Assimilation of AMSR-E snow water equivalent data in a lumped hydrological model

David Dziubanski
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

Recommended Citation
https://lib.dr.iastate.edu/etd/13614

This Thesis is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Assimilation of AMSR-E snow water equivalent data in a lumped hydrological model

by

David Joseph Dziubanski

A thesis submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Geology

Program of Study Committee:
Kristie J. Franz, Major Professor
William J. Gutowski
William W. Simpkins

Iowa State University
Ames, Iowa
2013

Copyright © David Joseph Dziubanski, 2013. All rights reserved
TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................................................ iii
LIST OF TABLES ........................................................................................................................................ vi
ABSTRACT ................................................................................................................................................... vii
CHAPTER 1. GENERAL INTRODUCTION .................................................................................................. 1
CHAPTER 2. METHODOLOGY ........................................................................................................................ 5
  2.1 Study Sites and Data .............................................................................................................................. 5
  2.2 Models ................................................................................................................................................... 8
    2.2.1 National Weather Service SNOW17 Model ...................................................................................... 8
    2.2.2 Sacramento Soil Moisture Accounting Model (SAC-SMA) .......................................................... 10
  2.3 The ensemble Kalman Filter ................................................................................................................ 11
  2.4 Observation Data .................................................................................................................................. 15
  2.5 Data Correction .................................................................................................................................... 18
  2.6 Ensemble Perturbation ........................................................................................................................ 21
  2.7 Model Simulations ............................................................................................................................... 22
  2.8 Model Verification ............................................................................................................................... 23
CHAPTER 3. RESULTS AND DISCUSSION ................................................................................................. 25
CHAPTER 4. CONCLUSIONS .................................................................................................................... 47
APPENDIX ................................................................................................................................................... 49
  A.1 Discharge Figures ............................................................................................................................... 49
    A.1.1 BCHW3 ........................................................................................................................................ 49
    A.1.2 SCHR4 .......................................................................................................................................... 52
    A.1.3 DARW3 ....................................................................................................................................... 55
    A.1.4 DANW3 ....................................................................................................................................... 58
    A.1.5 MMLM5 ....................................................................................................................................... 61
    A.1.6 PLUM5 ......................................................................................................................................... 64
    A.1.7 RAPM5 ......................................................................................................................................... 67
REFERENCES .............................................................................................................................................. 70
ACKNOWLEDGEMENTS .......................................................................................................................... 74
LIST OF FIGURES

Figure 1. (a) Map of the study sites located within the North Central River Forecast Center (NCRFC) across the Upper Midwest. (b) Map of major land cover types of the Upper Midwest from the 2006 National Land Cover Database (Fry et al. 2011). ............................................. 6

Figure 2. Structure of the SNOW17 model (Anderson 1973) ................................................................................................................................. 9

Figure 3. Structure of the EnKF. An ensemble of initial states is propagated forward in time until a measurement becomes available. Uncertainty associated with the measurement and the model is used to calculate the Kalman gain. The Kalman gain is used within the updating equation to optimally weight the measurement and measurement forecast. The updated states are then used as initial states to further propagate the model forward in time.................................................................................................................. 14

Figure 4. AMSR-E SWE bias at each corresponding NOHRSC airborne SWE observation for the 2005-2011 period. Increasingly negative bias values are displayed with increasingly larger solid circle. Increasingly positive bias values are displayed with increasingly larger open circle........................................................................................................................ 20

Figure 5. Comparison of SWE between model simulations and airborne observations for each of the 7 study sites. SWE values for each study site are shown with a different symbol. ........................................................................................................................................................................... 26

Figure 6. SWE (top) and stream discharge (bottom) for BCHW3, water year 2010. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line. ................................................................................................................. 28

Figure 7. SWE (top) and stream discharge (bottom) for SCRI4, water year 2008. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line. ................................................................................................................. 28

Figure 8. SWE (top) and stream discharge (bottom) for DARW3, water year 2009. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line. ................................................................................................................. 30

Figure 9. Average bias, normalized mean absolute error (N MAE), Person’s r, and containing ratio (CR) for each bi-weekly period for BCHW3 across all study years. .......... 30

Figure 10. Average bias, normalized mean absolute error (N MAE), Person’s r, and containing ratio (CR) for each bi-weekly period for SCRI4 across all study years............. 31

Figure 11. Average bias, normalized mean absolute error (N MAE), Person’s r, and containing ratio (CR) for each bi-weekly period for DARW3 across all study years. ........... 31
Figure 12. SWE (top) and stream discharge (bottom) for DANW3, water year 2006. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line. ......................................................... 33

Figure 13. Average bias, normalized mean absolute error (N MAE), Person’s r, and containing ratio (CR) for each bi-weekly period for DANW3 across all study years. .......... 33

Figure 14. SWE (top) and stream discharge (bottom) for MMLM5, water year 2006. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line. ............................................. 36

Figure 15. Average bias, normalized mean absolute error (N MAE), Person’s r, and containing ratio (CR) for each bi-weekly period for MMLM5 across all study years. ....... 37

Figure 16. Average Normalized MAE for the upper 1/3 percentile flows for each study. Normalized MAE error was calculated on all peak flows occurring during a time period between 1 February till 7 days past full snow melt. ................................................................. 38

Figure 17. SWE (top) and stream discharge (bottom) for DANW3, water year 2008. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line. ................................................................. 39

Figure 18. SWE (top) and stream discharge (bottom) for PLUM5, water year 2006. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line. ................................................. 40

Figure 19. SWE (top) and stream discharge (bottom) for PLUM5, water year 2007. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line. ................................................. 41

Figure 20. SWE (top) and stream discharge (bottom) for PLUM5, water year 2008. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line. ................................................. 42

Figure 21. Average bias, normalized mean absolute error (N MAE), Person’s r, and containing ratio (CR) for each bi-weekly period for MMLM5 across all study years. ....... 42

Figure 22. SWE (top) and stream discharge (bottom) for RAPM5, water year 2007. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line. ................................................. 44

Figure 23. Average bias, normalized mean absolute error (N MAE), Person’s r, and containing ratio (CR) for each bi-weekly period for RAPM5 across all study years. ....... 44
**Figure 24.** SWE (top) and stream discharge (bottom) for RAPM5, water year 2010. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line.

**Figure 25.** SWE (top) and stream discharge (bottom) for MMLM5, water year 2010. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line.
LIST OF TABLES

Table 1. Location ID, USGS gaging station number, size, and mean observations for February to September from 2005 to 2011 for the seven study sites. ............................................... 7

Table 2. Description of SNOW17 model parameters and RFC values used for each study site. ........................................................................................................................................................................... 10

Table 3. Description of SAC-SMA model parameters and RFC values used for each study site. ........................................................................................................................................................................... 11

Table 4. Average maximum SWE for each study site and number of AMSR-E grid cells located within the study site boundaries. ........................................................................................................................................................................... 16

Table 5. Mean absolute error of SWE between model simulations and airborne observations for each of the 7 study sites. ........................................................................................................................................................................... 26
ABSTRACT

Snow cover is a significant component of the hydrological cycle affecting stream discharge through snowmelt and soil moisture. Current operational streamflow forecasting is prone to error due to input data uncertainties and model biases, making it difficult to accurately forecast discharge during snow melt events. Data assimilation is a technique of weighting model estimates and observations based on uncertainties that allows optimal estimation of model states. In this study, we assimilate snow water equivalent (SWE) data from the Advance Microwave Scanning Radiometer – Earth Observing System (AMSR-E) instrument into a conceptual temperature index snow model, the US National Weather Service (NWS) SNOW17 model. This model is coupled with the NWS Sacramento Soil Moisture Accounting (SAC-SMA) model, which ultimately produces stream discharge. The objective of this study is to improve the SNOW17 estimate of SWE by integrating SWE observations and uncertainties associated with meteorological forcing data within the model. For the purpose of this study, 25 km AMSR-E SWE data is used. An ensemble Kalman filter (EnKF) assimilation framework performs assimilation on a daily cycle for a 6 year period, water years 2006-2011. This method is tested on seven watersheds in the Upper Mississippi River basin that are under the forecasting jurisdiction of the NWS North Central River Forecasting Center (NCRFC). Prior to assimilation, AMSR-E data is bias corrected using data from the National Operational Hydrologic Remote Sensing Center (NOHRSC) airborne snow survey program. Discharge output from the SAC-SMA is verified using observed discharge from the outlet of each study site. Improvements in discharge are evident for five sites, in particular for high discharge magnitudes associated with snow melt runoff. Evidence points to the SNOW17 having a consistent SWE
underestimation bias and error in snow melt rate. Overall results indicate that the EnKF is a viable and effective solution for integrating observations directly with operational models.
CHAPTER 1. GENERAL INTRODUCTION

Snowmelt runoff is an important component within the hydrologic cycle in mid-latitude regions. A snowpack may be melted rapidly during the spring snowmelt season, raising the potential for large stream discharge events. Additionally, snow alters surface energy and moisture fluxes by increasing albedo, decreasing surface roughness, and creating a barrier between the soil and atmosphere, thus limiting heat transfer (Armstrong and Brun, 2010). The northern Midwest of the United States is particularly susceptible to flooding events during periods when rapid snowmelt is coupled with rainfall events (Perry, 2000). This makes it critical to accurately estimate snow properties such as snow water equivalent (SWE) and snow covered area (SCA). Unlike the Western United States which has snow monitoring networks (e.g Natural Resources Conservation Service (NRCS) Snowpack Telemetry (SNOTEL) network), the northern Midwest has a very limited number of ground SWE observations. This is a persistent problem present across many areas. Thus, to overcome this limitation, remote sensing data is heavily relied upon to capture the snow pack state (Schmugge et al., 2002).

Satellite observations of SCA and SWE are widely available, and provide useful spatial and temporal information about the current snow pack state. The Moderate Resolution Imaging Spectroradiometer (MODIS) and the Advanced Very High Resolution Radiometer (AVHRR) provide visible measurements of SCA at the 1 km resolution. However, satellite measurements in the visible range are unable to capture data at night and are prone to error due to cloud cover and vegetation interference, thus misclassification errors typically exist (Maurer et al., 2003). Since SWE is the main variable of interest in hydrologic modeling, SCA data needs to be related to SWE using an areal depletion curve (Essery and Pomeroy, 2004).
SWE estimates using areal depletion curves usually contain a high degree of uncertainty. This is due to numerous factors affecting the spatial distribution of snow, particularly wind exposure, land cover type, and topography (Pomeroy et al., 1998). Due to this uncertainty, areal depletions curves are unable to capture an accurate magnitude of SWE under all possible conditions.

Observations from passive microwave sensors provide an alternative option to measure snow pack state without the limitations of visible measurements. Brightness temperatures measured by these sensors can be used to determine the depth of snow (Foster et al., 1984). SWE measurements can be derived from these observations through use of snow density, thus bypassing the need to use a method for deriving SWE from SCA. Several microwave sensors have been used extensively in snow studies, most notably the Scanning Multichannel Microwave Radiometer (SMMR), Spatial Sensor Microwave Imager (SSM/I), and recently, the Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E). However, microwave observations usually have a coarse 25 km resolution and are not without error, as they are usually associated with increases in error due to snow metamorphosis, liquid water, and forest cover (Cordisco et al., 2006; Dong et al., 2005; Foster et al., 2005; Kelly et al., 2003).

Obtaining the best estimate of the current snow state may be achieved by directly integrating passive microwave SWE observations with model simulations. This technique is known as data assimilation and has been widely used in operational meteorology to update model states (Reichle, 2008). Due to the availability of high resolution satellite observations, numerous studies have investigated assimilation of data into hydrologic models. A variety of assimilation techniques are available, one of the most commonly used techniques is the
Ensemble Kalman Filter (EnKF). Studies that have assimilated SWE from microwave observations using the EnKF have done so with some success for study sites located in the western United States. Dong et al. (2007) assimilated data from the SMMR sensor and found overall improvement in SWE estimates, particularly when data was omitted due to melting conditions and SWE above 100 mm. Konstantinos and Lettenmaier (2005) assimilated AMSR-E data for the Snake River Basin in Idaho and validated model results against SNOTEL SWE data. Results indicated small improvement in capturing SWE magnitude; however, results degraded with increasingly deeper snow packs. Additionally, a second experiment was performed which incorporated a SWE cut off value of 240 mm; that is, SWE data was not assimilated when modeled SWE values were greater than 240 mm. Subsequently, SWE results were marginally improved over the first assimilation. DeLannoy et al. (2009) utilized several complex versions of the EnKF, namely a spatial (3D) EnKF to assimilate course resolution (25 km) SWE observations into a fine scale (1 km) model for an area in north central Colorado. Using a spatial EnKF that assimilates data for a fine scale grid cell by incorporating surrounding course scale observations, a decrease in the RMSE of 60% was achieved when compared to the open loop simulation. Other studies have shown improvement in capturing SWE magnitude or discharge when using the EnKF (De Lannoy et al., 2012; Slater and Clark, 2006; Thirel et al., 2011).

In this study, we demonstrate the feasibility of using an EnKF framework to assimilate SWE data from the AMSR-E sensor into the SNOW17 model for seven study sites in the northern Midwest. The SNOW17 model is currently used by the National Weather Service (NWS) River Forecast Center (RFC) as a tool for predicting snowmelt outflow (Anderson, 1973). The SNOW17 model is coupled with the Sacramento Soil Moisture Accounting Model (SAC-
SMA) to produce total stream discharge. AMSR-E data is validated and corrected for bias using SWE data measured by the National Operational Hydrologic Remote Sensing Center (NOHRSC) airborne snow survey program. Subsequent assimilation of bias-corrected data is performed based on data availability, with the model propagating forward in time between observations. Simulations using data assimilation are evaluated against control runs for each study site using observed stream flow data and NOHRSC airborne SWE data.
CHAPTER 2. EXPERIMENT SETUP AND METHODOLOGY

2.1 Study Sites and Data

The study area consists of seven watersheds located in the Great Lakes and northern Plains region (Figure 1a). Criteria for choosing study sites included locations important to NCRFC operations and data availability. Site importance was based on key tributaries to larger river systems, recent flooding events, or forecasting difficulty. The selected watersheds vary in size from 572 km$^2$ to 6230 km$^2$ and have slight differences in characteristics and geology (Table 1).

The North Raccoon River (SCRI4), Blue Earth River (RAPM5), and Redwood River (MMLM5) watersheds are located within the Des Moines Lobe. This region is characterized by low relief and glacial till deposits varying between 30-120 meters thick (Alberts, 1995; Olson and Mossler, 1982; Prior, 1991). Based on the National Land Cover Database, the primary land use within these watersheds is agricultural (Figure 1b, Fry et al., 2011). Snowfall over these study sites generally averages less than 100 cm each season. Similar to the southern Minnesota basins, the parent material of the Clearwater River (PLUM5) watershed near Plummer, Minnesota consist of sands and gravels which are characteristic of glacial outwash. This area is contained within the Red River valley, and receives approximately 100-110 cm of snowfall each winter. PLUM5 is located in a transitional land cover region; forest cover primarily encompasses the eastern half of this watershed, whereas the western half is used for agricultural purposes. The Pecatonica (DARW3) and East Pecatonica (BCHW3) River watersheds are located within the Driftless Area of southwest Wisconsin. This area was not glaciated, and is therefore, characterized by a highly eroded landscape with moderate relief of 150-250 m. Due to the significant erosion, bedrock is highly exposed within the region. A mixed land cover of cultivated crops, pasture, and forest
cover exists within these watersheds. Average snowfalls typically range from 80 – 90 cm each winter. The St. Croix River watershed near Danbury, WI receives the greatest snowfall, with amounts ranging between 130 cm over the western end to 150 cm over the eastern end of the watershed. This area is primarily covered by glacier drift 100-600 ft thick and is heavily forested.

Figure 1. (a) Map of the study sites located within the North Central River Forecast Center (NCRFC) across the upper Midwest. (b) Map of major land cover types of the upper Midwest from the 2006 National Land Cover Database (Fry et al., 2011).
Precipitation, temperature, evaporation, and discharge data were obtained from the North Central River Forecast Center (NCRFC). Approximately 7 years of 6-hourly operational mean areal precipitation (MAP), 6-hourly operational mean areal temperature (MAT), 6-hourly climatological evaporation, and daily discharge are available for each basin from 2005-2011 (Table 1). Mean areal precipitation during the study period for all seven study sites was 846 mm, and ranged from 655 mm over the northwestern part to 1037 mm over the southeastern part of the upper Midwest. Mean daily discharge for the study sites varied significantly from 3.7 m$^3$/day to 52.0 m$^3$/day based on watershed size.

<table>
<thead>
<tr>
<th>Basin Name</th>
<th>Location ID</th>
<th>USGS Gauging Station</th>
<th>Size (km$^2$)</th>
<th>Mean Daily Discharge (m$^3$/day)</th>
<th>Mean Annual Precipitation (mm/year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clearwater River at Plummer, MN</td>
<td>PLUM5</td>
<td>05078000</td>
<td>1395</td>
<td>6.3</td>
<td>655.2</td>
</tr>
<tr>
<td>Redwood River at Marshall, MN</td>
<td>MMLM5</td>
<td>05315000</td>
<td>682</td>
<td>3.7</td>
<td>726.1</td>
</tr>
<tr>
<td>Blue Earth River at Rapidan, MN</td>
<td>RAPM5</td>
<td>05320000</td>
<td>6230</td>
<td>52.0</td>
<td>905.1</td>
</tr>
<tr>
<td>St. Croix River at Danbury, WI</td>
<td>DANW3</td>
<td>05333500</td>
<td>3986</td>
<td>31.8</td>
<td>760.2</td>
</tr>
<tr>
<td>Pecatonica River at Darlington, WI</td>
<td>DARW3</td>
<td>05432500</td>
<td>710</td>
<td>8.0</td>
<td>1037.2</td>
</tr>
<tr>
<td>East Branch Pecatonica at Blanchardville, WI</td>
<td>BCHW3</td>
<td>05433000</td>
<td>572</td>
<td>6.0</td>
<td>985.5</td>
</tr>
<tr>
<td>North Raccoon River at Sac City, IA</td>
<td>SCRI4</td>
<td>05482300</td>
<td>1855</td>
<td>15.8</td>
<td>855.0</td>
</tr>
</tbody>
</table>

*Table 1.* Location ID, USGS gaging station number, size, and mean observations for February to September from 2005 to 2011 for the seven study sites.
2.2 Models

2.2.1 National Weather Service SNOW17 Model

The SNOW17 (Anderson, 1973) is a conceptual snow accumulation and ablation model that is currently used for operational purposes by the River Forecast Center (RFC) as part of their river forecasting system. Snow is modeled as a single layer, with the two inputs into the model being temperature and precipitation. Empirically-based equations approximate heat deficit of the snow pack, snow pack density, water retention and transmission. Air temperature inputs are used to determine snow pack energy exchange with the atmosphere, as well as snow accumulation, snowmelt, and heat content within the snow pack (Figure 2). The gradient between the antecedent air temperature and the current air temperature determines heat deficit changes within the snow pack (Anderson, 2006). Snow melt occurs when energy exchange forces the heat deficit to reach zero. If the liquid water holding capacity of the snow pack has been reached, excess water is present. This excess water is lagged and attenuated, which subsequently turns into snow melt outflow. Final output from the SNOW17 includes a rain-snowmelt time series and basin average snow water equivalent. When the SNOW17 is applied at a basin average scale, 11 model parameters are required, including an areal depletion curve (Anderson, 1973, 2006). For purposes of this study, RFC SNOW17 model parameters were used (Table 2). The SNOW17 model uses the areal depletion curve to determine the percentage of the basin that is covered by snow and contributing to snow melt outflow. Using the areal extent of snow cover, total outflow for each time interval is computed by:

\[
M = (M_s \cdot A_s) + \left[ (1.0 - A_s) \cdot (P \cdot f_r) \right]
\]  

(1)
where $M$ is the total outflow (mm), $M_s$ is snow cover outflow (mm), $A_s$ is areal extent of snowcover, $P$ is total precipitation (mm), and $f_r$ is the fraction of precipitation in rain form. Precipitation falling on bare ground is added to the total outflow, while precipitation falling on the snow cover is added to snow cover outflow. The SNOW17 model within this study was applied at a basin-wide scale and 6-hour time step.

![Figure 2. Structure of the SNOW17 model (Anderson, 1973)](image-url)
Table 2. Description of SNOW17 model parameters and RFC values used for each study site.

2.2.2 Sacramento Soil Moisture Accounting Model (SAC-SMA)

The SACSMA (Burnash et al., 1973) is a rainfall-runoff model connected with the SNOW17 model. Input into the SAC-SMA includes precipitation data as well as rain-snowmelt time series output from the SNOW17. Two soil layers are used within the model to represent different subsurface processes. The upper layer represents unsaturated flow, which includes interception storage, whereas the lower zone represents groundwater flow, particularly that which contributes to stream flow. All major hydrologic processes are represented within the SAC-SMA, including soil moisture, percolation, drainage, and evapotranspiration (ET). ET calculation within the model requires an ET demand curve, which is based on climatological data. Final output from the SCA-SMA is basin average runoff depth, which is represented as
stream flow discharge for each study site. 16 basin specific parameters are required for the SAC-SMA (Burnash, 1995). RFC SAC-SMA parameters are used within this study (Table 3).

<table>
<thead>
<tr>
<th>SAC-SMA Parameters</th>
<th>Parameter Description</th>
<th>PLUM5</th>
<th>MMLM5</th>
<th>RAPM5</th>
<th>DANW3</th>
<th>DARW3</th>
<th>BCHW3</th>
<th>SCRI4</th>
</tr>
</thead>
<tbody>
<tr>
<td>UZTWM</td>
<td>Upper-zone tension water maximum storage (mm)</td>
<td>20</td>
<td>83.00</td>
<td>60.00</td>
<td>30</td>
<td>30.00</td>
<td>20.00</td>
<td>110.00</td>
</tr>
<tr>
<td>UZFWM</td>
<td>Upper-zone free water maximum storage (mm)</td>
<td>80</td>
<td>40.00</td>
<td>40.00</td>
<td>80</td>
<td>27.00</td>
<td>25.00</td>
<td>50.00</td>
</tr>
<tr>
<td>U2K</td>
<td>Upper-zone free water lateral depletion rate (day-1)</td>
<td>0.4</td>
<td>0.34</td>
<td>0.30</td>
<td>0.25</td>
<td>0.25</td>
<td>0.40</td>
<td>0.30</td>
</tr>
<tr>
<td>PCTIM</td>
<td>Impervious fraction of the watershed (decimal fraction)</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.015</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>ADIMP</td>
<td>Additional impervious area (decimal fraction)</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>RIVA</td>
<td>Riparian vegetation (decimal fraction)</td>
<td>0</td>
<td>0.005</td>
<td>0.030</td>
<td>0</td>
<td>0.000</td>
<td>0.000</td>
<td>0.005</td>
</tr>
<tr>
<td>ZPERC</td>
<td>Maximum percolation rate (dimensionless)</td>
<td>80</td>
<td>82.00</td>
<td>40.00</td>
<td>325</td>
<td>170.00</td>
<td>60.00</td>
<td>40.00</td>
</tr>
<tr>
<td>REXP</td>
<td>Exponent of the percolation equation (dimensionless)</td>
<td>2</td>
<td>1.80</td>
<td>1.40</td>
<td>2.8</td>
<td>2.00</td>
<td>1.20</td>
<td>2.10</td>
</tr>
<tr>
<td>LZTWM</td>
<td>Lower-zone tension water maximum storage (mm)</td>
<td>85</td>
<td>170.00</td>
<td>150.00</td>
<td>140</td>
<td>140.00</td>
<td>150.00</td>
<td>150.00</td>
</tr>
<tr>
<td>LZFSM</td>
<td>Lower-zone free water supplementary maximum storage (mm)</td>
<td>60</td>
<td>100.00</td>
<td>50.00</td>
<td>55</td>
<td>25.00</td>
<td>60.00</td>
<td>60.00</td>
</tr>
<tr>
<td>LZFPM</td>
<td>Lower-zone free water primary maximum storage (mm)</td>
<td>70</td>
<td>25.00</td>
<td>40.00</td>
<td>725</td>
<td>400.00</td>
<td>410.00</td>
<td>50.00</td>
</tr>
<tr>
<td>LZSK</td>
<td>Lower-zone supplementary free water depletion rate (day-1)</td>
<td>0.08</td>
<td>0.070</td>
<td>0.050</td>
<td>0.07</td>
<td>0.050</td>
<td>0.035</td>
<td>0.100</td>
</tr>
<tr>
<td>LZPK</td>
<td>Lower-zone primary free water depletion rate (day-1)</td>
<td>0.005</td>
<td>0.004</td>
<td>0.005</td>
<td>0.0015</td>
<td>0.002</td>
<td>0.003</td>
<td>0.020</td>
</tr>
<tr>
<td>PFREE</td>
<td>Fraction of water percolating from upper zone directly to lower zone free water storage (decimal fraction)</td>
<td>0.1</td>
<td>0.30</td>
<td>0.20</td>
<td>0.2</td>
<td>0.01</td>
<td>0.10</td>
<td>0.25</td>
</tr>
<tr>
<td>SIDE</td>
<td>Ratio of deep recharge to channel base flow (decimal fraction)</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>RSERV</td>
<td>Fraction of lower-zone free water not transferable to lower-zone tension water (decimal fraction)</td>
<td>0.3</td>
<td>0.30</td>
<td>0.30</td>
<td>0.3</td>
<td>0.00</td>
<td>0.30</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Table 3. Description of SAC-SMA model parameters and RFC values used for each study site.

2.3 The ensemble Kalman Filter

The assimilation technique that we implemented is the ensemble Kalman filter (EnKF).

This technique was first developed by Evenson (1994) as a means to overcome computational limitations associated with the standard Kalman filter (KF) and the extended Kalman filter (EKF) (Evensen, 1994; Reichle et al., 2002; Evensen, 2003). The KF is a technique that integrates modeled estimates with observations to minimize estimation error. The EKF is a nonlinear
implementation of the KF which updates a model state estimate whenever an observation is available. Optimal weighting is a key feature behind the updating process which allows optimal estimation based on errors in the model and observations. The EnKF works in the same manner; however, it captures the model forecast error from an ensemble of model trajectories at the time of an update. Each ensemble member is produced by perturbed model forcing fields, which include precipitation and temperature. This technique has computational advantages over the KF and EKF in that it does not propagate model error information through time with a dynamic equation, but rather through use of ensemble members. The main EnKF processes are given below. Consider the vector \( Y \) composed of model state variables of interest (e.g. SWE). The model equation can be written as

\[
\frac{dY}{dt} = F(Y, \alpha, u) + w
\]  

(2)

where the model operator \( F \) is the nonlinear hydrologic model. The change in the state variable of interest through time is dependent on the initial state \( Y \), model parameters \( \alpha \), and time-dependent forcing data \( u \) (e.g. observed temperature). Uncertainties in the forcing data or the model are associated with the term \( w \). If all observations of the state variable of interest taken during a certain time interval are input into a measurement vector \( Z \), then the measurement process can be defined by a measurement model equation as

\[
Z = H[Y] + v
\]  

(3)

where \( Z \) represents the measurement vector, \( H \) represents an observation operator which indicates the state variable that was observed in the update, and \( v \) is the measurement error. Three main steps are performed within the EnKF (Figure 3):
1) An ensemble of initial states \( Y_i, i = 1, \ldots, n \), are input into the EnKF at \( t = 0 \). These initial states are generated based on a mean \( Y \) with a covariance of \( C \). The nonlinear hydrologic model is propagated forward in time, generating an ensemble of states at each time step, until the next measurement becomes available.

2) Compute the Kalman gain \( K \):

\[
K = C_{yz}(C_z + C_v)^{-1}
\]

(4)

where \( C_z \) is the error covariance of the measurement forecast, \( C_{yz} \) is the cross covariance between the measurement forecast and model state variables, and \( C_v \) is the error covariance of the measurements. The Kalman gain optimally weights the errors associated with the measurements and the measurement forecast, allowing for updating of the state variables of interest within the model.

3) Update the state variables of interest \( Y \) for each ensemble member (Evensen, 1994):

\[
Y^i_+ = Y^i_- + K((Z + v^i) - H(Y^i_-))
\]

(5)

The best estimate of the model states of interest \( Y^i_+ \) is based on updating the previous forecast state for each ensemble member \( Y^i_- \). This update is based on the weighted difference between the observations \( Z \) and forecast \( H(Y^i_-) \). In each update step within the EnKF, an ensemble of perturbed observations is generated (Burgers et al., 1997). A random measurement error \( v^i \) is generated for each ensemble member from a normal distribution with mean equal zero and covariance equal to \( C_v \). The measurement errors are then added to the original observations \( Z + v^i \).
The main advantage of the EnKF is its ability to incorporate model and forcing data uncertainty; however, appropriately determining the uncertainty associated with observational data is a difficult task in practice, particularly for a localized region. Therefore, it is common practice to prescribe synthetic errors to the assimilation data. In this study, we attempt to derive errors associated with the assimilation data, but prescribe synthetic errors to the model forcing data.

**Figure 3.** Structure of the EnKF. An ensemble of initial states is propagated forward in time until a measurement becomes available. Uncertainty associated with the measurement and the model is used to calculate the Kalman gain. The Kalman gain is used within the updating equation to optimally weight the measurement and measurement forecast. The updated states are then used as initial states to further propagate the model forward in time.
2.4 Observation Data

The data source we assimilate in this study is the AMSR-E Level-3 daily SWE product (AE_DySno, V09, 2005 – 2011, [Tedesco et al., 2004; Kelly et al., 2003; R. Kelly, 2009]) from the NASA Aqua satellite made available through the National Snow and Ice Data Center (NSIDC). The AMSR-E instrument is a 6 frequency passive microwave radiometer system that measures brightness temperatures ranging from 6.9 GHz to 89 GHz. A snow mapping algorithm developed by Kelly et al. (2003) uses differences in brightness temperature at different frequencies and forest cover to determine total snow depth (SD). The total SD is based on the equation:

\[
SD = (ff \times (SD_f)) + ((1 - ff) \times (SD_o))
\]

where \(SD_f\) is the snow depth from the forested component of the instantaneous field of view (IFOV), \(SD_o\) is the snow depth from the non-forested component of the IFOV, and \(ff\) is the forest fraction (Kelly, 2009). \(SD_f\) is determined by using differences in brightness temperature at the 18 and 36 GHz channels. \(SD_o\) is determined by using differences in brightness temperature at the 10 and 36 GHz channels for moderate snow depth and the 10 and 18 GHz channels for deeper snow packs. SWE is retrieved from SD based on a mean monthly global snow density map, since SWE is directly related to snow density (Sturm et al., 1995). The AMSR-E SWE is then averaged to a 25 km resolution Equal Area Scalable Earth Grid (EASE-grid) projection. Mean maximum SWE values for our study basins ranged from 70 - 100 mm for the study period (Table 4).
Based on a number of studies, typical error standard deviations for the AMSR-E SWE range from 10 - 50 mm (Tong and Velicogna, 2010; Kelly et al., 2003; Foster et al., 2005). These errors can be attributed to a number of detection issues. Microwave brightness temperatures of a snowpack are affected by the number of snow grains along an emission path, the size of the snow grains, and the density of the snow pack (Kelly et al., 2003). These factors can change rapidly due to periods of snow metamorphism, thus causing variations in the microwave signal. Vegetation canopy interference and highly variable topography can also produce complex microwave signals leading to greater error (Foster et al., 2005; Dong et al., 2005; Tong et al., 2010). Periods of snow melt and refreezing will lead to a presence of ice or free water. This creates an overestimation in brightness temperature because water within the snow pack does not scatter microwave radiation, but rather absorbs and reemits microwave radiation, thus affecting the brightness temperature signal (Foster et al., 2005; Derksen et al., 2000). Further,
assuming a constant snow density across an area will subsequently lead to error since overall snow density can vary based on conditions at the time of snowfall, metamorphism, time period, and new snowfall over old snow pack (Tedesco and Narvekar, 2010). In this study, we derive AMSR-E SWE bias for the entire Upper Midwest region. This will be discussed in further detail in section 2.5.

The data source that is used for generating AMSR-E SWE bias and verifying modeled SWE output is National Operational Hydrologic Remote Sensing Center (NOHRSC) airborne snow survey data. Limited in situ SWE observations in the region of interest makes it difficult to verify modeled SWE output. NOHRSC performs airborne surveys of SWE several times during a winter season across the entire Midwest region. Similar to the AMSR-E SWE, the airborne observations use a gamma radiation sensor to estimate SWE (Carroll, 2001). Potassium, uranium, and thorium radioisotopes within the upper zone of soil naturally emit gamma radiation. The radiation signal is attenuated by any water mass present on the soil surface, regardless of phase. Measurement of the signal over bare ground can be compared to the signal over a snow pack. Differences in the radiation signal can be used to estimate the SWE. Airborne surveys are conducted in predefined flight lines approximately 16 km long and 152 m above the ground surface. Gamma measurements make it possible to estimate SWE over approximately a 5 sq. km area. NOHRSC has conducted verification research of the gamma SWE observations over agricultural and heavily forested areas. One thousand depth and density ground measurement were collected for each flight line verification in the study. 15 flight lines over agricultural areas and 70 flight lines over forested areas were verified. Mean absolute error (MAE) was 0.762 cm for agricultural areas; however, the MAE value over forested areas
was higher at 1.87 cm, primarily due to canopy interference with the radiation signal. All watersheds in this study with the exception of DANW3 are located in agricultural areas.

Because the SNOW17 model used in this study is lumped, basin average daily AMSR-E SWE values were derived for each watershed using ArcGIS. The number of AMSR-E grid cells contained within one watershed varied widely between watersheds due to variation in watershed size (Table 4). Twenty AMSR-E grid cells were contained within the boundaries of RAPM5, whereas MMLM5 contained the fewest number of AMSR-E grid cells at 4. Certain grid cells were only partially contained within the watershed boundaries, therefore, a weighted average SWE value was found.

2.5 Data Correction

Due to the factors that create uncertainty in passive microwave and gamma radiation measurements previously outlined, outlier removal and bias correction was performed on the data using the modified Z-score test. This was necessary in order to prevent skewing of AMSR-E bias values that were calculated for the entire upper Midwest and to prevent assimilation of erroneous data values. This test was introduced by Iglewicz and Hoaglin (1993) as a more robust outlier detection technique compared to the traditional Z score. The modified Z score test was chosen over other formal outlier tests due to certain limitations associated with those latter methods such as having to specify the number of outliers or the test being designed for detection of a single outlier. The modified Z score can be written as:

$$M_i = \frac{0.6745(x_i - \bar{x})}{MAD}$$  \hspace{1cm} (7)

$$MAD = median\{|x_i - \bar{x}|\}$$  \hspace{1cm} (8)
where MAD is the median absolute deviation, $\bar{x}$ is the sample median, and $x_i$ is the particular data point being tested. The modified Z score is a more robust detection technique over the standard Z score because it uses the median for calculation of the Z score, whereas the standard Z score uses the mean of the data set; therefore, the modified Z score is less susceptible to influence by outliers. This test uses units to determine potential outliers. The goal of using the modified Z score test was to establish an appropriate mean bias value for the AMSR-E data without influence by outliers, and to eliminate potential outliers from the assimilation data. For purposes of AMSR-E SWE bias calculation, a total of 1582 NOHRSC airborne SWE observations were used from 482 predefined flight lines for the 2005-2011 period. SWE error was calculated between each NOHRSC airborne observation and corresponding AMSR-E grid cell based on:

$$e = F - A$$

(9)

where $F$ is the AMSR-E SWE value, and $A$ is the NOHRSC airborne observation taken to be as the true value. Both data sets are assumed to be normally distributed. The modified Z score test was then applied to the NOHRSC airborne SWE points, the corresponding AMSR-E SWE values, and the calculated error. This test was applied to the error to eliminate points where extreme differences occur between the NOHRSC airborne SWE and AMSR-E SWE. If points were eliminated in any dataset, corresponding points in the other data sets were eliminated. Iglewicz and Hoaglin (1993) suggested using a modified Z-score value of 3.5 or greater in determining potential outliers. This Z score value was initially used; however, values that appeared to be erroneous were not eliminated. Therefore, a stricter Z score of 3.0 was used. This value was chosen based on the Three Sigma Rule which states that 99.73% of observations fall within 3
standard deviations of the mean. Figure 4 displays AMSR-E SWE bias at each corresponding NOHRSC airborne SWE flight line for the entire 2005-2011 period. The general bias pattern across the upper Midwest region indicates an overall negative bias. Mean AMSR-E bias after outlier removal was computed to be -17.91 mm with a standard deviation of 29.73 mm. All AMSR-E SWE data was bias corrected by using the mean value of 17.91 mm before input into the EnKF. The standard deviation value of 29.73 mm was used to calculate the covariance $C_v$ of the observations, which was subsequently used in computing the Kalman gain $K$ and measurement error $v^i$ within the EnKF update equation (4). Certain areas do indicate a heavy negative bias less than the mean of -17.91 mm, especially near RAPM5, MMLM5, and DANW3; however, for purposes of simplicity, the mean bias value was used for each site.

![Figure 4. AMSR-E SWE bias at each corresponding NOHRSC airborne SWE observation for the 2005-2011 period. Increasingly negative bias values are displayed with increasingly larger solid circle. Increasingly positive bias values are displayed with increasingly larger open circle.](image-url)
2.6 Ensemble Perturbation

The EnKF has the advantage over other assimilation methods in that it is able to represent uncertainties associated with model inputs. Our main state variable of interest, SWE, is most sensitive to the time-dependent forcing variables temperature and precipitation. In our application, we accounted for forcing error by stochastically perturbing these forcing variables from their nominal value with continuous probability distributions. This generates an ensemble of model precipitation and temperature inputs, which in turn produces an ensemble of model states at each time step. The ensemble of model states allows us to estimate the model uncertainty through calculations of $C_z$ and $C_{yz}$ (3), which are then input into calculation of the Kalman gain. Perturbed precipitation values were assumed to follow multiplicative log-normal distribution:

$$P = \bar{P} \gamma$$

where $\bar{P}$ is the raw observed precipitation, and $\gamma$ is a precipitation error drawn from a lognormal distribution with a mean value of one. We assume that there is temporal error correlation with regards to precipitation error at sequential time steps. A lag-1 autoregressive model (AR(1)) is applied to the precipitation error term:

$$\varepsilon_{t+1} = \rho \varepsilon_t + w_t$$

where $\varepsilon_t$ is the error at time step $t$, $\varepsilon_{t+1}$ is the error at time step $t + 1$, $\rho$ is a lag-1 autocorrelation coefficient that represents an exponential function, and $w_t$ is the error white noise with a mean of zero. The air temperature is assumed to follow a normal distribution:

$$T' = T + T_e$$
where $\overline{T}$ is the raw observed temperature, $T_e$ is the error drawn from a normal distribution with mean of zero and standard deviation 0.5°C. It is worth noting that these forcing errors are synthetic approximations, and therefore, do not fully encompass all possible sources of error that can affect these forcing variables. In this application, these errors are used to generate ensemble members. Many data assimilation studies have used the same strategy with positive results (Konstantinos et al., 2005; Slater and Clark, 2005; De Lannoy et al., 2012).

### 2.7 Model Simulations

Assimilation is performed for seven basins in hope of improving SWE estimation during periods of high SWE variability and rapid snow melt. Results suggest that producing 100 EnKF ensemble members was sufficient to capture a complete model error structure. Two separate simulations were performed for each study basin for six water years, 1 October 2005 to 30 September 2011. An initial spin up period of 1 February 2005 to 30 September 2005 was needed to bring model states into equilibrium with precipitation and temperature trends during water year 2005. Each simulation utilized basin-specific SNOW17 and SAC-SMA parameters from the North Central River Forecast Center (NCRFC). An initial control simulation with no EnKF updating was performed for each basin. This simulation was forced with the raw precipitation and temperature observations without taking into account any uncertainties associated with these variables. The second simulation involved EnKF updating, and was forced with perturbed temperature and precipitation fields. The variable of interest, SWE, was updated on a daily time step, or whenever AMSR-E SWE data was available. The AMSR-E captures snow depth on all descending passes of the NASA AQUA satellite. Descending pass over times over the Upper Midwest range between 0700Z and 0900Z. Therefore, AMSR-E SWE observations were
assimilated at the nearest time step 0600Z or hour 0 within the model. During periods when
snow was not detected by the AMSR-E, updating of the model SWE state occurred every three
days. In each simulation, snowmelt outflow from the SNOW17 is input into the SAC-SMA, which
produces a time series of daily river discharge. Since our main period of interest is the melt
period, an analysis of river discharge was completed on the 16 Feb – 31 May period of each
water year. Verification statistics are outlined in the following section.

2.8 Model Verification

The verification period used for each study site was 1 Feb. 1 – 31 May, water years 2006-2011.
Each month within the verification period was subdivided into 2 parts: day 1 – day 15, and day
16 – end of month. Each model discharge simulation was evaluated at the daily time step for
each monthly subdivision through use of forecast bias, mean absolute error (MAE), Pearson’s r ,
and the Containing Ratio (CR):

\[
\text{Bias} = \frac{1}{N} \sum_{t=1}^{N} x_t - y_t
\]

\[
\text{MAE} = \frac{1}{N} \sum_{t=1}^{N} |x_t - y_t|
\]

\[
\text{r} = \frac{\sum_{t=1}^{N}(x_t-\bar{x})(y_t-\bar{y})}{\sqrt{\sum_{t=1}^{N}(x_t-\bar{x})^2} \sqrt{\sum_{t=1}^{N}(y_t-\bar{y})^2}}
\]

where \(x_t\) is the ensemble mean at time \(t\) and \(y_t\) is the observation at time \(t\). The containing ratio
is a measure of ensemble accuracy (Xiong and O’Connor, 2008):

\[
\text{CR} = \frac{1}{N} \sum_{t=1}^{N} I[y_t]
\]

where \(I[\cdot]\) is an indicator function as follows:
I \{y_t\} = \begin{cases} \frac{1}{x_{1(t)} < y_t < x_{2(t)}} & \text{if the observation falls within the upper and lower bounds of the ensemble members and} \\ 0, \text{otherwise} & \text{if the observation falls outside of the ensemble bounds.} \end{cases} \tag{17}

I \{y_t\} equals one if the observation falls within the upper and lower bounds of the ensemble members and I \{y_t\} equals zero if the observation falls outside of the ensemble bounds. It is important to note that all statistics and analysis were calculated using the arithmetic mean of the EnKF ensemble members with the exception of the containing ratio.
CHAPTER 3. RESULTS AND DISCUSSION

Simulated discharge results with no data assimilation are referred to as baseline, and results with daily EnKF updating are referred to as EnKF. Daily discharge data obtained from the NCRFC for each basin outlet was used to compute the aforementioned statistics. In addition, SWE from the baseline and EnKF model simulations was evaluated against airborne SWE observation. For each study site, airborne SWE observations were extracted from an area contained within 15 km of the basin boundary. These observations were used for SWE analysis. We felt that a 15 km zone was an appropriate distance, as this is the approximate lower boundary of mesoscale weather pattern, which are most prevalent across the northern Midwest (Orlanski, 1975).

Figure 5 displays a comparison of SWE between model simulations and airborne observations. The left and right figures show comparisons of airborne SWE with the baseline and EnKF simulations respectively. There is some evidence that the EnKF improves capture of observed SWE for certain basins; however, this result is not seen across all basins. Table 5 contains SWE MAE values for each site. Four of the seven sites displayed improved SWE MAE values: PLUM5, DANW3, BCHW3, SCR14. RAPM5 displayed the poorest results, with the EnKF MAE value (34.2 mm) being 13.2 mm higher over the baseline simulation (21.0 mm). For this particular basin, the EnKF underestimates SWE, particularly for years with larger snow accumulation. The baseline simulation matches SWE magnitude above 80 mm with more accuracy over the EnKF, which could translate to better peak performance during large melt periods. This signal, however, was not seen for SWE values below 80 mm for RAPM5. The four basins that showed lower SWE MAE values for the EnKF, improved on average by 5.9 mm. In
general, it can be seen that there is a larger clustering of points near the 1-1 line for the EnKF simulation for observed SWE values between 40 and 100 mm. This indicates overall improvement in matching SWE magnitudes in some cases. The baseline simulation displays a slightly more negative bias for this same range of observed SWE values. It is important to note that strong conclusions cannot be drawn from Figure 5 due to the general lack of airborne SWE observations near the basins for all time periods. NOHRSC performs airborne SWE surveys usually once or twice during any particular accumulation season near a study site. As will be discussed further, certain basins, in particular MMLM5, exhibit poor SWE results when making comparisons with limited airborne observations during the simulation period, but display overall improved discharge.

![Figure 5. Comparison of SWE between model simulations and airborne observations for each of the 7 study sites. SWE values for each study site are shown with a different symbol.](image)

<table>
<thead>
<tr>
<th></th>
<th>SWE MAE (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PLUM5</td>
</tr>
<tr>
<td>Baseline</td>
<td>25.1</td>
</tr>
<tr>
<td>EnKF</td>
<td>21.4</td>
</tr>
</tbody>
</table>

**Table 5.** Mean absolute error of SWE between model simulations and airborne observations for each of the 7 study sites.
In a large number of snow melt cases, improved SWE magnitude capture allowed the EnKF ensemble mean to match the observed discharge more closely in contrast to the baseline simulation. The EnKF displayed stronger skill for most basins, in particular, BCHW3, SCRI4, DARW3, DANW3, and MMLM5 (Appendix A.1.1, A.1.2, A.1.3, A.1.4, A.1.5). However, improvements were not limited to these basins. Basins that had mixed results in general showed at least one case where EnKF updating significantly improved discharge. The most common pattern that was present across basins with strong EnKF results was improvement in peak discharge associated with snow melt events. BCHW3, SCRI4, and DARW3 showed the best overall EnKF performance in this study. Simulation results for water year 2010 for BCHW3 display the baseline and EnKF having similar SWE magnitudes through late January 2010 (Figure 6). After this period, the baseline SWE contains a large negative bias. The EnKF mean and ensemble range were able to capture airborne SWE magnitude with greatest accuracy. EnKF SWE during this year was approximately 20-40 mm higher than baseline SWE. The divergence of EnKF SWE from the baseline SWE translates to greater EnKF skill in mid-March 2010. The EnKF displayed an improvement in maximum peak discharge by approximately 5 cms. Most other years for BCHW3 show the EnKF as being a feasible tool to improve discharge simulation. The EnKF simulation for SCRI4, water year 2008, displays a similar trend in results to BCHW3, and is one of the strongest examples of improvement using EnKF updating (Figure 7). The EnKF exhibits a large and progressively more significant SWE deviation from the baseline SWE as the snow pack evolves over time. Due to the low SWE values, the baseline discharge entirely missed the largest peak during the March snow melt period by approximately 70 cms, whereas the EnKF simulation contains comparable discharge values to the observations. However, water
year 2008 does also show evidence for AMSR-E SWE error. Discussion of AMSR-E SWE error will be presented in a later section within the results.

Figure 6. SWE (top) and stream discharge (bottom) for BCHW3, water year 2010. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line.

Figure 7. SWE (top) and stream discharge (bottom) for SCRI4, water year 2008. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line.
Baseline SWE values were lower than the EnKF SWE values across most years for BCHW3 and SCRI4. Airborne SWE observations taken during water years 2008, 2010 and 2011 near BCHW3 support the increase in EnKF SWE values over the baseline SWE values (Appendix A.1.1). Likewise, airborne SWE observations near SCRI4 indicate a similar pattern. During years when airborne SWE observations were available (2007, 2010), a comparison of EnKF SWE to baseline SWE shows the EnKF capturing the current snow pack state more accurately (Appendix A.1.2). The EnKF showed a continuing trend of discharge improvement due to increased SWE for DARW3, with the best EnKF results seen for water year 2009 (Figure 8). Snow pack melt characteristics for 2009 for DARW3 were identical to BCHW3, with 3 main discharge peaks associated with 3 significant decreases in SWE. As results for BCHW3 showed, the EnKF simulation for DARW3 displayed improved discharge for all three events. Verification statistics support the results seen for these watersheds. Normalized MAE, bias, and R values for BCHW3, SCRI4, and DARW3 show a general improvement in the EnKF over the baseline simulation during most periods, with the largest reductions in MAE during frequent melt periods (Figure 9, 10, 11). SWE and temperature uncertainty allow the upper EnKF ensemble members to capture the observed peak discharge even when the mean EnKF value showed overall underestimation, as is most evident in 2009 for DARW3. This translates to containing ratio values of 0.7-0.9 during the most frequent melt period 3/1-3/15 for DARW3. Similar containing ratios were achieved for BCHW3 and SCRI4 during the March time frame (3/1 – 3/15, 3/16 – 3/31), thus displaying confidence that the EnKF can capture a range of possible discharge magnitudes.
Figure 8. SWE (top) and stream discharge (bottom) for DARW3, water year 2009. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line.

Figure 9. Average bias, normalized mean absolute error (N MAE), Pearson’s r, and containing ratio (CR) for each bi-weekly period for BCHW3 across all study years.
Figure 10. Average bias, normalized mean absolute error (N MAE), Pearson’s r, and containing ratio (CR) for each bi-weekly period for SCRI4 across all study years.

Figure 11. Average bias, normalized mean absolute error (N MAE), Pearson’s r, and containing ratio (CR) for each bi-weekly period for DARW3 across all study years.
Results for DANW3 correlated similarly with results seen for BCHW3, SCRI4, and DARW3; however, the EnKF did not display the same degree of skills. Simulation results for water year 2006 show the EnKF SWE having a larger overall SWE magnitude than the baseline (Figure 12). One airborne SWE observation in late January clearly indicates both the baseline and EnKF simulations having a negative bias of 10 mm of SWE during that time period. It is important to note that the upper EnKF ensemble bounds do capture this airborne SWE observation. Since the EnKF simulation contains a larger SWE value later in the accumulation season, it is plausible to assume that the baseline simulation may be affected by a larger negative bias. This is supported by EnKF results showing improvement in peak discharge by 20 cms over the baseline during a rapid snowmelt event in late March and early April 2006. Comparison of the EnKF and baseline SWE simulations with a SWE observation in 2011 reveals the same signal seen in 2006, supporting the conclusion that the higher EnKF SWE tends to capture the true SWE value with greater accuracy (Appendix A.1.4). Other water years for DANW3, in particular 2008 and 2010, display the same divergent trends between the baseline and EnKF SWE simulations. Verification statistics do not show the EnKF having the strong skill seen for the previously discussed basins; however, normalized MAE was minimized during certain periods (Figure 13).
Figure 12. SWE (top) and stream discharge (bottom) for DANW3, water year 2006. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line.

Figure 13. Average bias, normalized mean absolute error (N MAE), Pearson’s r, and containing ratio (CR) for each bi-weekly period for DANW3 across all study years.
The progressive deviation of EnKF SWE from baseline SWE evident in BCHW3, SCRI4, DARW3, and DANW3 may be due to several critical model parameters associated with the SNOW17 model, most notably the snow correction factor (SCF). The SCF parameter is a multiplier within the SNOW17 used to control the amount of new snow that has accumulated before it is added to the existing snowpack (Anderson, 2006). This parameter is used to account for losses that have occurred during the accumulation period due to redistribution by wind and sublimation. Most airborne SWE observations near these study basins point to the SNOW17 being prone to SWE underestimation. This SWE error may be cumulative as the snow accumulation season progresses. Assimilation of AMSR-E data corrects for this underestimation to a certain extent in most cases. Based on the divergent trends in SWE, it is possible that the low baseline SWE accumulation may be due to poor RFC SCF values.

Precipitation analysis was performed for BCHW3 on water year 2010 from the time of first significant SWE accumulation through the time of complete snowmelt (3 December – 8 March). During this time frame, 79.6 mm of precipitation fell below the temperature threshold for snow (PXTEMP). Based on an RFC SCF value of 1.2 for BCHW3, approximately 95.5 mm of SWE should have accumulated due to new snowfall. It is important to note that 70.5 mm of liquid precipitation also fell on a significant snow pack during this period when temperatures were marginally above PXTEMP. A total of 11.5 days had temperature above PXTEMP. An analysis of accumulated SWE reveals that the baseline and EnKF simulations accumulated 96.5 mm and 170.09 mm of SWE respectively. It is reasonable to assume that the RFC SCF value of 1.2 may be too low; however, strictly taking into account accumulated precipitation below
PXTEMP, the EnKF indicates an SCF value of 2.13, which is considered to be unreasonably high. Therefore, we can conclude that the baseline simulation may contain not only error in the SCF value, but also in the percent liquid water holding capacity of the snow pack (PLWHC), or PXTEMP. If PLWHC is too small, the model may underestimate the total amount of SWE. Considering 70.5 mm of liquid precipitation fell during periods when a significant snowpack existed, this parameter can introduce significant error into the total SWE value. Uncertainty in PXTEMP may also factor in on snow accumulation. Precipitation that fell in a liquid state during this time frame usually fell with temperatures 1-2°C above PXTEMP. For instance, late December 2009 shows a period were AMSR-E SWE increases by a substantial amount close to 40 mm. This increase in AMSR-E SWE corresponds to a period of precipitation. Most notably, 31.8 mm of precipitation fell on 24 December 2009 with temperatures during that day ranging between -0.7°C and 1.2°C. Variations in the true PXTEMP is highly dependent on atmospheric conditions (Dingman, 2002). Therefore, the SNOW17 model may accumulate a smaller SWE amount than observed if the model PXTEMP is lower than the true PXTEMP. In most cases, a combination of errors in the SCF, PXTEMP, and PLWHC parameters may influence the baseline simulation.

MMLM5 displayed results similar to the previously discussed study basins; however, AMSR-E error is clearly evident during certain years, which reveals another possible model error. Water year 2006 shows corresponding trends in SWE to the previously discussed sites, with EnKF SWE diverging from baseline SWE in mid-December 2005 (Figure 14). Airborne observations in mid-January support the EnKF SWE value, with upper and lower ensemble members capturing the range of airborne SWE. SWE observations during 2007, 2010, and 2011
all point to the SNOW17 being prone to SWE underestimation (Appendix A.1.5). Due to the main land use within MMLM5 being agricultural, blowing snow can increase precipitation losses at the observing station, which translates to lower SWE amounts within the SNOW17 model. Based on the airborne observations, true SCF values may range between 1.3 – 2.0, higher than the RFC value of 1.1. However, as was discussed earlier, the error in the baseline simulation may be due to numerous model parameters. The increased SWE amount allows the EnKF to capture the snow melt discharge peak in early April 2006 with significant accuracy. Overall statistics show decreased bias and normalized MAE for most bi-weekly periods with the exception of 3/1 – 3/15 (Figure 15). The poor normalized MAE value for 3/1 – 3/15 however, is a result of the large overestimation in discharge during March 2006 due to AMSR-E overestimation error.

Figure 14. SWE (top) and stream discharge (bottom) for MMLM5, water year 2006. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line.
Overall discharge results for BCHW3, SCRI4, DARW3, MMLM5, and DANW3 displayed significant improvements in upper 1/3 percentile discharge associated with snow melt outflow (Figure 16). However, questions arise as to whether some improvement may have actually been introduced by AMSR-E error. Even with good peak discharge performance during water year 2006 for MMLM5, error within the AMSR-E SWE measurement is evident during this year after the first major snowmelt in late January during which temperatures reached 6.2° C (Figure 14). As was previously mentioned, microwave brightness temperature is sensitive to changes in the structure and composition of snow, and also to ice or free water present within the snow pack, which may lead to overestimation in SWE (Foster et al., 2005, Dong et al., 2005, Derksen et al.,
2000). The airborne SWE observation in early February 2006 indicates the EnKF overestimating SWE by approximately 40 mm, which causes substantial overestimation of low flow discharge during the first half of March 2006.

Figure 16. Average Normalized MAE for the upper 1/3 percentile flows for each study. Normalized MAE error was calculated on all peak flows occurring during a time period between 1 February till 7 days past full snow melt.

AMSR-E error is also evident for other study sites after snowmelt, in particular SCRI4 2008 and DANW3 2008. Analysis of AMSR-E SWE for SCRI4 water year 2008 shows an increase in SWE value by approximately 40 mm in early January (Figure 7). This period featured a maximum temperature of 3.8°C. Measureable precipitation not recorded until 17 January 2008. This is a clear example of AMSR-E error due to snow metamorphosis and liquid water present within the snow pack. A similar error also occurred toward the end of January 2008. Due to the erroneous AMSR-E SWE values, EnKF SWE values are increased. Therefore, it is feasible to assume that the heightened EnKF discharge magnitude during this year may be
specifically tied to error. AMSR-E SWE values for late January 2008 for DANW3 follow this same general error pattern. Subsequently, EnKF discharge was greater than baseline discharge, showing improvement (Figure 17). Due to limited SWE observations, it is difficult to pinpoint specific model error. In certain cases, AMSR-E overestimation error might be inducing better SNOW17 performance, thus indicating that the model needs to contain artificially high SWE values to show improved discharge for the largest peak magnitudes. However, in cases where airborne SWE observations are available, EnKF SWE values seem to capture the true magnitude of SWE values more accurately, which would indicate that the model may have a SWE underestimation error due to specific parameters or model structure.

![Graph of SWE and stream discharge](image)

**Figure 17.** SWE (top) and stream discharge (bottom) for DANW3, water year 2008. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line.

In contrast to results seen for the majority of the study sites, mixed result were obtained using the EnKF for PLUM5. During certain years, the EnKF exhibited strong skill, but
these results were offset by an equal number of periods with poor performance. Water year 2006 for PLUM5 follows a similar pattern compared to the sites with good overall performance. The EnKF simulation shows SWE being approximately 40 mm higher at the end of March than the baseline simulation (Figure 18). This allows the EnKF to parallel the magnitude of the peak discharge in early April, thus displaying an increase in discharge over the baseline by 45 cms. However, a comparison of AMSR-E SWE values with precipitation and temperature data shows signs of AMSR-E overestimation error, particularly after a melt event in late January 2006. This provides further supporting evidence that the SNOW17 structure may require artificially high SWE values to capture the observed discharge magnitude in certain cases.

![Plummer, MN 2006](image)

**Figure 18.** SWE (top) and stream discharge (bottom) for PLUM5, water year 2006. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line.

Water year 2007 represents another case where the largest EnKF peak discharge corresponds to the largest observed discharge better than the baseline discharge (Figure 19). It
is worthy to point out that model melt error can also be seen during this year. In mid-March 2007, the first significant snow melt of the accumulation season occurs. The maximum EnKF SWE is greater than 20 mm higher than the baseline SWE. Snow melt in mid-March does not show a corresponding increase in discharge however, with both the EnKF and baseline simulations underestimating peak discharge by the same magnitude. This error may be a result of poor model melt factor parameters (MFMAX and MFMIN in the SNOW17 model), or the model lagging snow melt runoff for too significant of a time period. Water year 2008 is an example of poor performance by the EnKF for PLUM5 (Figure 20). It is clearly evident that EnKF SWE is overestimating the true amount of SWE and is therefore, greatly overestimating peak discharge in April. Due to the mixed results, certain periods show improved normalized MAE values, particularly for February and March, whereas both bi-weekly periods in April show higher normalized MAE values for the EnKF (Figure 21).

Figure 19. SWE (top) and stream discharge (bottom) for PLUM5, water year 2007. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line.
Figure 20. SWE (top) and stream discharge (bottom) for PLUM5, water year 2008. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line.

Figure 21. Average bias, normalized mean absolute error (N MAE), Pearson's r, and containing ratio (CR) for each bi-weekly period for MMLM5 across all study years.
Water year 2008 for PLUM5 may indicate that an AMSR-E bias correction of 17.91 mm might not be applicable to all study sites due to EnKF SWE overestimation in this case. A mixed pattern of negative and positive AMSR-E SWE bias is shown near PLUM5, with a higher number of positive bias points (Figure 4). Therefore, the positive bias correction may cause large SWE overestimation in certain years for PLUM5. Results for water year 2007 for RAPM5 mimic this conclusion, but in contrast to results for PLUM5, underestimation is prevalent in the case of RAPM5 (Figure 22). Analysis of SWE shows the EnKF underestimating SWE by approximately 40-60 mm in March based on airborne SWE observations. This underestimation is greater than baseline SWE underestimation, thus introducing a maximum error of 300 cms into the simulated discharge. Large EnKF SWE underestimation is also present across most other years for RAPM5, which does not allow EnKF discharge results to match the greater magnitude of observed discharge (Appendix A.1.7). This creates higher normalized MAE for most periods (Figure 23). The largest EnKF normalized MAE and bias is associated with the most frequent melting period (3/16 – 3/31). In contrast to PLUM5, a significant negative AMSR-E SWE bias is shown over RAPM5, with most points displaying a bias of -40 to -100 mm (Figure 4). These negative SWE bias values are also present over SCRI4 and MMLM5. These sites displayed positive discharge results due the bias-corrected AMSR-E SWE adjusting EnKF SWE upward in most cases in contrast to RAPM5. However, water year 2007 for SCRI4 and MMLM4 showed comparable EnKF SWE underestimation to RAPM5 (Appendix A.1.2 and A.1.5). RAPM5, SCRI4, and MMLM5 are located in a region that typically sees frequent, short melt periods during the winter period. This induces rapid snow metamorphosis such that snow grain size increases. The result is a reduction of snow grains within the snow pack over a large region, causing a
reduction in the brightness temperature signal (Kelly et al., 2003). Therefore, it may be necessary to derive basin-specific AMSR-E SWE bias values.

Figure 22. SWE (top) and stream discharge (bottom) for RAPM5, water year 2007. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line.

Figure 23. Average bias, normalized mean absolute error (N MAE), Pearson’s r, and containing ratio (CR) for each bi-weekly period for RAPM5 across all study years.
It is evident from SWE and discharge results that SNOW17 model error is not limited to SWE underestimation, but is also present in snow melt rates. This melt error is a pattern that is consistent across all study basins for water year 2010, and is most prevalent in years characterized by large SWE amount and rapid snow melt. Figure 24 and Figure 25 display results for RAPM5 and MMLM5 for water year 2010. Results support conclusions previously made, most notably that AMSR-E bias is basin-specific, and in most cases, airborne SWE observations validate the general assumption that large negative SWE bias is present within model. Further, a comparison of model simulations against AMSR-E observations during the melt period point to a rather slow melt rate within the SNOW17 model. In the case of RAPM5, the baseline simulation matches observed SWE with greater accuracy. However, the AMSR-E melts the snow pack within 4 days, whereas the simulations show 22 days of melting. Comparable melting time frames are also seen for MMLM5 during this year. Considering that the baseline simulation underestimates the peak discharge by approximately 400 cms for RAPM5, it can be assumed that the melt rate may need to be larger in order to match the rapid increase in discharge magnitude. It is possible that errors exist within the SNOW17 minimum and maximum melt factor parameters (MFMIN, MFMAX) that may be inducing poor capture of discharge. Based on these results, MFMIN and MFMAX may need to be adjusted upward.
Figure 24. SWE (top) and stream discharge (bottom) for RAPM5, water year 2010. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line.

Figure 25. SWE (top) and stream discharge (bottom) for MMLM5, water year 2010. The EnKF ensemble spread is shown with a gray shaded area, and the mean of the ensemble members is shown with a solid white line.
CHAPTER 4. CONCLUSIONS

Flood forecasting due to snowmelt relies on accuracy of initial snow states within a hydrologic model. Remote sensing provides a means of observing these model states directly; however limitations exist primarily due to errors associated with land cover types and weather conditions. Combining complimentary information from model simulations and passive microwave observations can provide optimal estimates of model states. In this study, we explore the use of the EnKF framework to assimilate remotely sensed SWE data from the AMSR-E sensor into the lumped RFC SNOW17 model. Prior to assimilation, AMSR-E SWE data was verified using data from the NOHRSC airborne snow survey program. Outliers were removed from each dataset and AMSR-E bias was calculated. A general average AMSR-E SWE bias of -17.91 mm was found for the Upper Midwest. For simplicity purposes, this bias was applied to AMSR-E SWE data for each basin prior to assimilation. Assimilation was performed at the daily time step for seven study basins in the NCRFC. 100 EnKF ensemble members were used to capture a complete model error structure. Each ensemble member was forced with perturbed precipitation and temperature fields in order to introduce potential forcing data uncertainty that may exist. An initial control run was performed for each basin to use as a base comparison for the EnKF simulation.

In this study, we demonstrate that assimilation of AMSR-E data produces improved results when compared to the baseline simulations, with overall improvement seen for five study sties: BCHW3, SCRI4, DARW3, MMLM5, and DANW3. The EnKF demonstrated the highest skill on peak discharge associated with snowmelt, as can been seen in the normalized MAE values in Figure 16. This is due to a general pattern of higher EnKF SWE values seen across
most sites. Airborne SWE observations confirm the validity of increased EnKF SWE values, thereby demonstrating that the SNOW17 model has a tendency to underestimate SWE values. This may be due to error in model structure or several model parameters, most specifically SCF, PXTEMP, and PLWHC. However, AMSR-E SWE error was determined to be a factor during and after periods of snowmelt due to the different scattering properties associated with liquid water. This has complicated analysis, making it difficult to pinpoint whether certain improvements may have actually been caused by AMSR-E error. This leads to a second general conclusion that the model may need to overestimate SWE in certain cases in order to achieve a more realistic discharge simulation.

Assimilation of AMSR-E SWE data for the NCRFC basins introduces several challenges that will need to be addressed in future studies. A more detailed AMSR-E bias analysis needs to be performed for each study basin to determine regional differences in bias potentially due to differing land cover types. Additionally, SWE needs to be analyzed for different periods to determine variations in bias due to wet and dry years, new snowfall, and refreezing of snowpack. It may be useful to investigate more sophisticated bias reduction methods such as that found in Reichle and Koster (2004). Furthermore, narrowing down assimilation to study sites with many airborne SWE observations would allow a more in depth combined precipitation and SWE analysis to be performed. This may reveal a more definitive picture of patterns associated with AMSR-E error. Since the model may contain parameter error, manually adjusting parameters within the baseline simulation could help pinpoint specific errors associated with the model. Testing the assimilation of AMSR-E SWE data could further be completed with a different snow model to see if similar improvement can be achieved.
APPENDIX

A.1 Discharge Figures

A.1.1 BCHW3
A.1.2 SCRI4

Sac City, IA 2006

Sac City, IA 2007
A.1.3 DARW3

Darlington, WI 2006

Darlington, WI 2007
A.1.4 DANW3
A.1.5 MMLM5

Marshall, MN 2006

Marshall, MN 2007
A.1.6 PLUM5

Plummer, MN 2006

Plummer, MN 2007
A.1.7 RAPM5

Rapidan, MN 2006

Rapidan, MN 2007
REFERENCES


ACKNOWLEDGEMENTS

Financial support for this work was provided by NASA grant #NNX10AQ77G S01. I would like to sincerely thank numerous people who guided me through this project and helped me achieve my goal. I would like to thank my advisor, Dr. Kristie Franz, for providing me with the opportunity to perform some excellent research. Dr. Franz guided me through many steps of this project and helped me gain skills that will be useful in completion of my Ph.D.

I would like to thank my committee members, Dr. Bill Simpkins and Dr. Bill Gutowski, for their excellent feedback and ideas.

Thanks to my fellow ISU colleagues, Ryan Spies and Angela Bowman, for their assistance and entertainment over the last two years. Special thanks to my family for their encouragement and support through the many years of my education. I would not be at the point where I currently am without their guidance. Lastly, I would like to thank Ms. Amanda Black for all of her help over the last several months.