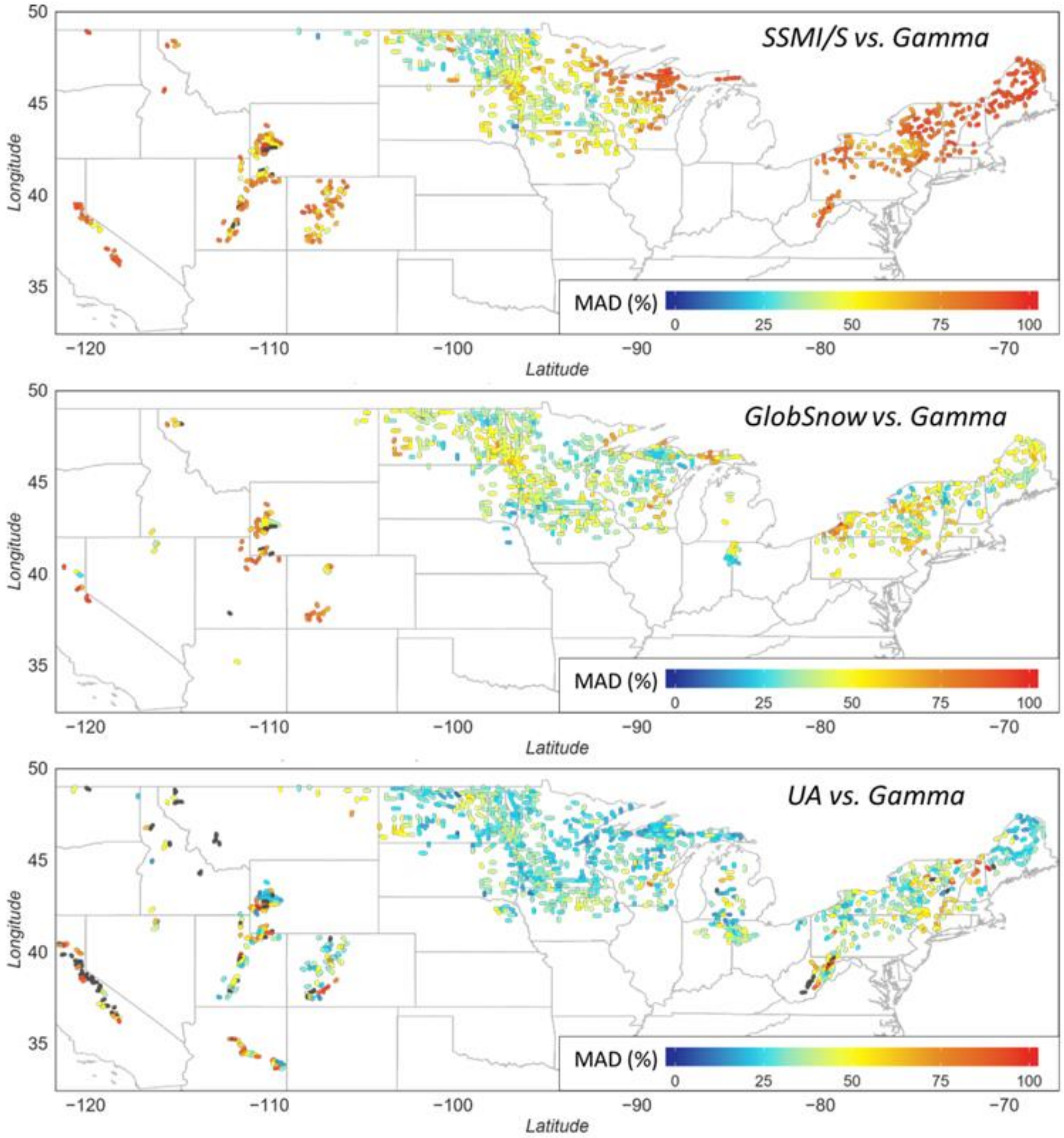


Figure 3. Mean absolute difference (MAD) maps of between daily SSMI/S, GlobSnow-2, and UA snow water equivalent with daily NOAA airborne gamma radiation snow water equivalent for each gamma flight line from 1982 to 2017 (Black color represents that the MAD (%) values are larger than 100%).

5.5.2 Differences in SWE agreements by seasonal snow classification and land cover types

The historical gamma SWE was compared to the three SWE products by Sturm's seasonal snow cover classification (Figure 4). The SSMI/S SWE extensively underestimates gamma SWE for all values of SWE. Because the SSMI/S SWE upper bound is about 172, 131, 84, 184, and 102 mm in *tundra*, *taiga*, *maritime*, *prairie*, and *warm forest*, respectively, none of the larger gamma values are captured. The best agreement between SSMI/S and gamma SWE is for the *prairie* region.

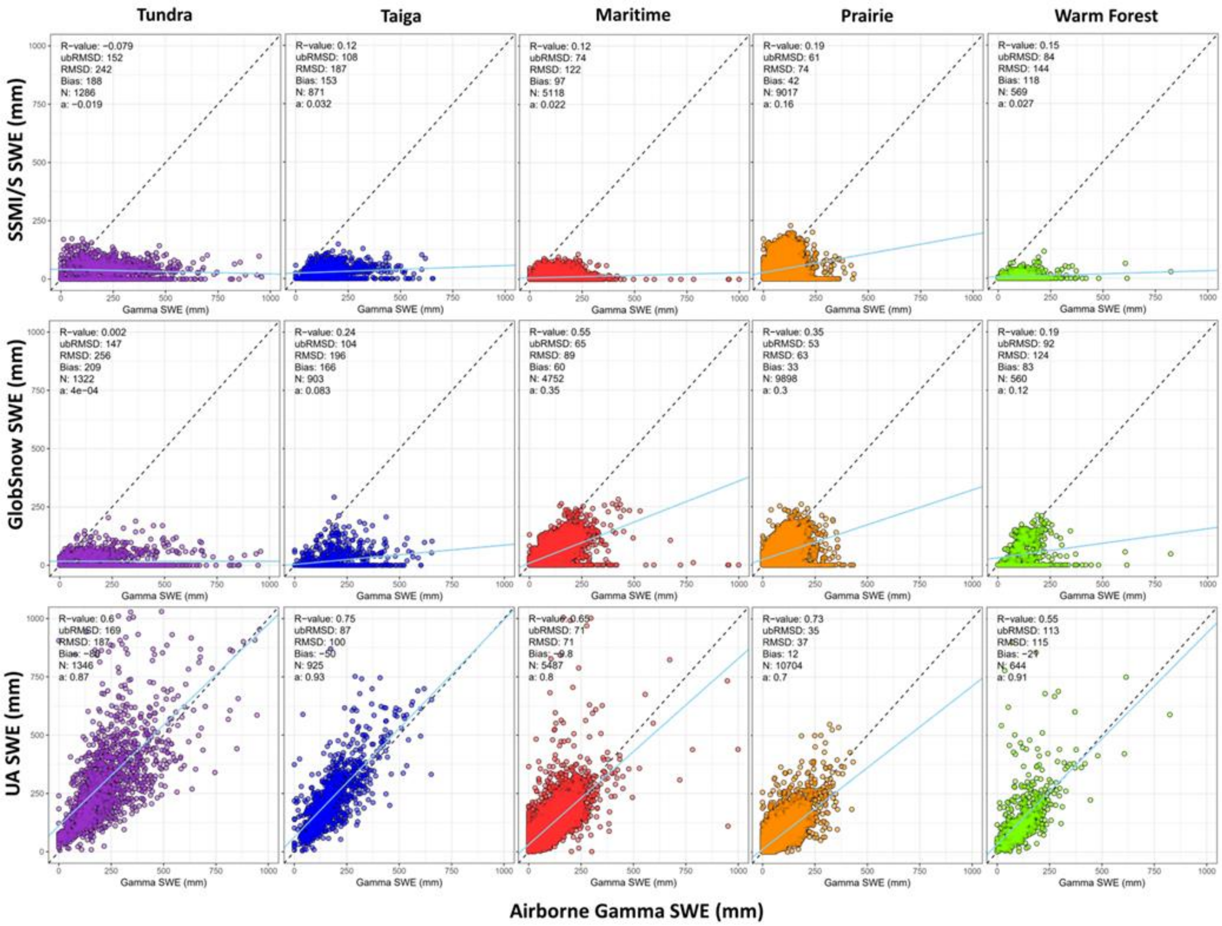
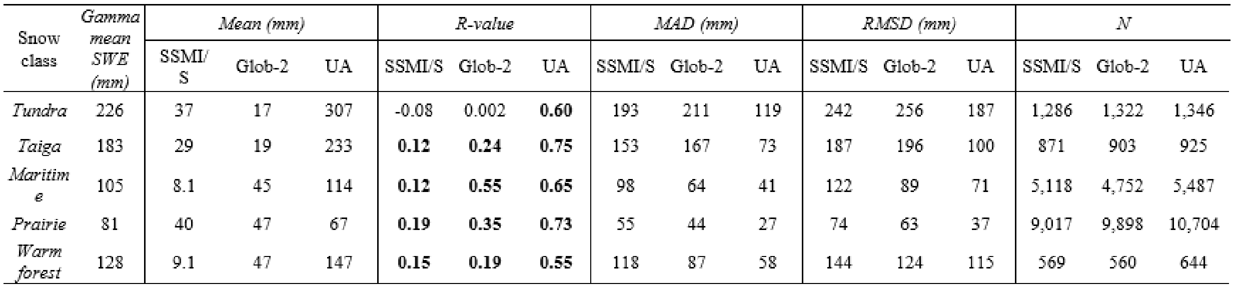


Figure 4. Comparison between daily SSMI/S (top), GlobSnow-2 (middle), and UA (bottom) snow water equivalent with daily NOAA airborne gamma radiation snow water equivalent observations from 1982 to 2017 by the Sturm’s seasonal snow classification

Table 2 Agreement between daily SSMI/S, GlobSnow-2 (Glob-2), and UA SWE and airborne gamma SWE by the Sturm et al. snow cover classification (N is a total number of valid SWE values by the snow class; R-values with **bold** indicate significant, $p < 0.05$)



Snow class	*Gamma mean SWE (mm)*	*Mean (mm)*			*R-value*			*MAD (mm)*			*RMSD (mm)*			*N*		
		SSMI/S	Glob-2	UA	SSMI/S	Glob-2	UA	SSMI/S	Glob-2	UA	SSMI/S	Glob-2	UA	SSMI/S	Glob-2	UA
Tundra	226	37	17	307	-0.08	0.002	**0.60**	193	211	119	242	256	187	1,286	1,322	1,346
Taiga	183	29	19	233	**0.12**	**0.24**	**0.75**	153	167	73	187	196	100	871	903	925
Maritime	105	8.1	45	114	**0.12**	**0.55**	**0.65**	98	64	41	122	89	71	5,118	4,752	5,487
Prairie	81	40	47	67	**0.19**	**0.35**	**0.73**	55	44	27	74	63	37	9,017	9,898	10,704
Warm forest	128	9.1	47	147	**0.15**	**0.19**	**0.55**	118	87	58	144	124	115	569	560	644

Compared to the SSMI/S SWE, GlobSnow-2 SWE has relatively better agreement with gamma SWE, especially in the *maritime, prairie,* and *warm forest* (Table 2). Correlations of the GlobSnow-2 SWE for the three classes are 0.55, 0.35, and 0.19 which are better than those of SSMI/S SWE (0.12, 0.19, and 0.15, respectively). However, GlobSnow-2 SWE is also unable to capture high SWE values (> 250 mm) in the three classes. In *tundra* and *taiga*, there are even larger errors (*MAD*: 211 and 167 mm) with gamma SWE similar to those of SSMI/S SWE (*MAD*: 193 and 153 mm).

UA SWE has notably stronger agreement with gamma SWE with small biases for all snow classes. The slopes are closer to 1:1 line (0.70 to 0.93), R-values are high (0.60 to 0.75), and *MAD* values are moderate (119 and 73 mm) in *tundra* and *taiga* classes. The UA product tends to overestimate SWE in some cases, even though the overestimations are generally small compared to the underestimations for the other products. There are high biases in most regions with a notable number of outliers above the 1:1 line. In *tundra*, UA SWE has heteroscedastic errors where the differences between UA and gamma SWE gradually increase with increasing gamma SWE.

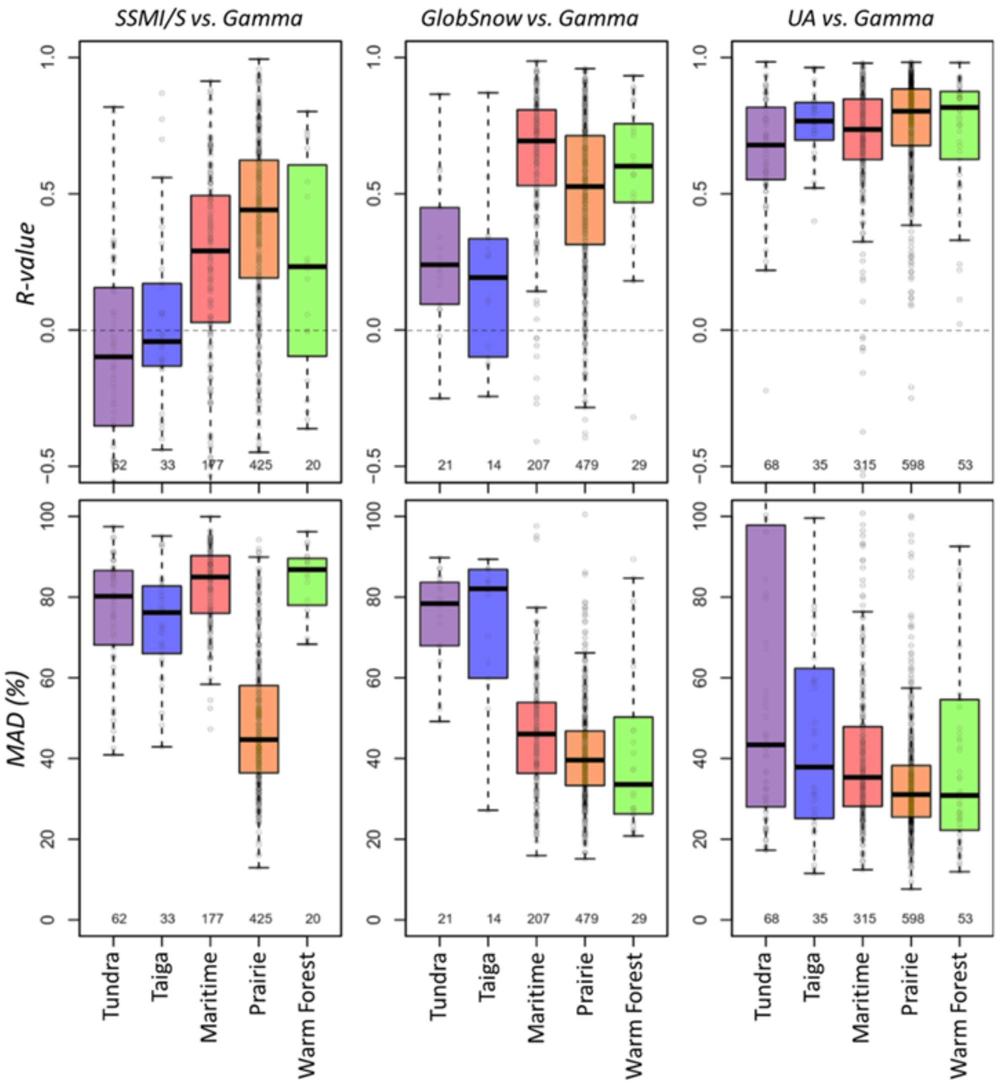


Figure 5. Boxplots of correlation coefficient (R-value) and mean absolute difference (MAD) between daily SSMI/S, GlobSnow-2, and UA snow water equivalent and daily NOAA airborne gamma radiation snow water equivalent for each gamma flight line by five snow classes (Tundra, Taiga, Maritime, Prairie, and Warm forest). The number below each boxplot is a total valid number of the statistic for each class.

Figure 5 summarizes the spatial statistics from Figures 2 & 3 by the five seasonal snow cover classes. For SSMI/S SWE, there is a moderate overall agreement in *prairie* only (median R: 0.44 and MAD: 45%) with *maritime* and *warm forest* classes having moderate correlations, but very high MAD values (85%). *Tundra* and *taiga* classes have no correlations and very high

MADs. GlobSnow-2 has much better correlations and smaller MADs in the *maritime* and *warm forest* as compared to SSMI/S. However, in *tundra* and *taiga*, the GlobSnow-2 MADs are still extremely high (median: 78 and 82% for the two classes), while their correlations (median: 0.24 and 0.19) are slightly better than that of SSMI/S SWE. The UA SWE has strong correlations for all snow classes, ranging from 0.68 (median value) in *tundra* regions to 0.82 at the *warm forest*, and small MADs from 31% in *tundra* regions to 43% at the *warm forest*. As compared to SSMI/S and GlobSnow-2 products, the UA SWE and gamma SWE agreement is quite strong in *tundra* and *taiga* even though there are greater ranges of MAD values.

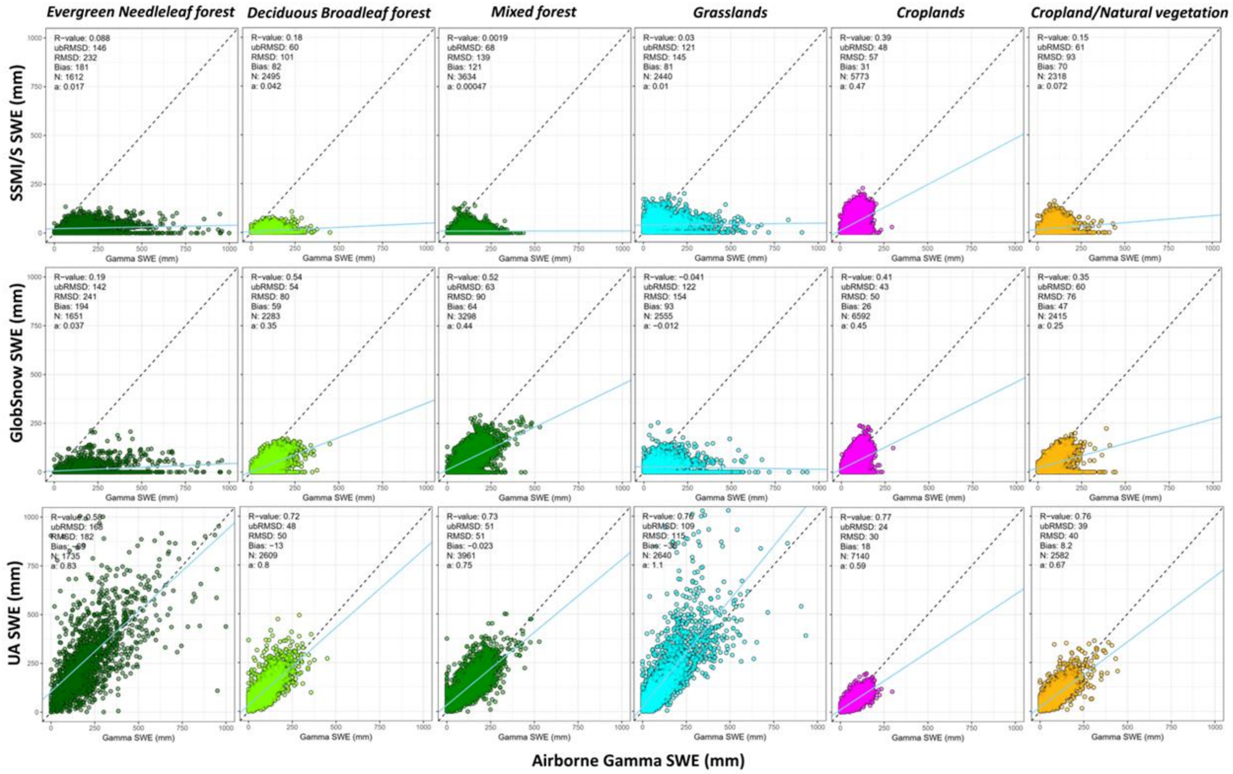
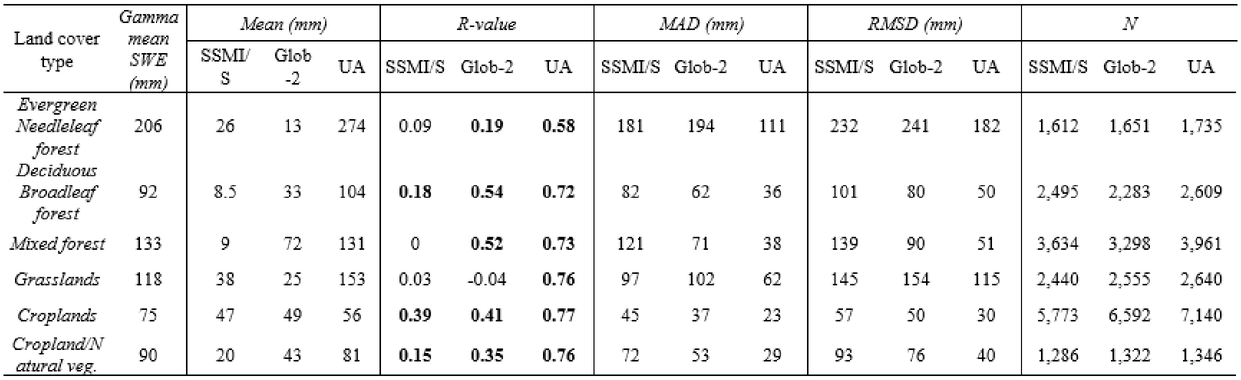


Figure 6. Comparison between daily SSMI/S (top), GlobSnow-2 (middle), and UA (bottom) snow water equivalent with daily NOAA airborne gamma radiation snow water equivalent using all available data from 1982 to 2017 by six land cover types

Table 3 Same as Table 2, but by the IGBP land cover types (N is a total number of valid SWE values by the land cover type; R-values with **bold** indicate significant, $p < 0.05$)



Land cover type	*Gamma mean SWE (mm)*	*Mean (mm)*			*R-value*			*MAD (mm)*			*RMSD (mm)*			*N*		
		SSMI/S	Glob-2	UA	SSMI/S	Glob-2	UA	SSMI/S	Glob-2	UA	SSMI/S	Glob-2	UA	SSMI/S	Glob-2	UA
Evergreen Needleleaf forest	206	26	13	274	0.09	**0.19**	**0.58**	181	194	111	232	241	182	1,612	1,651	1,735
Deciduous Broadleaf forest	92	8.5	33	104	**0.18**	**0.54**	**0.72**	82	62	36	101	80	50	2,495	2,283	2,609
Mixed forest	133	9	72	131	0	**0.52**	**0.73**	121	71	38	139	90	51	3,634	3,298	3,961
Grasslands	118	38	25	153	0.03	-0.04	**0.76**	97	102	62	145	154	115	2,440	2,555	2,640
Croplands	75	47	49	56	**0.39**	**0.41**	**0.77**	45	37	23	57	50	30	5,773	6,592	7,140
Cropland/Natural veg.	90	20	43	81	**0.15**	**0.35**	**0.76**	72	53	29	93	76	40	1,286	1,322	1,346

Because SWE products performance can be affected by land surface characteristics (e.g. vegetation types and homo/heterogeneity), the above comparison was repeated based on land cover types, *evergreen needleleaf forest*, *deciduous broadleaf forest, mixed forest, grasslands, croplands,* and *cropland/natural vegetations* (Figure 6; Table 3). Overall, there is similarly poor performance for the SSMI/S and GlobSnow-2 SWE products and UA SWE has very strong agreement with gamma SWE for all land cover types. As compared to the SSMI/S SWE, the GlobSnow-2 SWE has a better agreement with the gamma SWE in the *deciduous broadleaf forest* and *mixed forest*. However, deeper SWE values (> 250 mm) were not measured by both SSMI/S and GlobSnow-2 SWE regardless of land cover types even though they are observed by the gamma SWE. In forested regions, the UA SWE has a much stronger agreement with gamma SWE as compared to SSMI/S and GlobSnow-2 SWE. To provide insights for watershed management based on the land cover within the watershed, these comparisons were repeated for the eight NOAA RFCs and are provided in the Appendix (Figures S1 & S2). The results indicate that the UA product provides reliable SWE values for all RFCs with the exception of the California Nevada RFC.

Figure 7 summarizes the spatial statistics from Figures 2 & 3 by land cover types. SSMI/S SWE agrees best with gamma SWE in croplands with moderate correlation (median: 0.47) and MAD values (33 mm, 42%) as compared to the other classes. Negative or very weak correlations are found in the forest classes. The GlobSnow-2 SWE has much better correlations (median: 0.70, 0.65, and 0.69) and smaller MADs (median: 47, 57, and 33 mm; 48, 41, and 38%, respectively) with gamma SWE for *deciduous broadleaf forest*, *mixed forest*, and *cropland/natural vegetation*. The differences between GlobSnow-2 and SSMI/S SWE in *evergreen needleleaf forest*, *grasslands*, and *croplands* are minimal. The *evergreen needleleaf forest*, typically located in the northern Rocky Mountain and the Sierra Nevada regions, had the weakest agreement among the six classes. The UA SWE shows very high correlations with gamma SWE for all land covers. While there are fairly large MADs in the *evergreen needleleaf forest* and *grasslands* and large interquartile ranges (median: 87 mm, 55% and 35 mm, 49%, respectively), the UA SWE MAD values are much smaller than those of the other SWE products.

In summary, SSMI/S SWE has a modest agreement only for the *prairie* snow class and the *cropland* land cover. GlobSnow-2 SWE agrees well in the *maritime* and *warm forest* snow classes and *deciduous broadleaf forest*, *mixed forest*, and *cropland/natural vegetation* land cover. UA SWE has strong correlations regardless of class or land cover with larger MADs in *tundra* and *taiga* snow classes and the *evergreen needleleaf forest* and *grasslands* land cover. For further details about the agreements by subgroup, a series of box plots for the correlations where each of the snow (or land cover) classifications is subdivided into the six land cover types (or five snow classes) appears in the Appendix (Figure S4).

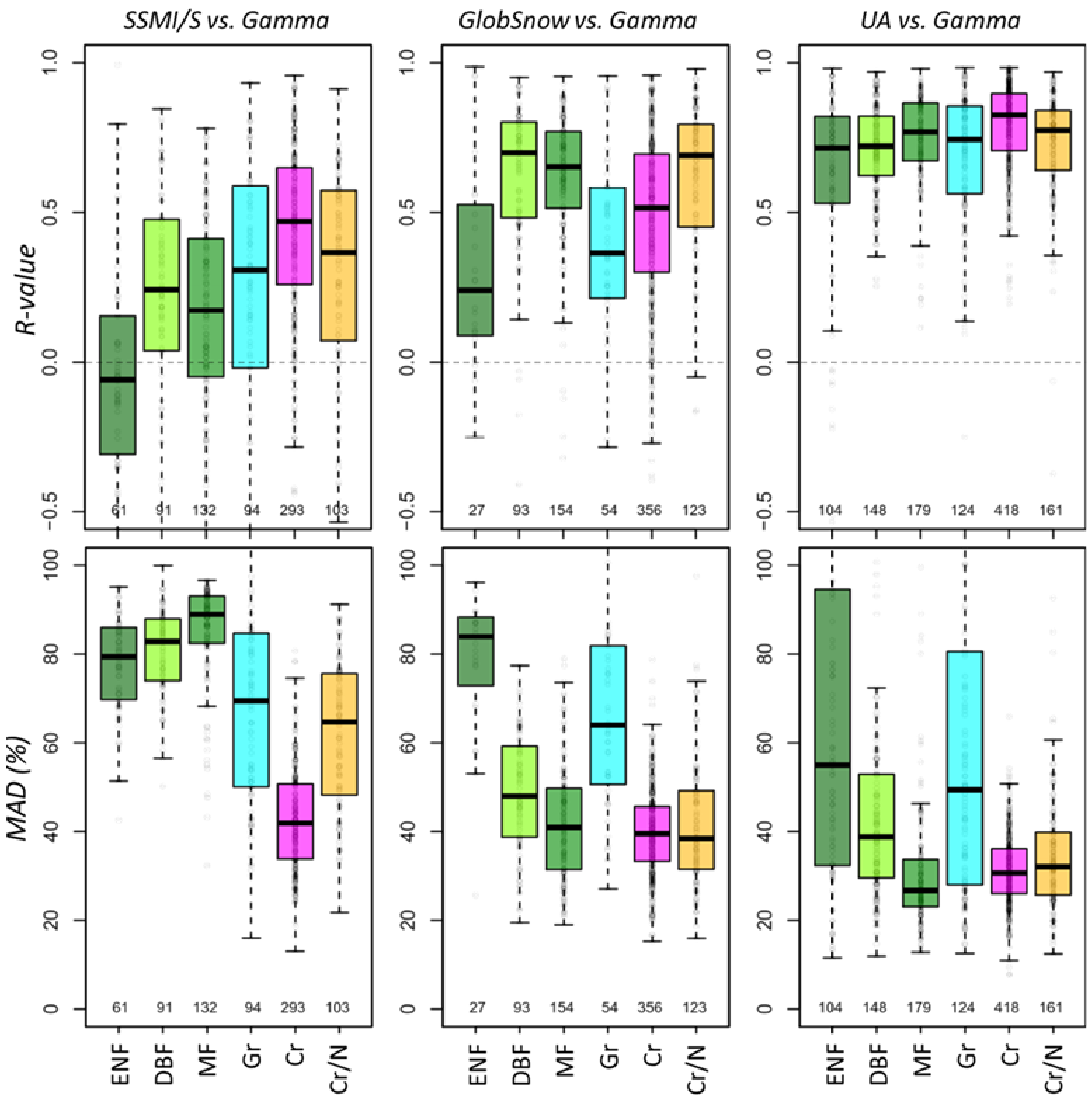


Figure 7. Boxplots of correlation coefficient (R-value) and mean absolute difference (MAD, %) between daily SSMI/S, GlobSnow-2, and UA snow water equivalent and daily NOAA airborne gamma radiation snow water equivalent for each gamma flight line by six land cover types (Evergreen Needleleaf forest, ENF, Deciduous Broadleaf forest, DBF, Mixed forest, MF, Croplands, Cr, Cropland/Natural veg., Cr/N, and Grasslands, Gr). The number below each boxplot is a total valid number of the statistic for each class.

5.5.3 Effect of tree cover fraction and topographic heterogeneity

The three historical SWE data were analyzed by tree cover fraction to determine its effect on the differences between the SWE products (Figure 8). The difference between SSMI/S and gamma SWE increases with increasing tree fraction up to 100%. The difference between GlobSnow-2 SWE and gamma SWE increased to about 70 mm (median) with increasing tree fraction up to 50%. Above 50%, the differences remain approximately constant. The differences between UA SWE and gamma SWE with forest fraction are minimal as compared to those of the other SWE products. The UA SWE slightly underestimates the gamma SWE from 0 to 80% tree fraction and then overestimates above 80%. In tree fractions below 30%, UA SWE has positive outliers (overestimation) while the SSMI/S and GlobSnow-2 SWE have negative outliers.

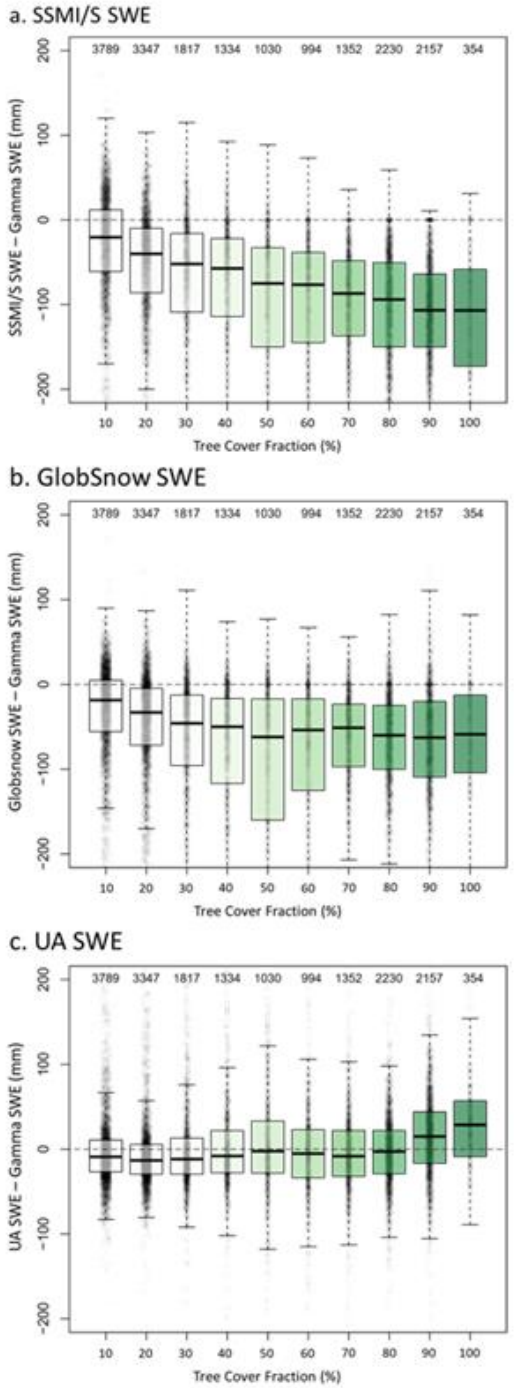


Figure 8. SWE differences between SSMI/S, GlobSnow-2, and UA products and airborne gamma radiation data by fractional tree cover (%). The white circles indicate every points in the bin.

The impacts of three topographic characteristics (slope, elevation range, and elevation) on SWE differences were examined using the same method as the preceding tree cover analysis. Figure 9 shows that slope and elevation range (surface heterogeneity) are clearly related to the SWE difference between the three gridded products and gamma SWE. For SSMI/S SWE, slope increases from 0.7 to 1.1° show increasing differences from about 48 mm to 85 mm (median) then remain constant for slopes steeper than 1.1°. The differences between GlobSnow-2 and gamma SWE are also relatively constant up to a slope of 0.7 °, then increase linearly with slope. In contrast, UA SWE differences increase monotonically from -20 mm (at 0.1 ° slope) to 24 mm (at 4.0 ° slope) with increasing slope. The elevation range has a similar impact on the SWE differences to that of the slope. However, elevation itself does not have a coherent impact on the SWE differences. At the highest elevation (2500 m), all SWE products have the widest interquartile range and the UA SWE exceeds the gamma SWE.

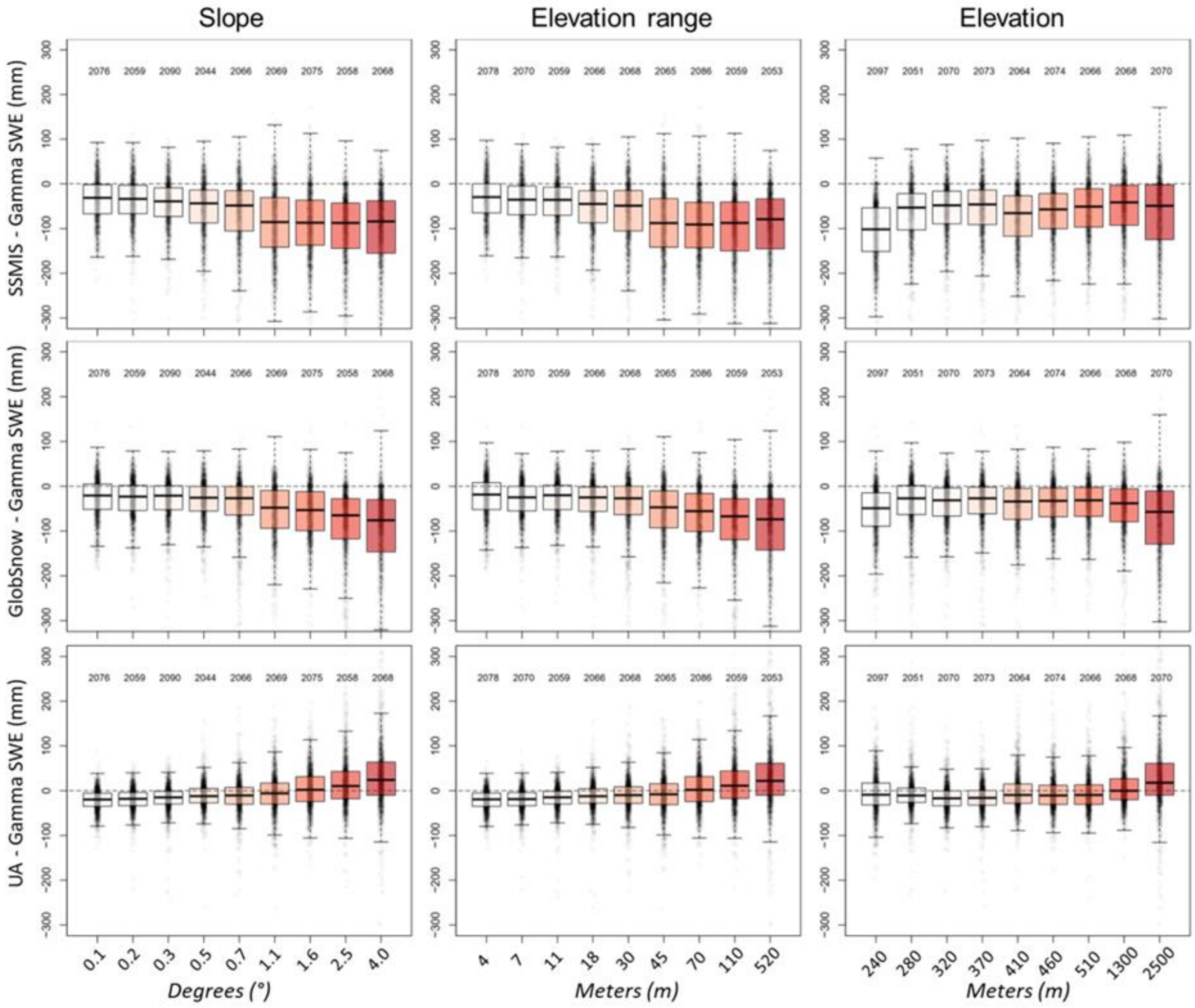
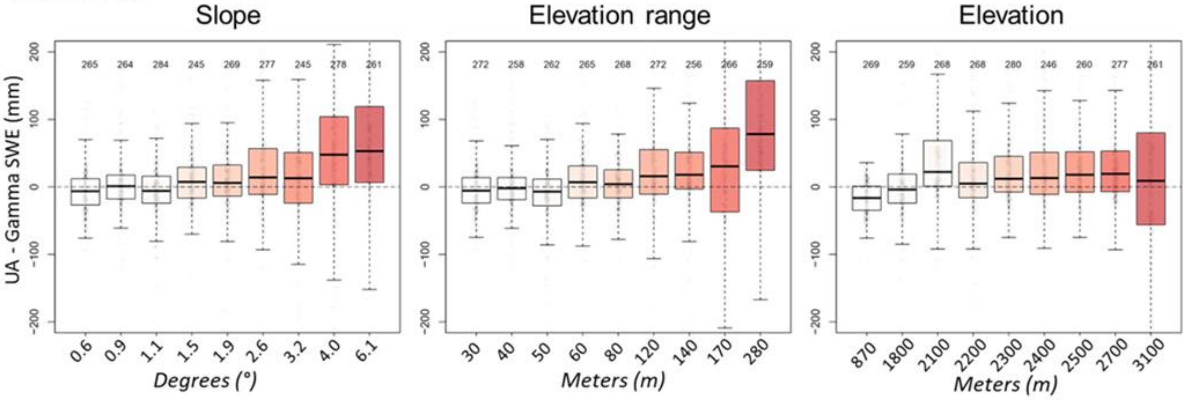


Figure 9. SWE differences between SSMI/S, GlobSnow-2, and UA products and airborne gamma radiation data by slope (degree), elevation range (m), and elevation (m). The gray circles indicate every points in the bin.

Over 90% of the SWE values in the *grasslands* are found in the western U.S. and all of the SWE values in *evergreen needleleaf forest* is in the western U.S. (Table S1). The SWE differences were analyzed by the topographic characteristics for these two western land cover types. In *grasslands*, the slope and elevation range (surface heterogeneity) are clearly related to the SWE difference between UA and gamma products. The SWE difference slightly increases with an increasing slope up to 3.2°, then much greater increases at higher slopes. For elevation range, there is a similar pattern where the SWE difference remains approximately constant up to 80 m of elevation range and exponentially increases by 78 mm with increasing the elevation range up to 280 m. Compared to the slope and elevation range, elevation has little effect on the SWE difference even though the interquartile range at 3000 m elevation is very wide. There is no consistent change in the SWE differences evident with either increasing slope or elevation range in *evergreen needleleaf forest*. The UA SWE consistently overestimates the gamma SWE for all ranges of the three topographic features.

Gamma SWE exceeds the UA SWE for tree cover fraction up to 80% but is lower than UA SWE for denser canopies (Figure 8). Over 80% of these gamma flight lines with dense canopies are in the north-central U.S. and northeastern U.S. (e.g. *deciduous broadleaf forest* and *mixed forest*; Table S2). The heavily forested areas (> 80% tree cover) were further stratified by topographic characteristics for the UA and gamma SWE differences. Figure 11 shows that slope and elevation range impact the SWE difference in the heavily forested areas. The SWE difference increases by 40 mm and 45 mm (median) with increasing slope and elevation range up to 2.5 °and 120 m, respectively. Based on the result, heterogeneous terrains, represented as slope and roughness, may partially explain the difference between UA and gamma SWE in addition to the dense tree cover.

a) Grasslands



b) Evergreen needleleaf forest

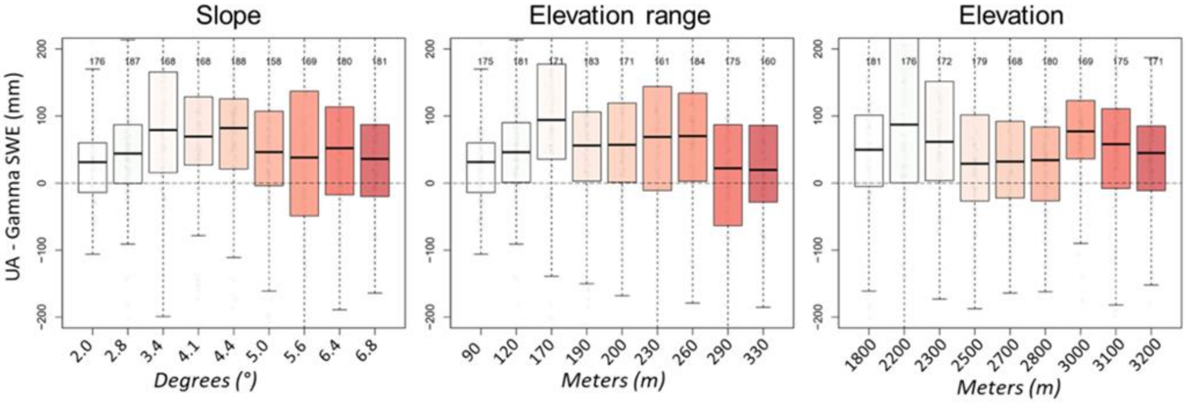


Figure 10. Same as Figure 9, but for the SWE difference between UA and gamma SWE in grassland and evergreen needleleaf forest types only

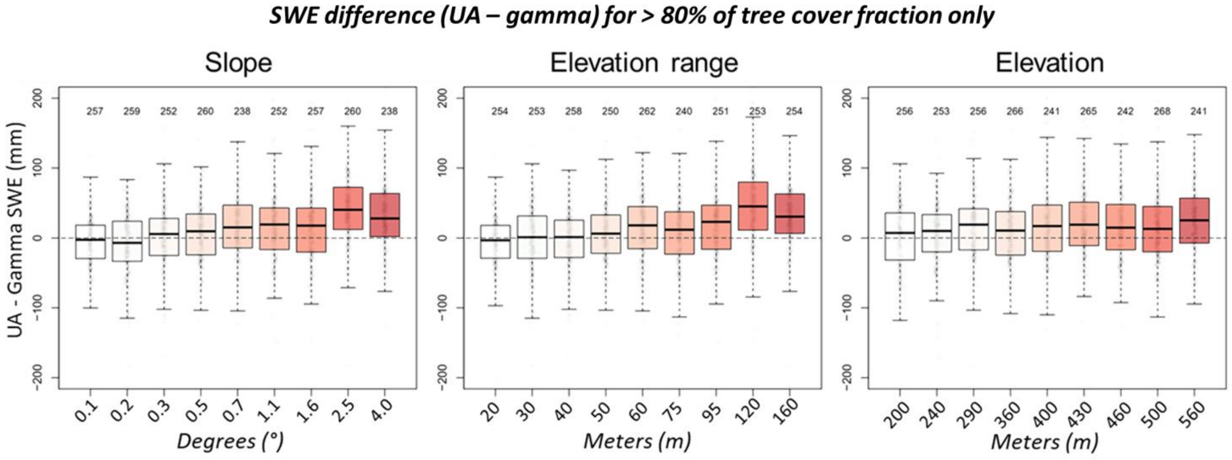

SWE difference (UA – gamma) for > 80% of tree cover fraction only

Figure 11. Same as Figure 10, but for the SWE difference between UA and gamma SWE in areas with more than 80% of tree cover fraction only

5.6 Discussion

5.6.1 Comparison of three SWE products

Deep SWE was not captured by SSMI/S retrievals regardless of snow class and land cover type. This is due to a known limitation of passive microwave signal “saturation effect” for deep snowpack (approximately 1 m snow depth) (Dong et al., 2005; Vuyovich et al., 2014). In the deep snowpack, soil emissions are no longer detectable at the higher microwave frequency (e.g. 37 GHz). Since the GlobSnow-2 SWE product is a fusion of ground measurements of snow depth and remotely sensed data from passive microwave SSMI/S instruments, the product also appears to inherit the saturation effect (Dawson et al., 2018; Takala et al., 2011).

Both SSMI/S and GlobSnow-2 largely underestimate SWE in forested land cover types and the *tundra* and *taiga* snow classes. This may be due to the known impact of forest cover on passive microwave SWE retrievals (Foster et al., 2005; Vuyovich et al., 2014). However, similar patterns were found in grasslands, even though the average percent tree cover in grassland flight lines is typically small (mean: 9.2%). Vuyovich et al. (2014) found a disagreement between microwave and modeled SWE products in the Upper Powder Basin, Wyoming (grasslands), even though the areas have little vegetation and modest elevation ranges. This could be due to the limited availability of ground observations to inform modeled products, or it may be due to other physical effects on the microwave signal such as spatial variability and elevation gradients (Mätzler & Standley, 2000). Cai et al. (2017) mentioned that the elevation-dependent factors, besides forest fraction, affect the relationship between SWE and Tb. Another possibility is that the redistribution of SWE by wind drifting of the snowpack frequently occur in *tundra* and *taiga* snow classes, likely resulting in larger discrepancies between the SWE products (Clow et al., 2012; Winstral et al., 2002).

The UA SWE had better performance with fewer errors in all snow and land cover classes. This may be partly because the UA SWE product (4 km × 4 km) has higher spatial resolution than the SSMI/S and GlobSnow-2 SWE (25 km × 25 km), However, there were fewer random errors (e.g. the degree of dispersion in Figures 4 & 5) as well as lower systematic biases (e.g. sustained over/underestimation) between UA and gamma SWE. The superiority of the UA SWE data may be primarily due to the unique methodologies such as a new interpolation approach and snow density parameterization as well as the use of the reliable ground-based observations from SNOTEL SWE, COOP snow depth network, and PRISM precipitation and temperature data as inputs (Zeng et al., 2018). The UA SWE is produced by interpolating "normalized SWE" (SWE divided by accumulated snowfall), rather than SWE itself. The normalized quantities of SWE were more spatially consistent and scale-independent than original SWE quantities (Broxton et al., 2016a). A new snow density parameterization with seasonal evolutions may have contributed to the better performance compared to the other SWE products (Dawson et al., 2017). Broxton et al. (2016b) mentioned that the features of snow density parameters helped overcome a common deficiency in land surface models and reanalysis products of too much ablation at near-freezing temperatures.

The gamma SWE surveys have the highest accuracy in the north central U.S. (e.g. NCRFC; croplands/prairie snow class) and are extremely critical for that region because of its paucity of in situ observations (Schroeder et al., 2019; Tuttle et al., 2017). Because the spatial extent of the gamma survey is larger than most in situ observations, it also adds economic and social values in most regions that are vulnerable to spring snowmelt floods including those with established observation networks (Simonovic, 1999).

5.6.2 Effect of tree fraction

The underestimation of SSMI/S SWE is similar to the previous findings of Foster et al. (2005). In order to overcome the uncertainty, they developed a bias-corrected SWE algorithm with the forest information defined as “Forest factor”. For each 10th percentile of the fractional forest cover, different forest factors ranging from 1 (no forest) to 2 (100% fractional forest cover) were multiplied by the original SWE algorithm to correct the underestimation error. To test how much the uncertainty could be reduced, the corrected SSMI/S SWE was computed by considering the fractional tree cover and compared to the original SWE (Figure S5). Based on the result, the forest factor can partially reduce errors in the SSMI/S SWE. But there are still large underestimates as compared to gamma SWE, probably due to the saturation effect of passive microwave signal on the deep snowpack. GlobSnow-2 SWE performs relatively better than SSMI/S SWE in areas with dense tree cover but not in areas with sparser tree cover. SSMI/S SWE is impacted by the attenuation of microwave signal by tree canopy cover. In contrast, the GlobSnow-2 algorithm includes ground-based snow depth measurements (Luojus et al., 2017) and forest canopy information in the HUT Snow Emission Model (Pulliainen & Hallikainen, 2001) which likely reduces errors below those expected when only the microwave emissions are considered. Regardless, GlobSnow-2 still underestimates SWE for all tree cover fractions. This could be partly related to the use of a constant low snow density (0.24 kg/m^2) to estimate GlobSnow-2 SWE without considering the seasonal evolution of snow density. The constant snow density generally represents early winter conditions (0.22, 0.26, 0.23, and 0.24 kg/m^2 on October 1 for *alpine*, *maritime*, *prairie*, and *tundra*, respectively; Sturm et al., 2010). However, approximately 85% of the gamma SWE and GlobSnow-2 SWE values are obtained in mid- and late-winter from February 1 and April 30. There appears to be larger snow density due to compaction and snowpack metamorphism (Anderson et al., 2006; Sturm et al., 2010; Dawson et al., 2017; Hill et al., 2019). In this period, the typical snow densities in most snow classes from in-situ measurements are larger than 0.24 kg/m^2 (e.g. Maritime: 0.30 – 0.38 kg/m^2 and Prairie: 0.24 – 0.32 kg/m^2; Dawson et al., 2017). Also, the underestimation of GlobSnow-2 SWE could be due to the saturation effect of deep snow on the microwave signal as mentioned in the preceding section. For example, while the mean value of the deep gamma SWE (> 250 mm of gamma SWE; N = 1,077) is 355 mm, and UA SWE is 388 mm, the corresponding GlobSnow-2 and SSMI/S SWE was 41 mm and 25 mm, respectively. This indicates that the GlobSnow-2 SWE still has a similar limitation due to the saturation effect of the passive microwave signal for the deep snow (Dawson et al., 2018).

Results show that UA SWE performs well for a variety of forest covers. While the passive microwave measurements are adversely affected by forest cover, forest cover does not affect UA SWE. The superiority of the UA SWE product can be explained by the new snow density parameterization and reliable input data from SNOTEL and PRISM data in forested-type regions. Unlike the other two products, the new snow density parameters used in the UA product likely reflect physical processes with an SWE evolution for all snow classes (Dawson et al., 2017). Also, the PRISM precipitation and temperature data assimilates nearly 13,000 precipitation and 10,000 temperature point stations over the conterminous U.S. including a few thousand stations over the northeastern U.S. (*deciduous broadleaf forest* and *mixed forest types*) and western U.S. (*evergreen needleleaf forest)* (Daly et al., 2008). This suggests that UA SWE may include the accuracy of PRISM precipitation data over the heavily forested regions if the partitioning of snowfall and rainfall was fairly accurate. A recent study from Dawson et al. (2018) presented that UA SWE had strong agreement with ASO SWE in California’s Sierra Nevada.

However, they evaluated the UA SWE using 32 ASO SWE measurements only and the forest cover fractions in the study area were less than 35%. In addition to their results, our finding strongly supports the reliable accuracy of the UA SWE data in forested regions.

5.6.3 Potential sources of error in gamma SWE

While the operational NOAA airborne snow survey using a gamma radiation technique has successfully provided SWE observations over the last 40 years, this method can generate errors when estimating SWE. The potential sources of error have been identified and well-documented in previous studies (Carroll &Carroll., 1989a, b; Glynn et al., 1989; Offenbacher & Colbeck, 1991). An uncertainty in baseline fall SM measurement can potentially introduce biases into gamma SWE estimates. The operational standard approach assumes that soil moisture conditions remains constant following the baseline fall survey. However, early-winter snowmelt and rainfall events can change the soil moisture after the fall survey, resulting in under/overestimation of the gamma SWE estimates. Our recent finding in Cho et al. (2018) showed that standard gamma SWE values were improved by capturing the changes in soil moisture using daily Soil Moisture Active Passive (SMAP) satellite data. An effect of forest biomass on the accuracy of airborne SWE measurements has been explored over forested watersheds (Vogel et al., 1985; Carroll &Vose, 1984; Carroll & Carroll, 1989a). Carroll &Vose (1984) showed there was 23 mm of RMSE between airborne gamma SWE and ground-based SWE with the moderate snowpack that ranged from 20 mm to 470 mm in forest environments over Lake Superior and Saint John Basins. Glynn et al. (1988) found that underestimates of airborne gamma SWE (RMSE: 63 mm) could occur when gamma flight flew on forested areas with the very deep snowpack (in-situ SWE: 460 mm) in the Saint John basin over the provinces of Quebec and New Brunswick and the state of Maine.

The physical basis for a difference between UA SWE and gamma SWE in the regions with very high tree cover fractions (Figure 8) could be either an overestimation of UA SWE or an underestimation of gamma SWE or possibly both. High forest biomass may cause an underestimation of gamma SWE. Glynn et al. (1988) and Carroll & Carroll (1989) mentioned gamma radiation emitted by the forest biomass above the snowpack as a source of error. There is a considerable amount of potassium, ^{40}K, and a minor amount of thorium in the forested-type land cover. This impact increases in a very deep snowpack because when gamma radiation coming from the ground is attenuated by the snow, the radiation emitted by the forest biomass above the deep snow can become a large portion of the radiation detected in the gamma aircraft.

Among the underestimated SWE values with the dense tree fraction, the larger underestimations occurred over flight lines with high slopes and a range of elevations within the flight footprint (Figure 9). Previous studies found that heterogeneous characteristics within a flight line, so-called “uneven effect”, can cause underestimates of gamma SWE (Cork and Loijens, 1980; Carroll & Carroll, 1989b). This heterogeneity is commonly caused by snow drifting or mountainous environment. A mean areal gamma SWE is calculated by integrating measures of gamma attenuation rate by an optimal counting interval during a flying (Carroll, 2001). If a large spatial variability in the snowpack exists along a flight line, the measurements of the attenuation of the gamma count rate are systematically underestimated (Cork and Loijens, 1980). For example, consider an alternating shallow and deep snowpack over a flight path. During a first counting interval, the gamma detector would measure gamma radiation from a uniform deep snowpack. Then the flight would measure gamma radiations from snow-covered areas with different proportions of deep and shallow snowpack until the gamma detector is centered above a uniform shallow snowpack. In the transition area between deep and shallow snowpack, the larger gamma count rates tend to be measured and the gamma rates may generate underestimates of SWE. Carroll & Carroll (1989b) found that the degree of the underestimation is related to a variance of the distribution of ground-based SWE measurements within a flight footprint. If the variance in the gamma SWE is known, SWE underestimates can be corrected. Because our results use the standard gamma radiation SWE product without manual corrections, results could be improved by updating gamma SWE products in regions with heterogeneous characteristics. The previous studies primarily focused on the spatial variation of snowpack itself and did not consider terrain. The results in this study were not able to determine if an "uneven snow effect" was present but terrain heterogeneity may have value as a proxy or may serve to add additional heterogeneous characteristics that also impact the gamma retrievals.

5.7 Conclusion and Future Perspectives

In the snow science community, observation-based long-term gridded SWE products have been developed for hydrological and climate research. However, an evaluation of the currently available SWE products has been limited due to the lack of independent SWE datasets at a continental scale. Furthermore, as land surface models and regional climate models continue to evolve at a rapid pace, independent and reliable SWE data is required to evaluate SWE outputs from the models to identify potential limitations of snow physical processes involved in each model. The historical 40-year and ongoing NOAA airborne gamma SWE record can be used as reference long-term reliable SWE across the U.S. and southern Canada. Even though the record has limited spatial and temporal coverages compared to the gridded satellite and reanalysis products, the record may be useful for snow hydrologists and modelers in providing accurate SWE values in various environments.

In this study, three long-term daily SWE products (more than 30 years), SSMI/S, GlobSnow-2, and UA, were evaluated using an airborne gamma radiation SWE record collected by NOAA NOHRSC. Comparisons were made by seasonal snow cover classes and land cover types from 1982 to 2017 in the conterminous U.S., and this provided several interesting insights. We found that SSMI/S and GlobSnow-2 SWE products showed similar patterns against the airborne gamma SWE: modest performances in *croplands* and *grasslands* land cover types and *prairie* of snow class over the north-central U.S. and poor performances (extremely underestimated SWE) in *evergreen needleleaf forest* and *grasslands,* and *tundra* and *taiga* in mountainous regions over the western U.S. These may correspond with weaknesses of inherent microwave satellite-driven signals. However, compared to SSMI/S SWE, GlobSnow-2 SWE had better agreement with gamma SWE in some forested land cover types, *mixed forest*, *deciduous broadleaf forest*, and *warm forest* and *maritime* of seasonal snow classes in the northeastern U.S. UA SWE has better agreement with gamma SWE as compared to SSMI/S and GlobSnow-2 SWE in all land cover types and snow classes, while there are relatively weak agreements in the *evergreen needleleaf forest* and *grasslands* land cover, and *tundra* snow class, which are likely due to potential limitations of UA products as well as gamma SWE in the mountainous regions (e.g. spatial heterogeneity). Through the sensitivity analysis to forest fraction, we found UA SWE is much less affected by forest fraction, while the SSMI/S and GlobSnow-2 SWE have increasing negative biases with increasing tree fraction within the gamma flight footprint. The effect of tree cover fraction to GlobSnow-2 SWE is less than in larger fractions (> 60%). With a known challenge of measuring SWE in *evergreen needleleaf forest* and *tundra* classes, unexpected weak agreement of UA SWE with gamma SWE in *grasslands* will be the focus of our future research to better understand the physical factors impacting on the results.

Additional use of gamma SWE products for validation of remote sensing observations and modeled SWE will likely necessitate additional inquiry into the gamma observation capabilities. Studies have shown that the improvements to the operational gamma SWE products are possible by minimizing the errors, even though potential sources of errors were identified before the 1990s (e.g. changes in fall soil moisture, the spatial variance within the footprint, and dense forest effect). Future studies can utilize state-of-the-art high-resolution earth observation products (e.g. lidar, synthetic aperture radar, and multi-spectrometer, etc.) to quantify snow or land characteristics within a gamma flight footprint to improve this valuable resource.

CHAPTER 6

Trend Analysis and Extreme Value Snow Water Equivalent and Snowmelt for Infrastructure Design over the Contiguous United States[5]

6.1 Introduction

Snow and snowmelt driven extreme events can have large societal and economic consequences. Extreme snow can damage infrastructure and buildings (American National Standard Institute, 1972; American Society of Civil Engineers, 2017; Sack, 2015). Heavy snow loads on roofs can cause structural failures (Geis et al., 2011; Bean et al., 2019). From 1989 to 2009, 1,029 snow-induced building collapse incidents in the U.S. caused 19 fatalities and 146 injuries with each incident costing up to $200 million (Geis et al., 2011). In many parts of the U.S., snow meltwater is a dominant driver of severe spring flooding (Berghuijs et al., 2016). Snowmelt floods routinely impact the north-central and -eastern U.S. (Wazney and Clark, 2015; Stadnyk et al., 2016; https://www.weather.gov/dvn/summary_SpringFlooding_2019). The Red River of the North Basin's (RRB) 1997 snowmelt flood caused more than $5 billion of damage in Fargo and Grand Forks, North Dakota and other communities (Todhunter, 2001).

Civil engineers and water resources managers rely on historical precipitation data when making hydrologic estimates of design floods to size infrastructures (e.g., water management facilities, bridges, and other hydraulic control structures). The ability of infrastructure to withstand environmental stressors depends on the quality of input data. The National Oceanic and Atmospheric Administration's National Weather Service Precipitation-Frequency Atlas 14

---

[5] Cho, E., J.M. Jacobs (2020) Trend Analysis and Extreme Value Snow Water Equivalent and Snowmelt for Infrastructure Design over the Continental United States *(In preparation)*

(NOAA Atlas 14) series are the U.S. government standard to use in designing infrastructure with adequate capacity for flood events.

In the western U.S., the magnitude of snowmelt exceeded 24-hr precipitation design values at 79 of the 379 mountain locations (22%) (Harpold and Kohler, 2017). Across the southern Rocky Mountains, the 24-hr snowmelt intensity exceeded the 24-hr rainfall intensity by 53 and 38% for 10- and 100-year events, respectively (Fassnacht and Records, 2015). Even though spring snowmelt is a dominant driver for snowmelt-driven floods in the north central and northeastern U.S. and western mountainous regions of the U.S., NOAA Atlas 14 does not typically consider snow accumulation and melt events (Bonnin et al., 2006; Berghuijs et al., 2016). In a recent NOAA Atlas 14 document (Volume 10 Version 3.0), the authors attempted to exclude snowfall from the precipitation in order to calculate frequency estimates based on rainfall (i.e., liquid-only precipitation) only. From their limited analysis of three sites in the northeastern U.S., Perica et al. (2015, revised 2019) they concluded that differences between design estimates using rainfall only versus all precipitation estimates were inconsequential and thus no liquid-only precipitation frequency analysis was conducted.

To accurately estimate the magnitude of snowmelt events over the U.S., a reliable, spatially distributed SWE record is required. Spatially distributed SWE products have been developed using assimilation techniques by ingesting in-situ snow station networks and remotely sensed snow products (Barrett, 2003; Carroll et al., 2006; Broxton et al. 2016a; Takala et al., 2011). The SNOw Data Assimilation System (SNODAS) operated by the National Weather Service (NWS) National Operational Hydrologic Remote Sensing Center (NOHRSC) provides a near real-time 30 arc-second grid (about 1 km^2) of spatially distributed SWE throughout the continental U.S. [CONUS] (Barrett 2003). The SNODAS data provide the only real-time spatially distributed estimate of snowpack conditions across the CONUS and are used operationally by hydrologists and flood forecasters in regional NOAA river forecast centers. Unfortunately, the SNODAS SWE data record, September 2003 to current (16 winters), is too short to be used to estimate extreme snowmelt events.

Recently, the University of Arizona (UA) released a long-term gridded (4-km) daily SWE dataset (hereafter UA SWE) over the CONUS for the period October 1981 to May 2017 through the NASA National Snow and Ice Data Center (Broxton et al. 2019). UA SWE was produced by assimilating the Parameter-elevation Regressions on Independent Slopes Model (PRISM) daily precipitation and temperature data developed by the PRISM Climate Group at Oregon State University (Daly et al., 2008), SWE and snow depth data from the Snowpack Telemetry (SNOTEL) network (Serreze et al., 1999), and SWE and snow depth data from the NWS Cooperative Observer Program (COOP) network. Broxton et al. (2016b) found that UA SWE has similar spatial and temporal variability as SNODAS SWE, while the reanalysis and Global Land Data Assimilation System (GLDAS) products substantially underestimate SWE compared to the two SWE products. Dawson et al. (2018) showed that UA SWE had strong agreement with Airborne Snow Observatory (ASO) SWE products over the Toulumne basin in California. Cho et al. (2019) demonstrated that the UA SWE had the best agreement with historical airborne gamma radiation SWE observations (N= 20,738 observations from 1982 to 2017) as compared to GlobSnow-2 and microwave satellite SWE, regardless of snow classification and land cover type.

The availability of the complementary and well-vetted SNODAS and UA SWE products provides a unique opportunity to inform engineering design of structures and facilities that must withstand snow loads or snowmelt runoff. Here we seek to draw from the strengths of these two

SWE products, UA SWE and SNODAS SWE, to develop design values for SWE and snowmelt over the CONUS that are needed to support a range of engineering design needs. This is achieved by, first, analyzing the systematic differences (i.e., bias) in extreme SWE and snowmelt between UA and SNODAS datasets; second, bias correcting and detrending UA SWE using the SNODAS data via empirical cumulative distribution function (ECDF) method (McGinnis et al., 2015) and traditional detrending techniques (Sen, 1968); and finally, conducting an extreme value frequency analysis in which historical annual maximum are used to fit the generalized extreme value (GEV) probability distribution (Bonnin et al., 2006; Perica et al., 2013, 2015; Cheng and AghaKouchak, 2014) and the fitted distribution is used to map annual maximum SWE and snowmelt for various return periods.

The subsequent sections are organized as follows. Section 2 explains the SNODAS and UA SWE data sets. Section 3 describes the methodologies including identification of the annual maximum 7-day snowmelt event, bias correcting and detrending, and application of the extreme value methods. Section 4 compares the annual maximum SWE and snowmelt from UA and SNODAS products, identifies annual maximum SWE and snowmelt trends, presents the 25- and 100-year design SWE and snowmelt maps, and compares these maps to the current NOAA Atlas 14 maps. Section 5 discusses results in light of previous work and presents potential limitations in the approach. Conclusion and future perspectives appear in Section 6.

6.2 Data

6.2.1 SNODAS SWE

In this study, SNODAS products are used as the benchmark. SNODAS products are well vetted by previous research throughout the US (Vuyovich et al., 2014; Broxton et al., 2016b;

Schroeder et al., 2019) and have been used operationally by hydrologists and flood forecasters in the regional river forecast centers for the entire U.S. (Carroll et al., 2006; personal communications with Mike DeWeese and Pedro Restrepo, NOAA North Central River Forecast Center). NOAA's SNODAS integrates model results with ground observations, airborne gamma SWE, and satellite snow cover estimates, to generate the best 1 km gridded estimate of the snow characteristics that minimizes errors associated with any individual method (Carroll et al., 2006). The three main procedures in the SNODAS ingest and downscale numerical weather prediction (NWP) output, simulate the snowpack using a physically based energy and mass balance model, and assimilate independent satellite, airborne, and ground-based observations of snow cover area (SCA) and SWE to adjust model results. Forcing data come from the Rapid Update Cycle 2 (RUC2) NWP output generated by the National Centers for Environmental Prediction (NCEP) and are downscaled using a digital elevation model. The snow model is an energy and mass-balance, multilayer model and consists of three snow layers and two soil layers (Barrett, 2003). For the assimilation procedure, snow observations include ground station-based SWE and airborne gamma radiation SWE as well as satellite-based SCA information.

SNODAS V1 data are freely available from the NASA National Snow and Ice Data Center from 1 October 2003 to the present (https://nsidc.org/data/g02158). In this study, the original, daily SNODAS SWE data (1 km x 1 km spatial grid) were used from October 1, 2003 to May 31, 2017 (14 water years). The original SNODAS data were aggregated/resampled to spatially match the UA SWE data (4 km x 4 km spatial grid) using the nearest neighbor method, then annual maximum SWE values were obtained for each available water year.

6.2.2 UA SWE

The observation-based 4 km UA SWE combines station-based SWE and snow depth observations with a background modeled SWE using an empirical temperature-index snow model over the CONUS (Broxton et al., 2016a; 2019). The station-based SWE and snow depth observations are from Snowpack Telemetry (SNOTEL) network and the National Weather Service Cooperative Observer Program (COOP) network, respectively. The background modeled SWE was generated using the PRISM precipitation and temperature data as forcing data into an empirical temperature-index snow model (Broxton et al., 2016a). The quality of the UA SWE data was evaluated with the current reanalysis and the Global Land Data Assimilation System (GLDAS)-based land surface model products (Broxton et al., 2016b). A summary of recent updates in the UA product with trend/driver analysis of the data was provided in Zeng et al. (2018). The UA SWE data (Version 1) were recently released and freely available at NASA National Snow and Ice Data Center from 1 October 1981 to 30 September 2017 (https://nsidc.org/data/nsidc-0719). In this study, daily 4 km gridded UA SWE values were used from October 1981 to May 2017 (36 water years).

6.2.3 NOAA Atlas 14 Precipitation Frequency Estimates

The NOAA Atlas 14 includes design precipitation estimates for selected durations and frequencies as well as the lower and upper bounds of the 90% confidence interval for the U.S. Atlas 14 also contains supplementary information on methodologies and results of trend and seasonal analyses. The Atlas 14 precipitation frequency estimates are determined from long-term precipitation records from a regional network of rainfall gauges (Bonnin et al., 2006; Perica et al., 2013, 2015). In this study, the recent Atlas 14 precipitation frequency CONUS maps are compared to this study’s design snowmelt maps, except for northwestern U.S. and Ohio River

Basin. The NOAA Atlas 14 maps are not available over the northwestern U.S. region. While Atlas14 Volume 2 provides design precipitation values for the Ohio River Basin and surrounding states, the annual maximum series for that region are not available. The Atlas 14 precipitation spatial maps at a 1 km spatial grid are available from the Hydrometeorological Design Studies Center (HDSC) within the Office of Water Prediction (OWP) of the NOAA's National Weather Service (NWS) (https://hdsc.nws.noaa.gov/hdsc/pfds/pfds_gis.html). The 1 km Atlas 14 maps were aggregated/resampled to spatially match the UA SWE data (4 km x 4 km spatial grid) using nearest neighbor method.

6.3 Methodology

6.3.1 Annual maximum SWE and 7-day snowmelt

All available gridded, daily time series of UA and SNODAS SWE data were used to obtain the annual maximum SWE and 7-day snowmelt values for 36 and 14 water years, respectively. Annual maximum SWE values are the one-day maximum value determined for each pixel using the October 1 to May 30 SWE daily time series. annual maximum 7-day snowmelt event ($Melt_{max,7d}$) for each pixel is defined as

$$Melt_{max,7d} = \max\,(SWE_i - SWE_{i+7}) \qquad (1)$$

where i is a date from 1 October to 31 May for each year and SWE_i and SWE_{i+7} is daily SWE at dates, i and i+7, respectively.

The study period is divided into two periods: the period where both UA and SNODAS data are available (14 years; WY2004 – WY2017) and the period during which only UA SWE is available (22 years; WY1982 – WY2003). The annual maximum snowmelt data for the 14 years period with concurrent data were used to determine a statistical relationship between UA and SNODAS.

6.3.2 Empirical cumulative distribution function (ECDF)

The SNODAS SWE is widely considered to be the most reliable gridded SWE product over the CONUS (Barrett, 2003; Broxton et al., 2016). The empirical cumulative distribution function (ECDF) method was applied to bias correct the UA annual maximum SWE and snowmelt data using SNODAS annual maximum SWE and snowmelt. The ECDF is widely used in climate modeling to correct systematic model bias based on observation data (McGinnis et al., 2015). In this study, CDFs of the UA annual maximum SWE and snowmelt values, $CDF_{UA_{14}}$, for the period where both UA and SNODAS data are available (14 years; WY2004 – WY2017) are used to determine probabilities associated with the UA annual maximum SWE or snowmelt values. The bias correction technique transforms these probabilities back into snowmelt values using the inverse CDF of the SNODAS distribution developed for the 14 years, $CDF^{-1}_{SNODAS_{14}}$, as follows

$$ECDF\_Melt_{UA,i} = CDF^{-1}_{SNODAS_{14}}\left(CDF_{UA_{14}}\left(Melt_{UA_i}\right)\right),\ if\ Melt_{UA,i} \leq max\left(Melt_{UA_{14}}\right)$$

(2) where $ECDF\_Melt_{UA,i}$ and $Melt_{UA_i}$ are given as the ECDF transformed annual maximum snowmelt event and the original annual maximum snowmelt at year *i*, respectively, and $max\left(Melt_{UA_{14}}\right)$ is the maximum snowmelt in the overlapping 14 years.

If a UA value at year i, $Melt_{UA_i}$, exceeded the maximum UA of the 14 years, $max\left(Melt_{UA_{14}}\right)$, a difference between the UA value the maximum UA multiplied by a ratio of standard deviation of the UA snowmelt, $\mathrm{std}(Melt_{UA_{14}})$, to that of SNODAS snowmelt,

$\mathrm{std}(Melt_{SNODAS_{14}})$, was added in the maximum SNODAS values of the 14 year periods, $\max\left(Melt_{SNODAS_{14}}\right)$, as follows

$$ECDF\_Melt_{UA,i} = \max\left(Melt_{SNODAS_{14}}\right) + \left[Melt_{UA,i} - max\left(Melt_{UA_{14}}\right)\right] \frac{\mathrm{std}\left(Melt_{SNODAS_{14}}\right)}{\mathrm{std}\left(Melt_{UA_{14}}\right)},$$

$$if\ Melt_{UA_i} > \ max\left(Melt_{UA_{14}}\right) \quad (3)$$

6.3.3 Trend identification and Detrend method

Because stationarity is an underlying assumption of most frequency analyses that develop estimates of extreme values (Khaliq et al., 2006), a trend analysis was conducted for each pixel's time series and time series having significant trends were detrended. The nonparametric Mann-Kendall test was used to identify statistically significant trends (5% confidence level) in the historical annual maximum SWE and snowmelt values (Kendall, 1938; Mann 1945). The Mann-Kendall test is resilient to outliers and can accommodate missing values, which occur frequently in an annual maximum time series of SWE in the more southern portion of the U.S. (Helsel & Hirsch, 1992) The Mann-Kendall statistic, *S*, is calculated as:

$$S = \sum_{k=1}^{n-1} \sum_{i=k+1}^{n} sgn\ (x_i - x_k) \qquad (4)$$

$$sgn\ (x_i - x_k) = \begin{cases} +1 & if\ \ x_i - x_k > 0 \\ 0 & if\ \ x_i - x_k = 0 \\ -1 & if\ \ x_i - x_k < 0 \end{cases} \qquad (5)$$

where x_i and x_k are the annual maximum SWE or snowmelt values that correspond to year *i* and *k*, *i* > *k*, respectively. The standard test statistic Z is computed using a standardization and normal approximation of the statistic S. If a significant trend in an annual maximum time series was found for a given pixel, each value in the time series was detrended using Sen's slope of the trend maintaining the time series' average (Sen, 1968; Fassnacht and Record, 2015) prior to fitting the extreme value distribution.

6.3.4 Generalized Extreme Value (GEV) Distribution

In this study, SWE and snowmelt magnitude-frequency estimates at individual grid cells are computed using the GEV-based frequency analysis approach based on L-moment statistics (Hosking et al., 1985). The GEV distribution is fitted using the annual maximum series of SWE and 7-day snowmelt and then used to estimate the 25- and 100-year return period SWE and 7-day snowmelt design values. The GEV distribution incorporates three types of extreme value distributions, Gumbel, Fréchet, and Weibull distribution. The cumulative distribution function of the GEV can be written as

$$F(x) = \exp\left\{-\left[1 - \frac{\kappa(x-\xi)}{\alpha}\right]^{\frac{1}{\kappa}}\right\} \; for \;\; \kappa \neq 0 \tag{6}$$

where the location parameter (ξ) represents the center of the GEV distribution, the scale parameter (α) specifies the deviation around α, and the shape parameter (κ) determines the tail behavior of the distribution. The Gumbel distribution is obtained when $\kappa = 0$. For $\kappa > 0$, the GEV corresponds to the Fréchet distribution. For $\kappa < 0$, the GEV leads to the Weibull distribution which has a thicker right-hand tail. The parameters of the GEV distribution are

$$\kappa = 7.8590c + 2.9554c^2, \; \mathrm{c} = \frac{2\lambda_2}{\lambda_3 + 3\lambda_2} - \frac{\ln(2)}{\ln(3)} \tag{7}$$

$$\alpha = \frac{\kappa\lambda_2}{\Gamma(1+\kappa)(1-2^{-\kappa})} \tag{8}$$

$$\xi = \lambda_1 + \frac{\alpha}{\kappa}\;[\Gamma(1+\kappa) - 1] \tag{9}$$

where λ_1, λ_2, and λ_3 are the first, second, and third L-moments, respectively, and $\Gamma()$ is the gamma function (Hosking et al., 1985; Stedinger et al., 1993). L-moments provide an alternative approach of describing the shape of probability distributions to conventional moments or maximum likelihood approach (Hosking, 1990). Because sample estimators of L-moments are linear combinations of ranked data, the L-moments are less susceptible to outliers than conventional moments (Vogel and Fennessey, 1993). They are also well suited for analyzing the data with significant skewness.

For pixels having some years with zero values, the zero values were considered to be NA values (so called "censored data") and a conditional probability model was employed. In this case, only the positive, nonzero values were included in the L-moments calculations used to fit the GEV distribution G(x). An additional parameter (P_o) was included to describe the probability that the SWE is zero. For pixels with the censored data, the non-exceedance probability of the GEV distribution was adjusted by a portion of the censored data (Stedinger et al., 1993) such that the unconditional cumulative distribution function F(x) for any value of SWE greater than zero is give an

$$\mathrm{F(x)} = P_0 + (1 - P_0) \cdot \mathrm{G(x)} \tag{10}$$

For example, if there were zero values in nine years out of 36 years, the non-exceedance probability, G(x), was multiplied by a portion of the censored data (1-P_0, 9/36 = 0.75 in this case) and then added P_0. If non-exceedance probabilities for 25- and 100-year return levels are 0.96 and 0.99, and the unconditional non-exceedance probabilities, F(x), are 0.97 and 0.9925. If 50% of more of a pixel's dataset was zero values, then the pixel was excluded from this study.

The goodness–of–fit of the GEV model was examined by the probability plot correlation coefficient (PPCC) test using R package. The PPCC test, developed by Filliben (1975) and

Looney and Gulledge (1985), is based on the moment correlation coefficient between the ordered data and the order statistic medians. If the data are drawn from the hypothesized GEV distribution, the correlation coefficient will be near to one and the plot is expected to be nearly linear (Chowdhury et al., 1991). In this study, the annual maximum SWE and 7-day snowmelt time series data excluding the censored data were tested for each pixel. When a pixel's time series was not rejected as being from the GEV distribution based on the PPCC test with 0.05 significant level, the pixel was used for developing GEV distribution and extreme values with 25- and 100-year return levels. If the GEV distribution was rejected, the extreme SWE and 7-day snowmelt values were excluded in this analysis.

6.4 Results

6.4.1 Comparison of annual maximum SWE and snowmelt between UA and SNODAS

Before developing the design extreme values, the UA-based annual maximum SWE and snowmelt were compared to that of SNODAS for the overlapping period of record (October 2003 to May 2017). **Figure 1** shows maps of the mean annual maximum SWE and snowmelt produced from the UA and SNODAS products. To support this figure, **Table 1** provides statistics of the mean annual maximum SWE and snowmelt by state (spatial mean, standard deviation, 99% quantiles). The two annual maximum SWE products show similar spatial patterns over the CONUS with the SNODAS product. The difference maps reveal that SNODAS annual maximum SWE is generally somewhat higher than the UA annual maximum SWE values in the north central U.S. and non-mountain western U.S. However, in the Great Lakes and northeastern regions, UA annual maximum SWE occasionally exceed SNODAS annual maximum SWE.

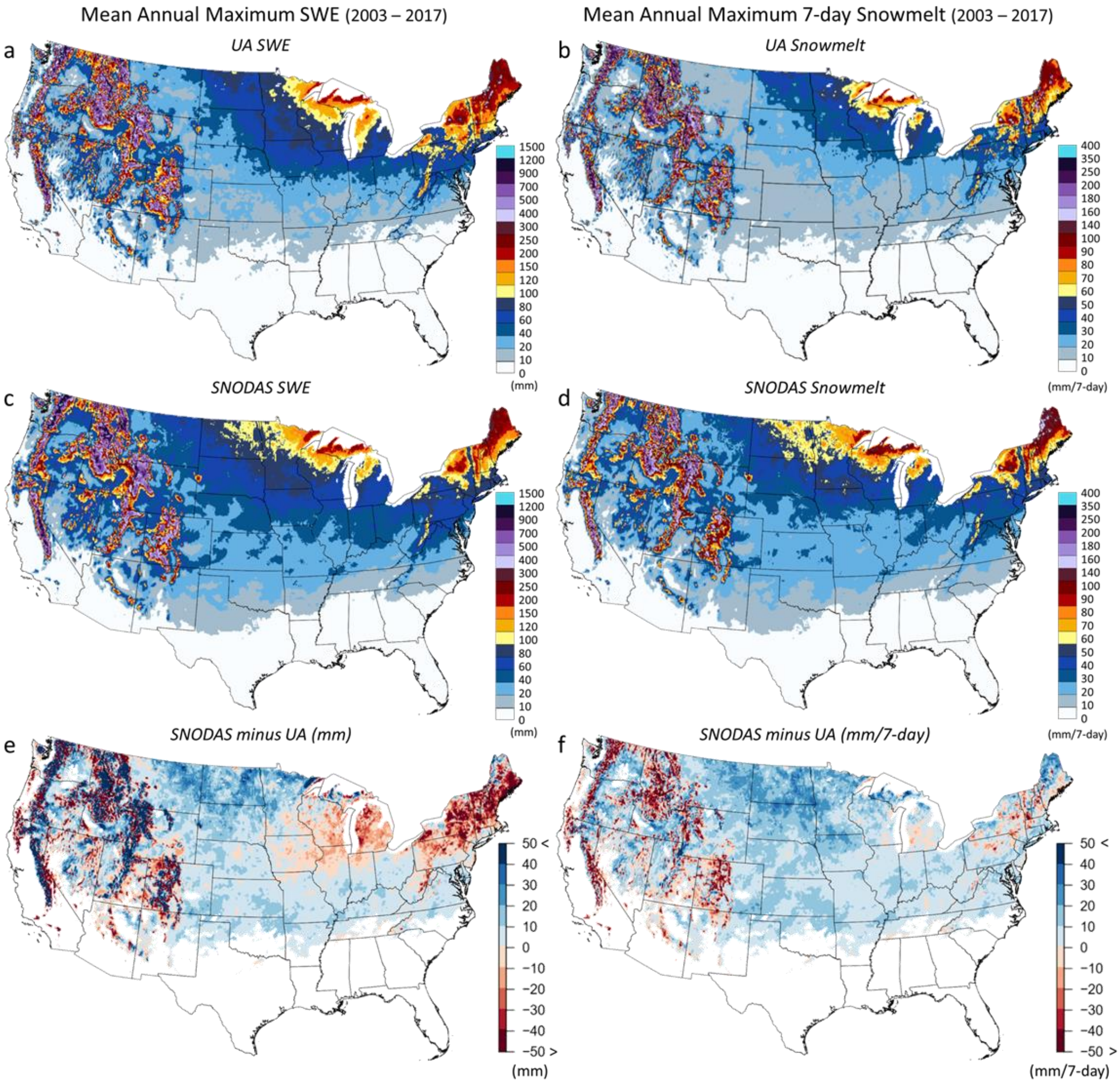


Figure 1. Mean annual maximum SWE and 7-day snowmelt maps from (a, b) UA and (c, d) SNODAS products, and (e, f) mean difference (SNODAS minus UA) maps in the SWE and 7-day snowmelt for the overlapping 14 years from October 2003 to May 2017

In the western U.S., both annual maximum SWE datasets have high mean values and large standard deviations due to the spatially heterogeneous regions. Based on the difference map, the western U.S. differences are mixed especially in the mountainous areas. For example, state-mean values of UA are 189, 206, and 111 mm, and SNODAS, 202, 213, and 121 mm [Standard deviations: 321, 211, and 176 mm for UA and 322, 220, and 179 mm for SNODAS] in

Washington, Idaho, and Wyoming, respectively (**Table 1**). The 99% quantile values, characterizing SWE in those states' high mountain areas, are 1439, 886, and 797 mm from UA product and 1363, 943, and 753 mm from SNODAS. While the state-mean values from UA are slightly lower than the values from SNODAS, the 99% quantile UA values are higher than the SNODAS values in most western U.S. states (e.g. Washington, Wyoming, Montana, Colorado, and Utah).

Table 1. Summary of the mean annual maximum SWE and 7-day snowmelt from UA and SNODAS products from 2003 to 2017 by U.S. states. The states were arranged from largest (top) to smallest (bottom) mean UA Annual Maximum SWE.

State	UA Annual Maximum SWE (mm)			SNODAS Annual Maximum SWE (mm)			UA Annual Maximum 7-day Snowmelt (mm/7 day)			SNODAS Annual Maximum 7-day Snowmelt (mm/7 day)		
	Mean	Std	99%	Mean	Std	99%	Mean	Std	99%	Mean	Std	99%
Idaho	206	211	886	213	220	943	88	69	263	82	50	211
Maine	203	46	337	177	57	289	94	21	167	103	26	148
Washington	189	321	1439	202	322	1363	60	70	260	58	47	202
New Hampshire	167	61	345	132	49	268	83	32	171	80	24	142
Vermont	166	53	312	131	44	235	87	30	173	81	23	131
New York	115	53	280	93	39	216	63	25	141	65	21	129
Wyoming	111	176	797	121	179	753	51	53	216	56	44	192
Colorado	111	160	722	107	137	592	52	51	207	50	37	159
Montana	106	184	929	118	175	855	48	59	253	54	40	192
Oregon	99	175	918	93	141	722	43	50	238	44	34	154
Michigan	99	43	219	89	48	228	56	20	108	62	27	132
Massachusetts	97	31	178	78	23	153	52	18	105	54	11	90

Utah	83	111	532	89	119	559	45	44	194	46	39	171
Wisconsin	80	20	137	74	22	140	49	13	85	57	16	101
Minnesota	73	21	162	83	26	176	46	12	89	61	13	100
Connecticut	71	20	139	68	9	99	40	9	69	51	6	70
California	67	175	894	67	172	861	28	58	262	24	44	186
Rhode Island	64	13	84	61	10	75	38	6	49	48	5	55
Pennsylvania	63	21	114	55	13	88	38	12	72	42	9	68
North Dakota	52	12	74	70	12	94	36	9	52	56	9	75
New Jersey	47	17	96	43	13	69	31	8	57	36	9	50
Iowa	46	11	70	45	11	71	27	5	38	38	9	58
Nevada	43	55	277	43	52	250	28	27	134	30	21	109
West Virginia	41	27	142	40	18	99	29	14	88	33	12	73
South Dakota	37	12	71	49	14	84	26	6	46	43	9	67
Maryland	36	20	118	40	15	96	25	10	65	33	10	71
Ohio	34	11	75	38	7	68	24	5	42	32	5	53
Indiana	32	10	60	35	5	49	23	5	37	30	4	40
Illinois	31	13	61	32	6	51	21	6	36	28	4	41
Nebraska	26	5	39	33	5	44	20	3	27	30	4	39
Delaware	25	6	43	30	5	41	20	4	29	25	3	34
Virginia	24	9	57	29	8	48	19	6	40	26	6	40
New Mexico	22	4	34	28	3	34	15	19	106	19	16	88
Missouri	22	41	212	25	38	209	19	2	23	25	3	31
Kansas	20	3	28	28	3	34	16	2	21	26	2	32

Kentucky	20	4	30	26	4	34	18	3	25	24	4	31
Arizona	15	31	159	15	28	149	10	18	92	11	15	75
Tennessee	12	5	25	17	6	30	11	4	22	14	4	28
Arkansas	12	4	26	15	5	29	12	5	23	17	6	28
Oklahoma	11	4	21	21	5	29	10	3	18	21	4	29
North Carolina	11	7	42	16	7	33	10	6	32	16	6	29
Texas	4	3	13	8	5	19	4	4	15	7	7	24
South Carolina	4	4	17	7	7	25	4	3	12	8	5	19
Mississippi	3	2	12	5	3	11	3	2	11	5	3	11
Alabama	3	2	9	4	3	10	3	2	8	4	3	10
Georgia	2	3	13	4	5	21	2	3	12	4	5	20
Louisiana	1	1	4	3	2	7	1	1	4	3	2	7
Florida	0	0	1	0	0	1	0	0	1	0	0	1

In the northeastern regions, UA annual maximum SWE is much higher than that from SNODAS with UA's state-mean SWE values of 203, 167, and 166 mm as compared to SNODAS values of 177, 132, and 131 mm, for Maine, New Hampshire, and Vermont, respectively. Not surprisingly, the standard deviations in the northeast are much lower than the western regions with UA and SNODAS having similar variability (UA: 46, 61, and 53 mm and SNODAS: 57, 49, and 44 mm). The SNODAS 99% quantile values, 289, 268, and 235 mm, are similar to the state-mean values indicating that there is relatively little spatial variation in the eastern U.S.

In the north central U.S., the annual maximum SWE from both products are relatively low compared to the other regions. The state-mean values of SNODAS range from 49 to 89 mm with the 99% quantile values from 84 to 228 mm, while UA state-mean values ranges from 37 to 99 mm and 99% quantile values range from 71 to 219 mm. The SNODAS annual maximum SWE is generally higher than UA product in these regions (e.g. Northern Great Plains).

The annual maximum 7 day-snowmelt spatial patterns are similar to the annual maximum SWE patterns where the largest snowmelt events occur in the western mountainous regions, followed by northeastern, then the north central U.S. SNODAS annual maximum snowmelt values are generally higher than UA values in the non-mountainous areas over the CONUS. UA snowmelt is much higher than SNODAS in the western mountainous areas where the differences between UA and SNODAS snowmelt values reach up to 300 mm.

At a state level, the northeastern states have the largest state-mean snowmelt (e.g. UA: 94, 87, and 83 mm; SNODAS: 103, 81, and 80 mm for Maine, Vermont, and New Hampshire, respectively). In the north central U.S., the state-mean SNODAS snowmelt values, 61, 56, and 43 mm, are consistently larger than the UA snowmelt values, 46, 36, and 26 mm for Minnesota, North Dakota, and South Dakota, respectively. The 99% quantile values have similar differences between two products to the state-mean snowmelt values in the regions (11, 23, and 21 mm for the states, respectively). In the western U.S., the average values by state are similar but the extremes differ. While the state-mean values of UA snowmelt, 88, 60, and 51 mm, are slightly larger than SNODAS, 82, 58, and 56 mm, for Idaho, Washington, and Wyoming, the 99% quantile values are UA snowmelt values are consistently much larger than SNODAS values (UA: 263, 260, and 216 mm; SNODAS: 211, 202, and 192 mm, respectively).

Figure 2's agreement statistics (R-value) of annual maximum SWE and snowmelt between the two data products reflect the consistency between the products' low and high snow and snowmelt years from 2004 to 2017. For the annual maximum SWE, the spatial R-value map typically shows strong agreement. In the north-central and northeastern U.S., the correlation is relatively high (> 0.8) compared to the other regions. There is less agreement over parts of Montana, Wyoming, Nebraska, and Kansas, as well as the western Mountainous regions. The annual maximum snowmelt correlation patterns are similar, but the agreement is not as strong as that for the annual maximum SWE comparison. The Great Lakes and Northeast are notable for their poor agreement between annual maximum snowmelt values that were not evident in the annual maximum SWE agreement.

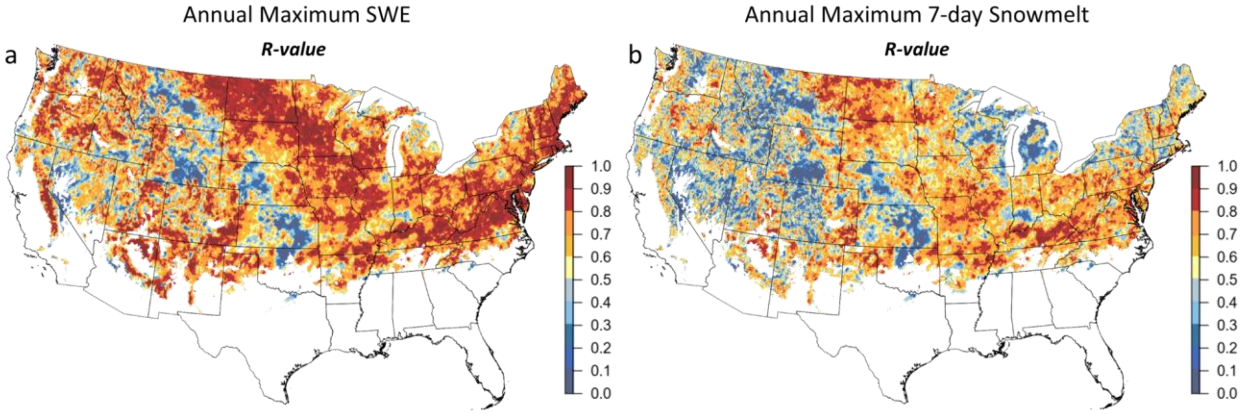


Figure 2. Pearson correlation (R-value) maps of the annual maximum SWE and 7-day snowmelt between SNODAS and UA products for the overlapping 14 years from October 2003 to May 2017

6.4.2 The annual maximum SWE and 7-day snowmelt trends

To demonstrate how the ECDF-transformations update the original UA data, an annual time series of the annual maximum 7-day snowmelt from the original UA and the ECDF-transformed UA data are shown for eight states (**Figure 3;** four western states, two north central states, and two northeastern states in the U.S.). In North Dakota the original UA snowmelt is much lower than the SNODAS snowmelt in 2009, 2010, and 2011, but the ECDF-transformed UA snowmelt maintains the interannual variations that are consistent with the SNODAS snowmelt. In New York, the SNODAS snowmelt is higher than the original UA snowmelt. The ECDF-transformed UA snowmelt has greater interannual variation compared to the original UA snowmelt.

The total amount of annual maximum snowmelt is calculated by multiplying the spatial mean annual maximum snowmelt of significant pixels (mm/7-day) by the total area of the pixels. In some cases, the ECDF-transformed UA resulted in large differences in the water yield over the 7-day annual maximum as compared to the original UA. For example, the water yield from snowmelt in North Dakota over the 36-year period from the ECDF-transformed UA melt is nearly double (2,746 x 10^6 m^3) that of the original UA yield of 1,494 x 10^6 m^3. Similarly, in Maine, the water yield from the original and ECDF-transformed UA snowmelt are 283 x 10^6 m^3 and 538 x 10^6 m^3, respectively. There are a few states that the ECDF-transformed UA resulted in smaller differences in the water yield. In Montana, the water yield from the transformed UA snowmelt decreases by 586 x 10^6 m^3 over the 36-year period while decrease in the water yield from the original UA-based snowmelt was 943 x 10^6 m^3.

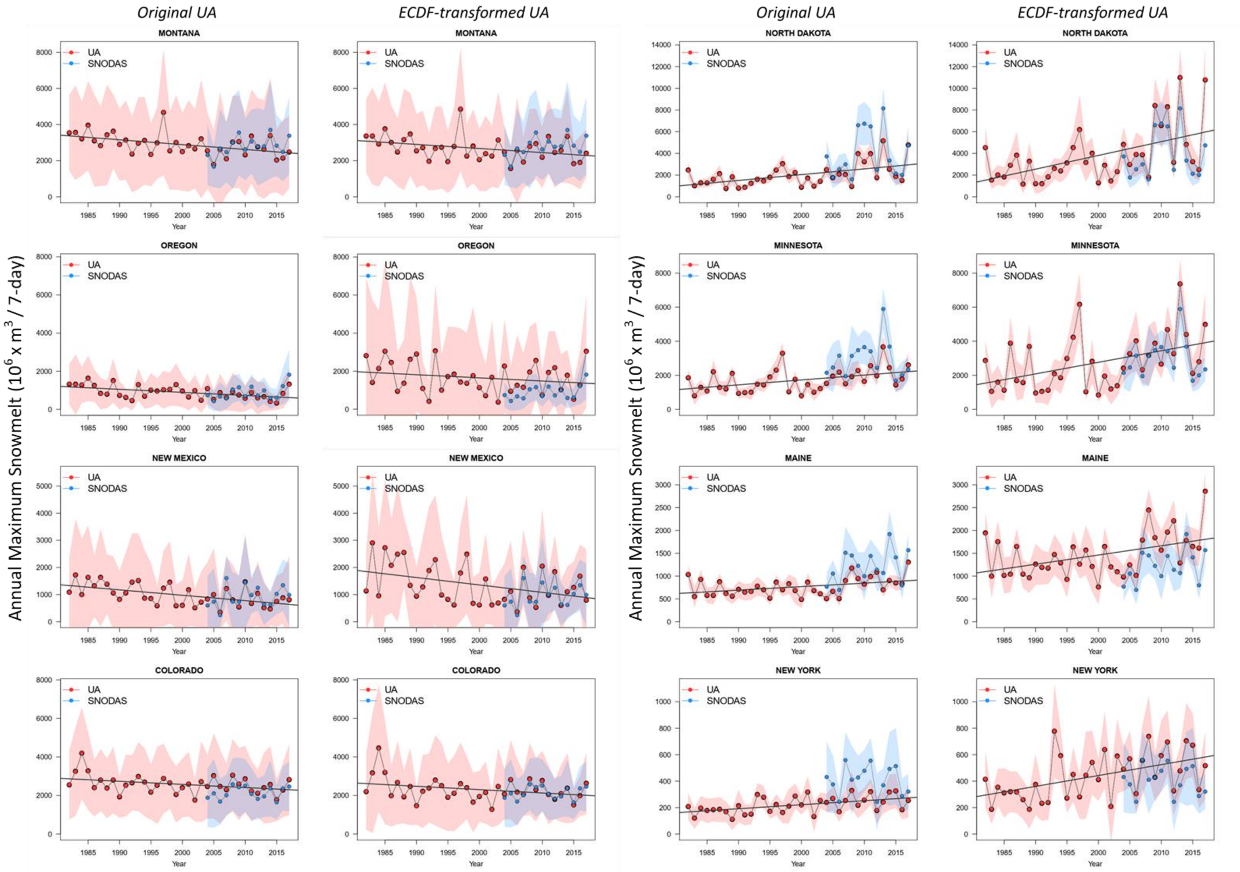


Figure 3. Examples of annual maximum 7-day snowmelt time series for eight U.S. states based on the original UA and the ECDF-transformed UA data using SNODAS from water years 1982 to 2017. Each point is the spatial mean of the annual maximum 7-day snowmelt for pixels with significant trend only (p-value < 0.05) multiplied by a total area of the pixels in a state, and the light-colored range shows the mean plus (upper boundary) or minus (lower boundary) one spatial standard deviation multiplied by a total area of the significant pixels. The solid line is the linear trend line for the 36-year period.

A trend analysis of the UA annual maximum SWE and 7-day snowmelt values was conducted for each pixel using the Mann-Kendall test to identify significant trends (p-value < 0.05) from 1982 to 2017. Figure 4 shows the results for both the original and ECDF-transformed UA data, which is merged with SNODAS annual maximum data using ECDF method, to identify similarities/differences of the spatial trend maps. There are negative trends in most mountainous regions of western states (**Figures 4a & b**). For several states, nearly 10% of the state had significant annual maximum SWE decreases including New Mexico (16.1%), Colorado (12.9%), and Nevada (9.8%). The greatest state-average annual maximum SWE decrease rates were -3.3 mm/year in Colorado followed by Utah (-2.8 mm/year), Idaho (-2.5 mm/year), and Oregon (-2.4 mm/year). Positive trends in annual maximum SWE were rare, but present in 5.1% of North Dakota and 1% of Minnesota. In North Dakota and Minnesota, 29% and 9.7% of the states had statistically significant snowmelt increases (0.66 and 0.68 mm/7-day/year), respectively. Maine snowmelt also increased at an even higher rate 0.94 mm/7-day/year.

Trend maps using the ECDF-transformed UA differ slightly from the maps using the original UA product (**Figures 4c & d**). In the mountainous western U.S., the area with significant decreases in both annual maximum SWE and snowmelt was generally reduced. However, the annual maximum SWE and snowmelt trends strengthened in the north central and northeastern regions. Considering that SNODAS SWE and snowmelt values were generally larger than UA products for those regions, the transformed UA values may be larger than original UA values leading to increases in the trends' magnitude.

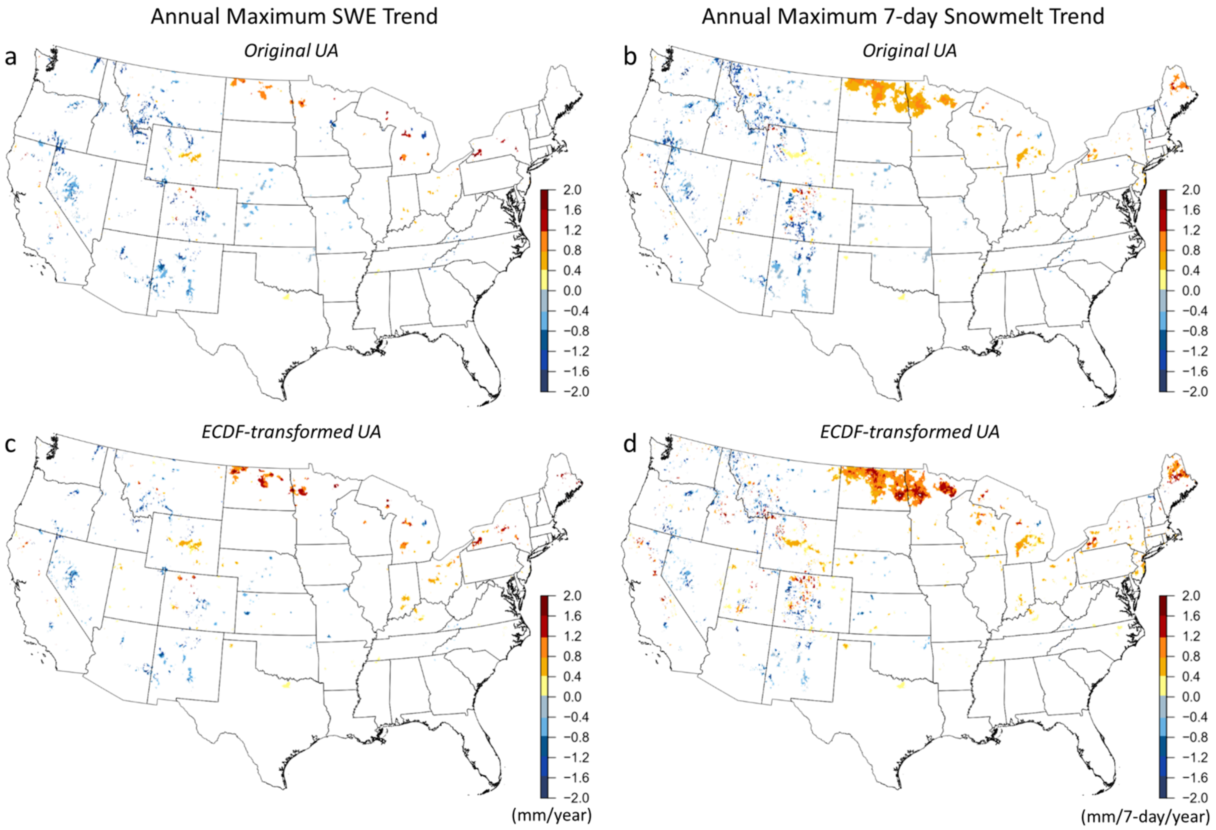
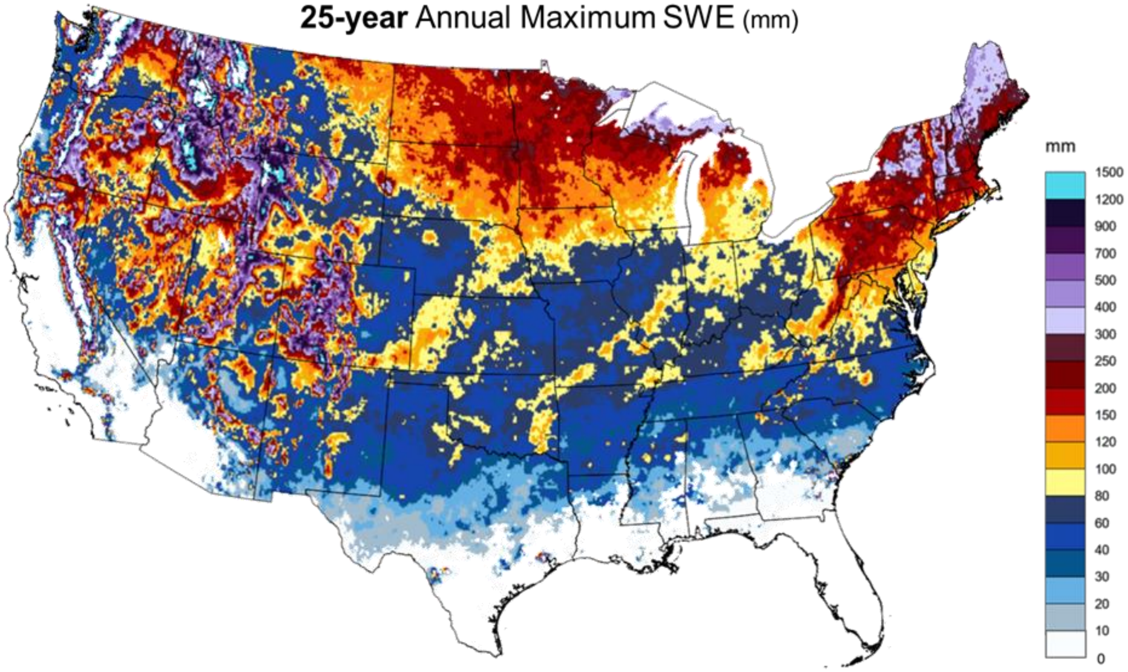


Figure 4. Trend maps of the annual maximum SWE and 7-day snowmelt from the original UA and ECDF-transformed UA data using nonparametric Mann-Kendall test with Sen's slope from 1982 to 2017. Only pixels where there is a significant positive/negative trend ($p < 0.05$) are shown.

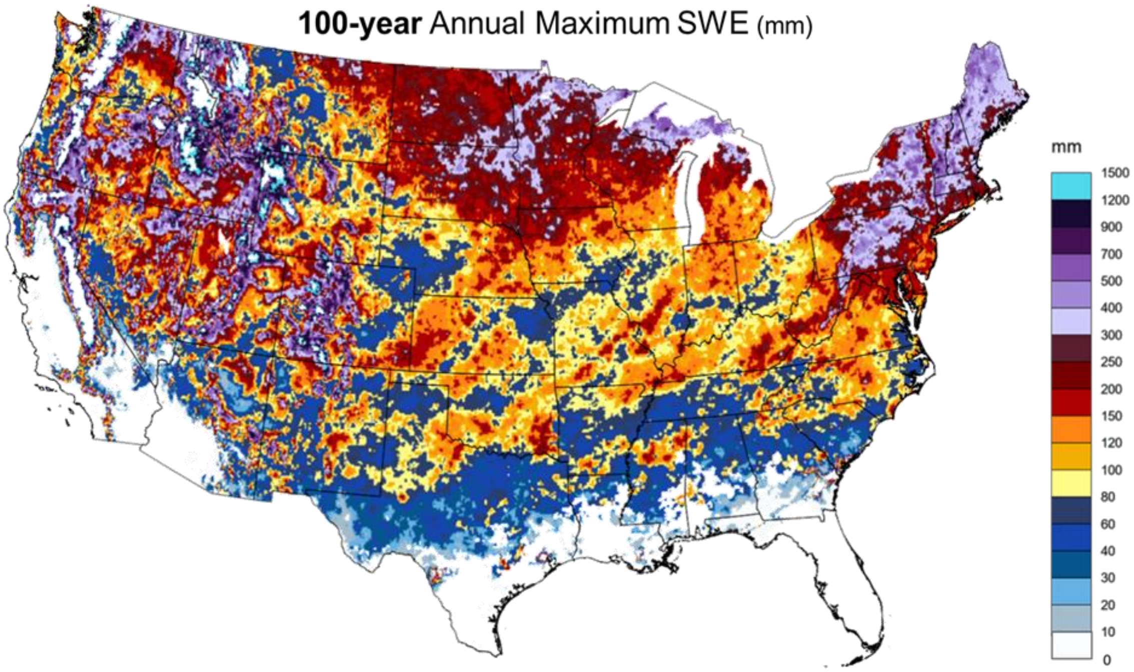
6.4.3 Design SWE and snowmelt maps over the CONUS

The spatial patterns of the 25- and 100-year annual maximum SWE values are similar but differ in magnitude (**Figures 5**). The largest design SWE values, defined here as top 1% (99% quantile) values for each state, were found in western mountainous regions. For example, in Washington the top 1% 25- and 100-year return level SWE values are 2,713 mm and 3,542 mm, respectively. In the northeastern U.S., the top 1% SWE values, 25-year values range from 443 to 639 mm, and, for 100-year value from 627 to 905 mm, are much lower than those in the mountainous western U.S. However, the state-median values in the northeastern U.S. are higher than the western U.S. states. In the north-central U.S., the greatest design SWE values occur in the areas near Lake Michigan and the Red River of the North Basin in Minnesota. The top 1% values are 452 and 346 mm for the 25-year return level (654 and 512 mm for 100-year return level) in Michigan and Minnesota, respectively.

The 25- and 100-year snowmelt maps generally have similar spatial patterns to those of the annual maximum SWE maps (**Figure 6**). Regions with high SWE are likely to also have large snowmelt events. The largest 7-day snowmelt values (top 1% for state) for 25 and 100-year return levels were also found in the western U.S. The top 1% snowmelt values for 25-year return level range from 396 to 541 mm, and, from 611 to 843 mm for 100-year return level over the western mountainous regions including Washington, California, Idaho, and Oregon. In the northeastern U.S., the magnitudes of the 25- and 100-year snowmelt are lower than that of the western U.S. The northeastern's top 1% values for 25-year range from 236 to 382 mm and, for 100-year, from 347 to 563 mm over the northeastern U.S. states. The design snowmelt values in the north-central U.S. are typically lower than that of the northeastern region. For 25- and 100-year, the top 1% snowmelt values are from 176 to 262 mm and from 311 to 397 mm, respectively, over the north-central U.S. states including Michigan, Minnesota, South Dakota, and North Dakota.



25-year Annual Maximum SWE (mm)



100-year Annual Maximum SWE (mm)

Figure 5. 25- and 100-year return level design SWE maps using the detrended, ECDF-transformed annual maximum SWE.

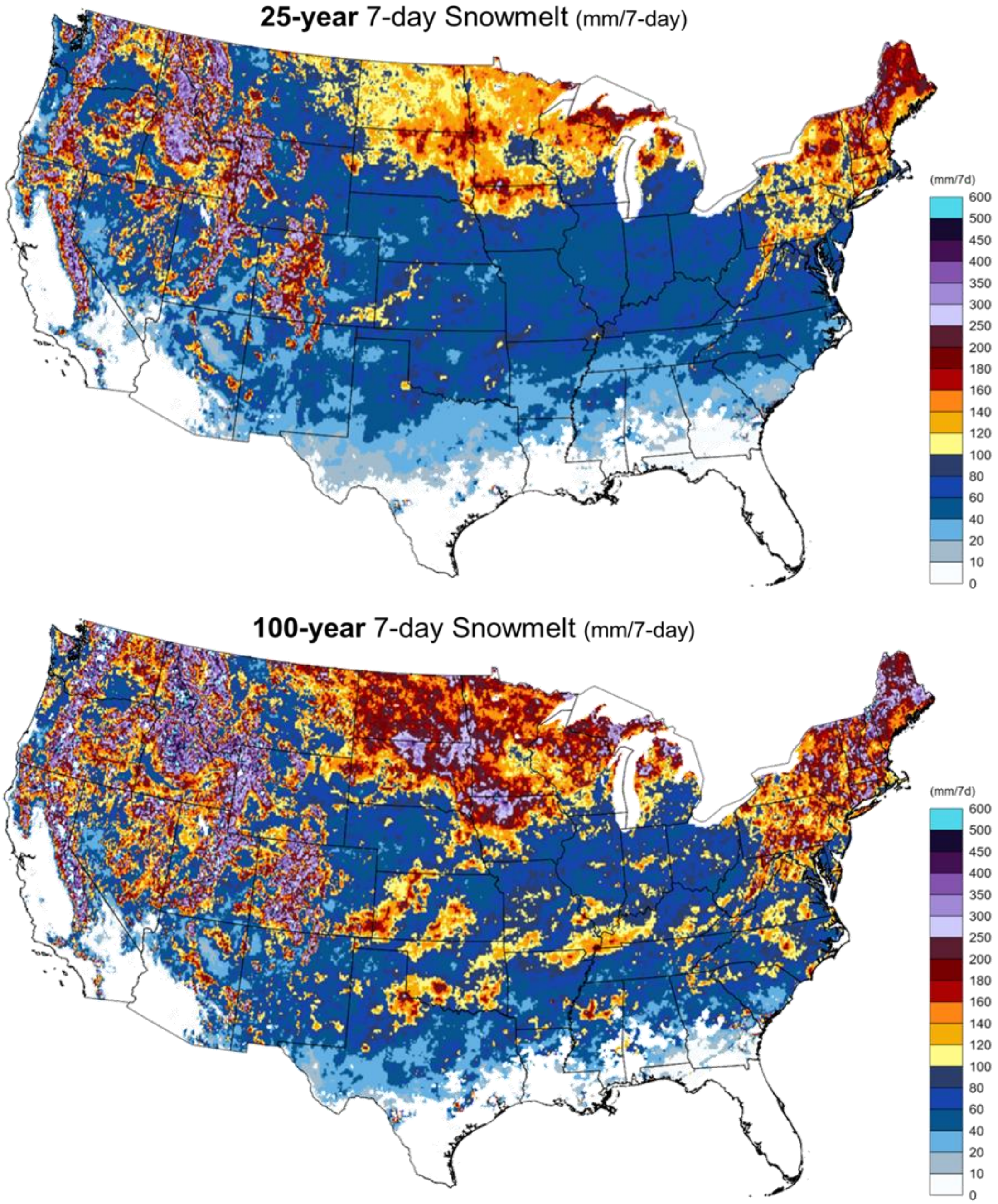


Figure 6. 25- and 100-year return level design 7-day snowmelt maps using the detrended, ECDF-transformed annual maximum snowmelt.

6.4.4 Design snowmelt versus NOAA Atlas 14 precipitation

To understand the contribution of snowmelt to runoff, the current snowmelt results were considered in light of the NOAA Atlas 14 7-day duration precipitation maps for 25- and 100-year recurrence levels in the northeastern U.S. region (**Figure 7**). With notable exceptions, the design precipitation values are typically considerably larger than the snowmelt values. In the mountain western U.S. and north central U.S., snowmelt can exceed precipitation. In the mountain western U.S., this typically occurs in regions where the 7-day precipitation values are relatively low, while the snowmelt is extremely high. The 25- and 100-year difference maps consistently identify those parts of the western U.S. where design snowmelt values are higher than design precipitation values. In the northern Great Plains and upper Midwest, design

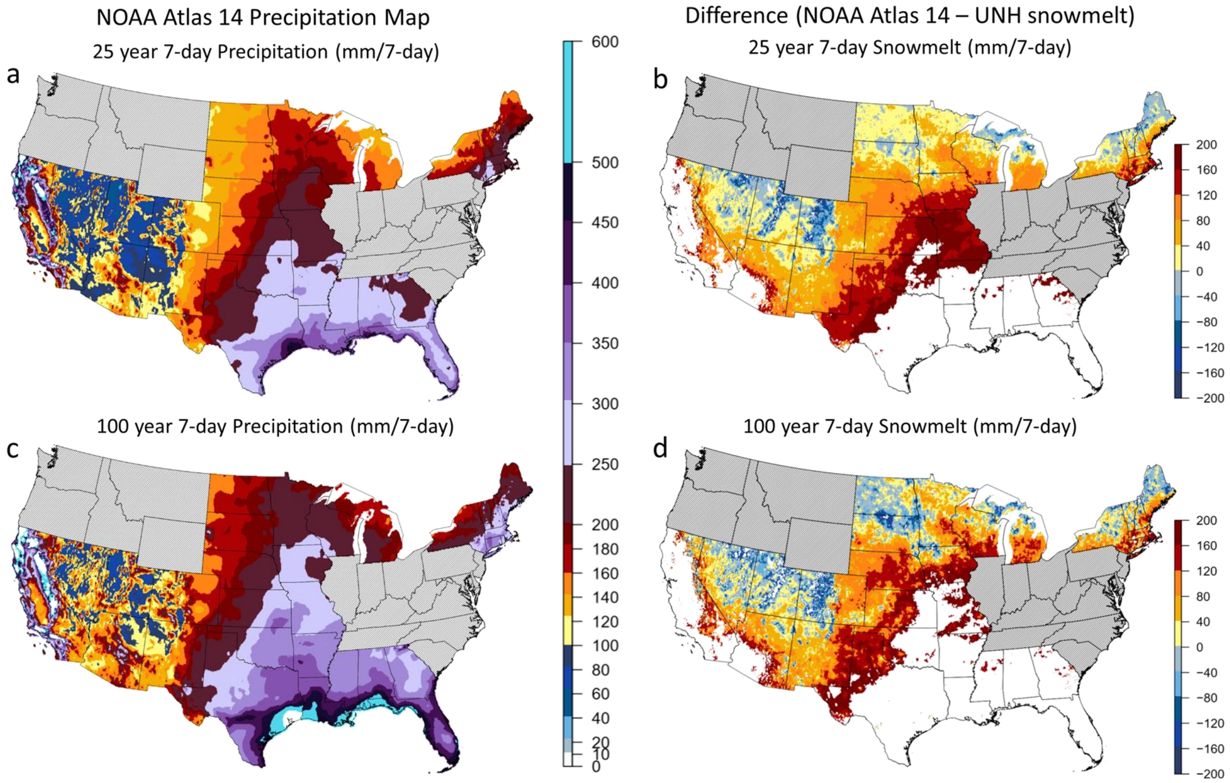


Figure 7. (a, c) The NOAA Atlas 14 25- and 100-year 7-day precipitation maps, (b, d) the difference maps (Atlas 14 minus corresponding snowmelt maps) over the CONUS. Cool colors indicate regions where the snowmelt values exceed the Atlas 14 precipitation values. The gray regions indicate U.S. states where annual maximum Atlas 14 data are not available. The white areas are out of range.

snowmelt and design precipitation values are typically similar for 25-year return periods, but the 100-year snowmelt design values can exceed the precipitation values in many parts of those regions. The inland portion of the Northeastern states also has some regions where snowmelt values are modestly higher than precipitation values.

A more detailed comparison of design snowmelt and design precipitation values was conducted for two regions that are historically vulnerable to snowmelt-driven floods (Berghuijs et al., 2016; Stadnyk et al., 2016; Slater and Villarini, 2016), the Northeastern states (Atlas 14 Volume 10; **Figure 8**) and Midwestern states (Atlas 14 Volume 8; **Figure 9**). In the Northeast, the Atlas 14 design precipitation values gradually decrease from the southeast to the northwest with the highest values in western Massachusetts and Connecticut as well as around Mount Washington in New Hampshire (**Figures 8a & d**). For the 25-year and 100-year return periods, the precipitation depths in the southeastern areas range from about 200 to 280 mm and 250 to 370 mm, respectively, while in the northwestern areas precipitation values range from 120 to 200 mm and 250 to 370 mm for the 25-year and the 100-year values, respectively. Although the snowmelt maps have larger spatial variability than the Atlas 14 precipitation maps (**Figures 8b & e**), design snowmelt values, as well as maximum SWE values, gradually increase from the southeast to the northwest. These design snowmelt gradients are perpendicular to the design precipitation gradients. The combination of higher SWE and more moderate design rainfall amplifies the importance of snowmelt driven runoff in northern Vermont and Maine near the Canadian border, where design snowmelt can exceed design rainfall by up to 140 and 200 mm for the 25- and 100-year return periods, respectively.

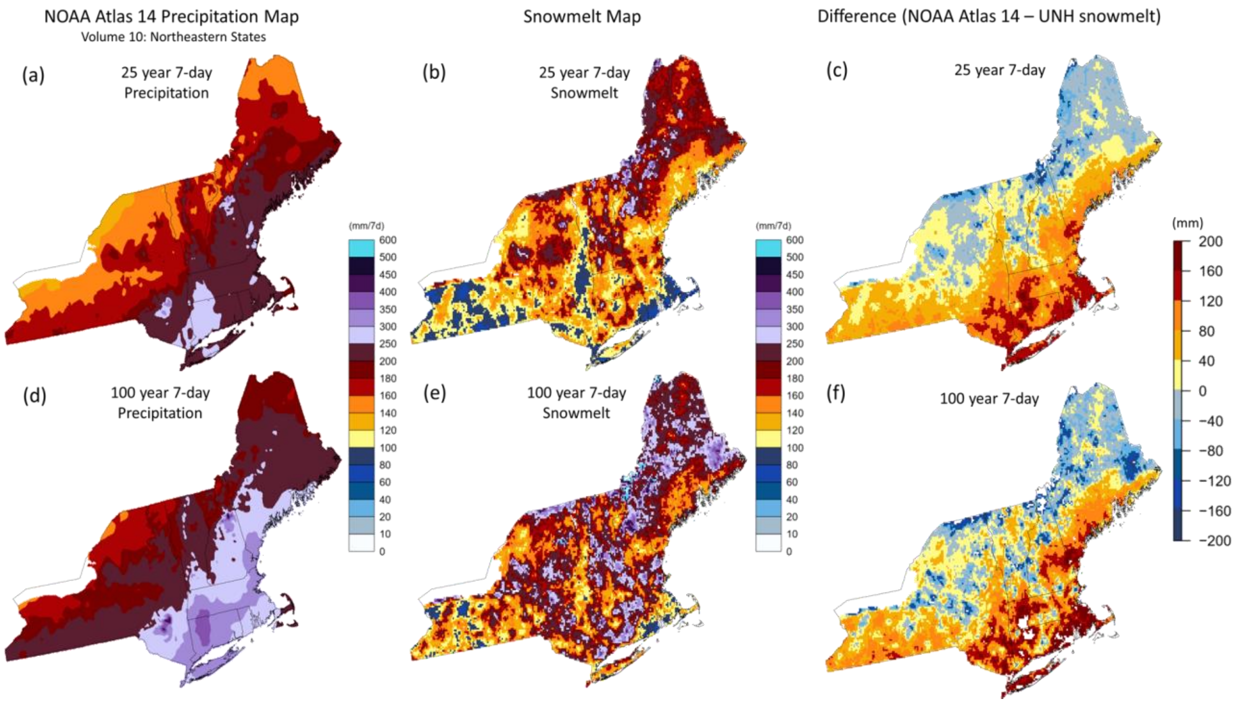


Figure 8. (a, d) The NOAA Atlas 14 25- and 100-year 7-day precipitation maps, (b, e) the corresponding snowmelt maps, and these differences (Atlas 14 minus snowmelt) over the northeastern U.S. including 7 states (Massachusetts, New York, Vermont, New Hampshire, Maine, Rhode Island, and Connecticut).

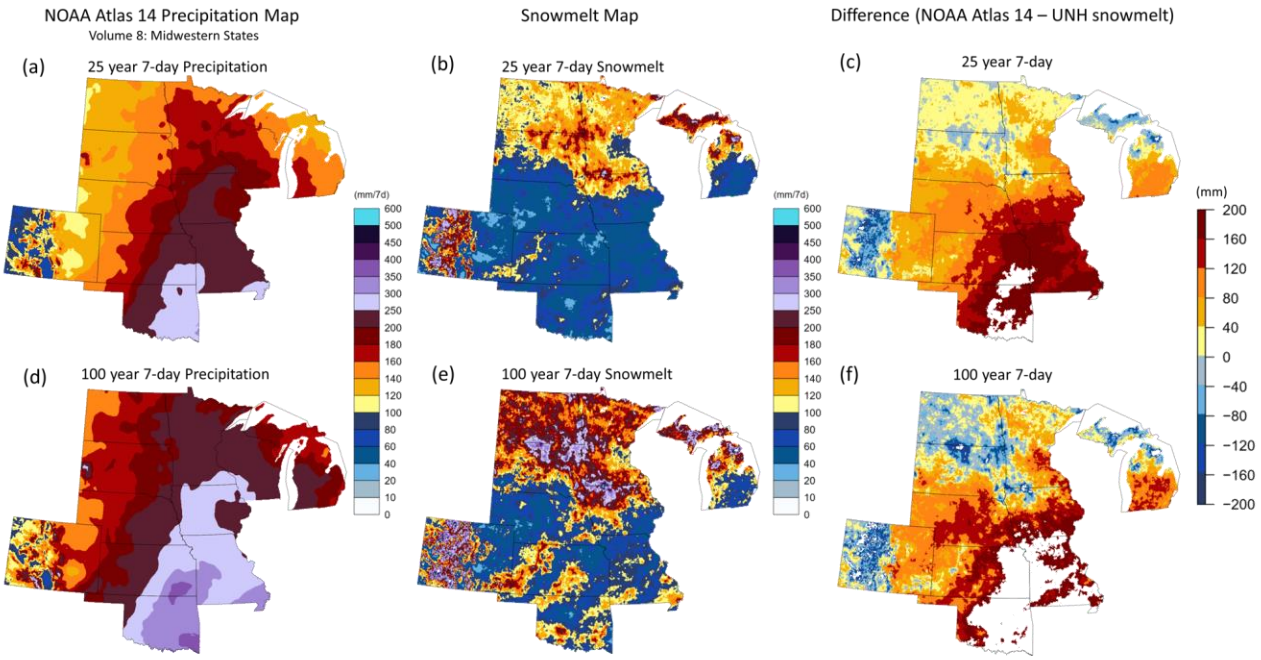


Figure 9. Same as Figure 8, but in the Midwest U.S. including 7 states (North Dakota, Minnesota, South Dakota, Nebraska, Colorado, Iowa, Missouri, Michigan, Kansas, and Oklahoma)

In the Midwest U.S., the NOAA Atlas 14 precipitation values are typically higher than the snowmelt values (**Figure 9**). In this region, the precipitation intensities decrease considerably to northwest with ranges from 60 to 140 mm (25-year) and 80 to 230 mm (100-year). Western Colorado has the lowest design precipitation values. The snowmelt maps in this region largely increase from south to north. In locations with localized extremes at North and South Dakotas' eastern borders, the border between Iowa and Minnesota, and Michigan's Upper Peninsula, the snowmelt magnitudes are higher than precipitation depth by up to 160 and 240 mm for the 25- and 100-year recurrent intervals, respectively (**Figures 9e & f)**. In the Rocky Mountain region of the western Colorado, the snowmelt magnitudes can be markedly higher than the precipitation with difference up to 190 and 310 mm for the 25- and 100-year return periods, respectively.

6.5 Discussion

6.5.1 Comparison between UA and SNODAS data

To date, no studies have directly compared the UA and SNODAS products, because the UA SWE dataset was only recently publicly released (March 2019) via National Snow and Ice Data Center (Broxton et al., 2019). Broxton et al. (2016b) showed that the reanalysis and GLDAS SWE products have considerably lower SWE than both UA and SNODAS SWE data. Their finding that SNODAS generally has higher SWE (and annual maximum SWE for 2008) than the UA product in Washington and Idaho as well as northern Great Plains, primarily for regions where the SWE is low, is consistent with our results. Dawson et al.'s (2017) study of snow density parameterization for developing UA SWE product found that SNODAS snow densities have low biases compared to SNOTEL observations, particularly in ephemeral and maritime classes (11.1% and 16.2% of relative mean absolute error, respectively). They assert that these biases could be due to the assimilation of snow depth and SWE observations across different platforms and spatial scales (i.e., airborne gamma SWE, point-based snow depth and SWE observations, and satellite-based snow covered area) without ingesting snow density observations.

Our findings that SNODAS snowmelt exceeded UA snowmelt over the north central and the western non-mountain U.S. may stem from the larger SNODAS annual maximum SWE compared to UA annual maximum SWE. The increased availability of SWE for melt may generate large snowmelt during the spring. However in most western mountain regions, the UA annual maximum snowmelt exceeded the SNODAS despite the larger SNODAS SWE (**Figure 10**). Differences in the UA and SNODAS products' snow ablation procedures are likely responsible. The UA product calculates accumulated snow ablation using a simple cumulative degree day above 0 °C method during the snow-covered period and generates SWE by subtracting accumulated snow ablation from accumulated snowfall (Broxton et al., 2016a). In contrast, SNODAS estimates snow ablation using the snow thermal model (SNTHERM.89) which tracks two variables, liquid water in the snowpack (state variable) and melt runoff rate at the base of the snowpack (diagnostic variable) (Barrett, 2003; Carroll, 2001). SNODAS considers both solid water and the liquid water in the snowpack as SWE. Thus, it is possible that UA snowmelt may be underestimated, in part, due to the lack of liquid water storage in the snowpack.

In the mountainous western U.S., the negative trends in annual maximum SWE correspond with numerous previous findings showing declines in SWE since 1980 (Mote et al., 2018; Pederson et al., 2013; Pierce et al., 2008). In the north central U.S., there were limited previous findings probably due to the lack of the ground-based long-term SWE measurements over the regions (Cho et al., 2019). Our finding of considerable increases in the annual maximum snowmelt in North Dakota and Minnesota may be related to dramatic increases in spring streamflow and a frequency of flood events found by Hirsch and Ryberg (2012) and Peterson et al. (2013). Hirsch and Ryberg (2012) suggested that the increases in streamflow might be due to a large increase in precipitation since 1980 as a result of changing climate. This indicates that the increase in the amount of snowmelt in a short period is likely due to the increases in precipitation as snowfall and springtime warming (Kunkel et al., 2013), resulting in the amplified streamflow and frequent spring floods. Projected winter precipitation (2070 – 2099) is expected to increase by over 20 % relative to the current precipitation (1976 – 2005) in this region (Easterling et al., 2017). The positive trend in snowmelt could last and intensify the frequency and severity of snowmelt floods for future decades.

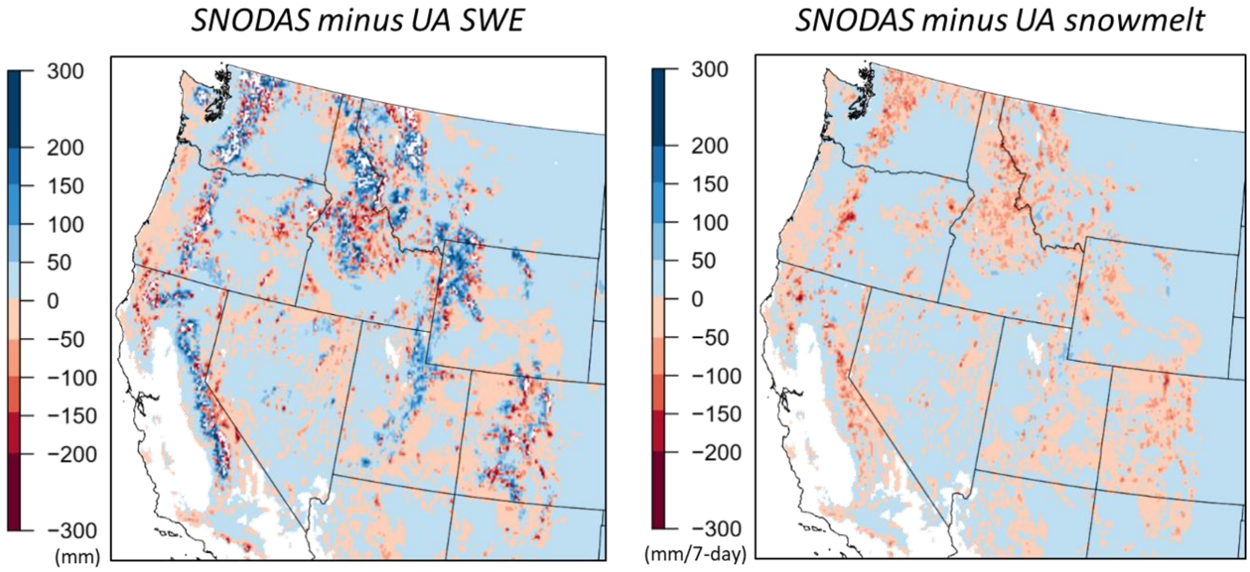


Figure 10. Same as Figure 1e & f (Mean difference maps of the annual maximum SWE and 7-day snowmelt), but changed the color bars ranging from -300 to 300 mm. Blue (Red) color areas indicate that SNODAS is larger (smaller) than UA product.

6.5.2 Comparison to the NOAA Atlas 14 and the National Engineering Handbook

The current NOAA Atlas 14 Volumes provides very limited guidance about snowmelt events (Perica et al., 2015, revised 2019). They focus on liquid precipitation (rainfall). Snowmelt events were not considered in the frequency estimates because this is a different runoff production mechanism. Unlike liquid precipitation, snowfall goes into storage (i.e., snowpack) and delays runoff until it melts. Our snowmelt estimates from this study complement the rainfall-based frequency estimates from the NOAA Atlas 14 series. From an engineering design perspective, the Atlas 14 design precipitation values are reasonable in regions where the Atlas 14 precipitation estimates equal or exceed our snowmelt estimates. However, in the areas where our snowmelt estimates are larger than the Atlas 14 precipitation values, infrastructure designed without consideration of snowmelt floods may underestimate design runoff and undersize structures. The maps presented in this study can provide guidance when developing flood defense structures in areas where major snowmelt floods have occurred over the past few decades (Todhunter, 2001; Changnon et al., 2001).

The National Engineering Handbook from United States Department of Agriculture Natural Resources Conservation Service (USDA NRCS) mapped snowmelt runoff volumes in the eastern Montana region (USDA, 2004). Their 25-year, 7-day snowmelt runoff volumes ranged from 10 to 45 mm, which are much lower than the range of our snowmelt maps (40 to 140 mm) for the same region. The handbook also provided an example of a 25-year maximum March 16–31 SWE map in the north central U.S., which was taken from a technical paper from U.S. Department of Commerce Weather Bureau (U.S. Department of Commerce, 1964). Their SWE values, ranging from 38 mm in the eastern Nebraska and South Dakota) to 305 mm near Lake Michigan, are somewhat lower than our SWE values for the same region (40 – 450 mm).

The differences between these earlier maps are probably due, at least in part, to different time periods and datasets. Those maps used SWE or snow depth data obtained from several snow station networks with 15 to 55 year records (U.S. Department of Commerce, 1964) with a typical measurement period of 35 years from 1930 to 1964 as compared to UA's 1981 to 2017 record. There are not only differences in the SWE magnitude, but also the SWE spatial patterns. While the earlier maps showed that SWE gradually increased from southwestern (Nebraska and South Dakota) to northeastern parts of the region (e.g. Wisconsin and Michigan), our map shows very high SWE values in the headwaters of the Red River of the North Basin (near western Minnesota and southeastern North Dakota). This could be due to data limitations during the earlier time periods or to regional changes in snowfall under the changing climate (Hirsch and Ryberg, 2012; Byun et al., 2018). While there is little documentation about long-term SWE increases in that region, there is consensus that the Red River of the North Basin has experienced more frequent snowmelt flooding since the late 1990s (e.g. 1997, 2006, 2009, 2011 and 2019; Tuttle et al., 2017).

6.5.3 Limitations

Despite our efforts to combine the UA and SNODAS SWE products, regions where the annual maximum values from the two products have clear differences suggest inherent uncertainties. In this study, SNODAS SWE is assumed to be the most reliable data source over the CONUS based on previous studies considering (Broxton et al., 2016a, 2016b; Vuyovich et al., 2014). Furthermore, the SNODAS data have been widely vetted by the NWS regional river forecast centers for their use in operational flood forecasting (Barrett, 2003). However, recent studies have found that SNODAS SWE or snow depth has reduced performance in some regions as compared to independent data. Clow et al. (2012) showed that SNODAS SWE had relatively poorer agreement ($R^2 = 0.30$) with in-situ snow surveys in alpine areas, while SNODAS SWE performed well in forested areas ($R^2 = 0.77$) in Colorado Rocky Mountains, U.S. They indicated that wind redistribution of snow in alpine terrain may not be fully considered in a snow model in SNODAS, even though the SNODAS model is run with surface zonal wind as a driving variable (Barrett, 2003). While not yet studied, because the interpolation and assimilation processes for UA SWE do not account for wind effects, it is reasonable that wind redistribution would also impact the UA SWE estimates. Anderson (2011) also found that SNODAS underestimated snow depths in forested alpine terrain. Hedrick et al.'s (2015) evaluation of SNODAS snow depth using lidar-based snow depth measurements during the 2007 Colorado Cold Lands Processes Experiment (CLPX-2) also found regional differences between the two snow depth products, especially in the areas with dense low sagebrush where high winds scour the snow throughout the winter. Boniface et al. (2015), conducting a comparison study of SNODAS snow depth with the Global Positioning System Interferometric Reflectometry (GPS-IR)-based snow depth observations over the western U.S., found that there were clear differences between the two snow depth products in areas with complex terrain or strong vegetation heterogeneities.

In the upper Tuolumne River Basin in the California's Sierra Nevada, comparison results of the ASO SWE with SNODAS SWE showed that SNODAS overestimated SWE during the melt phase (Bair et al. 2016). However, they could not determine the mechanism that caused the errors because there are few publications that address the details of the snow model structure and assimilation scheme in SNODAS. Dozier et al. (2016) suggested a potential cause for the overestimates is that the snow pillows, whose measurements are assimilated in SNODAS, could hold more SWE than the average of the surrounding terrain because they hinder drainage of melted water to the underlying soil. Given that UA SWE mainly ingests snow pillow measurements from the SNOTEL network over the western U.S. (Zeng et al., 2018), snow pillows may also be a source of uncertainties in UA SWE product.

6.6 Conclusion

The current engineering practice (e.g. NOAA Atlas 14) provides very limited guidance on designing infrastructure to accommodate snowmelt driven floods in the CONUS. In this study, we leverage two vetted, long-term CONUS snow products from the University of Arizona and NOAA to develop 25- and 100-year return level design SWE and snowmelt maps. Extreme value statistical methods are used to fit the GEV distribution to annual maximum UA SWE and snowmelt values from a 36-year record (water years 1982 to 2017). Somewhat surprisingly, trend analyses, showed few trends in annual maximum UA SWE or snowmelt with localized increases in the north central U.S. and modest decreases in isolated regions in the Southwestern U.S.

Despite their use of different methods to estimate SWE, the UA and SNODAS annual maximum SWE products are strongly correlated indicating that year-to-year variations in annual maximum SWE values are readily distinguished. There is notably less agreement between the two products' year-to-year snowmelt patterns. This suggests a need for additional study regarding reliable approaches for estimating snowmelt at time and space scales that are appropriate for design. Much of the study region had reasonable agreement for the magnitude of the annual maximums. SNODAS has somewhat higher annual maximum SWE and snowmelt in regions where annual SWE is typically low and may benefit from its ability to assimilate snow observations from a network of ground observers. For regions with deeper snowpack, including the western U.S. more differences were evident.

The snowmelt frequency estimates from this study complement the NOAA Atlas 14 design precipitation and may provide additional guidance on infrastructure design for snowmelt-flooding over snow-dominant regions in the continental U.S. The SWE and 7-day snowmelt design maps show similar patterns and, as anticipated, regions having larger annual maximums values typically also have greater snowmelt. In most regions the NOAA Atlas 14 design precipitation values exceed design snowmelt. However, in the northeastern U.S. near the Canadian border, the north central U.S. where just 10 cm of SWE can cause flooding, and the western mountainous U.S., design snowmelt substantially exceeds the NOAA Atlas 14 design precipitation.

CHAPTER 7

7.1 Summary and Major Findings

The objective of this dissertation was to improve current and future flood forecasting models by understanding SSD impacts on hydrologic responses, updating operational gamma SWE estimates, and identifying extreme SWE and snowmelt trends in the north-central U.S.

To improve the flood forecasting model with the SSD spatial information, in **Chapter 2**, the SSD maps were developed using multiple satellite “big” data and a Random Forest Machine Learning (RFML) technique. The maps agreed well with SSD permit records (overall accuracies of 76.9–87.0%) and corresponded with subwatershed-level statistics (r = 0.77–0.96). The RFML classifier identified soil properties and land surface temperature to be the strongest predictors of SSD. Predictor variables differed between the two spatial scales, suggesting that SSD models are sensitive to the spatial scale. In **Chapter 3**, the spatial SSD maps developed in Chapter 2 were used to quantify hydrological responses by the SSD expansion using the Noah-MP LSM. The inclusion of SSD in Noah-MP causes the surface soil moisture in SSD to decrease faster compared to UD conditions. There are remarkable differences in total evapotranspiration between SSD and UD conditions in the summer. We found that decreasing soil water storage directly results in decreased total evapotranspiration. The magnitude of peak runoff events (as well as the amount of surface runoff) in SSD tends to be larger than that of UD.

With the limitation of the model’s physic, accurate observations are critical to determining flood forecasting capacity in the north-central U.S. However, uncertainties in current airborne gamma snow survey hamper operational flood forecasting model. As described in **Chapter 4**, the operational gamma SWE observations were improved by updating antecedent soil moisture conditions using the SMAP satellite. The SMAP-updated gamma SWE had better agreement with ground-based SWE and microwave satellite SWE as compared to the operational gamma SWE. A sensitivity analysis was conducted to quantify how much SWE updates can increase streamflow in RRB using the flood forecasting model. Based on the analysis, the peak flow in 2013 can increase by 130% with a 25% (43 mm of SWE) increase in annual maximum SWE. Considering that the current gamma SWE observations were updated up to 41 mm by antecedent soil moisture, the improvement of SWE by SMAP soil moisture can contribute to accurate flood forecasting in the north-central U.S.

From a long-term perspective, flood forecasters and state governments need knowledge of historical changes in snowpack and snowmelt to develop flood management plans to adapt to climate changes. However, historical snowmelt trends are little known due to the limited long-term snow observations. In **Chapter 5**, three available long-term SWE products were evaluated to find the most reliable SWE data. As compared to SSMI/S and GlobSnow-2 SWE, UA SWE has much better agreement with gamma SWE in all land cover types and snow classes. This suggests that UA SWE is reliable for trend analysis of historical snowpack and snowmelt over the continental U.S including the northern great plains region. Tree cover and topographic heterogeneity affect the agreement between the gamma and gridded SWE and the accuracy of gamma SWE itself. Using the well-vetted UA SWE with the national snow product from SNODAS, described in **Chapter 6**, a trend analysis of extreme SWE and snowmelt events for the last 40 years was provided using the long-term UA SWE data over the U.S. The trend analyses showed negative trends in most mountainous regions of western states but positive trends in annual maximum snowmelt in the north-central U.S. For several states in the western U.S., there exist regions with decreases in annual maximum SWE (e.g. New Mexico, Colorado, and Nevada). Positive trends in annual maximum SWE were rare but remarkably present in North Dakota and Minnesota. North Dakota and Minnesota also had regions with statistically significant snowmelt increases, respectively.

7.2 Contributions

7.2.1 Current Flood Forecasting in Practice

This dissertation contributes to the state of practice by improving the current operational NOAA NWS flood forecasts. For example, in spring 2019, updated gamma SWE observations using the method developed in Chapter 4 were requested by and provided to the flood forecasters at the NCRFC and the NOAA's Office of Water Prediction. That year's original gamma SWE observations were updated by to 5 – 10 mm when corrected using the antecedent fall soil moisture conditions captured by SMAP. Due to abrupt increases in soil moisture by rainfall events in the late fall 2018 over the southern RRB, the original gamma survey overestimated SWE. This overestimate would have caused flood forecasts to overpredict the magnitude of spring floods. The practical use showed that the method can enhance the capability of the current flood forecasting model.

For year-by-year forecasting, the reliable SWE information may be more important than other factors over the north-central U.S. In a general comparison between the given results in Chapters 3 and 4, the improvement in gamma SWE has a larger influence on the peak flow than the inclusion of the current SSD condition. Given the snowmelt flood-generating mechanism, even small change in SWE can impact snowmelt-driven peak flow (personnel communication, Mike DeWeese NCRFC) due to land surface characteristics in this region (e.g. the extremely flat and very low-permeable clay soils).

While not immediately and directly relevant to the flood forecast models, from a long-term perspective, the inclusion of the SSD scheme in the NWM will become critical. Currently, flood forecast modelers in the region use their knowledge of soil water retention to modify the soil-related parameters (e.g. ZPERC [Maximum Percolation Rate Coefficient] and REXP [Percolation Equation Exponent]) to match flood forecasts into the observations for the regions where the models are not performing well. The manual adjustments by modelers may work in the current condition (the current percentage of the SSD area in RRB is less than 10% only). However, the SSD impact on the model capacity could have non-linear impacts on the flood response as the SSD system is exponentially expanded over the north-central U.S., particularly in the Northern Great Plains (Lark et al., 2015; Otto et al., 2016; US Farm Bill conservation programs were reduced). Using the future flood forecasting model (Noah-MP), this work provided valuable insights on potential hydrological changes due to the SSD expansion (Chapter 3). They will help to improve the next-generation flood forecast system capacity.

The significant snowmelt trends in the north-central U.S. found in Chapter 6 demand attention by communities and governments to allocate capacity and budgets to help flood management and to develop strategies to adapt to climate changes. The trend results may also urge flood forecasters to improve melting physics in the current and future flood forecasting model. The SNOW-17 used in the current flood model uses depletion curves that relate the areal snow cover versus the mean areal SWE. However, the depletion curves cannot capture anomalous patterns of accumulation and melt nor are they appropriate when new snow occurs on partially depleted surfaces. The north-central region may not be well suited for depletion curve input to runoff estimation due to wind redistribution and storm patterns. The Noah-MP snowmelt process in the future flood forecasting model has more sophisticated physics compared to the

SNOW-17 and uses multi-layer snowpack and parameterization schemes (e.g., snow surface albedo, snow cover fraction, liquid water retention and refreezing; Niu et al., 2011). However, my ongoing study found that in the melting phase, Noah-MP snowmelt rates are still too fast compared to UA and SNODAS SWE data. This indicates that there are still areas requiring improvement in the melting physics in the Noah-MP (e.g. accurate simulations of the diurnal cycle of snow temperature).

7.2.2 Research Advance in Hydrology

In addition to the practical use for enhanced flood forecasting, this work contributes to research advances in hydrology and agricultural water management. This work provides a novel SSD identification framework that can be widely used for hydrological changes in drainage-dominant regions worldwide. Currently, SSD remains largely unregulated throughout the north-central U.S. and Canada. The U.S. drainage census data are inconsistent with errors. In many regions, the SSD system had been already installed without historical records. The SSD RFML method is a powerful tool to track existing SSD locations which is required for multiple-scale hydrological studies (Kelly et al., 2017). Also, emerging remote sensing platforms such as CubeSat and Unmanned Aerial Vehicles can easily be used in the current SSD identification method as inputs with ultrahigh resolutions via the Google Earth Engine computing platform (McCabe et al., 2017; NASA CubeSat Launch Initiative, 2018). More than 130 CubeSats launched by Planet currently provide daily visible and near-infrared imagery with ultrahigh resolutions (e.g., 3 m and 72 cm), capturing daily near-global coverage (Planet Team, 2018).

The RFML method with SSD spatial maps will help to address the water quality issue. For example, phosphorus loss from agricultural landscapes has been an important water quality issue for decades because phosphorus transport plays in eutrophication. In an agricultural watershed, there are two primary pathways for nutrients to enter river flow or reservoir: surface runoff and subsurface flow. Nutrients transport via SSD pathways was often deemed “unknown” or inaccurate assumption due to the lack of spatial SSD information (King et al., 2014). With reliable SSD maps, researchers will be able to understand variations in nutrients delivery to rivers and lakes across the study area and to quantify the role of SSD in nutrient transport at a watershed scale.

The updated SSD information can also directly improve coupled-land surface/climate models by providing reliable ancillary information about recent land-use change conditions. For example, recent studies investigated the effects of human modifications (e.g. irrigation, deforestation, and urbanization) on atmospheric feedback at local to regional scales (Mahmood et al., 2014; Pei et al., 2016; Zhang et al., 2017). However, the widely used land surface models (e.g. Variable Infiltration Capacity, Community Land Model, and Noah-MP) currently do not consider the SSD expansion. Thus, the inclusion of SSD information might lead to better characterizing surface hydrologic processes and states (e.g. surface moisture and temperature, infiltration, surface, and subsurface runoff) as well as local weather and climate feedback in the current land surface model systems.

7.2.3 Research Advances in Snow Science

This work contributes to the snow science community and future snow satellite mission. In the snow science community, the evaluation of global SWE products has been limited due to the lack of reliable, independent SWE datasets over various snow and land cover classifications. The results in Chapter 5 provided important findings including the (1) limitations of GlobSnow and microwave satellite SWE products, (2) the reliability of the UA SWE products regardless land environments, and (3) the value of the gamma radiation technique to measure SWE, especially in forested regions.

As land surface models (e.g. Snow Ensemble Uncertainty Project [SEUP]; Kim et al., 2020) and regional climate models (e.g. Coordinated Regional Downscaling Experiment [CORDEX]) continue to evolve, independent and reliable SWE data are required to evaluate SWE outputs from the models. These evaluations will identify potential limitations of snow physical processes involved in each model and target snow classes and snow phase for improvement. The historical 40-year and ongoing NOAA airborne gamma SWE record will be a valuable reference for snow hydrologists and modelers, even though the record has limited spatial and temporal coverages compared to the gridded satellite and reanalysis products. For example, NASA scientists and the SEUP group are considering using the historical gamma SWE observations as a calibration reference to develop an assimilated optimal SWE ensemble for an observing system simulation experiment (OSSE).

This work also contributes to the future snow satellite mission. Currently, efforts are underway to determine which sensors should be included in the next generation snow satellite mission (NASA SnowEx plan, Durand et al. 2019; Kim et al., 2020). The NASA Terrestrial Hydrology Program has initiated an airborne and ground campaign (SnowEx) to collect SWE observations for snow mission design. As a well-vetted technique, the NOAA airborne gamma snow survey team has participated in SnowEx 2020 to test SWE algorithms by providing multi-sensor observations of a seasonal snow-covered landscape. The historical gamma SWE record and its accuracy may provide insights to refine the observational requirements of a future anticipated snow mission.

7.3 Future Direction and Research Needs

NWM flood forecasting model The high-resolution NWM will replace the thirteen individual RFC flood forecast models. Because the NWM uses the Noah-MP (without SSD scheme) to simulate land surface processes, the SSD scheme developed in this study would enhance surface hydrological processes in the NWM. As the snowpack variable was updated in the current NWS flood forecasting model, the NWM-simulated snowpack can also be evaluated in various environments and updated with the updated gamma SWE and UA SWE for enhanced forecasting. With advancing high-performance computing, LSMs including NWM will evolve their capacity to create hyperspatial and temporal resolution.

Airborne gamma SWE The historical 40-year and ongoing NOAA airborne gamma SWE records are useful for snow hydrologists and modelers in providing accurate SWE values in various environments. To maximize an additional use of gamma SWE products, further inquiry into the gamma observation capabilities will be necessary. Studies have shown that the improvements to the operational gamma SWE products are possible by minimizing the errors, even though potential sources of errors were identified before the 1990s (e.g., the spatial variance within the footprint, and dense forest effect). Future studies can utilize state-of-the-art high-resolution earth observation products (e.g., lidar, synthetic aperture radar, and multi-spectrometer) to quantify snow or land characteristics within a gamma flight footprint to improve this valuable resource. Additionally, the use of state-of-the-art assimilation techniques instead of the linear regression model may potentially help to expand the reliability of the updated gamma SWE despite its limited spatial and temporal coverage on the spatial distributed SWE (Margulis et al., 2016).

Compound impacts of change Understanding the compound impact of climate change and evolving agricultural practices (e.g. irrigation and SSD expansion) on the regional water cycle is critical to predicting extreme events and future water availability. Future research can quantify the combined effects using land surface hydrological models with regional climate models (RCMs) outputs with different future emission scenarios (e.g. Representative Concentration Pathway; RCP 2.6, 4.5, and 8.5) from the Coupled Model Intercomparison Project (CMIP) 5 & 6. This would improve our understanding of future snowpack and melting phases and quantify consequent impacts on water availability and flood timing and magnitude in light of human modifications.

7.4 Concluding Remarks

My dissertation seeks to overcome flood forecasting operational and scientific challenges that result from human modification to the landscape and climate change and to enhance snowmelt flood predictions. The dissertation statement is *human and climate impacts, as well as limited and noisy data, cause uncertainties in flood prediction in the great plains, but integrated approaches using remote sensing, big data analytics, and modeling can quantify the hydrological changes and reduce the uncertainties*. Through its five main chapters, this dissertation contributes to improving flood forecasting in practice and research advances in hydrology and snow science.

REFERENCES

Chapter 1

Anderson, E. A. (1976). A point of energy and mass balance model of snow cover. NOAA Technical Reports NWS, 19, 1–150

Brasnett B. (1999) A global analysis of snow depth for numerical weather prediction. J. Appl. Meteorol. 38(6):726–740

Barrett, A. P. (2003). National operational hydrologic remote sensing center snow data assimilation system (SNODAS) products at NSIDC (p. 19). Boulder, CO: National Snow and Ice Data Center, Cooperative Institute for Research in Environmental Sciences.

Burnash, R. J. C. (1995). The NWS river forecasting system. In V. P. Singh (Ed.), Computer models of watershed hydrology (pp. 395–442). Highlands Ranch, CO: Water Resources Publications, LLC.

Burnash, R. J. C., Ferral, R. L., & McGuire, R. A. (1973). A generalized streamflow simulation system—Conceptual modeling for digital computers. Sacramento, CA: Joint Federal and State River Forecast Center, US National Weather Service/California State Department of Water Resources.

Carroll, T. (2001). Airborne Gamma Radiation Snow Survey Program: A User's Guide, Version 5.0. National Operational Hydrologic Remote Sensing Center (NOHRSC), Chanhassen, 14

Carroll, T. R., Schaake Jr, J. C. (1983). Airborne snow water equivalent and soil moisture measurement using natural terrestrial gamma radiation. In Optical Engineering for Cold Environments (Vol. 414, pp. 208-214). International Society for Optics and Photonics.

Carroll, T. R. & Vose, G. D. (1984). Airborne snow water equivalent measurements over a forested environment using terrestrial gamma radiation. *In Proceedings of the Eastern Snow Conference*, 29, 101-115.

Cohen, S., Praskievicz, S., & Maidment, D. R. (2018). Featured Collection Introduction: National Water Model. JAWRA Journal of the American Water Resources Association, 54(4), 767-769.

Cooper, M.G., Nolin, A.W. & Safeeq, M. (2016). Testing the recent snow drought as an analog for climate warming sensitivity of Cascades snowpacks. *Environmental Research Letters*, 11(8), p.084009.

Dickinson, R. E., M. Shaikh, R. Bryant, and L. Graumlich (1998), Interactive canopies for a climate model, J. Clim., 11(11), 2823–2836.

Dollinger, D., Lundeen, B., Stroom, K., Anderson, P., Monson, B., Nelson, S., Parson, K., Butzer, A., & Streitz, A. (2013). Bois de Sioux River Watershed Monitoring and Assessment Report, Minnesota Pollution Control Agency, Saint Paul, MN.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley (2003), Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, J. Geophys. Res., 108(D22, 8851), doi:10.1029/2002JD003296.

Finocchiaro, R. G. (2014), Agricultural Subsurface Drainage Tile Locations by Permits in South Dakota, U.S. Geological Survey data release, doi:10.5066/F7KS6PNW.

Finocchiaro, R. G. (2016), Agricultural Subsurface Drainage Tile Locations by Permits in North Dakota, U.S. Geological Survey data release, doi:10.5066/F7QF8QZW.

Glynn, J. E., Carroll, T. R., Holman, P. B., & Grasty, R. L. (1988). An airborne gamma ray snow survey of a forested covered area with a deep snowpack, *Remote Sensing of Environment*, 26(2), 149-160.

Goodison, B. E., Ferguson, H. L., & McKay, G. A. (1981). Measurement and data analysis. Handbook of snow, 191-274.

Hirsch, R.M., Ryberg, K.R. (2012). Has the magnitude of floods across the USA changed with global CO 2 levels. *Hydrological Sciences Journal*, 57, 37–41. https://doi.org/10.1080/02626667.2011.621895.

Maidment, D. R. (2017). Conceptual framework for the national flood interoperability experiment. JAWRA Journal of the American Water Resources Association, 53(2), 245-257.

Mote, P. W., Li, S., Lettenmaier, D. P., Xiao, M., & Engel, R. (2018). Dramatic declines in snowpack in the western US. *Npj Climate and Atmospheric Science*, 1(1), 2. doi:10.1038/s41612-018-0012-1

Niu, G.-Y., Z.-L. Yang, R. E. Dickinson, and L. E. Gulden (2005), A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models, J. Geophys. Res., 110, D21106, doi:10.1029/2005JD006111

Niu, G.-Y., Z.-L. Yang, R. E. Dickinson, L. E. Gulden, and H. Su (2007), Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data, J. Geophys. Res., 112, D07103, doi:10.1029/2006JD007522.

Peck, E. L., Bissell, V. C., Jones, E. B., & Burge, D. L. (1971). Evaluation of snow water equivalent by airborne measurement of passive terrestrial gamma radiation. *Water Resources Research*, 7(5), 1151–1159.

Peterson, T. C., Heim Jr, R. R., Hirsch, R., Kaiser, D. P., Brooks, H., Diffenbaugh, N. S., ... & Katz, R. W. (2013). Monitoring and understanding changes in heat waves, cold waves, floods, and droughts in the United States: state of knowledge. *Bulletin of the American Meteorological Society*, 94(6), 821-834.

Pierce, D.W., Barnett, T.P., Hidalgo, H.G., Das, T., Bonfils, C., Santer, B.D., Bala, G., Dettinger, M.D., Cayan, D.R., Mirin, A. & Wood, A.W. (2008). Attribution of declining western US snowpack to human effects. *Journal of Climate*, 21(23), 6425-6444.

Rannie, W. (2015), The 1997 flood event in the Red River basin: Causes, assessment and damages. Canadian Water Resources Journal / Revue canadienne des ressources hydriques, 41(1-2), 45-55. https://doi.org/10.1080/07011784.2015.1004198

Rasmussen, P. F. (2016), Assessing the impact of climate change on the frequency of floods in the Red River basin, Can. Water Resour. J., 41(1-2), 331-342.

Schwert D. P. (2003) A geologist's perspective on the Red River of the North: history, geography, and planning/ management issues. In: Proceedings of 1st International Water Conference, Red River Basin Institute, Moorhead, MN

Todhunter P.E. (2001) A hydroclimatological analysis of the Red River of the North Snowmelt Flood Catastrophe of 1997. J Am Water Resour Assoc 37:1263–1278

Tuttle, S. E., Cho, E., Restrepo, P. J., Jia, X. Vuyovich, C. M., Cosh, M. H., & Jacobs, J. M. (2017). Remote Sensing of Drivers of Spring Snowmelt Flooding in the North Central U.S, 21-45. https://doi.org/10.1007/978-3-319-43744-6_2

Yang, Z.-L., and G.-Y. Niu (2003), The versatile integrator of surface and atmosphere processes: Part 1. Model description, Global Planet. Change, 38(1–2), 175–189, doi:10.1016/S0921-8181(03)00028-6.

Chapter 2

Abatzoglou, J. T. (2013). Development of gridded surface meteorological data for ecological applications and modelling. International Journal of Climatology, 33(1), 121-131. https://doi.org/10.1002/joc.3413

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing, 114, 24-31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8), 1798–1828. https://doi.org/10.1109/TPAMI.2013.50

Blann, K. L., Anderson, J. L., Sands, G. R., & Vondracek, B. (2009). Effects of Agricultural Drainage on Aquatic Ecosystems: A Review, Critical Reviews in Environmental Science and Technology, 39(11), 909-1001. https://doi.org/10.1080/10643380801977966

Breiman, L. (2001), Random forests. Machine learning, 45(1), 5-32. https://doi.org/10.1023/A:1010933404324

Chaney, N. W., Wood, E. F., McBratney, A. B., Hempel, J. W., Nauman, T. W., Brungard, C. W., & Odgers, N. P. (2016). POLARIS: A 30-meter probabilistic soil series map of the contiguous United States. Geoderma, 274, 54-67. doi:10.1016/j.geoderma.2016.03.025

Chaney, N. W., Minasny, B., Herman, J. D., Nauman, T. W., Brungard, C. W., Morgan, C. L. S., et al. (2019). POLARIS Soil Properties: 30-m Probabilistic Maps of Soil Properties Over the Contiguous United States. Water Resources Research, 2018WR022797. https://doi.org/10.1029/2018WR022797

Deines, J. M., Kendall, A. D., & Hyndman, D. W. (2017). Annual Irrigation Dynamics in the U.S. Northern High Plains Derived from Landsat Satellite Data. Geophysical Research Letters, 44(18), 9350-9360. doi:10.1002/2017gl074071

Dollinger, D., Lundeen, B., Stroom, K., Anderson, P., Monson, B., Nelson, S., Parson, K., Butzer, A., & Streitz, A. (2013). Bois de Sioux River Watershed Monitoring and Assessment Report, Minnesota Pollution Control Agency, Saint Paul, MN.

Eastman, M., Gollamudi, A., Stämpfli, N., Madramootoo, C.A. & Sarangi, A. (2010). Comparative evaluation of phosphorus losses from subsurface and naturally drained agricultural fields in the Pike River watershed of Quebec, Canada. Agricultural Water Management, 97(5), 596-604. doi:10.1016/j.agwat.2009.11.010

Finocchiaro, R. G. (2014). Agricultural Subsurface Drainage Tile Locations by Permits in South Dakota. U.S. Geological Survey data release. doi:10.5066/F7KS6PNW

Finocchiaro, R. G. (2016). Agricultural Subsurface Drainage Tile Locations by Permits in North Dakota. U.S. Geological Survey data release. doi:10.5066/F7QF8QZW

Foufoula-Georgiou, E., Takbiri, Z., Czuba, J. A., & Schwenk, J. (2015). The change of nature and the nature of change in agricultural landscapes: Hydrologic regime shifts modulate ecological transitions. Water Resources Research, 51(8), 6649-6671. https://doi.org/10.1002/2015WR017637

Frans, C., Istanbulluoglu, E., Mishra, V., Munoz-Arriola, F., & Lettenmaier, D. P. (2013). Are climatic or land cover changes the dominant cause of runoff trends in the Upper Mississippi River Basin? Geophysical Research Letters, 40(6), 1104-1110. https://doi.org/10.1002/grl.50262

Gökkaya, K., Budhathoki, M., Christopher, S. F., Hanrahan, B. R., & Tank, J. L. (2017). Subsurface tile drained area detection using GIS and remote sensing in an agricultural watershed. Ecological Engineering, 108, 370-379. https://doi.org/10.1016/j.ecoleng.2017.06.048

Gao, B. C. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sensing of Environment, 58(3), 257–266. https://doi.org/10.1016/S0034-4257(96)00067-3

Ge, Y., Hu, S., Ren, Z., Jia, Y., Wang, J., Liu, M., ... & Bai, H. (2019). Mapping annual land use changes in China's poverty-stricken areas from 2013 to 2018. Remote Sensing of Environment, 232, 111285.

Gitelson, A. A., Viña, A., Ciganda, V., Rundquist, D. C., & Arkebauer, T. J. (2005). Remote estimation of canopy chlorophyll content in crops. Geophysical Research Letters, 32, L08403. https://doi.org/10.1029/2005GL022688

Gómez, C., White, J. C., & Wulder, M. A. (2016). Optical remotely sensed time series data for land cover classification: A review. ISPRS Journal of Photogrammetry and Remote Sensing, 116, 55-72. https://doi.org/10.1016/j.isprsjprs.2016.03.008

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 202, 18-27. https://doi.org/10.1016/j.rse.2017.06.031

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sensing of Environment, 83, 195–213. https://doi.org/10.1016/S0034-4257(02)00096-2

Jacobs, J. M., Cho, E., & Jia, X. (2017, December). Tile Drainage Expansion Detection using Satellite Soil Moisture Dynamics. In AGU Fall Meeting Abstracts.

Jin, Z., Azzari, G., You, C., Di Tommaso, S., Aston, S., Burke, M., & Lobell, D. B. (2019). Smallholder maize area and yield mapping at national scales with Google Earth Engine. Remote Sensing of Environment, 228, 115-128.

Kelly, S. A., Takbiri, Z., Belmont, P., & Foufoula-Georgiou, E. (2017). Human amplified changes in precipitation-runoff patterns in large river basins of the Midwestern United States, Hydrology and Earth System Sciences, 21, 5065-5088, https://doi.org/10.5194/hess-21-5065-2017

Kerr, Y. H., Waldteufel, P., Wigneron, J. P., Delwart, S., Cabot, F., Boutin, J., ... & Juglea, S. E. (2010), The SMOS Mission: New Tool for Monitoring Key Elements ofthe Global Water Cycle. Proceedings of the IEEE, 98(5), 666-687. doi:10.1109/jproc.2010.2043032

King, K. W., Fausey, N. R., & Williams, M. R. (2014). Effect of subsurface drainage on streamflow in an agricultural headwater watershed. Journal of hydrology, 519, 438-445. doi:10.1016/j.jhydrol.2014.07.035

Kladivko, E. J., Frankenberger, J. R., Jaynes, D. B., Meek, D. W., Jenkinson, B. J., & Fausey, N. R. (2004). Nitrate leaching to subsurface drains as affected by drain spacing and changes in crop production system. Journal of Environmental Quality, 33(5), 1803-1813. doi:10.2134/jeq2004.1803.

Krapu, C., Kumar, M., & Borsuk, M. (2018). Identifying Wetland Consolidation Using Remote Sensing in the North Dakota Prairie Pothole Region. Water Resources Research. https://doi.org/10.1029/2018WR023338

Lenhart, C. F., Peterson, H., & Nieber, J. (2011). Increased streamflow in agricultural watersheds of the Midwest: implications for management. Watershed science bulletin, 2, 25-31.

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. R news, 2(3), 18-22.

Ma, L., Li, M., Ma, X., Cheng, L., Du, P., & Liu, Y. (2017). A review of supervised object-based land-cover image classification. ISPRS Journal of Photogrammetry and Remote Sensing, 130, 277-293. https://doi.org/10.1016/j.isprsjprs.2017.06.001

McCabe, M. F., Rodell, M., Alsdorf, D. E., Miralles, D. G., Uijlenhoet, R., Wagner, W., et al. (2017). The future of Earth observation in hydrology. Hydrology and Earth System Sciences, 21(7), 3879–3914. https://doi.org/10.5194/hess-21-3879-2017

Møller, A. B., Beucher, A., Iversen, B. V., & Greve, M. H. (2018). Predicting artificially drained areas by means of a selective model ensemble. Geoderma, 320, 30-42. https://doi.org/10.1016/j.geoderma.2018.01.018

Nakagaki, N. & Wieczorek, M. E. (2016). Estimates of subsurface tile drainage extent for 12 Midwest states, 2012: U.S. Geological Survey data release, http://dx.doi.org/10.5066/F7W37TDP

NASA CubeSat Launch Initiative (2018), https://www.nasa.gov/directorates/heo/home/CubeSats_initiative (Last access: 5 August 2019).

NASA (2017), CubeSat 101: Basic Concepts and Processes for First-Time CubeSat Developers, https://www.nasa.gov/sites/default/files/atoms/files/nasa_csli_cubesat_101_508.pdf (Last access: 5 August 2019)

Naz, B. S., & Bowling, L. C. (2008). Automated identification of tile lines from remotely sensed data. Transactions of the ASABE, 51(6), 1937-1950. https://doi.org/10.13031/2013.25399

Naz, B. S., Ale, S., & Bowling, L. C. (2009). Detecting subsurface drainage systems and estimating drain spacing in intensively managed agricultural landscapes. Agricultural water management, 96(4), 627-637. https://doi.org/10.1016/j.agwat.2008.10.002

Northcott, W. J., Verma, A. K., & Cooke, R. A. (2000). Mapping subsurface drainage systems using remote sensing and GIS. Mapping subsurface drainage systems using remote sensing and GIS., 1-10.

Ok, A. O., Akar, O., & Gungor, O. (2012). Evaluation of random forest method for agricultural crop classification. European Journal of Remote Sensing, 45(1), 421-432. https://doi.org/10.5721/EuJRS20124535

Petty, T. R., & Dhingra, P. (2018). Streamflow hydrology estimate using machine learning (SHEM). JAWRA Journal of the American Water Resources Association, 54(1), 55-68. https://doi.org/10.1111/1752-1688.12555

Planet Team (2018). Planet Application Program Interface: In Space for Life on Earth. San Francisco, CA. https://api.planet.com.

Rahman, M. M., Lin, Z., Jia, X., Steele, D. D., & DeSutter, T. M. (2014). Impact of subsurface drainage on streamflows in the Red River of the North basin. Journal of Hydrology, 511, 474-483. https://doi.org/10.1016/j.jhydrol.2014.01.070

Rannie, W. (2015), The 1997 flood event in the Red River basin: Causes, assessment and damages. Canadian Water Resources Journal / Revue canadienne des ressources hydriques, 41(1-2), 45-55. https://doi.org/10.1080/07011784.2015.1004198

Raymond, P. A., Oh, N. H., Turner, R. E., & Broussard, W. (2008). Anthropogenically enhanced fluxes of water and carbon from the Mississippi River. Nature, 451(7177), 449. https://doi.org/10.1038/nature06505

Rijal, I., Jia, X., Zhang, X., Steele, D. D., Scherer, T. F., & Akyuz, A. (2012). Effects of subsurface drainage on evapotranspiration for corn and soybean crops in southeastern North Dakota. Journal of Irrigation and Drainage Engineering, 138(12), 1060-1067. https://doi.org/10.1061/(asce)ir.1943-4774.0000508

Rodgers, M., Mulqueen, J., & McHale, J. (2003). A model study of mole drain spacing and performance, Agricultural Water Management, 60(1), 33-42, doi:10.1016/s0378-3774(02)00153-1.

Rogger, M., Agnoletti, M., Alaoui, A., Bathurst, J. C., Bodner, G., Borga, M., ... & Holden, J. (2017). Land use change impacts on floods at the catchment scale: Challenges and opportunities for future research. Water Resources Research, 53(7), 5209-5219. https://doi.org/10.1002/2017WR020723

Sadeghi, M., Jones, S. B., & Philpot, W. D. (2015). A linear physically-based model for remote sensing of soil moisture using short wave infrared bands. Remote Sensing of Environment, 164, 66-76. https://doi.org/10.1016/j.rse.2015.04.007

Schilling, K. E., Chan, K. S., Liu, H., & Zhang, Y. K. (2010). Quantifying the effect of land use land cover change on increasing discharge in the Upper Mississippi River. Journal of Hydrology, 387(3-4), 343-345. https://doi.org/10.1016/j.jhydrol.2010.04.019

Schottler, S. P., Ulrich, J., Belmont, P., Moore, R., Lauer, J. W., Engstrom, D. R., & Almendinger, J. E. (2014). Twentieth century agricultural drainage creates more erosive rivers. Hydrological Processes, 28(4), 1951-1961. https://doi.org/doi:10.1002/hyp.9738

Shen, C. (2018). A trans-disciplinary review of deep learning research and its relevance for water resources scientists. Water Resources Research, 54(11), 8558-8593. https://doi.org/10.1029/2018WR022643

Shokri, A. & Bardsley, W. E. (2015). Enhancement of the Hooghoudt Drain-Spacing Equation, Journal of Irrigation and Drainage Engineering, 141(6), 04014070, doi:10.1061/(asce)ir.1943-4774.0000835.

Shortridge, J. E., Guikema, S. D., & Zaitchik, B. F. (2016). Machine learning methods for empirical streamflow simulation: a comparison of model accuracy, interpretability, and uncertainty in seasonal watersheds. Hydrology and Earth System Sciences, 20(7), 2611-2628. https://doi.org/10.5194/hess-20-2611-2016

Sugg, Z. (2007). Assessing U.S. farm drainage: Can GIS lead to better estimates of subsurface drainage extent. World Resources Institute, Washington, DC, 20002

Tao, Y., Gao, X., Hsu, K., Sorooshian, S., & Ihler, A. (2016). A Deep Neural Network Modeling Framework to Reduce Bias in Satellite Precipitation Products. Journal of Hydrometeorology, 17, 931–945. https://doi.org/10.1175/JHM-D-15-0075.1

Tetzlaff, B., Kuhr, P., Vereecken, H., & Wendland, F. (2009a). Aerial photograph-based delineation of artificially drained areas as a basis for water balance and phosphorus modelling in large river basins. Physics and Chemistry of the Earth, Parts A/B/C, 34(8-9), 552-564. https://doi.org/10.1016/j.pce.2009.02.002

Tetzlaff, B., Kuhr, P. & Wendland, F. (2009b). A new method for creating maps of artificially drained areas in large river basins based on aerial photographs and geodata. Irrigation and Drainage: The journal of the International Commission on Irrigation and Drainage, 58(5), 569-585. https://doi.org/10.1002/ird.426

Tlapáková, L., Žaloudík, J., Kulhavý, Z., & Pelíšek, I. (2015). Use of remote sensing for identification and description of subsurface drainage system condition. Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis, 63(5), 1587-1599.

Todhunter, P. E. (2001), A hydroclimatological analysis of the Red River of the North snowmelt flood catastrophe of 1997. JAWRA Journal of the American Water Resources Association, 37(5), 1263-1278. https://doi.org/10.1111/j.1752-1688.2001.tb03637.x

Tuttle, S. E., Cho, E., Restrepo, P. J., Jia, X., Vuyovich, C. M., Cosh, M. H., & Jacobs, J. M. (2017). Remote sensing of drivers of spring snowmelt flooding in the North Central US. In Remote sensing of hydrological extremes (pp. 21-45). Springer, Cham. https://doi.org/10.1007/978-3-319-43744-6_2

United States Department of Agriculture National Agricultural Statistics Service (2014). Special tabulation: county-level land drained by tile and ditches, 2012 Census Agric.

Varner B. L., Gress, T., Copenhaver, K., White, S. (2002). The effectiveness and economic feasibility of image based agricultural tile maps. Inst. of Tech., Champaign, IL. Final Report to NASA ESAD 2002.

Verma, A. K., Cooke, R. A., Wendte, L. (1996). Mapping subsurface drainage systems with color infrared aerial photographs. American Water Resource Association’s 32nd Annual Conference and Symposium ‘‘GIS and Water Resources’’, September 22–26, Ft. Lauderdale, Florida.

Wang, Z., Lai, C., Chen, X., Yang, B., Zhao, S., & Bai, X. (2015). Flood hazard risk assessment model based on random forest. Journal of Hydrology, 527, 1130-1141. https://doi.org/10.1016/j.jhydrol.2015.06.008

Wieczorek, M. (2004). Subsurface Drains on Agricultural Land in the Conterminous United States, 1992: National Resource Inventory Conservation Practice 606, raster digital data, Version 1.1, http://water.usgs.gov/lookup/getspatial?nri92_cp606

Williams, M. R., King, K. W., & Fausey, N. R. (2015). Contribution of tile drains to basin discharge and nitrogen export in a headwater agricultural watershed. Agricultural water management, 158, 42-50.

Xie, Z., Phinn, S. R., Game, E. T., Pannell, D. J., Hobbs, R. J., Briggs, P. R., & McDonald-Madden, E. (2019). Using Landsat observations (1988–2017) and Google Earth Engine to detect vegetation cover changes in rangelands-A first step towards identifying degraded lands for conservation. Remote Sensing of Environment, 232, 111317.

Youngs, E. (1975), The effect of the depth of an impermeable barrier on water-table heights in drained homogeneous soils, Journal of Hydrology, 24(3-4), 283-290.

Zhang, C., Walters, D., & Kovacs, J. M. (2014). Applications of low altitude remote sensing in agriculture upon farmers' requests–a case study in northeastern Ontario, Canada. PloS one, 9(11), e112894. https://doi.org/10.1371/journal.pone.0112894

Chapter 3

Arnold, J.G., R. Srinivasan, R.S. Muttiah, and J.R. Williams. (1998). Large-area hydrologic modeling and assessment: Part I. Model development. Journal of American Water Resources Association 34(1):73-89.

Arnold, J.G., and N. Fohrer. 2005. SWAT2000: Current capabilities and research opportunities in applied watershed modeling. Hydrological Processes 19(3):563-572

Baldocchi, D., et al. (2001), FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities, Bull. Am. Meteorol. Soc., 82(11), 2415–2434

Brasnett B. (1999) A global analysis of snow depth for numerical weather prediction. J. Appl. Meteorol. 38(6):726–740

Blann, K. L., J. L. Anderson, G. R. Sands, and B. Vondracek (2009), Effects of Agricultural Drainage on Aquatic Ecosystems: A Review, Critical Reviews in Environmental Science and Technology, 39(11), 909-1001, doi:10.1080/10643380801977966.

Blöschl, G., S. Ardoin-Bardin, M. Bonell, M. Dorninger, D. Goodrich, D. Gutknecht, D. Matamoros, B. Merz, P. Shand, and J. Szolgay (2007), At what scales do climate variability and land cover change impact on flooding and low flows?, Hydrol. Processes, 21(9), 1241-1247.

Bolten, J., W.T. Crow, X. Zhan, T.J. Jackson, and C.A. Reynolds (2010). Evaluating the Utility of Remotely Sensed Soil Moisture Retrievals for Operational Agricultural Drought Monitoring, IEEE Transactions on Geoscience and Remote Sensing, 3(1): 57-66. DOI 10.1109/JSTARS.2009.2037163

Bowman, A. L., K. J. Franz, T. S. Hogue, and A. M. Kinoshita (2015), MODIS-Based Potential Evapotranspiration Demand Curves for the Sacramento Soil Moisture Accounting Model, Journal of Hydrologic Engineering, 21(1), 04015055.

Cai, X., Yang, Z.-L., Xia, Y., Huang, M., Wei, H., Leung, L. R., et al. (2015). Assessment of simulated water balance from Noah, Noah-MP, CLM, and VIC over CONUS using the NLDAS test bed. Journal of Geophysical Research: Atmospheres, 119, 13751–13770. https://doi.org/10.1002/2014JD022113

Chen et al. (2016). Assessment of the SMAP level 2 passive soil moisture product. IEEE Trans. Geosci. Remote Sens. 54 (8), 4994–5007. https://doi.org/10.1109/TGRS.2016. 2561938.

Cho, E., Jacobs, J. M., Schroeder, R., Tuttle, S. E., & Olheiser, C. (2020). Improvement of operational airborne gamma radiation snow water equivalent estimates using SMAP soil moisture. Remote Sensing of Environment, 240, 111668.

Cho, E., Jacobs, J. M., Jia, X., & Kraatz, S. (2019). Identifying Subsurface Drainage using Satellite Big Data and Machine Learning via Google Earth Engine. Water Resources Research, 55. https://doi.org/10.1029/2019WR024892

Colliander, A., Jackson, T.J., Bindlish, R., Chan, S., Das, N.N., Kim, S., Cosh, M.H., Dunbar, R.S., Dang, L., Pashaian, L., 2017. Validation of SMAP surface soil moisture products with core validation sites. Remote Sens. Environ. 191, 215–231.

Cohen, S., Praskievicz, S., & Maidment, D. R. (2018). Featured Collection Introduction: National Water Model. JAWRA Journal of the American Water Resources Association, 54(4), 767-769.

Dickinson, R. E., M. Shaikh, R. Bryant, and L. Graumlich (1998), Interactive canopies for a climate model, J. Clim., 11(11), 2823–2836.

Drury, C. F., Tan, C. S., Reynolds, W. D., Welacky, T. W., Oloya, T. O., & Gaynor, J. D. (2009). Managing tile drainage, subirrigation, and nitrogen fertilization to enhance crop yields and reduce nitrate loss. Journal of Environmental Quality, 38(3), 1193-1204.

Du, B., Arnold, J. G., Saleh, A., & Jaynes, D. B. (2005). Development and application of SWAT to landscapes with tiles and potholes. Transactions of the ASAE, 48(3), 1121-1133.

Eastman, M., A. Gollamudi, N. Stämpfli, C. A. Madramootoo, and A. Sarangi (2010), Comparative evaluation of phosphorus losses from subsurface and naturally drained agricultural fields in the Pike River watershed of Quebec, Canada, Agricultural Water Management, 97(5), 596-604, doi:10.1016/j.agwat.2009.11.010.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley (2003), Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, J. Geophys. Res., 108(D22, 8851), doi:10.1029/2002JD003296.

Finocchiaro, R. G. (2014), Agricultural Subsurface Drainage Tile Locations by Permits in South Dakota, U.S. Geological Survey data release, doi:10.5066/F7KS6PNW.

Finocchiaro, R. G. (2016), Agricultural Subsurface Drainage Tile Locations by Permits in North Dakota, U.S. Geological Survey data release, doi:10.5066/F7QF8QZW.

Frans, C., E. Istanbulluoglu, V. Mishra, F. Munoz-Arriola, and D. P. Lettenmaier (2013), Are climatic or land cover changes the dominant cause of runoff trends in the Upper Mississippi River Basin?, Geophys. Res. Lett., 40(6), 1104-1110, doi:10.1002/grl.50262.

Guo, T., Cibin, R., Chaubey, I., Gitau, M., Arnold, J. G., Srinivasan, R., ... & Engel, B. A. (2018a). Evaluation of bioenergy crop growth and the impacts of bioenergy crops on streamflow, tile drain flow and nutrient losses in an extensively tile-drained watershed using SWAT. Science of the total environment, 613, 724-735.

Guo, T., Gitau, M., Merwade, V., Arnold, J., Raghavan, S., Hirschi, M., & Engel, B. (2018b). Comparison of performance of tile drainage routines in SWAT 2009 and 2012 in an extensively tile-drained watershed in the Midwest. Hydrology and Earth System Sciences, 22(1), 89.

Hooghoudt, S.B. (1940). Bijdragen tot de kenins van enige natuurkundige grootheden van de ground. No. 7. Versl. Landbouwk. Onderz. (Contributions to the knowledge of some physical constants of the soil. No.7. Report Agricultural Research) 46: 515-707.

International Joint Commission (2000), Living with the Red: A report to the governments of Canada and the United States on reducing flood impacts in the Red River Basin, Ottawa, Canada and Washington, DC.

Kelly, S. A., Z. Takbiri, P. Belmont, and E. Foufoula-Georgiou (2017), Human amplified changes in precipitation-runoff patterns in large river basins of the Midwestern United States, Hydrol. Earth Syst. Sci., 21, 5065-5088, doi:/10.5194/hess-21-5065-2017.

King, K. W., N. R. Fausey, and M. R. Williams (2014), Effect of subsurface drainage on streamflow in an agricultural headwater watershed, J. Hydrol., 519, 438-445, doi:10.1016/j.jhydrol.2014.07.035.

Kirkham, D. (1957). Theory of land drainage. In "Drainage of Agricultural Lands Agronomy Monograph No. 7. Madison, WI: American Society of Agronomy

Kladivko, E. J., J. R. Frankenberger, D. B. Jaynes, D. W. Meek, B. J. Jenkinson, and N. R. Fausey (2004), Nitrate Leaching to Subsurface Drains as Affected by Drain Spacing and Changes in Crop Production System, Journal of Environment Quality, 33(5), 1803, doi:10.2134/jeq2004.1803.

Kumar, S. V., Peters-Lidard, C. D., Tian, Y., & Houser, P. R. (2006). Land information system: An interoperable framework for high resolution land surface modeling. Environmental Modelling & Software, 21(10), 1402–1415. https://doi.org/10.1016/j.envsoft.2005.07.004

Lenhart, C. F., H. Peterson, and J. Nieber (2011), Increased streamflow in agricultural watersheds of the Midwest: implications for management, Watershed science bulletin, 2(1), 25-31.

Maidment, D. R. (2017). Conceptual framework for the national flood interoperability experiment. JAWRA Journal of the American Water Resources Association, 53(2), 245-257.

Miller J. E., Frink D. L. (1984) Changes in flood response of the Red River of the North basin, North Dakota-Minnesota. United States Government Printing Office, Washington, DC

Moriasi, D. N., Rossi, C. G., Arnold, J. G., & Tomer, M. D. (2012). Evaluating hydrology of the Soil and Water Assessment Tool (SWAT) with new tile drain equations. Journal of soil and water conservation, 67(6), 513-524.

Moody, W.T. 1966. Nonlinear differential equation of drain spacing. Journal of the Irrigation and Drainage Division, American Society of Civil Engineers 92(IR2):1-9.

Murray, S. J., I. M. Watson, and I. C. Prentice (2013), The use of dynamic global vegetation models for simulating hydrology and the potential integration of satellite observations, Prog. Phys. Geogr., 37(1), 63–97, doi:10.1177/0309133312460072.

Ng, H. Y. F., Tan, C. S., Drury, C. F., & Gaynor, J. D. (2002). Controlled drainage and subirrigation influences tile nitrate loss and corn yields in a sandy loam soil in Southwestern Ontario. Agriculture, ecosystems & environment, 90(1), 81-88.

Niu, G.-Y., Z.-L. Yang, R. E. Dickinson, and L. E. Gulden (2005), A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models, J. Geophys. Res., 110, D21106, doi:10.1029/2005JD006111

Niu, G.-Y., Z.-L. Yang, R. E. Dickinson, L. E. Gulden, and H. Su (2007), Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data, J. Geophys. Res., 112, D07103, doi:10.1029/2006JD007522.

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

Ozdogan, M., M. Rodell, H. K. Beaudoing, and D. L. Toll (2010), Simulating the effects of irrigation over the United States in a land surface model based on satellite-derived agricultural data, J. Hydrometeorol., 11(1), 171–184, doi:10.1175/2009jhm1116.1.

Rahman, M. M., Z. Lin, X. Jia, D. D. Steele, and T. M. DeSutter (2014), Impact of subsurface drainage on streamflows in the Red River of the North basin, J. Hydrol., 511, 474-483, doi:10.1016/j.jhydrol.2014.01.070.

Randall, G., J. Vetsch, and J. Huffman (2003), Nitrate losses in subsurface drainage from a corn–soybean rotation as affected by time of nitrogen application and use of nitrapyrin, J. Environ. Qual., 32(5), 1764-1772.

Rannie, W. (2015), The 1997 flood event in the Red River basin: Causes, assessment and damages, Canadian Water Resources Journal / Revue canadienne des ressources hydriques, 41(1-2), 45-55, doi:10.1080/07011784.2015.1004198.

Rasmussen, P. F. (2016), Assessing the impact of climate change on the frequency of floods in the Red River basin, Can. Water Resour. J., 41(1-2), 331-342.

Rodgers, M., J. Mulqueen, and J. McHale (2003), A model study of mole drain spacing and performance, Agricultural Water Management, 60(1), 33-42, doi:10.1016/s0378-3774(02)00153-1.

Rijal, I., X. Jia, X. Zhang, D. D. Steele, T. F. Scherer, and A. Akyuz (2012), Effects of Subsurface Drainage on Evapotranspiration for Corn and Soybean Crops in Southeastern North Dakota, J. Irrigat. Drain. Eng., 138(12), 1060-1067, doi:10.1061/(asce)ir.1943-4774.0000508.

Sands, G. R. (2012), Tile Drainage: Hydrology and the RRB, A Presentation of the Tiling and Drainage Workshop, August 28-29, 2012, Hankinson, ND.

Sands, G. R., I. Song, L. M. Busman, and B. J. Hansen (2008), The effects of subsurface drainage depth and intensity on nitrate loads in the northern cornbelt, Transactions of the ASABE, 51(3), 937-946.

Sazib, N., Mladenova, I., & Bolten, J. (2018). Leveraging the google earth engine for drought assessment using global soil moisture data. Remote Sensing, 10(8), 1265.

Schilling, K. E., Gassman, P. W., Arenas-Amado, A., Jones, C. S., & Arnold, J. (2019). Quantifying the contribution of tile drainage to basin-scale water yield using analytical and numerical models. Science of The Total Environment, 657, 297-309.

Schottler, S. P., J. Ulrich, P. Belmont, R. Moore, J. W. Lauer, D. R. Engstrom, and J. E. Almendinger (2014), Twentieth century agricultural drainage creates more erosive rivers, Hydrol. Processes, 28(4), 1951-1961, doi:10.1002/hyp.9738.

Schwert D. P. (2003) A geologist’s perspective on the Red River of the North: history, geography, and planning/ management issues. In: Proceedings of 1st International Water Conference, Red River Basin Institute, Moorhead, MN

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., ... & Teuling, A. J. (2010). Investigating soil moisture–climate interactions in a changing climate: A review. Earth-Science Reviews, 99(3-4), 125-161.

Shokri, A., and W. E. Bardsley (2015), Enhancement of the Hooghoudt Drain-Spacing Equation, J. Irrigat. Drain. Eng., 141(6), 04014070, doi:10.1061/(asce)ir.1943-4774.0000835.

Skaggs, R.W. 1980. Methods for design and evaluation of drainage-water management systems for soils with high water tables. Drainmod Reference Report. Fort Worth, TX: USDA Soil Conservation Society South National Technical Center.

Skaggs, R. W., M. Breve, and J. Gilliam (1994), Hydrologic and water quality impacts of agricultural drainage∗, Critical reviews in environmental science and technology, 24(1), 1-32.

Tan, C. S., Drury, C. F., Soultani, M., Van Wesenbeeck, I. J., Ng, H. Y. F., Gaynor, J. D., & Welacky, T. W. (1998). Effect of controlled drainage and tillage on soil structure and tile drainage nitrate loss at the field scale. Water Science and Technology, 38(4-5), 103-110.

Todhunter P.E. (2001) A hydroclimatological analysis of the Red River of the North Snowmelt Flood Catastrophe of 1997. J Am Water Resour Assoc 37:1263–1278

Williams, M., K. King, and N. Fausey (2015), Contribution of tile drains to basin discharge and nitrogen export in a headwater agricultural watershed, Agricultural Water Management, 158, 42-50.

Wood, E. F., et al. (2011), Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth’s terrestrial water, Water Resour. Res., 47, W05301, doi:10.1029/2010WR010090.

Yang, Z.-L., and G.-Y. Niu (2003), The versatile integrator of surface and atmosphere processes: Part 1. Model description, Global Planet. Change, 38(1–2), 175–189, doi:10.1016/S0921-8181(03)00028-6.

Yang, Z.-L., et al. (2011), The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins, J. Geophys. Res., 116, D12110, doi:10.1029/2010JD015140.

Youngs, E. (1975), The effect of the depth of an impermeable barrier on water-table heights in drained homogeneous soils, J. Hydrol., 24(3-4), 283-290.

Xia, Y., Ek, M., Wei, H., & Meng, J. (2012). Comparative analysis of relationships between NLDAS-2 forcings and model outputs. Hydrological Processes, 26(3), 467–474. https://doi.org/10.1002/hyp.8240

Zwieback, S., Colliander, A., Cosh, M.H., Martínez-Fernándezj, J., McNairn, H., Starks, P.J., Berg, A., 2018. Estimating time-dependent vegetation biases in the SMAP soil moisture product. Hydrol. Earth Syst. Sci. 22 (8), 4473–4489

Chapter 4

Ahl, A., Bieber, G., 2010. Correction of the attenuation effect of vegetation on airborne gamma-ray spectrometry data using laser altimeter data. Near Surf. Geophys. 8(4), 271-278.

Akbar, R., Short Gianotti, D., McColl, K. A., Haghighi, E., Salvucci, G. D., Entekhabi, D., 2018. Hydrological storage length scales represented by remote sensing estimates of soil moisture and precipitation. Water Resour. Res. 54(3), 1476-1492.

Al-Yaari, A., Wigneron, J.P., Ducharne, A., Kerr, Y.H., Wagner, W., De Lannoy, G., Reichle, R., Al Bitar, A., Dorigo, W., Richaume, P., Mialon, A., 2014. Global-scale comparison of passive (SMOS) and active (ASCAT) satellite based microwave soil moisture retrievals with soil moisture simulations (MERRA-Land). Remote Sens. Environ. 152, 614-626.

Armstrong, R., Knowles, K. Brodzik, M. Hardman, M., 1994. DMSP SSM/I-SSMIS pathfinder daily EASE-grid brightness temperatures. Version 2. NASA National Snow Ice Data Center Distributed Active Archive Center: Boulder, CO, USA. [accessed on 1 March, 2019]

Armstrong, R. L., Chang, A. Rango, A. Josberger, E., 1993. Snow depths and grain-size relationships with relevance for passive microwave studies. Ann. Glaciol., 17, 171-176

Armstrong, R.L.; Brodzik, M.J., 2001. Recent northern hemisphere snow extent: A comparison of data derived from visible and microwave satellite sensors. Geophys. Res. Lett. 28, 3673–3676.

Babaeian, E., Sadeghi, M., Jones, S. B., Montzka, C., Vereecken, H., & Tuller, M., 2019. Ground, Proximal and Satellite Remote Sensing of Soil Moisture. Rev. Geophys.

Barrett, A. P., 2003. National operational hydrologic remote sensing center snow data assimilation system (SNODAS) products at NSIDC (p. 19). Boulder, CO: National Snow and Ice Data Center, Cooperative Institute for Research in Environmental Sciences.

Bergeron, J. M., Trudel, M. Leconte, R., 2016. Combined assimilation of streamflow and snow water equivalent for mid-term ensemble streamflow forecasts in snow-dominated regions. Hydrol. Earth Syst. Sci., 20(10), 4375-4389. doi:10.5194/hess-20-4375-2016

Bilskie, J., 2001. Soil Water Status: Content and Potential. Campbell Scientific, Inc., Logan, UT

Bindlish, R., Cosh, M. H., Jackson, T. J., Koike, T., Fujii, H., Chan, S. K., ... & Collins, C. H., 2018. GCOM-W AMSR2 soil moisture product validation using core validation sites. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 11(1), 209-219. DOI: 10.1109/JSTARS.2017.2754293

Bland, W. L., Helmke, P. A. Baker, J. M. 1997. High-resolution snow-water equivalent measurement by gamma-ray spectroscopy. Agr. Forest Meteorol., 83(1-2), 27-36

Brodzik, M. J. 2014, F17 vs. F13 SWE Regression. Available online: http://cires1.colorado.edu/~brodzik/F13-F17swe/ [accessed on 27 March 2019].

Burnash, R. J. C., 1995, The NWS river forecast system–catchment modeling, in Computer Models of Watershed Hydrology, edited by V. P. Singh, pp. 311–366, Water Resour. Publ., Littleton, Colo.

Carroll, T. R., 1981. Airborne soil moisture measurement using natural terrestrial gamma radiation. Soil Sci. 132(5), 358-366.

Carroll, T., 2001. Airborne Gamma Radiation Snow Survey Program: A User's Guide, Version 5.0. National Operational Hydrologic Remote Sensing Center (NOHRSC), Chanhassen, 14

Carroll S. S. & Carroll T. R., 1989a. Effect of forest biomass on airborne snow water equivalent estimates obtained by measuring terrestrial gamma radiation. Remote Sens. Environ. 27(3), 313–319.

Carroll S. S., & Carroll T. R., 1989b. Effect of uneven snow cover on airborne snow water equivalent estimates obtained by measuring terrestrial gamma radiation. Water Resour. Res. 25 (7):1505–1510.

Carroll, T. R., Schaake Jr, J. C., 1983. Airborne snow water equivalent and soil moisture measurement using natural terrestrial gamma radiation. In Optical Engineering for Cold Environments (Vol. 414, pp. 208-214). International Society for Optics and Photonics.

Chan, S., Bindlish, R., O'Neill, P., Njoku, E.G., Jackson, T.J., Colliander, A., Chen, F., Bürgin, M., Dunbar, S., Piepmeier, J., Yueh, S., Entekhabi, D., Cosh, M., Caldwell, T., Walker, J., Wu, X., Berg, A., Rowlandson, T., Pacheco, A., McNairn, H., Thibeault, M., MartínezFernández, J., González-Zamora, Á., Seyfried, M., Bosch, D., Starks, P., Goodrich, D., Prueger, J., Palecki, M., Small, E., Calvet, J.-C., Crow, W., Kerr, Y., 2016. Assessment of the SMAP level 2 passive soil moisture product. IEEE Trans. Geosci. Remote Sens. 54 (8), 4994–5007. DOI:10.1109/TGRS.2016.2561938

Chan, S. K., Bindlish, R., O'Neill, P., Jackson, T., Njoku, E., Dunbar, S., ... & Colliander, A., 2018. Development and assessment of the SMAP enhanced passive soil moisture product. Remote Sens. Environ. 204, 931-941.

Chang, A., Foster, J. Hall, D. K., 1987. Nimbus-7 SMMR derived global snow cover parameters. Ann. Glaciol., 9(1), 39-44

Channan, S., Collins, K., Emanuel, W. R., 2014. Global mosaics of the standard MODIS land cover type data. University of Maryland and the Pacific Northwest National Laboratory, College Park, Maryland, USA, 30.

Cho, E., Jacobs, J. M., Tuttle, S. E., Schroeder, R., Olheiser, C., 2018. Improvement of airborne gamma radiation snow water equivalent estimations with spaceborne soil moisture observations, Proceedings of the 75th Annual Eastern Snow Conference: 5-8 June 2018, NOAA Center for Weather and Climate Prediction, College Park, Maryland, USA

Cho, E., Jacobs, J.M. and Vuyovich, C., 2019. The value of long-term (40 years) airborne gamma radiation SWE record for evaluating three observation-based gridded SWE datasets by seasonal snow and land cover classifications. Water Resour. Res. https://doi.org/10.1029/2019WR025813

Cho, E., Moon, H., Choi, M., 2015. First assessment of the advanced microwave scanning radiometer 2 (AMSR2) soil moisture contents in Northeast Asia. J. Meteorol. Soc. Jpn., 93(1), 117-129. https://doi.org/10.2151/jmsj.2015-008.

Cho, E., Tuttle, S. E., Jacobs, J. M., 2017. Evaluating Consistency of Snow Water Equivalent Retrievals from Passive Microwave Sensors over the North Central US: SSM/I vs. SSMIS and AMSR-E vs. AMSR2. Remote Sens. 9(5), 465.

Clark, M. P., Hendrikx, J., Slater, A. G., Kavetski, D., Anderson, B., Cullen, N. J., ... & Woods, R. A., 2011. Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A review. Water Resour. Res. 47(7).

Clow, D. W., Nanus, L., Verdin, K. L., Schmidt, J., 2012. Evaluation of SNODAS snow depth and snow water equivalent estimates for the Colorado Rocky Mountains, USA. Hydrol. Process. 26(17), 2583-2591.

Colliander, A., Jackson, T.J., Bindlish, R., Chan, S., Das, N.N., Kim, S., Cosh, M.H., Dunbar, R.S., Dang, L., Pashaian, L., 2017. Validation of SMAP surface soil moisture products with core validation sites. Remote Sens. Environ. 191, 215–231.

Dai, L., Che, T., Wang, J., Zhang, P., 2012. Snow depth and snow water equivalent estimation from AMSR-E data based on a priori snow characteristics in Xinjiang, China. Remote Sens. Environ. 127, 14-29.

Das, N., D. Entekhabi, S. Dunbar, M. J. Chaubell, A. Colliander, S. Yueh, T. Jagdhuber, F. Chen, W. T. Crow, P. E. O'Neill, J. Walker, A. Berg, D. Bosch, T. Caldwell, M. Cosh, C. H. Collins, E. Lopez-Baeza, and M. Thibeault. 2019. The SMAP and Copernicus Sentinel 1A/B microwave active-passive high resolution surface soil moisture product, Remote Sens. Environ. 233. 111380. https://doi.org/10.1016/j.rse.2019.111380

Derksen, C., LeDrew, E., Walker, A., Goodison, B. 2000. Influence of sensor overpass time on passive microwave-derived snow cover parameters. Remote Sens. Environ. 71, 297–308.

Derksen, C., Walker, A., Goodison, B., 2005. Evaluation of passive microwave snow water equivalent retrievals across the boreal forest/tundra transition of western Canada. Remote Sens. Environ. 96(3-4), 315-327.

De Roo, A. P., Gouweleeuw, B., Thielen, J., Bartholmes, J., Bongioannini-Cerlini, P., Todini, E. Bates, P.D., Horritt, M., Hunter, N., Beven, K., 2003. Development of a European flood forecasting system. Intl. J. River Basin Management 1(1), 49-59

Dobson, M.C., Ulaby, F.T., Hallikainen, M.T., El-rayes, M., 1985. Microwave dielectric behavior of wet soil-Part II: dielectric mixing models. IEEE Trans. Geosci. Remote Sens. GE-23, 35-46.

Dong, J., Crow, W. T., & Bindlish, R. (2018). The error structure of the SMAP single and dual channel soil moisture retrievals. Geophysical Research Letters, 45(2), 758-765.

Durand, M., Gatebe, C., Kim, E., Molotch, N., Painter, T., Raleigh, M., Sandells, M., Vuyovich, C., 2019. NASA SnowEx Science Plan: Assessing approaches for measuring water in Earth's seasonal snow, Version 1.6, [accessed on March 27, 2019] https://goo.gl/sFkxHc.

Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., ... & Tarpley, J. D., 2003. Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. J. Geophys. Res. Atmos., 108(D22).

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., ... & Kimball, J., 2010. The Soil Moisture Active Passive (SMAP) mission. Proc. IEEE 98 (5), 704–716. DOI: 10.1109/JPROC.2010.2043918

Escorihuela, M.J., Chanzy, A., Wigneron, J.P., Kerr, Y.H., 2010. Effective soil moisture sampling depth of L-band radiometry: a case study. Remote Sens. Environ. 114, 995–1001

Foster, J.L., Sun, C., Walker, J.P., Kelly, R., Chang, A., Dong, J., Powell, H., 2005. Quantifying the uncertainty in passive microwave snow water equivalent observations. Remote Sens. Environ. 2005, 94, 187–203.

Fournier, S., Reager, J. T., Lee, T., Vazquez-Cuervo, J., David, C. H., & Gierach, M. M., 2016. SMAP observes flooding from land to sea: The Texas event of 2015. Geophys. Res. Lett., 43(19), 10-338.

Glynn, J. E., Carroll, T. R., Holman, P. B., & Grasty, R. L., 1988. An airborne gamma ray snow survey of a forested covered area with a deep snowpack, Remote Sens. Environ., 26(2), 149-160.

Goodison, B. E., Ferguson, H. L., & McKay, G. A., 1981. Measurement and data analysis. Handbook of snow, 191-274.

Grasty, R., 1982. Direct snow-water equivalent measurement by air-borne gamma-ray spectrometry. J. Hydrol., 55(1-4), 213-235.

Gruber, A., Dorigo, W. A., Zwieback, S., Xaver, A., & Wagner, W., 2013. Characterizing coarse-scale representativeness of in situ soil moisture measurements from the International Soil Moisture Network. Vadose Zone J. 12(2).

Hedrick, A., Marshall, H. P., Winstral, A., Elder, K., Yueh, S., & Cline, D., 2015. Independent evaluation of the SNODAS snow depth product using regional-scale lidar-derived measurements. The Cryosphere, 9(1), 13-23.

Imaoka, K., Kachi, M., Kasahara, M., Ito, N., Nakagawa, K., & Oki, T., 2010. Instrument performance and calibration of AMSR-E and AMSR2. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, 38(8), 13-18.

Ishizaki, A., Sanada, Y., Mori, A. Imura, M., Ishida, M., Munakata, M., 2016. Investigation of snow cover effects and attenuation correction of gamma ray in aerial radiation monitoring. Remote Sens. 8(11), 892

Jackson, R. D., 1973. Diurnal Changes in Soil Water Content During Drying 1. Field soil water regime, (fieldsoilwaterr), 37-55.

Jackson, T. J., Schmugge, T. J., 1991. Vegetation effects on the microwave emission of soils. Remote Sens. Environ. 36(3), 203–212.

Jones, W. K., & Carroll, T. R., 1983. Error analysis of airborne gamma radiation soil moisture measurements. Agric. Meteorol. 28(1), 19-30.

Kelly, R., 2009. The AMSR-E snow depth algorithm: Description and initial results. J. Remote Sens. Soc. Jpn,, 29(1), 307-317

Kelly, R. E., Chang, A. T., Tsang, L., Foster, J. L., 2003. A prototype AMSR-E global snow area and snow depth algorithm. IEEE Trans. Geosci. Remote Sens. 41(2), 230-242.

Kim, H., & Lakshmi, V., 2019. Global Dynamics of Stored Precipitation Water in the Topsoil Layer from Satellite and Reanalysis Data. Water Resour. Res., 55(4), 3328-3346.

Kim, H., Parinussa, R., Konings, A. G., Wagner, W., Cosh, M. H., Lakshmi, V., Zohaib, M., Choi, M., 2018. Global-scale assessment and combination of SMAP with ASCAT (active) and AMSR2 (passive) soil moisture products. Remote Sens. Environ. 204, 260-275.

Koike, T., 2013. Description of the GCOM-W1 AMSR2 Soil Moisture Algorithm, Technical Report–Description of the GCOM-W1 Level 1R and Level 2 Algorithms. Japan Aerospace Exploration Agency (JAXA) Earth Observation Research Center, pp. 8.1–8.13 NDX-120015A, Chapter 8.

Koster, R. D., Guo, Z., Yang, R., Dirmeyer, P. A., Mitchell, K., & Puma, M. J. (2009). On the nature of soil moisture in land surface models. Journal of Climate, 22(16), 4322-4335.

Koster, R. D., Suarez, M. J., 1996. Energy and water balance calculations in the Mosaic LSM. NASA Tech. Memo. 104606, Vol. 9, 60 pp. [Available online at http://gmao.gsfc.nasa.gov/pubs/docs/Koster130.pdf.]

Lawston, P. M., Santanello, J. A., & Kumar, S. V. (2017). Irrigation signals detected from SMAP soil moisture retrievals. Geophys. Res. Lett., 44(23).

Liang, X., Lettenmaier, D. P. Wood, E. F., Burges, S. J., 1994. A simple hydrologically based model of land surface water and energy fluxes for GCMs, J. Geophys. Res., 99, 14,415–14,428, doi:10.1029/94JD00483.

Liu, Y., Weerts, A., Clark, M., Hendricks Franssen, H. J., Kumar, S., Moradkhani, H., ... & Van Velzen, N., 2012. Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities. Hydrol. Earth Syst. Sci. 16(10), 3863-3887. doi:10.5194/hess-16-3863-2012

Loew, A., 2008. Impact of surface heterogeneity on surface soil moisture retrievals from passive microwave data at the regional scale: the upper Danube case. Remote Sens. Environ. 112 (1), 231–248.

Luojus, K., Pullianen, J., Takala, M., Lemmetyinen, J., Kangwa, M., Eskelinen, M., et al., 2014. GlobSnow-2 Final Report, Global Snow Monitoring for Climate Research, European Space Agency, http://www.globsnow.info/docs/GlobSnow_2_Final_Report_release.pdf.

Ma, H., Zeng, J., Chen, N., Zhang, X., Cosh, M. H., & Wang, W. (2019). Satellite surface soil moisture from SMAP, SMOS, AMSR2 and ESA CCI: A comprehensive assessment using global ground-based observations. Remote Sens. Environ. 231, 111215. doi:10.1016/j.rse.2019.111215

Mladenova, I. E., Jackson, T. J., Njoku, E., Bindlish, R., Chan, S., Cosh, M. H., Holmes, T.R.H., de Jeu, R.A.M., Jones, L., Kimball, J., Paloscia, S., Santi, E., 2014. Remote monitoring of soil moisture using passive microwave-based techniques—Theoretical basis and overview of selected algorithms for AMSR-E. Remote Sens. Environ. 144, 197-213.

Mote, T. L., Grundstein, A. J., Leathers, D. J., & Robinson, D. A., 2003. A comparison of modeled, remotely sensed, and measured snow water equivalent in the Northern Great Plains. Water Resour. Res., 39(8), 1209. https://doi.org/10.1029/2002WR001782

McColl, K. A., Alemohammad, S. H., Akbar, R., Konings, A. G., Yueh, S., Entekhabi, D., 2017. The global distribution and dynamics of surface soil moisture. Nat. Geosci. 10(2), 100.

Mishra, A., Vu, T., Veettil, A. V., & Entekhabi, D. (2017). Drought monitoring with soil moisture active passive (SMAP) measurements. J. Hydrol., 552, 620-632.

Njoku, E. G., Kong, J. A., 1977. Theory for passive microwave remote sensing of near-surface soil moisture. J. Geophys. Res. 82 (20) (July).

O'Neill, P. E., Chan, S. Njoku, E. G. Jackson, T., Bindlish. R., 2018. SMAP Enhanced L3 Radiometer Global Daily 9 km EASE-Grid Soil Moisture, Version 2. Boulder, Colorado

USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. doi: https://doi.org/10.5067/RFKIZ5QY5ABN.

Owe, M., De Jeu, R.A.M., Holmes, T.R.H., 2008. Multi-sensor historical climatology of satellite derived global land surface moisture. J. Geophys. Res.: Atmos. 113:F1 F01002. http://dx.doi.org/10.1029/2007JF000769

Painter, T. H., et al., 2016. The Airborne Snow Observatory: Fusion of scanning lidar, imaging spectrometer, and physically-based modeling for mapping snow water equivalent and snow albedo. Remote Sens. Environ. 184, 139-152. doi:10.1016/j.rse.2016.06.018

Peck, E. L., Bissell, V. C., Jones, E. B., & Burge, D. L., 1971. Evaluation of snow water equivalent by airborne measurement of passive terrestrial gamma radiation. Water Resour. Res. 7(5), 1151–1159.

Peck, E. L., Carroll, T. R., VanDemark, S. C., 1980. Operational aerial snow surveying in the United States/Etude de neige aérienne effectuée aux Etats Unis. Hydrol. Sci. J., 25(1), 51–62

Pulliainen, J., Hallikainen, M., 2001. Retrieval of regional snow water equivalent from space-borne passive microwave observations. Remote Sens. Environ. 75(1), 76-85.

Pulliainen, J., 2006. Mapping of snow water equivalent and snow depth in boreal and sub-arctic zones by assimilating space-borne microwave radiometer data and ground-based observations. Remote Sens. Environ. 101(2), 257-269.

Rannie, W., 2015. The 1997 flood event in the Red River basin: Causes, assessment and damages. Canadian Water Resources Journal / Revue canadienne des ressources hydriques, 41(1-2), 45-55. doi:10.1080/07011784.2015.1004198

Schroeder, R., Jacobs, J.M., Cho, E., Olheiser, C., DeWeese, M., Connelly, B., Cosh, M., Jia, X., Vuyovich, C., Tuttle. S.E., 2019. Comparison of satellite passive microwave with modeled snow water equivalent estimates in the Red River of the North basin. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. doi: 10.1109/JSTARS.2019.2926058

Schetselaar E.M., Rencz A.N., 1997. Reducing the effects of vegetation cover on airborne radiometric data using Landsat TM data. Int. J. Remote Sens. 18, 1503–1515.

Stadnyk, T., Dow, K., Wazney, L., Blais, E.-L., 2016. The 2011 flood event in the Red River Basin: causes, assessment and damages. Can. Water Resour. J., 41(1-2), 65-73

Takala, M., Luojus, K., Pulliainen, J., Derksen, C., Lemmetyinen, J., Kärnä, J. P., Koskinen, J., & Bojkov, B., 2011. Estimating northern hemisphere snow water equivalent for climate 811 research through assimilation of space-borne radiometer data and ground-based 812 measurements. Remote Sens. Environ. 115(12), 3517-3529.

Tait, A. B., 1998. Estimation of snow water equivalent using passive microwave radiation data. Remote Sens. Environ. 64(3), 286-291.

Todhunter, P. E., 2001. A hydroclimatological analysis of the Red River of the North snowmelt flood catastrophe of 1997. JAWRA J. Am. Water Resour. Assoc. 37(5), 1263-1278

Tuttle, S. E., Cho, E., Restrepo, P. J., Jia, X. Vuyovich, C. M., Cosh, M. H., and Jacobs, J. M., 2017. Remote Sensing of Drivers of Spring Snowmelt Flooding in the North Central U.S, 21-45. doi:10.1007/978-3-319-43744-6_2

Tuttle, S. E., Jacobs, J. M., Vuyovich, C. M., Olheiser, C., Cho, E., 2018. Intercomparison of snow water equivalent observations in the Northern Great Plains. Hydrol. Process. 32(6), 817-829.

Vittucci, C., Ferrazzoli, P., Kerr, Y., Richaume, P., Guerriero, L., Rahmoune, R., Vaglio Laurin, G., 2016. SMOS retrieval over forests: exploitation of optical depth and tests of soil moisture estimates. Remote Sens. Environ. 180, 115–127.

Vuyovich, C., Jacobs, J. M., 2011. Snowpack and runoff generation using AMSR-E passive microwave observations in the Upper Helmand Watershed, Afghanistan. Remote Sens. Environ., 115(12), 3313-3321. doi:10.1016/j.rse.2011.07.014

Vuyovich, C. M., Jacobs, J. M., Hiemstra, C. A., Deeb, E. J., 2017. Effect of spatial variability of wet snow on modeled and observed microwave emissions. Remote Sens. Environ. 198, 310-320. doi:10.1016/j.rse.2017.06.016

Wang, J. R., Mcmurtrey, J. E., III, Engman, E. T., Jackson, T. J., Schmugge, T. J., Gould, W. I., et al., 1982. Radiometric measurements over bare and vegetated fields at 1.4-GHz and 5-GHz frequencies. Remote Sens. Environ. 12(4), 295–311.

Wazney, L., Clark, S.P., 2015. The 2009 flood event in the Red River Basin: Causes, assessment and damages. Can. Water Resour. J. / Revue canadienne des ressources hydriques, 41(1-2), 56-64. doi:10.1080/07011784.2015.1009949

Wigneron, J. P., Calvet, J. C., Pellarin, T., Van de Griend, A. A., Berger, M., & Ferrazzoli, P., 2003. Retrieving near-surface soil moisture from microwave radiometric observations: current status and future plans. Remote Sens. Environ. 85(4), 489-506.

Wood, E. F., Lettenmaier, D., Liang, X., Nijssen, B., & Wetzel, S. W., 1997. Hydrological modeling of continental-scale basins. Annual Review of Earth and Planetary Sciences, 25(1), 279-300.

Woods F.W., Hough W.A., O'Nea D. Barnett J., 1965. Gamma ray attenuation by Loblolly pine wood: An investigation of integral counting. For. Sci. 11, 341–345.

Xia, Y., Sheffield, J., Ek, M. B., Dong, J., Chaney, N., Wei, H., Meng, J., Wood, E. F., 2014. Evaluation of multi-model simulated soil moisture in NLDAS-2. J. Hydrol. 512, 107-125. https://doi.org/10.1016/j.jhydrol.2014.02.027

Xia, Y., Ek, M. B., Wu, Y., Ford, T., Quiring, S. M., 2015. Comparison of NLDAS-2 simulated and NASMD observed daily soil moisture. Part I: Comparison and analysis. J. Hydrometeorol. 16(5), 1962-1980. https://doi.org/10.1175/JHM-D-14-0096.1

Zhang, R., Kim, S., & Sharma, A., 2019. A comprehensive validation of the SMAP Enhanced Level-3 Soil Moisture product using ground measurements over varied climates and landscapes. Remote Sens. Environ. 223, 82–94. doi:10.1016/j.rse.2019.01.015

Zotimov, N. V., 1968. Investigation of a method of measuring snow storage by using the gamma radiation of the earth. Sov. Hydrol. Sel. Pap., (3), 254-266.

Zwieback, S., Colliander, A., Cosh, M. H., Martínez-Fernándezj, J., McNairn, H., Starks, P. J., ... & Berg, A., 2018. Estimating time-dependent vegetation biases in the SMAP soil moisture product. Hydrol. Earth Syst. Sci., 22(8), 4473-4489.

Chapter 5

Adams, R. M., Houston, L. L, & Weiher, R. F. (2004). The Value of Snow and Snow Information Services. Report prepared for NOAA's Office of Program, Planning and Integration under contract DG1330-03-SE-1097.

Anderson, E. A. (2006). Snow accumulation and ablation model—SNOW-17. User's manual, *NOAA National Weather Service*, pp. 61 http://www.nws.noaa.gov/oh/hrl/nwsrfs/users_manual/part2/_pdf/22snow17.pdf

André, C., Ottlé, C., Royer, A., & Maignan, F. (2015). Land surface temperature retrieval over circumpolar Arctic using SSM/I-SSMIS and MODIS data. *Remote Sensing of Environment*, 162, 1–10.

Armstrong, R., Knowles, K. Brodzik, M. & Hardman, M. (1994). DMSP SSM/I-SSMIS pathfinder daily EASE-grid brightness temperatures. Version 2. NASA National Snow Ice Data Center Distributed Active Archive Center: Boulder, CO, USA. [accessed on 1 March, 2019]

Armstrong, R.L. & Brodzik, M. J. (2001). Recent northern hemisphere snow extent: A comparison of data derived from visible and microwave satellite sensors. *Geophysical Research Letters*, 28, 3673–3676.

Ashfaq, M., Ghosh, S., Kao, S.C., Bowling, L.C., Mote, P., Touma, D., Rauscher, S.A. & Diffenbaugh, N.S. (2013). Near-term acceleration of hydroclimatic change in the western US. *Journal of Geophysical Research: Atmospheres*, 118(19), 10-676.

Barrett, A. P. (2003). National operational hydrologic remote sensing center snow data assimilation system (SNODAS) products at NSIDC (p. 19). Boulder, CO: National Snow and Ice Data Center, Cooperative Institute for Research in Environmental Sciences.

Brodzik, M. J. (2014). F17 vs. F13 SWE Regression. Available online: http://cires1.colorado.edu/~brodzik/F13-F17swe/ [accessed on 27 March 2019].

Broxton, P., Dawson, N., & Zeng, X. (2016a). Linking snowfall and snow accumulation to generate spatial maps of SWE and snow depth. *Earth and Space Science*, 3, 246–256. https://doi.org/10.1002/2016EA000174

Broxton, P., Zeng, X., & Dawson, N. (2016b). Why do global reanalyses and land data assimilation products underestimate snow water equivalent? *Journal of Hydrometeorology*, 17(11), 2743–2761. https://doi.org/10.1175/JHM-D-16-0056.1

Carroll, T. R. (1981). Airborne soil moisture measurement using natural terrestrial gamma radiation. *Soil Science*, 132(5), 358-366.

Carroll, T. R. (2001). Airborne Gamma Radiation Snow Survey Program: A User's Guide, Version 5.0. National Operational Hydrologic Remote Sensing Center (NOHRSC), Chanhassen, 14

Carroll S. S. & Carroll T. R. (1989a), Effect of forest biomass on airborne snow water equivalent estimates obtained by measuring terrestrial gamma radiation. *Remote Sensing of Environment*, 27(3), 313–319.

Carroll S. S., & Carroll T. R. (1989b), Effect of uneven snow cover on airborne snow water equivalent estimates obtained by measuring terrestrial gamma radiation. *Water Resources Research*, 25 (7):1505–1510.

Carroll, T. R. & Schaake Jr, J. C. (1983). Airborne snow water equivalent and soil moisture measurement using natural terrestrial gamma radiation. *In Optical Engineering for Cold Environments* (Vol. 414, pp. 208-214). International Society for Optics and Photonics.

Carroll, T. R. & Vose, G. D. (1984). Airborne snow water equivalent measurements over a forested environment using terrestrial gamma radiation. *In Proceedings of the Eastern Snow Conference*, 29, 101-115.

Channan, S., Collins, K., & Emanuel, W. R. (2014). Global mosaics of the standard MODIS land cover type data. University of Maryland and the Pacific Northwest National Laboratory, College Park, Maryland, USA, 30.

Cai, S., Li, D., Durand, M., & Margulis, S. A. (2017). Examination of the impacts of vegetation on the correlation between snow water equivalent and passive microwave brightness temperature. *Remote Sensing of Environment*, 193, 244-256. https://doi.org/10.1016/j.rse.2017.03.006

Cho, E., Jacobs, J. M., Tuttle, S. E., Schroeder, R., & Olheiser, C. (2020). Improvement of Airborne Gamma Radiation Snow Water Equivalent Measurements Using SMAP Soil Moisture, *Remote Sensing of Environment.*

Cho, E., Tuttle, S. E., & Jacobs, J. M. (2017). Evaluating Consistency of Snow Water Equivalent Retrievals from Passive Microwave Sensors over the North Central US: SSM/I vs. SSMIS and AMSR-E vs. AMSR2. *Remote Sensing*, 9(5), 465. https://doi.org/10.3390/rs9050465

Clow, D. W., Nanus, L., Verdin, K. L., & Schmidt, J. (2012). Evaluation of SNODAS snow depth and snow water equivalent estimates for the Colorado Rocky Mountains, USA. *Hydrological Processes*, 26(17), 2583-2591.

Cooper, M.G., Nolin, A.W. & Safeeq, M. (2016). Testing the recent snow drought as an analog for climate warming sensitivity of Cascades snowpacks. *Environmental Research Letters*, 11(8), p.084009.

Cork, H. F., & Loijens, H. S. (1980). The effect of snow drifting on gamma snow survey results. *Journal of Hydrology*, 48(1-2), 41-51.

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., et al. (2008). Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology: a Journal of the Royal Meteorological Society*, 28(15), 2031-2064.

Dawson, N., Broxton, P., & Zeng, X. (2017). A new snow density parameterization for land data assimilation. *Journal of Hydrometeorology*, 18(1), 197–207. https://doi.org/10.1175/JHM-D-16-0166.1

Dawson, N., Broxton, P., & Zeng, X. (2018). Evaluation of remotely-sensed snow water equivalent and snow cover extent over the contiguous United States. *Journal of Hydrometeorology*. https://doi.org/10.1175/JHM-D-18-0007.1

Derksen, C., Walker, A., & Goodison, B. (2005). Evaluation of passive microwave snow water equivalent retrievals across the boreal forest/tundra transition of western Canada. *Remote Sensing of Environment*, 96(3-4), 315-327.

Derksen, C., LeDrew, E., Walker, A., & Goodison, B. (2000). The influence of sensor overpass time on passive microwave retrieval of snow cover parameters. *Remote Sensing of Environment*, 71(3), 297–308.

Derksen, C. & Walker, A. (2003). Identification of systematic bias in the cross-platform (SMMR and SSM/I) EASE-Grid brightness temperature time series. *IEEE Transactions on Geoscience and Remote Sensing*, 41(4), 910–915.

Dong, J., Walker, J., & Houser, P. (2005), Factors affecting remotely sensed snow water equivalent uncertainty, *Remote Sensing of Environment*, 97, 68–82, doi:10.1016/j.rse.2005.04.010

Foster, J. L., Sun, C., Walker, J. P., Kelly, R., Chang, A., Dong, J., & Powell, H. (2005). Quantifying the uncertainty in passive microwave snow water equivalent observations. *Remote Sensing of Environment*, 2005, 94, 187–203.

Georgakakos, A., Fleming, P., Dettinger, M., Peters-Lidard, C., Richmond, T. T. C., Reckhow, K., White, K. & Yates, D. (2014), Ch. 3: Water Resources, *Climate Change Impacts in the United States: The Third National Climate Assessment*, edited by J. M. Melillo, T. C. Richmond and G. W. Yohe, 69-112, U.S. Global Change Research Program, doi:10.7930/J0G44N6T.

Glynn, J. E., Carroll, T. R., Holman, P. B., & Grasty, R. L. (1988). An airborne gamma ray snow survey of a forested covered area with a deep snowpack, *Remote Sensing of Environment*, 26(2), 149-160.

Goodison, B.E., Banga, A., & Halliday, R.A. (1984). Canada—United States Prairie Snow Cover Runoff Study, *Canadian Water Resources Journal*, 9:1, 99-107, https://doi.org/10.4296/cwrj0901099

Hancock, S., Baxter, R., Evans, J., & Huntley, B. (2013). Evaluating global snow water equivalent products for testing land surface models. *Remote sensing of environment*, 128, 107–117. https://doi.org/10.1016/j.rse.2012.10.004

Hansen, M. & Song, X. P. (2018). Vegetation Continuous Fields (VCF) Yearly Global 0.05 Deg. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MEaSUREs/VCF/VCF5KYR.001. [accessed on 1 March, 2019]

Hill, D. F., Burakowski, E. A., Crumley, R. L., Keon, J., Hu, J. M., Arendt, A. A., ... & Wolken, G. J. (2019). Converting snow depth to snow water equivalent using climatological variables. *The Cryosphere*, 13(7), 1767-1784.

Horn, B.K.P. (1981). Hill shading and the reflectance map. *Proceedings of the IEEE* 69:14-47

Kang, D.H., Barros, A.P., & Dery, S.J. (2014). Evaluating passive microwave radiometry for the dynamical transition from dry to wet snowpacks. *IEEE Transactions on Geoscience and Remote Sensing*, 52, 3–15. https://doi.org/10.1109/TGRS.2012.2234468

Larue, F., Royer, A., De Sève, D., Langlois, A., Roy, A., & Brucker, L. (2017). Validation of GlobSnow-2 snow water equivalent over Eastern Canada. *Remote sensing of environment*, 194, 264-277. https://doi.org/10.1016/j.rse.2017.03.027

Loveland, T. R., & Belward, A. S. (1997). The IGBP-DIS global 1km land cover data set, DISCover: first results. *International Journal of Remote Sensing*, 18(15), 3289-3295.

Luojus, K., Pullianen, J., Takala, M., Lemmetyinen, J., Kangwa, M., Eskelinen, M., et al. (2014) GlobSnow-2 Final Report, *Global Snow Monitoring for Climate Research, European Space Agency*, http://www.globsnow.info/docs/GlobSnow_2_Final_Report_release.pdf.

Mätzler, C. & Standley, A. (2000), Relief effects for passive microwave remote sensing, Int. J. Remote Sens., 21, 2403–2412, doi:10.1080/ 01431160050030538.

Molotch, N. P., & Bales, R. C. (2005). Scaling snow observations from the point to the grid element: Implications for observation network design. *Water Resources Research*, 41, W11421. https://doi.org/10.1029/2005WR004229

Mote, P.W., Li, S., Lettenmaier, D.P., Xiao, M. & Engel, R. (2018). Dramatic declines in snowpack in the western US. *npj Climate and Atmospheric Science*, 1(1), 2. doi:10.1038/s41612-018-0012-1

Mudryk, L. R., Derksen, C., Kushner, P. J., & Brown, R. (2015). Characterization of Northern Hemisphere snow water equivalent datasets, 1981–2010. *Journal of Climate*, http://dx.doi.org/10.1175/JCLI-D-15-0229.1.

Painter, T. H., Berisford, D. F., Boardman, J. W., Bormann, K. J., Deems, J. S., Gehrke, F., et al. (2016). The airborne snow observatory: Fusion of scanning lidar, imaging spectrometer, and physically-based modeling for mapping snow water equivalent and snow albedo. *Remote Sensing of Environment*, 184, 139–152. https://doi.org/10.1016/j.rse.2016.06.018

Peck, E. L., Bissell, V. C., Jones, E. B., & Burge, D. L. (1971). Evaluation of snow water equivalent by airborne measurement of passive terrestrial gamma radiation. *Water Resources Research*, 7(5), 1151–1159.

Peck, E. L., Carroll, T. R., & VanDemark, S. C. (1980). Operational aerial snow surveying in the United States/Etude de neige aérienne effectuée aux Etats Unis. *Hydrological Sciences Journal*, 25(1), 51–62

Pierce, D.W., Barnett, T.P., Hidalgo, H.G., Das, T., Bonfils, C., Santer, B.D., Bala, G., Dettinger, M.D., Cayan, D.R., Mirin, A. & Wood, A.W. (2008). Attribution of declining western US snowpack to human effects. *Journal of Climate*, 21(23), 6425-6444.

Pulliainen, J., & Hallikainen, M. (2001). Retrieval of regional snow water equivalent from space-borne passive microwave observations. *Remote Sensing of Environment*, 75, 76−85.

Pulliainen, J. (2006). Mapping of snow water equivalent and snow depth in boreal and sub-arctic zones by assimilating space-borne microwave radiometer data and ground-based observations. *Remote sensing of Environment*, 101(2), 257-269.

Royer, A. & Poirier, S. (2010). Surface temperature spatial and temporal variations in North America from homogenized satellite SMMR-SSM/I microwave measurements and reanalysis for 1979–2008. *Journal of Geophysical Research: Atmospheres*, 115, D08110.

Schroeder, R., Jacobs, J. M., Cho, E., Olheiser, C. M., DeWeese, M. M., Connelly, B. A., ... & Tuttle, S. E. (2019). Comparison of Satellite Passive Microwave with Modeled Snow Water Equivalent Estimates in the Red River of the North Basin. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(9), 3233-3246.

Serreze, M. C., Clark, M. P., Armstrong, R. L., McGinnis, D. A., & Pulwarty, R. S. (1999). Characteristics of the western United States snowpack from Snowpack Telemetry (SNOTEL) data. *Water Resources Research*, 35, 2145–2160, doi:10.1029/1999WR900090

Simonovic, S. P. (1999). Decision support system for flood management in the Red River Basin. *Canadian Water Resources Journal*, 24(3), 203-223.

Stadnyk, T., Dow, K., Wazney, L., & Blais, E.-L. (2016). The 2011 flood event in the Red River Basin: causes, assessment and damages. *Canadian Water Resources Journal*, 41(1-2), 65-73. https://doi.org/10.1080/07011784.2015.1009949

Sturm, M., Holmgren, J., & Liston, G. E. (1995). A seasonal snow cover classification system for local to global applications. *Journal of Climate*, 8(5), 1261-1283. https://doi.org/10.1175/1520-0442(1995)008<1261:ASSCCS>2.0.CO;2

Sturm, M., Taras, B., Liston, G. E., Derksen, C., Jonas, T., & Lea, J. (2010). Estimating snow water equivalent using snow depth data and climate classes. *Journal of Hydrometeorology*, 11(6), 1380-1394. https://doi.org/10.1175/2010JHM1202.1

Takala, M., Luojus, K., Pulliainen, J., Derksen, C., Lemmetyinen, J., Kärnä, J. P., Koskinen, J., & Bojkov, B. (2011). Estimating northern hemisphere snow water equivalent for climate research through assimilation of space-borne radiometer data and ground-based measurements. *Remote Sensing of Environment*, 115(12), 3517-3529.

Tait, A. B. (1998). Estimation of snow water equivalent using passive microwave radiation data. *Remote Sensing of Environment*, 64(3), 286-291.

Todhunter, P. E. (2001). A hydroclimatological analysis of the Red River of the North snowmelt flood catastrophe of 1997. *JAWRA Journal of the American Water Resources Association*, 37(5), 1263-1278. https://doi.org/10.1111/j.1752-1688.2001.tb03637.x

Tuttle, S. E., Cho, E., Restrepo, P. J., Jia, X. Vuyovich, C. M., Cosh, M. H., & Jacobs, J. M. (2017). Remote Sensing of Drivers of Spring Snowmelt Flooding in the North Central U.S, *Remote Sensing of Hydrological Extremes*, edited by V. Lakshmi. Springer International Publishing, 21-45. https://doi.org/10.1007/978-3-319-43744-6_2

Tuttle, S. E., Jacobs, J. M., Vuyovich, C. M., Olheiser, C., & Cho, E. (2018). Intercomparison of snow water equivalent observations in the Northern Great Plains. *Hydrological Processes*, 32(6), 817-829. https://doi.org/10.1002/hyp.11459

Vogel, R. M., Carroll, T. R., & Carroll, S. S. (1985), Simulation of airborne snow water equivalent measurement errors made over a forest environment, *Proceedings of the American Society of Civil Engineers Symposium*, Denver, CO, p. 9.

Vuyovich, C. M., Jacobs, J. M., Hiemstra, C. A., & Deeb, E. J. (2017). Effect of spatial variability of wet snow on modeled and observed microwave emissions. *Remote Sensing of Environment*, 198, 310-320. https://doi.org/10.1016/j.rse.2017.06.016

Walker, A.E. & Goodison, B.E. (1993). Discrimination of a wet snow cover using passive microwave satellite data. *Anuals of Glaciology*, 17, 307–311. https://doi.org/10.3189/S026030550001301X

Wazney, L. & Clark, S.P. (2015). The 2009 flood event in the Red River Basin: Causes, assessment and damages. *Canadian Water Resources Journal*, 41(1-2), 56-64. https://doi.org/10.1080/07011784.2015.1009949

Wilson, M.F.J., O'Connell, B., Brown, C., Guinan, J.C., Grehan, A.J. (2007). Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. Marine Geodesy 30: 3-35.

Winstral, A., Elder, K., & Davis, R. (2002). Spatial snow modeling of windredistributed snow using terrain-based parameters. *Journal of Hydrometeorology* 3, 524–538.

Zeng, X., Broxton, P., & Dawson, N. (2018). Snowpack change from 1982 to 2016 over conterminous United States. *Geophysical Research Letters*, 45, 12,940–12,947. https://doi.org/10.1029/2018GL079621

Chapter 6

Adams, R. M., Houston, L. L, & Weiher, R. F. (2004). The Value of Snow and Snow Information Services. Report prepared for NOAA’s Office of Program, Planning and Integration under contract DG1330-03-SE-1097.

American National Standards Institute (1972). Building code requirements for minimum design loads in buildings and other structures. ANSI A58.1-1972, New York.

American Society of Civil Engineers (2017). Minimum Design Loads and Associated Criteria for Buildings and Other Structures. ASCE/SEI 7-16, Reston, VA.

Anderson, B. T. (2011). Spatial Distribution and Evolution of a Seasonal Snowpack in Complex Terrain: An Evaluation of the SNODAS Modeling Product, *Masters thesis, Boise State University*, Boise, Idaho, USA.

Bair, E. H., Rittger, K., Davis, R. E., Painter, T. H., & Dozier, J. (2016). Validating reconstruction of snow water equivalent in California's Sierra Nevada using measurements from the NASA Airborne Snow Observatory. Water Resources Research, 52, 8437-8460. https://doi.org/10.1002/2016WR018704

Barrett, A. P. (2003). National operational hydrologic remote sensing center snow data assimilation system (SNODAS) products at NSIDC (p. 19). Boulder, CO: National Snow and Ice Data Center, Cooperative Institute for Research in Environmental Sciences.

Berghuijs, W. R., Woods, R. A., Hutton, C. J., & Sivapalan, M. (2016). Dominant flood generating mechanisms across the United States. *Geophysical Research Letters*, 43(9), 4382-4390. https://doi.org/10.1002/2016GL068070

Boniface, K., Braun, J. J., McCreight, J. L., & Nievinski, F. G. (2015). Comparison of snow data assimilation system with GPS reflectometry snow depth in the western United States. *Hydrological processes*, 29(10), 2425-2437. https://doi.org/10.1002/hyp.10346

Bonnin, G. M., Martin, D., Lin, B., Parzybok, T., Yekta, M., & Riley, D. (2006). NOAA Atlas 14 Volume 2 Version 3, *Precipitation-Frequency Atlas of the United States, Ohio River Basin and Surrounding States.* NOAA, National Weather Service, Silver Spring, MD.

Broxton, P., Dawson, N., & Zeng, X. (2016a). Linking snowfall and snow accumulation to generate spatial maps of SWE and snow depth. *Earth and Space Science*, 3, 246–256. https://doi.org/10.1002/2016EA000174

Broxton, P., Zeng, X., & Dawson, N. (2016b). Why do global reanalyses and land data assimilation products underestimate snow water equivalent? *Journal of Hydrometeorology*, 17(11), 2743–2761. https://doi.org/10.1175/JHM-D-16-0056.1

Broxton, P., X. Zeng, & N. Dawson. (2019). Daily 4 km Gridded SWE and Snow Depth from Assimilated In-Situ and Modeled Data over the Conterminous US, Version 1. Boulder, Colorado USA. *NASA National Snow and Ice Data Center Distributed Active Archive Center*. https://doi.org/10.5067/0GGPB220EX6A. [accessed on 1 May, 2019].

Byun, K., Chiu, C.-M., & Hamlet, A. F. (2018). Effects of 21st century climate change on seasonal flow regimes and hydrologic extremes over the Midwest and Great Lakes region of the US. *Science of The Total Environment*. doi:10.1016/j.scitotenv.2018.09.063

Carroll, T., D. Cline, C. Olheiser, A. Rost, A. Nilsson, G. Fall, C. Bovitz, and L. Li (2006), NOAA's National Snow Analyses, Proc. Western Snow Conference, Las Cruces, NM, vol. 74. [Available at www.westernsnowconference.org/sites/westernsnowconference.org/PDFs/2006Carroll.pdf.]

Changnon, S. A., Kunkel, K. E., & Andsager, K. (2001). Causes for record high flood losses in the central United States. *Water International*, 26(2), 223-230.

Cho E., Jacobs, J.M., Vuyovich, C. (2019). The value of long-term (40 years) airborne gamma radiation SWE record for evaluating three observation-based gridded SWE datasets by seasonal snow and land cover classifications, *Water Resources Research*. https://doi.org/10.1029/2019WR025813

Chowdhury, J. U., Stedinger, J. R., Lu, L.-H. (1991). Goodness-of-Fit Tests for Regional Generalized Extreme Value Flood Distributions, *Water Resources Research* 27, 1765-1776.

Clow, D. W., Nanus, L., Verdin, K. L., & Schmidt, J. (2012). Evaluation of SNODAS snow depth and snow water equivalent estimates for the Colorado Rocky Mountains, USA. *Hydrological Processes*, 26(17), 2583-2591.

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., et al. (2008). Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology: a Journal of the Royal Meteorological Society*, 28(15), 2031-2064.

Dawson, N., Broxton, P., & Zeng, X. (2017). A new snow density parameterization for land data assimilation. *Journal of Hydrometeorology*, 18(1), 197–207. https://doi.org/10.1175/JHM-D-16-0166.1

Dawson, N., Broxton, P., & Zeng, X. (2018). Evaluation of remotely-sensed snow water equivalent and snow cover extent over the contiguous United States. *Journal of Hydrometeorology*. https://doi.org/10.1175/JHM-D-18-0007.1

Dozier, J., Bair, E. H., & Davis, R. E. (2016). Estimating the spatial distribution of snow water equivalent in the world's mountains. WIREs Water, 3, 461-474. https://doi.org/10.1002/wat2.1140

Easterling, D.R., Kunkel, K.E., Arnold, J.R., Knutson, T., LeGrande, A.N., Leung, L.R., Vose, R.S., Waliser, D.E., and Wehner, M.F. (2017). Precipitation change in the United States. In: *Climate Science Special Report: Fourth National Climate Assessment*, Volume I [Wuebbles, D.J., D.W. Fahey, K.A. Hibbard, D.J. Dokken, B.C. Stewart, and T.K. Maycock (eds.)]. U.S. Global Change Research Program, Washington, DC, USA, pp. 207-230, doi: 10.7930/J0H993CC.

Fassnacht, S. R., and Records, R. M. (2015). Large snowmelt versus rainfall events in the mountains, *Journal of Geophysical Research: Atmospheres*, 120, 2375–2381, doi:10.1002/2014JD022753.

Filliben, J. J. (1975). The Probability Plot Correlation Coefficient Test for Normailty, *Technometrics*, 17, 111-117.

Hedrick, A., Marshall, H. P., Winstral, A., Elder, K., Yueh, S., & Cline, D. (2015). Independent evaluation of the SNODAS snow depth product using regional-scale lidar-derived measurements. *The Cryosphere*, 9(1), 13-23.

Helsel, D. R., & Hirsch, R. M. (1992). Statistical methods in water resources (Vol. 49). Elsevier.

Hirsch, R.M., Ryberg, K.R., 2012. Has the magnitude of floods across the USA changed with global CO 2 levels. *Hydrological Sciences Journal*, 57, 37–41. https://doi.org/10.1080/02626667.2011.621895.

Hodgkins, G. A., & Dudley, R. W. (2006). Changes in late-winter snowpack depth, water equivalent, and density in Maine, 1926–2004. *Hydrological Processes: An International Journal*, 20(4), 741-751.

Hosking, J. R. M. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society: Series B (Methodological)*, 52(1), 105-124.

Hosking, J. R. M., Wallis, J. R., & Wood, E. F. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, 27(3), 251-261.

Kendall, M. G. (1938), A new measure of rank correlation, *Biometrika*, 30(1/2), 81–93, doi:10.2307/2332226

Khaliq, M. N., Ouarda, T. B. M. J., Ondo, J. C., Gachon, P. and Bobée, B. (2006), Frequency analysis of a sequence of dependent and/or nonstationary hydro-meteorological observations: A review, *Journal of Hydrology*, 329(3–4), 534–552, doi:10.1016/j.jhydrol.2006.03.004

Kunkel, K. E., Stevens, L. E., Stevens, S. E., Sun, L., Janssen, E., Wuebbles, D., Hilberg, S., Timlin, M., Stoecker, L., and Westcott, N. (2013). *Regional climate trends and scenarios for the US national climate assessment: Part 4. Climate of the U.S. Great Plains*, 91 pp, NOAA Washington, D.C.

Looney, S. W., & Gulledge Jr, T. R. (1985). Use of the correlation coefficient with normal probability plots. *The American Statistician*, 39(1), 75-79.

Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica: Journal of the Econometric Society*, 245-259. doi:10.2307/1907187

McGinnis, S., Nychka, D., & Mearns, L. O. (2015). A new distribution mapping technique for climate model bias correction. In *Machine learning and data mining approaches to* climate science, 91-99, Springer, Cham.

Mote, P. W., Li, S., Lettenmaier, D. P., Xiao, M., & Engel, R. (2018). Dramatic declines in snowpack in the western US. *Npj Climate and Atmospheric Science*, 1(1), 2. doi:10.1038/s41612-018-0012-1

Ombadi, M., Nguyen, P., Sorooshian, S., & Hsu, K. L. (2018). Developing Intensity-Duration-Frequency (IDF) Curves From Satellite-Based Precipitation: Methodology and Evaluation. *Water Resources Research*, 54(10), 7752-7766.

Pederson, G. T., Betancourt, J. L., & McCabe, G. J. (2013). Regional patterns and proximal causes of the recent snowpack decline in the Rocky Mountains, U.S. *Geophysical Research Letters*, 40, 1811–1816, doi:10.1002/grl.50424.

Peterson, T. C., Heim Jr, R. R., Hirsch, R., Kaiser, D. P., Brooks, H., Diffenbaugh, N. S., ... & Katz, R. W. (2013). Monitoring and understanding changes in heat waves, cold waves, floods, and droughts in the United States: state of knowledge. *Bulletin of the American Meteorological Society*, 94(6), 821-834.

Perica, S., Martin, D., Pavlovic, S., Roy, I., St. Laurent, M., Trypaluk, C., Unruh, D., Yekta, M., Bonnin, G. (2013). NOAA Atlas 14 Volume 8 Version 2, *Precipitation-Frequency Atlas of the United States, Midwestern States*. NOAA, National Weather Service, Silver Spring, MD.

Perica, S., Pavlovic, S., St Laurent, M., Trypaluk, C., Unruh, D., Martin, D., Wilhite, O. (2015, revised 2019). NOAA Atlas 14 Volume 10 Version 3, *Precipitation-Frequency Atlas of the United States, Northeastern States.* NOAA, National Weather Service, Silver Spring, MD.

Pierce, D. W., Barnett, T. P., Hidalgo, H. G., Das, T., Bonfils, C., Santer, B. D., ... & Wood, A. W. (2008). Attribution of declining western US snowpack to human effects. *Journal of Climate*, 21(23), 6425-6444.

Sack, R. L. (2015). Ground snow loads for the Western United States: state of the art. *Journal of Structural Engineering*, 142(1), 04015082.

Schroeder, R., Jacobs, J. M., Cho, E., Olheiser, C. M., DeWeese, M. M., Connelly, B. A., ... & Tuttle, S. E. (2019). Comparison of Satellite Passive Microwave With Modeled Snow Water Equivalent Estimates in the Red River of the North Basin. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(9), 3233-3246.

Sen, P. K. (1968). Robustness of some nonparametric procedures in linear models. *The Annals of Mathematical Statistics*, 1913-1922.

Serreze, M. C., Clark, M. P., Armstrong, R. L., McGinnis, D. A., & Pulwarty, R. S. (1999). Characteristics of the western United States snowpack from Snowpack Telemetry (SNOTEL) data. *Water Resources Research,* 35, 2145–2160, doi:10.1029/1999WR900090

Slater, L. J., and Villarini, G. (2016), Recent trends in U.S. flood risk, *Geophysical Research Letters*, 43, 12,428–12,436, doi:10.1002/2016GL071199

Stadnyk, T., Dow, K., Wazney, L., & Blais, E.-L. (2016). The 2011 flood event in the Red River Basin: causes, assessment and damages. *Canadian Water Resources Journal*, 41(1-2), 65-73. https://doi.org/10.1080/07011784.2015.1009949

Stedinger, J. R., Vogel, R. M., & Foufoula-Georgiou, E. (1993). Chapter 18, Frequency analysis of extreme events, *Handbook of Hydrology, edited by Maidment, DR. DR McGraw-Hill, New York*.

Sturm, M., Holmgren, J., & Liston, G. E. (1995). A seasonal snow cover classification system for local to global applications. *Journal of Climate*, 8(5), 1261-1283. https://doi.org/10.1175/1520-0442(1995)008<1261:ASSCCS>2.0.CO;2

Sturm, M., Taras, B., Liston, G. E., Derksen, C., Jonas, T., & Lea, J. (2010). Estimating snow water equivalent using snow depth data and climate classes. *Journal of Hydrometeorology,* 11(6), 1380-1394. https://doi.org/10.1175/2010JHM1202.1

Takala, M., Luojus, K., Pulliainen, J., Derksen, C., Lemmetyinen, J., Kärnä, J. P., Koskinen, J., & Bojkov, B. (2011). Estimating northern hemisphere snow water equivalent for climate research through assimilation of space-borne radiometer data and ground-based measurements. *Remote Sensing of Environment*, 115(12), 3517-3529.

Todhunter, P. E. (2001). A hydroclimatological analysis of the Red River of the North snowmelt flood catastrophe of 1997. *JAWRA Journal of the American Water Resources Association*, 37(5), 1263-1278. https://doi.org/10.1111/j.1752-1688.2001.tb03637.x

Tuttle, S. E., Cho, E., Restrepo, P. J., Jia, X. Vuyovich, C. M., Cosh, M. H., & Jacobs, J. M. (2017). Remote Sensing of Drivers of Spring Snowmelt Flooding in the North Central U.S, 21-45. https://doi.org/10.1007/978-3-319-43744-6_2

United States Department of Commerce. (1964). Frequency of maximum water equivalent of March snow cover in North Central United States. Weather Bureau, Technical Paper, No. 50, Washington, DC. https://www.nws.noaa.gov/oh/hdsc/Technical_papers/TP50.pdf

Vogel, R. M. (1986). The Probability Plot Correlation Coefficient Test for the Normal, Lognormal, and Gumbel Distributional Hypotheses, *Water Resources Research*, 22, 587-590.

Vogel, R. M., & Fennessey, N. M. (1993). L moment diagrams should replace product moment diagrams. *Water resources research*, 29(6), 1745-1752.

Vuyovich, C. M., Jacobs, J. M., & Daly, S. F. (2014). Comparison of passive microwave and modeled estimates of total watershed SWE in the continental United States. *Water Resources Research*, 50(11), 9088-9102. https://doi.org/10.1002/2013WR014734

Wazney, L. & Clark, S.P. (2015). The 2009 flood event in the Red River Basin: Causes, assessment and damages. *Canadian Water Resources Journal*, 41(1-2), 56-64. https://doi.org/10.1080/07011784.2015.1009949

Zeng, X., Broxton, P., & Dawson, N. (2018). Snowpack change from 1982 to 2016 over conterminous United States. *Geophysical Research Letters*, 45, 12,940–12,947. https://doi.org/10.1029/2018GL079621

Chapter 7

Lark, T. J., J. Meghan Salmon, and H. K. Gibbs (2015), Cropland expansion outpaces agricultural and biofuel policies in the United States, Environ. Res. Lett., 10(4), 044003, doi:10.1088/1748-9326/10/4/044003.

King, K. W., N. R. Fausey, and M. R. Williams (2014), Effect of subsurface drainage on streamflow in an agricultural headwater watershed, J. Hydrol., 519, 438-445, doi:10.1016/j.jhydrol.2014.07.035.

Mahmood, R., et al. (2014), Land cover changes and their biogeophysical effects on climate, Int. J. Climatol., 34, 929–953, doi:10.1002/joc.3736

Margulis, S. A., Cortés, G., Girotto, M., & Durand, M. (2016). A Landsat-era Sierra Nevada snow reanalysis (1985–2015). Journal of Hydrometeorology, 17(4), 1203-1221.

McCabe, M. F., Rodell, M., Alsdorf, D. E., Miralles, D. G., Uijlenhoet, R., Wagner, W., et al. (2017). The future of Earth observation in hydrology. Hydrology and Earth System Sciences, 21(7), 3879–3914. https://doi.org/10.5194/hess-21-3879-2017

NASA CubeSat Launch Initiative (2018), https://www.nasa.gov/directorates/heo/home/CubeSats_initiative (Last access: 5 August 2019).

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

Otto, C. R., C. L. Roth, B. L. Carlson, and M. D. Smart (2016), Land-use change reduces habitat suitability for supporting managed honey bee colonies in the Northern Great Plains, Proceedings of the National Academy of Sciences, 113(37), 10430-10435.

Pei, L., N. J. Moore, S. Zhong, A. Kendall, Z. Gao, and D. W. Hyndman (2016), Effects of irrigation on summer precipitation over the United States, J. Clim., 29, 3541–3558, doi:10.1175/JCLI-D-15-0337.1

Planet Team (2018). Planet Application Program Interface: In Space for Life on Earth. San Francisco, CA. https://api.planet.com.

Zhang, X., Z. Xiong, and Q. Tang (2017), Modeled effects of irrigation on surface climate in the Heihe River Basin, Northwest China, J. Geophys. Res. Atmos., 122, doi:10.1002/2017JD026732.

APPENDIX

CHAPTER 2

Text S1. Variable equations

The following equations were used to calculate four vegetation indices, STR, and aridity:

- **NDVI** (Normalized Difference Vegetation Index)
  = (NIR – Red) / (NIR + Red)
- **EVI** (Enhanced Vegetation Index, *Huete et al.*, 2002)
  = 2.5 * (NIR -Red) / (NIR + 6*Red – 7.5 * Blue + 1)
- **GI** (Green Index, *Gitelson et al., 2005*)
  = NIR / Green
- **NDWI** (Normalized Difference Water Index, *Gao*, 1996)
  = (NIR – SWIR) / (NIR + SWIR)
- **STR** (Shortwave Infrared Transformed Reflectance, *Sadeghi et al.*, 2015)
  $= (1 – \text{SWIR})^2 / 2 * \text{SWIR}$
- **Aridity**
  = Precipitation / Potential Evapotranspiration (PET)

where NIR is the near-infrared band (0.77 - 0.90 µm), Red, Blue, and Green are the visible red (0.63 - 0.69 µm), blue (0.45 - 0.52 µm), and green (0.52 - 0.60 µm) band, respectively. SWIR is the shortwave Infrared Reflectance which is band 5 (SWIR1, 1.55 - 1.75 µm) and 7 (SWIR2, 2.08 - 2.35 µm) for Landsat 7 ETM+ and band 6 (SWIR1, 1.57 - 1.65 µm) and 7 (SWIR2, 2.11 - 2.29 µm) for Landsat 8 OLI/TIRS.

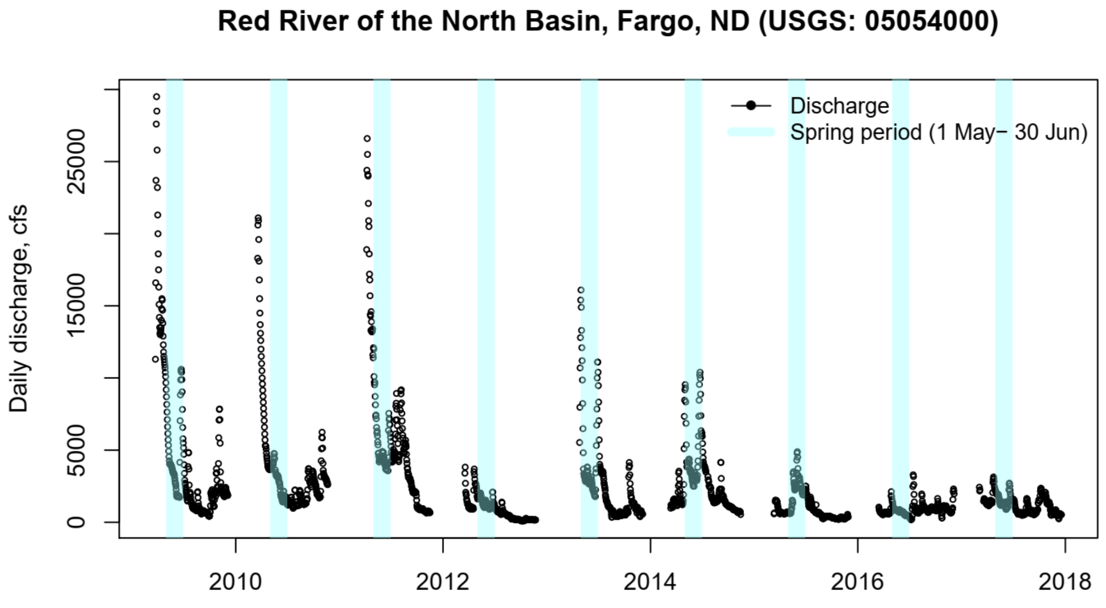


Figure S1. A hydrograph at Fargo, ND (USGS: 0505400), a major streamflow gage in the Red River of the North Basin

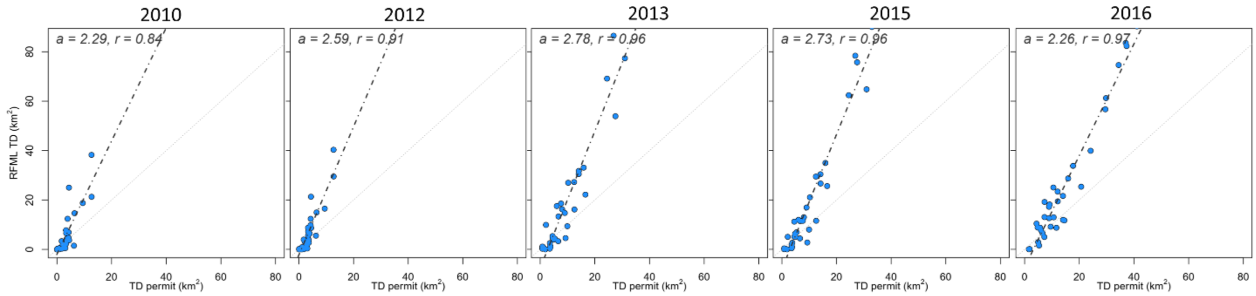


Figure S2. Subwatershed (HUC12)-level accuracy assessment over BdSW, Minnesota (N=34) in 2010, 2012, 2013, 2015, and 2016

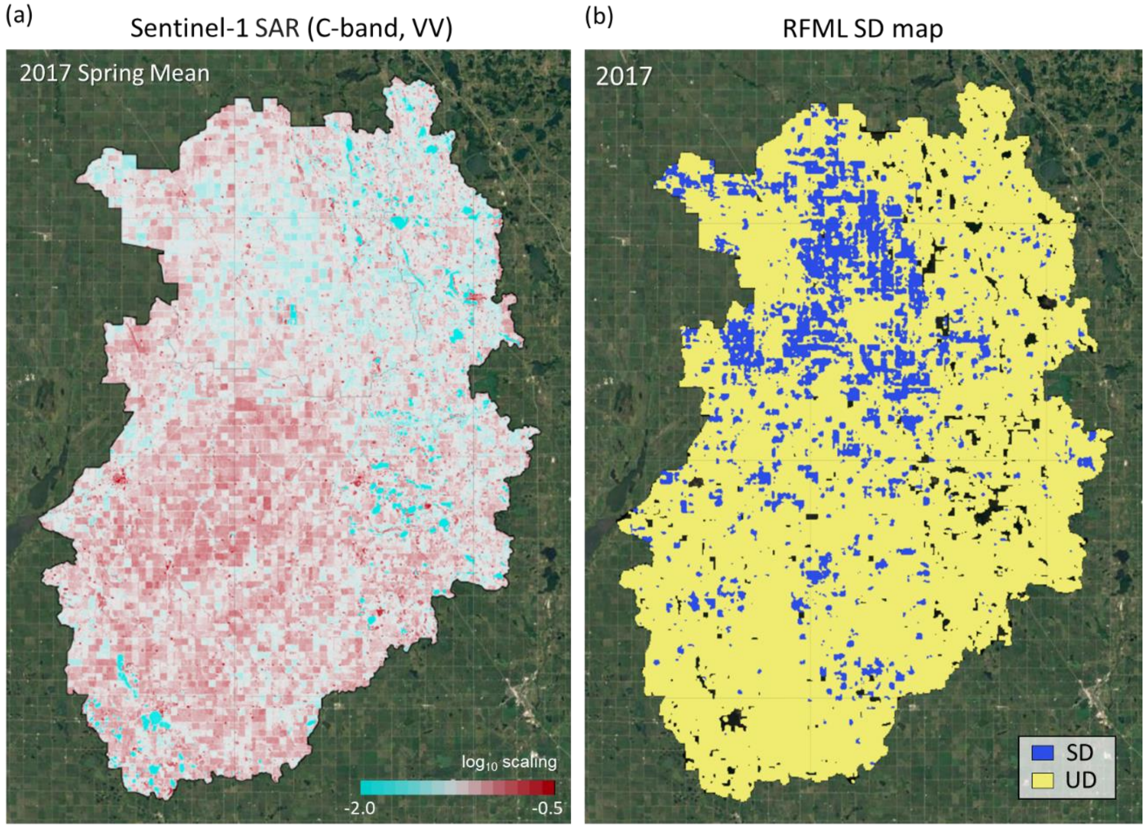


Figure S3. (a) A spring mean map of Sentinel-1 SAR data at VV polarization (as logarithm scale) and the RFML SD map generated by RFML including Sentinel-1 data (spring mean and range) as input variables over BdSW in 2017

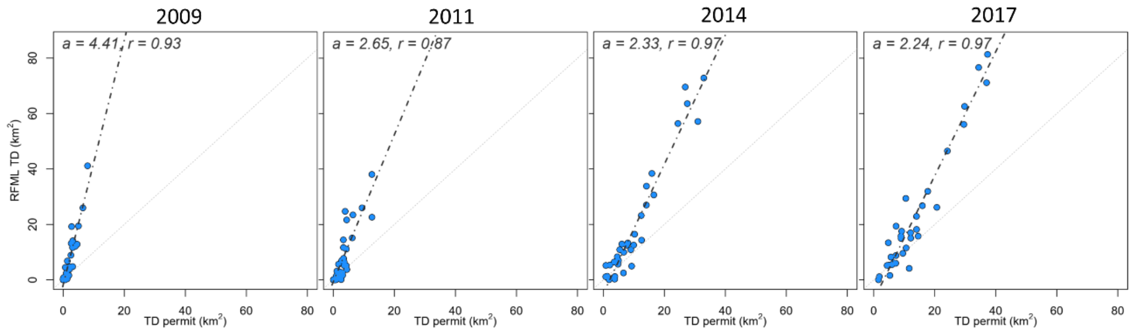


Figure S4. Subwatershed (HUC12)-level accuracy assessment of RFML SD maps (using 10 variables only) over BdSW, Minnesota (N=34). SD permit area from the BdSW district permit records compared with SD area from RFML classified maps against a 1:1 line (light dashed). Agreement between the two datasets was assessed with correlation coefficient (r) metrics from simple linear regression (trend line = thick dashed line, a = slope).

CHAPTER 4

Text S1. Statistical metrics

The agreement between airborne gamma survey and satellite/model SM (or SWE) products were quantified by the Pearson's linear correlation coefficient, *R*, the mean bias, *Bias*, the root mean square difference, *RMSD*, and the unbiased RMSD, *ubRMSD*, where additive bias is removed as follows:

$$R = \frac{\mathrm{cov}(\mathrm{V}_1, \mathrm{V}_2)}{std(\mathrm{V}_1) \cdot std(\mathrm{V}_2)}$$

$$Bias = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(V_{1,i} - V_{2,i})}$$

$$RMSD = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(V_{1,i} - V_{2,i})^2}$$

$$ubRMSD = \sqrt{\frac{1}{n}\sum_{t=1}^{n}[(V_{1,i} - \bar{V}_1) - (V_{2,i} - \bar{V}_2)]^2}$$

V_1 and V_2 refer to two different SM (or SWE) products. cov(•) and std(•) yield covariance and standard deviation statistics, respectively.

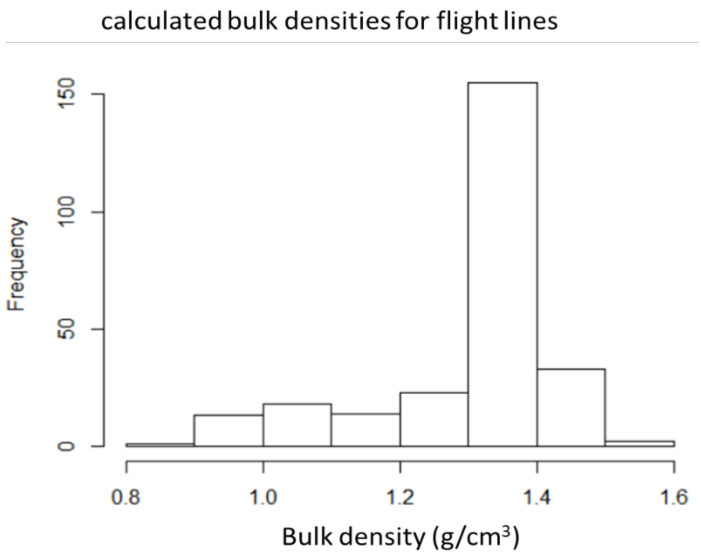


Figure S1. Histogram of bulk densities for gamma flight lines in U.S. calculated by selecting a majority value within the footprint using the 1-km POLARIS soil datasets (available at www.polaris.earth; Chaney et al., 2016)

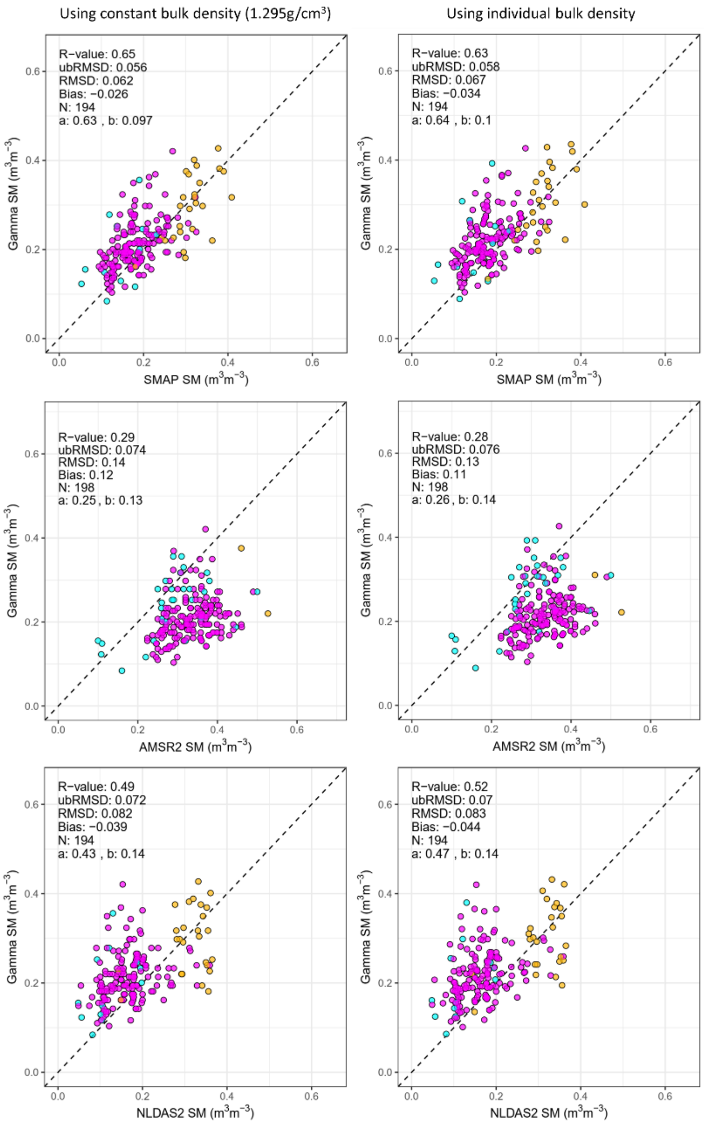


Figure S2. Comparison of agreements between NOAA airborne gamma soil moisture with constant bulk density, 1.295 g/cm^3 and individual bulk densities for each flight line, with Phase 2 of the North American Land Data Assimilation System (NLDAS-2) Mosaic, VIC, and Noah model SM products

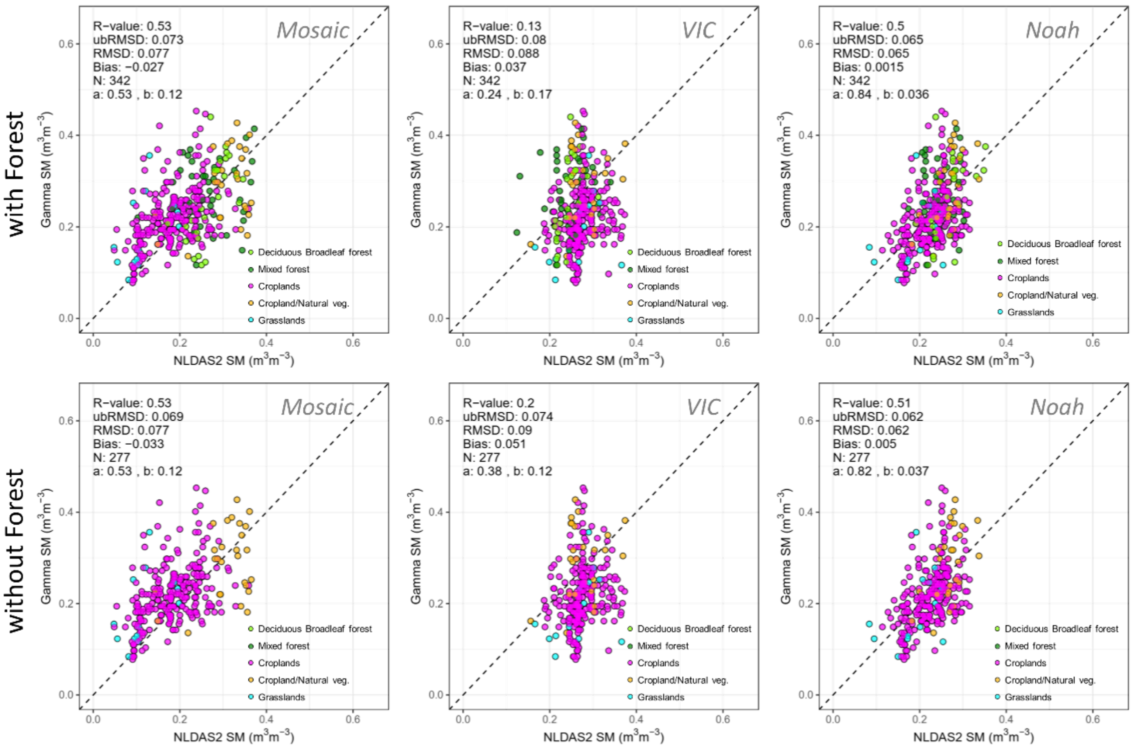


Figure S3. Comparison of NOAA airborne gamma soil moisture with Phase 2 of the North American Land Data Assimilation System (NLDAS-2) Mosaic, VIC, and Noah model SM products

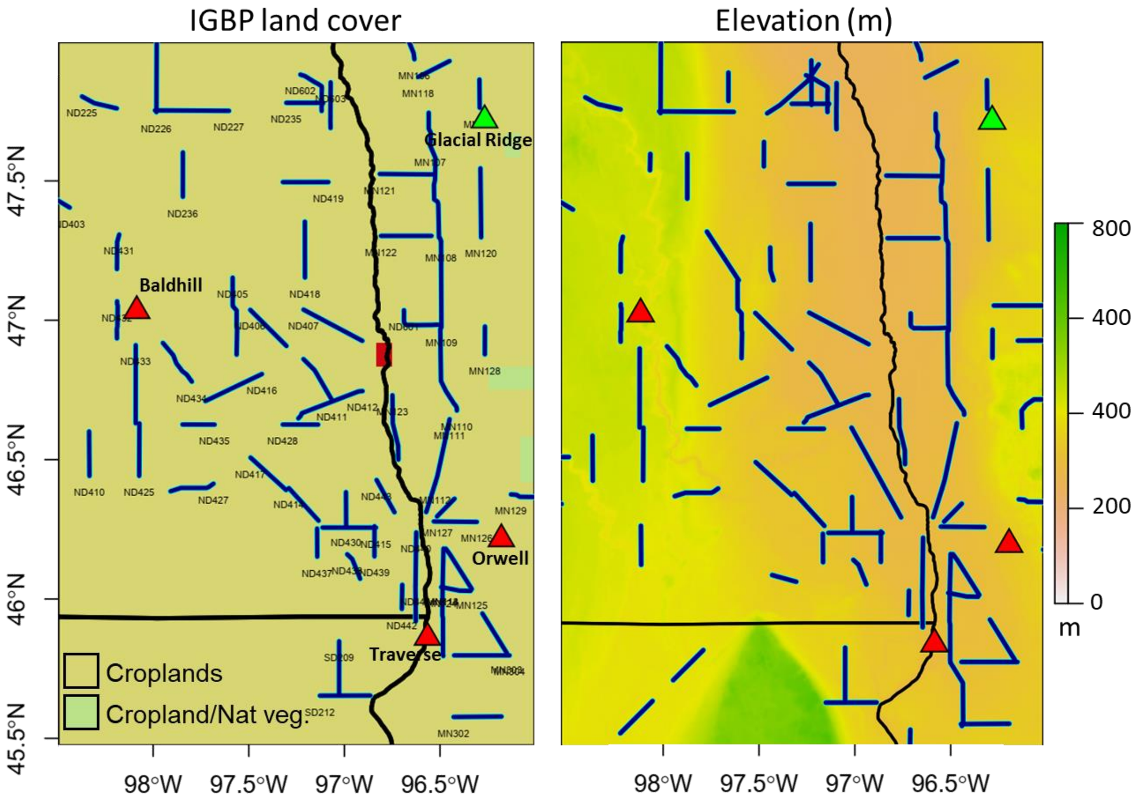


Figure S4. The ground-based SWE locations with the NOAA airborne gamma flight lines with IGBP land cover and elevation maps. The Glacial Ridge, Minnesota site operated by the United States Department of Agriculture (USDA) Soil Climate Analysis Network (SCAN) (green color) and three snow survey sites collected by the United States Army Corps of Engineer (USACE) St. Paul District (red color).

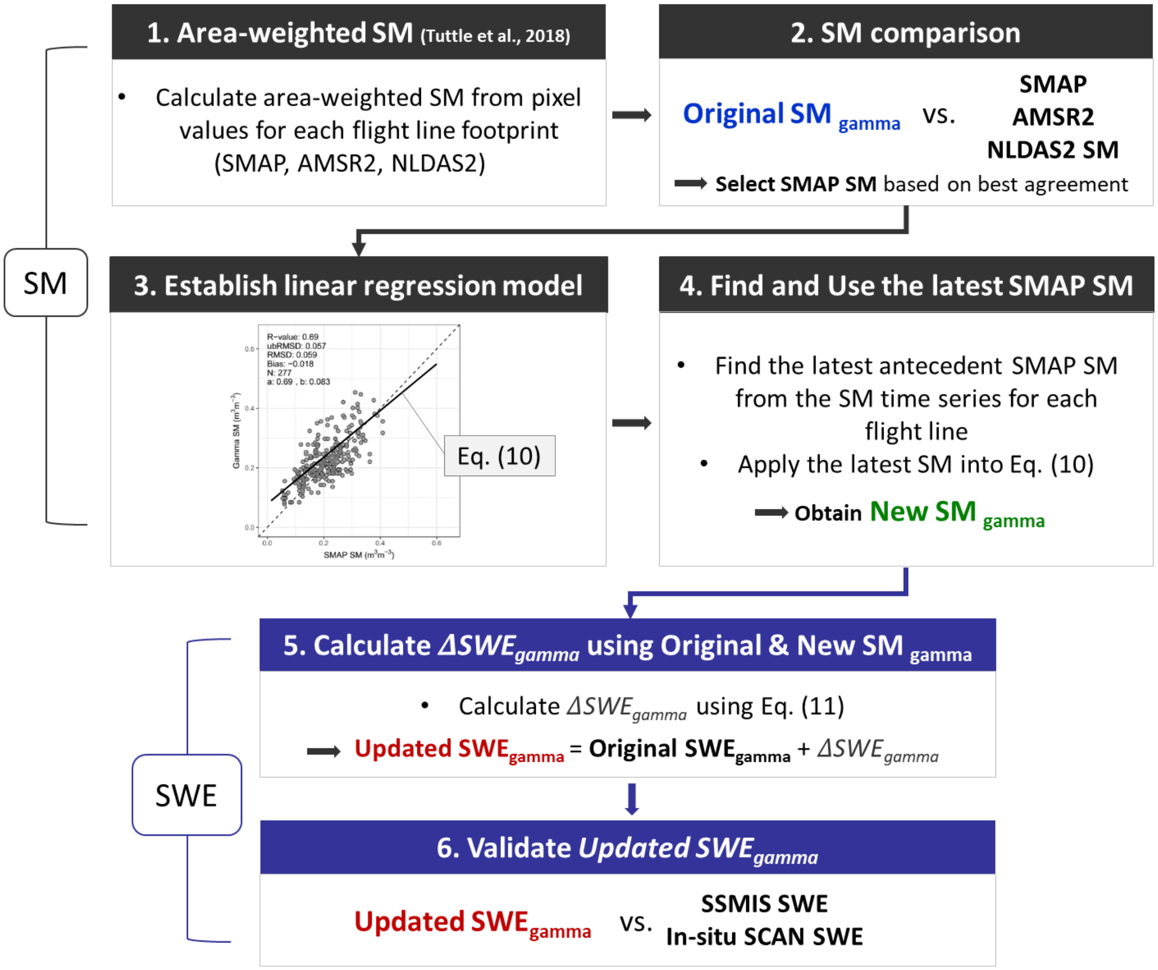


Figure S5. Schematic diagram of the entire process by separating the soil moisture (SM) and snow water equivalent (SWE) parts

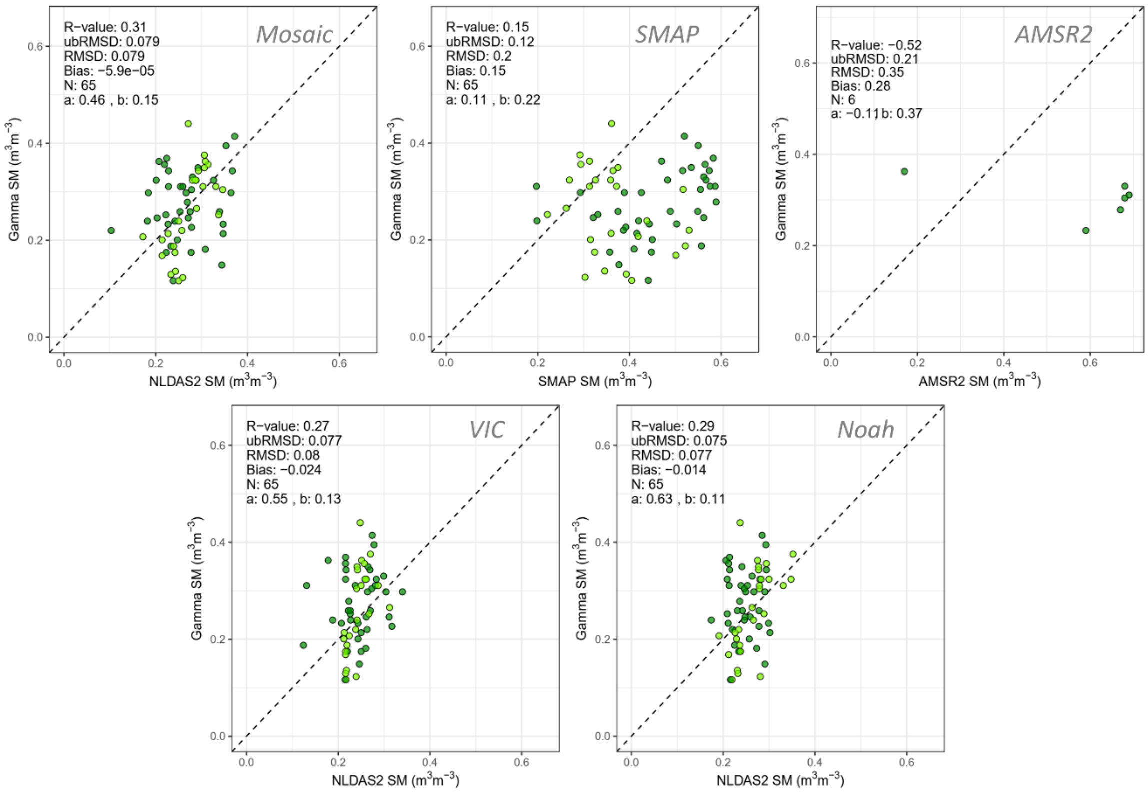


Figure S6. Comparison of NOAA airborne gamma soil moisture with NLDAS-2 Mosaic, SMAP, and AMSR2 SM products along with NLDAS-2 VIC and Noah model SM products for only forests

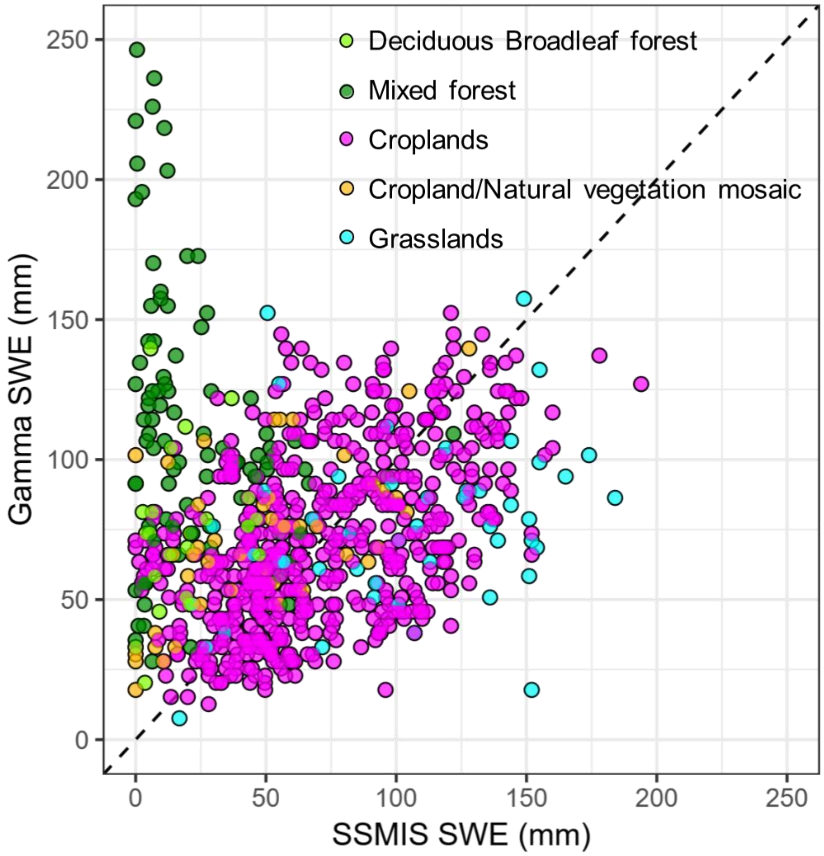


Figure S7. Agreements between operational NOAA airborne gamma SWE and Special Sensor Microwave Imager Sounder (SSMIS) SWE according to land cover type.

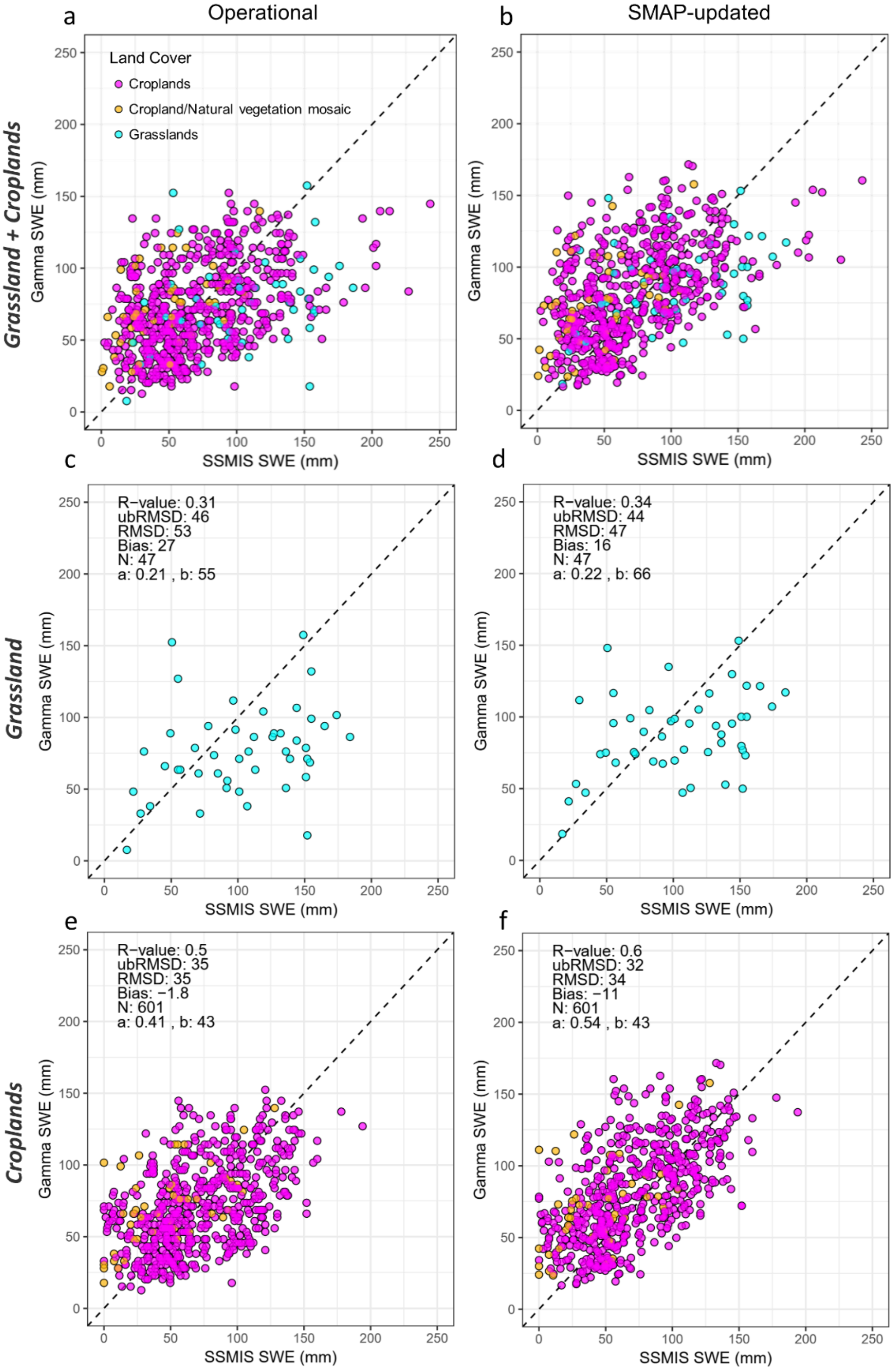


Figure S8. Same as Figure 7, but for grassland & croplands (a, b), grassland (c, d), and croplands (e, f), separately.

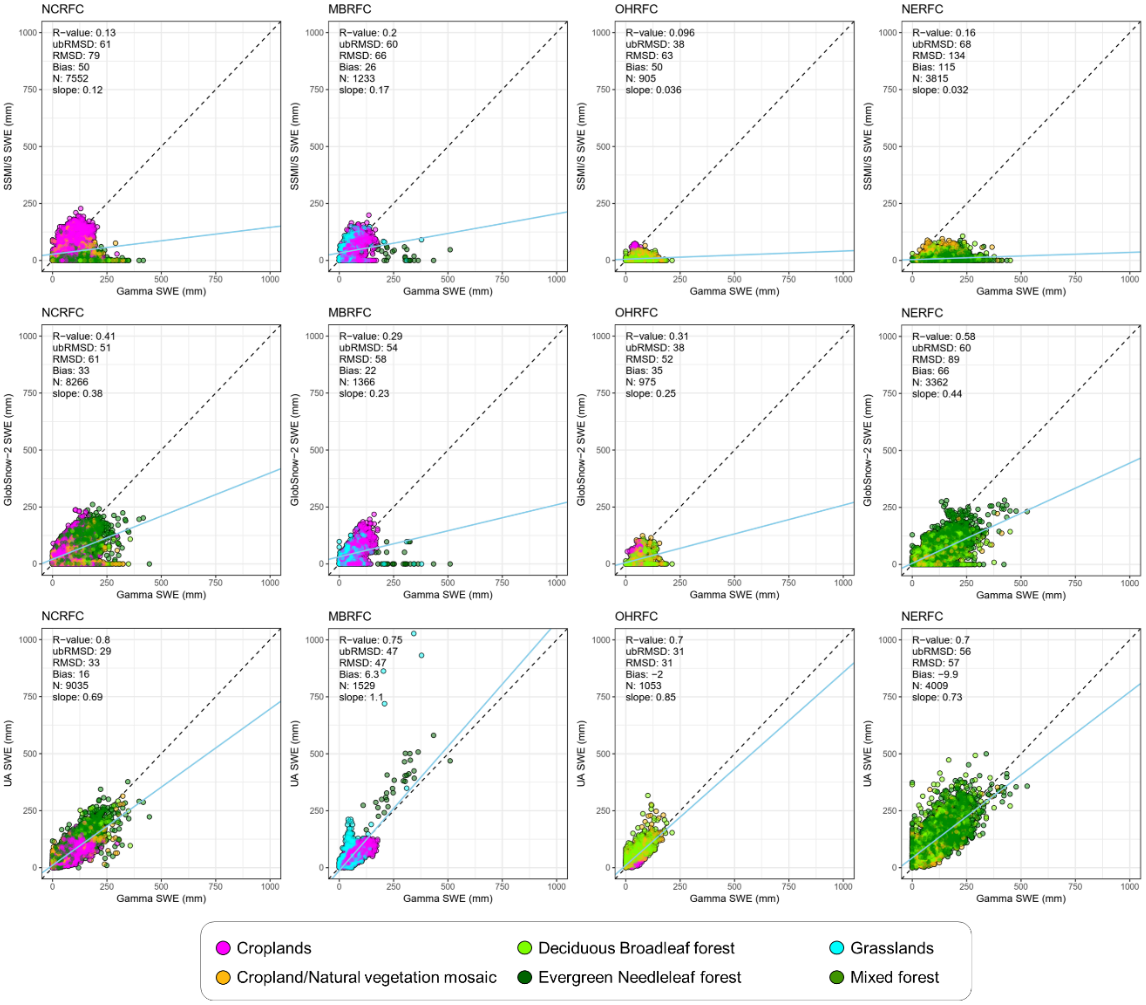


Figure S1. Comparison between daily SSMI/S, GlobSnow-2, and UA snow water equivalent with daily NOAA airborne gamma radiation snow water equivalent observations from 1982 to 2017 with coloring land cover type by NOAA river forecasting centers (RFCs) boundaries: North-Central (NC), Missouri Basin (MB), Ohio (OH), North-East (NE)

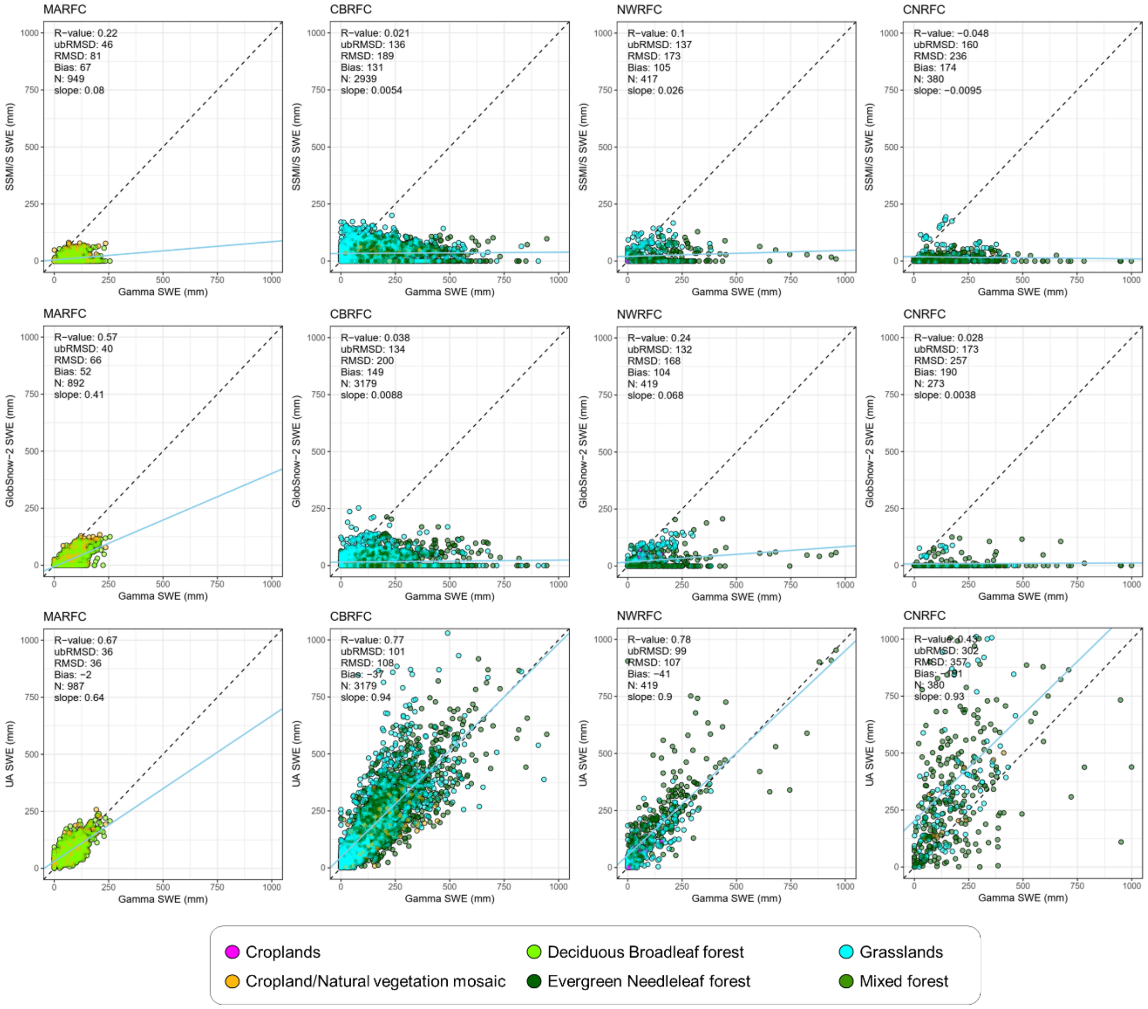


Figure S2. Same as Figure S1, but for Mid-Atlantic (MA), Colorado Basin RFC (CB), Northwest (NW), and California Nevada (CN)

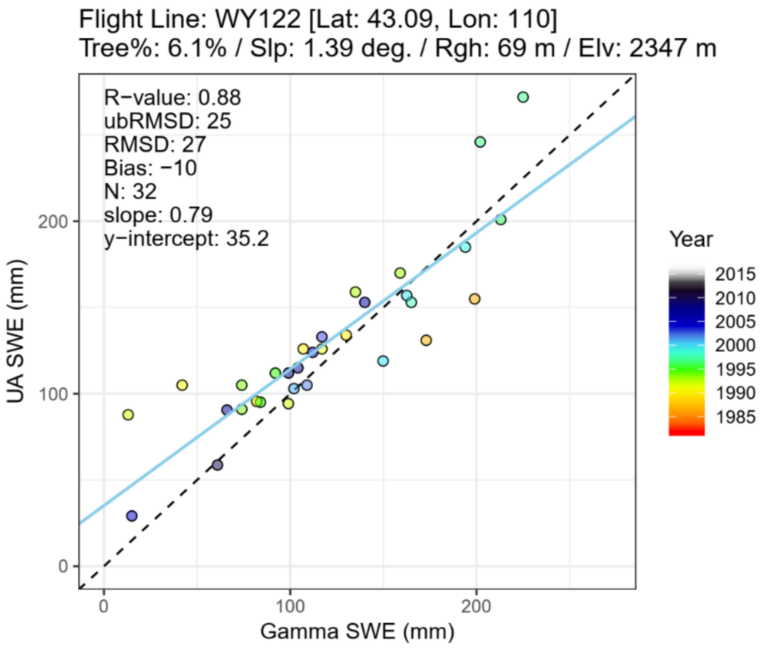


Figure S3. Scatterplot between gamma SWE and UA SWE for a gamma flight line "WY122" in Wyoming, U.S.

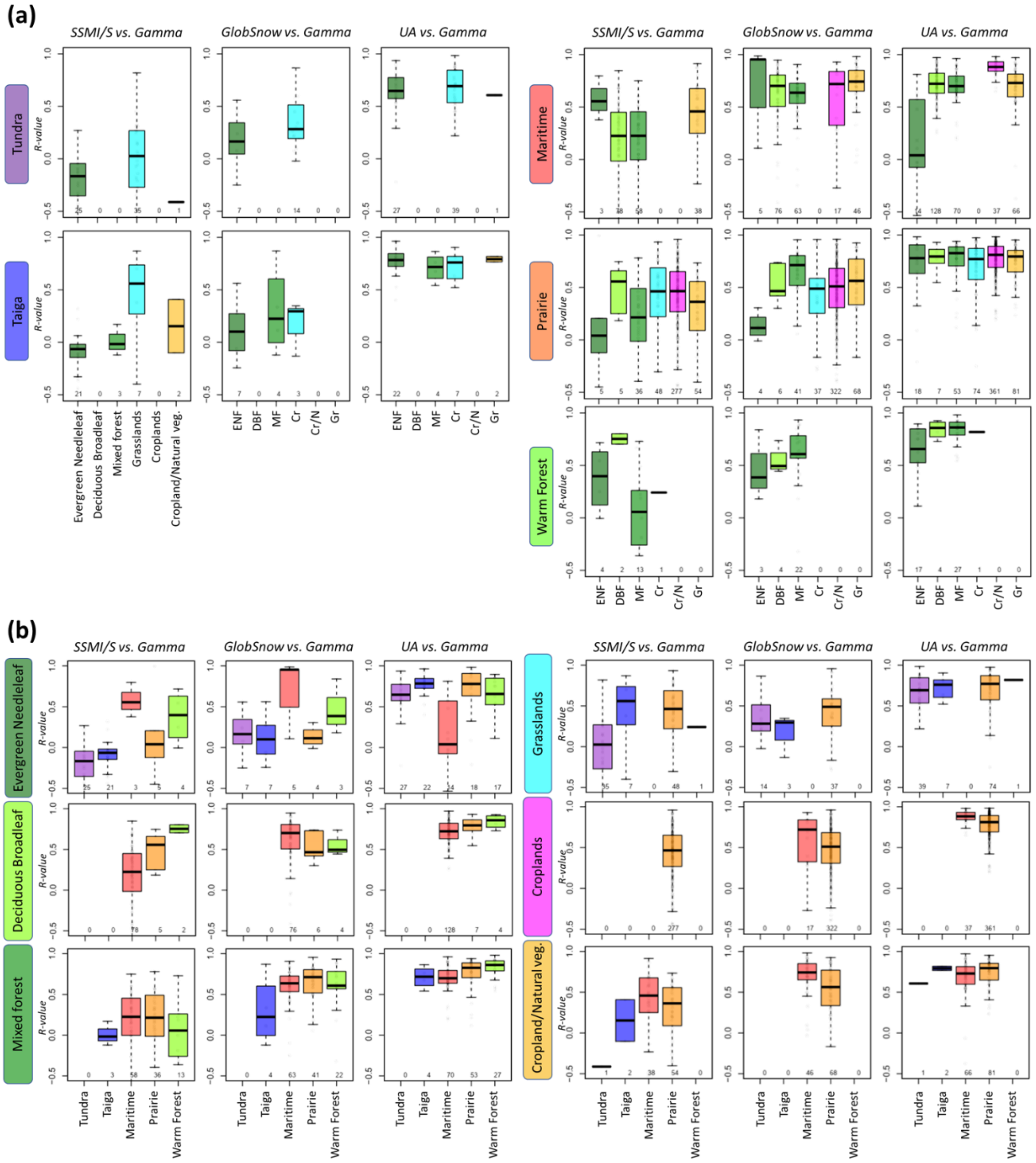


Figure S4. A series of boxplots for R-value between daily SSMI/S, GlobSnow-2, and UA snow water equivalent and daily NOAA airborne gamma radiation snow water equivalent where (a) each of the snow classifications is subdivided by six land cover types and (b) each of the land cover types is subdivided by five snow classes.

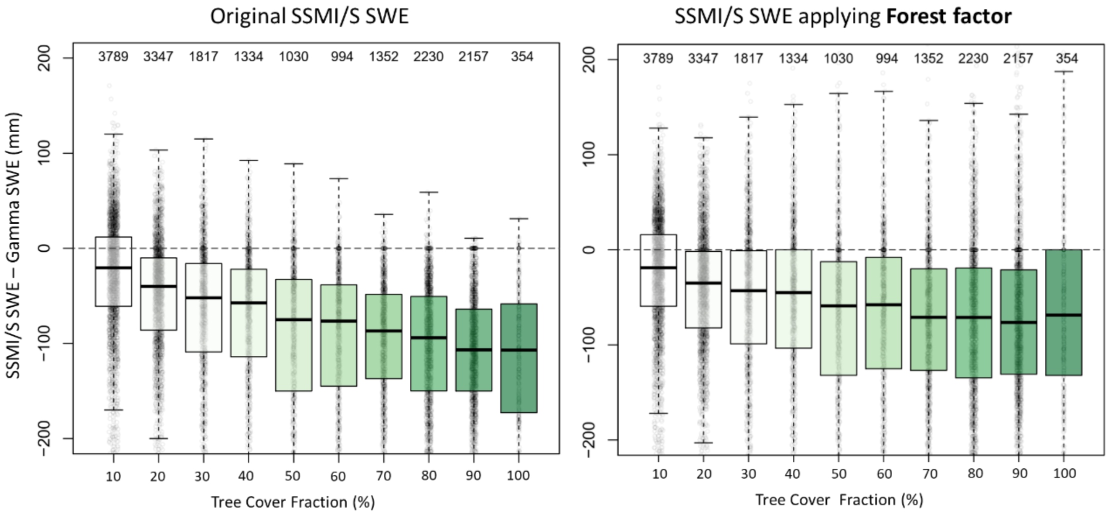


Figure S5. Comparison of SWE differences between original SSMI/S and modified SSMI/S by forest factor against airborne gamma radiation SWE data with fractional tree cover (%) (Foster et al., 2005). Forest factor is calculated by %tree cover ranging from 1 (no forest) to 2 (100% fractional forest cover).

Table S1. The number of airborne gamma radiation SWE flights in *grasslands* and *evergreen Needleleaf forest* land cover type by U.S. states

Grasslands			*Evergreen needleleaf forest*		
U.S. State	Number of SWE flights	Percent, %	U.S. State	Number of SWE flights	Percent, %
Wyoming	771	29.2	Colorado	897	51.7
Utah	650	24.6	California	262	15.1
Colorado	556	21.1	Utah	171	9.9
Montana	175	6.6	Arizona	164	9.5
North Dakota	160	6.1	Wyoming	66	3.8
Idaho	97	3.7	Idaho	65	3.7
California	92	3.5	Oregon	43	2.5
Nevada	49	1.9	Montana	41	2.4
Oregon	44	1.7	Washington	26	1.5
South Dakota	44	1.7			
Arizona	2	0.1			

Table S2. The number of airborne gamma radiation SWE flights in areas with more than 80% of tree cover fraction by U.S. states

> 80% of tree cover fraction		
U.S. State	Number of SWE flights	Percent, %
Maine	805	31.9
New York	622	24.6
Michigan	285	11.3
Vermont	188	7.4
New Hampshire	162	6.4
Wisconsin	159	6.3
Minnesota	74	2.9
West Virginia	73	2.9
Pennsylvania	70	2.8
Massachusetts	54	2.1
Idaho	13	0.5
Connecticut	11	0.4
California	4	0.2
Montana	3	0.1
Colorado	2	0.1
Washington	1	0.0