Hydrologic Data Assimilation: State Estimation and Model Calibration

by

Caleb Matthew DeChant

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science
in
Civil and Environmental Engineering

Thesis Committee:
Hamid Moradkhani, Chair
Dacian Daescu
David Jay

Portland State University
2010
Abstract

This thesis is a combination of two separate studies which examine hydrologic data assimilation techniques: 1) to determine the applicability of assimilation of remotely sensed data in operational models and 2) to compare the effectiveness of assimilation and other calibration techniques. The first study examines the ability of Data Assimilation of remotely sensed microwave radiance data to improve snow water equivalent prediction, and ultimately operational streamflow forecasts. Operational streamflow forecasts in the National Weather Service River Forecast Center are produced with a coupled SNOW17 (snow model) and SACramento Soil Moisture Accounting (SAC-SMA) model. A comparison of two assimilation techniques, the Ensemble Kalman Filter (EnKF) and the Particle Filter (PF), is made using a coupled SNOW17 and the Microwave Emission Model for Layered Snowpack model to assimilate microwave radiance data. Microwave radiance data, in the form of brightness temperature (TB), is gathered from the Advanced Microwave Scanning Radiometer-Earth Observing System at the 36.5GHz channel. SWE prediction is validated in a synthetic experiment. The distribution of snowmelt from an experiment with real data is then used to run the SAC-SMA model. Several scenarios on state or joint state-parameter updating with TB data assimilation to SNOW-17 and SAC-SMA models were analyzed, and the results show potential benefit for operational streamflow forecasting.

The second study compares the effectiveness of different calibration techniques in hydrologic modeling. Currently, the most commonly used methods for hydrologic model calibration are global optimization techniques. While these techniques have become very
efficient and effective in optimizing the complicated parameter space of hydrologic models, the uncertainty with respect to parameters is ignored. This has led to recent research looking into Bayesian Inference through Monte Carlo methods to analyze the ability to calibrate models and represent the uncertainty in relation to the parameters. Research has recently been performed in filtering and Markov Chain Monte Carlo (MCMC) techniques for optimization of hydrologic models. At this point, a comparison of the effectiveness of global optimization, filtering and MCMC techniques has yet to be reported in the hydrologic modeling community. This study compares global optimization, MCMC, the PF, the Particle Smoother, the EnKF and the Ensemble Kalman Smoother for the purpose of parameter estimation in both the HyMod and SAC-SMA hydrologic models.
Acknowledgements

In appreciation of the hard work of my thesis committee, research group and my advisor, I would like to extend my appreciation. For my advisor Dr. Moradkhani, you have spent countless hours teaching me the knowledge necessary to perform this research, provided me with necessary information to perform relevant research to the field, pushed me to achieve research objectives that seemed out of reach and helped to shape my research into the best it could be. Thank you for all of the time you have put into making this research possible. To my thesis committee, Dr. Jay and Dr. Daescu, I greatly appreciate your willingness to serve as members of my thesis committee in spite of the short notice and lack of involvement in my work. Your participation in editing my thesis and your suggestions for improving my research was quite valuable for my final thesis product and future work moving forward. To my research group, all of whom were always willing to help in improving my work as well as providing moral support to continue the hard work it has taken to complete my research, thank you all. Last but not least I would like to thank my parents for providing me with continued support throughout my studies. Your encouragement and guidance have been essential for my academic success. For the contributions from all of you, I am very grateful.
Table of Contents

Abstract ................................................................................................................................. i
Acknowledgements........................................................................................................... iii
List of Tables ....................................................................................................................... v
List of Figures .................................................................................................................... vi
List of Abbreviations ........................................................................................................ x
Chapter 1
Introduction ...................................................................................................................... 1
Chapter 2
Hydrologic Datasets ......................................................................................................... 11
Chapter 3
Hydrologic Models .......................................................................................................... 18
Chapter 4
Data Assimilation and Calibration Techniques .............................................................. 25
Chapter 5
Experimental Setup ......................................................................................................... 41
Chapter 6
Results and Discussion ................................................................................................. 44
Chapter 7
Conclusion ....................................................................................................................... 94
References ...................................................................................................................... 96
List of Tables

Table 1
HyMod States.....................................................................................................................18

Table 2
HyMod Parameters............................................................................................................ 18

Table 3
SAC-SMA States...............................................................................................................20

Table 4
SAC-SMA Parameters.......................................................................................................20

Table 5
SNOW-17 States................................................................................................................21

Table 6
SNOW-17 Parameters........................................................................................................22

Table 7
Data Assimilation Errors....................................................................................................58
List of Figures

Figure 1
East River Basin.................................................................................................................12
Figure 2
Leaf River Basin................................................................................................................ 13
Figure 3
EnKF Flowchart.................................................................................................................29
Figure 4
EnKS Flowchart.................................................................................................................31
Figure 5
Moving Batch Smoothing.................................................................................................32
Figure 6
PF Flowchart......................................................................................................................35
Figure 7
PS Flowchart......................................................................................................................37
Figure 8
Snow Prediction Performance Metrics.............................................................................. 45
Figure 9
Middle Elevation Band Snow Prediction Time Series......................................................46
Figure 10
Upper Elevation Band Snow Prediction Time Series.........................................................47
Figure 11
Winter Snow Prediction Performance Metrics.................................................................49
Figure 12
Middle Elevation Band Spatial Snow Prediction...............................................................51
Figure 13
Upper Elevation Band Spatial Snow Prediction.................................................................52
Figure 14
Streamflow Prediction Benchmark Efficiency.................................................................. 55
Figure 15
Streamflow Prediction RPSS ........................................................................................................56
Figure 16
Streamflow Prediction Time Series ..........................................................................................57
Figure 17
RMSE from HyMod ..................................................................................................................60
Figure 18
RMSE of Ensemble Prediction and Single Value Prediction ..................................................60
Figure 19
HyMod Calibration Box Plot ..................................................................................................62
Figure 20
Alpha Parameter Calibration .....................................................................................................63
Figure 21
Beta Parameter Calibration .......................................................................................................64
Figure 22
Cmax Parameter Calibration .....................................................................................................65
Figure 23
Rs Parameter Calibration .........................................................................................................66
Figure 24
Rq Parameter Calibration .........................................................................................................67
Figure 25
Adjusted EnKF and EnKS Results ..........................................................................................69
Figure 26
RMSE from SAC-SMA ..............................................................................................................71
Figure 27
RMSE from Ensemble Prediction and Single Value Prediction ..............................................72
Figure 28
UZTWM Parameter Calibration .............................................................................................73
Figure 29
UZFWM Parameter Calibration ...............................................................................................74
<table>
<thead>
<tr>
<th>Figure</th>
<th>Parameter Calibration</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>UZK Parameter Calibration</td>
<td>75</td>
</tr>
<tr>
<td>31</td>
<td>PCTIM Parameter Calibration</td>
<td>76</td>
</tr>
<tr>
<td>32</td>
<td>ADIMP Parameter Calibration</td>
<td>77</td>
</tr>
<tr>
<td>33</td>
<td>ZPERC Parameter Calibration</td>
<td>78</td>
</tr>
<tr>
<td>34</td>
<td>REXP Parameter Calibration</td>
<td>79</td>
</tr>
<tr>
<td>35</td>
<td>LZTWM Parameter Calibration</td>
<td>80</td>
</tr>
<tr>
<td>36</td>
<td>LZFSM Parameter Calibration</td>
<td>81</td>
</tr>
<tr>
<td>37</td>
<td>LZFPM Parameter Calibration</td>
<td>82</td>
</tr>
<tr>
<td>38</td>
<td>LZSK Parameter Calibration</td>
<td>83</td>
</tr>
<tr>
<td>39</td>
<td>LZPK Parameter Calibration</td>
<td>84</td>
</tr>
<tr>
<td>40</td>
<td>PFREE Parameter Calibration</td>
<td>85</td>
</tr>
<tr>
<td>41</td>
<td>Kq Parameter Calibration</td>
<td>86</td>
</tr>
<tr>
<td>42</td>
<td>SAC-SMA Calibration Box Plot 1</td>
<td>88</td>
</tr>
<tr>
<td>43</td>
<td>SAC-SMA Calibration Box Plot 2</td>
<td>89</td>
</tr>
<tr>
<td>44</td>
<td>SAC-SMA Calibration Box Plot 3</td>
<td>90</td>
</tr>
</tbody>
</table>
Figure 45
PF and EnKF Scenario Analysis........................................................................................................92
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMSR-E</td>
<td>Advanced Microwave Scanning Radiometer-Earth Observing System</td>
</tr>
<tr>
<td>BE</td>
<td>Benchmark Efficiency</td>
</tr>
<tr>
<td>EnKF</td>
<td>Ensemble Kalman Filter</td>
</tr>
<tr>
<td>EnKS</td>
<td>Ensemble Kalman Smoother</td>
</tr>
<tr>
<td>HDF-EOS</td>
<td>Hierarchical Data Format-Earth Observing System</td>
</tr>
<tr>
<td>MCMC</td>
<td>Markov Chain Monte Carlo</td>
</tr>
<tr>
<td>MEMLS</td>
<td>Microwave Emission Model for Layered Snowpack</td>
</tr>
<tr>
<td>MODIS</td>
<td>Moderate Resolution Imaging Spectroradiometer</td>
</tr>
<tr>
<td>NRCS</td>
<td>Natural Resources Conservation Service</td>
</tr>
<tr>
<td>NWS</td>
<td>National Weather Service</td>
</tr>
<tr>
<td>NWSRFC</td>
<td>National Weather Service River Forecast Center</td>
</tr>
<tr>
<td>PET</td>
<td>Potential Evapotranspiration</td>
</tr>
<tr>
<td>PF</td>
<td>Particle Filter</td>
</tr>
<tr>
<td>PS</td>
<td>Particle Smoother</td>
</tr>
<tr>
<td>RTM</td>
<td>Radiative Transfer Model</td>
</tr>
<tr>
<td>RPSS</td>
<td>Rank Probability Skill Score</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root Mean Square Error</td>
</tr>
<tr>
<td>SAC-SMA</td>
<td>Sacramento Soil Moisture Accounting Model</td>
</tr>
<tr>
<td>SCE-UA</td>
<td>Shuffle Complex Evolution-University of Arizona</td>
</tr>
<tr>
<td>SWE</td>
<td>Snow Water Equivalent</td>
</tr>
<tr>
<td>TB</td>
<td>Brightness Temperature</td>
</tr>
</tbody>
</table>
1. Introduction

1.1 Motivation

The management of water as a resource relies heavily on information about the current state of the water cycle, including characterization of the volume of water stored in and above the land surface, the amount of water being removed from the land surface and the amount of water being added to the land surface. Accurate quantification of hydrologically significant variables such as groundwater, soil moisture, snow, precipitation, temperature, evapotranspiration and streamflow, as well as a strong understanding of how these variables affect each other, is crucial to minimizing the risk of both drought and flood. Since both drought and flood can have catastrophic effects on society and the environment, it is necessary to reduce the risk of both events as much as possible. This can be achieved through increasing the accuracy to which the state of the hydrologic cycle can be estimated, as well as improving the accuracy of predictions from hydrologic models.

Of particular interest to this study is the characterization of snowpack properties. Seasonal snowpack accumulation and ablation is an important hydrologic process, significantly altering the energy and water balance at the land surface. Snow cover acts as a large reservoir over the winter accumulation season and decreases the thermal conductivity of the land surface, while increasing its albedo [49]. Seasonal snow cover has also been proven to affect numerical weather prediction and can even enhance North American climate anomalies. Due to the large quantities of water stored as snow during
winter months in mountainous river basins, snowpack ablation is the dominant factor impacting spring and early summer runoff, which in turn affects water resources management in these basins [31]. Since seasonal snow cover has such wide ranging impacts on land surface and atmospheric processes, accurate estimation of the volume of water stored as snow during the accumulation and melt season is necessary for water supply forecasting.

1.2 Objectives

The objective of this thesis is to show if there is potential for improving operational National Weather Service (NWS) river forecasts by improving snow melt estimates through remotely sensed microwave radiance data assimilation and to compare the effectiveness of optimization and various Monte Carlo Methods in hydrologic model calibration.

1.2.1 Microwave Radiance Data Assimilation

1. Compare the effectiveness of data assimilation techniques in assimilating remotely sensed microwave brightness temperature (TB), through the use of a radiative transfer model (RTM), to improve snow water equivalent estimates in the SNOW-17 model.

2. Determine the extent to which snow melt predictions, resulting from TB assimilation, can improve streamflow forecasts in the Sacramento Soil Moisture Accounting Model (SAC-SMA).
1.2.2 Comparison of Calibration Techniques

1. Demonstrate the usefulness of Monte Carlo techniques in calibrating hydrologic models.

2. Determine which calibration technique creates the most accurate prediction, in terms of expected value and from a probabilistic standpoint.

1.3 Literature Review

1.3.1 Hydrologic and Energy Balance Modeling

Prediction of hydrologic quantities is performed through several different types of models. In mountainous areas such as the western U.S., it is necessary to begin with a snowmelt model, since most of the runoff in these areas is due to the melt of winter snow accumulation. Snowmelt models currently used by researchers and operational hydrologists can take on a range of complexity from simple temperature index to a full energy and water balance model. Melt predictions from these models typically are used as inputs to a runoff model. A runoff model predicts the portion of incoming water that is infiltrated into the soil and the portion which is translated to runoff. The excess water is then routed to the watershed outlet.

In temperature index models, the complex energy balance of a snowpack is simplified to estimate snow accumulation and melt depending on the temperature and time of year. For example, if the temperature is above a certain threshold (somewhere around 0°C), all precipitation falls as rain and the snowpack melts according to a specified relationship depending on the time of the year [25]. If the temperature is below that threshold, all precipitation falls as snow and melt will only happen if it is late enough
in the year [1]. In the NWS-River Forecast Center (NWSRFC), the SNOW-17
(temperature index) model is coupled with the SAC-SMA model, to produce operational
streamflow predictions. Models such as the SNOW-17 can be very accurate for snow
prediction but rely heavily on calibration [25]. The main advantage of temperature index
models is their simplicity, which makes them quite efficient computationally.

More complex energy balance models can also be used in snow prediction. Several research models, including the Variable Infiltration Capacity model [8], Simple Atmosphere-Snow Transfer model [41], NOAH model [20], SNTHRM [26], Utah Energy Balance Model [42], and the Community Land Model [36] to name a small few, model snowpack evolution by solving complex energy balance equations for each snowpack layer. These can involve incoming shortwave radiation, incoming and outgoing longwave radiation, changes in temperature and energy lost or gained through phase changes [42]. These models often predict evolution of internal snow variables as well including snow grain size and vapor fluxes through layers [26]. The complexity of these models can be beneficial because it allows them to be less reliant on calibration but this complexity also increases the computational cost of the model.

Runoff models vary greatly in complexity and can be conceptual or deterministic. The term conceptual model refers to a system of equations to predict the amount of runoff from the land surface which have no physical meaning. These models are typically a system of conceptual reservoirs, in series and parallel, which fill and drain based on the amount of water applied to them, and the calibrated outflow parameter. For example, the SAC-SMA model six reservoirs represent water in different levels of the land surface but
are related by parameters which are not physically based [6]. Similar to temperature
index models, conceptual models can be very simple but rely heavily on calibration [12].
While variables in conceptual models have no physical meaning, deterministic models
aim to predict the exact state of the land surface. Deterministic models that solve both the
water and energy balance of the land surface are called land surface models. These
models use distributed datasets of forcing, land cover and soil types to parameterize
equations which are based on the physical process of the hydrologic cycle. Water can be
held in different layers of the soil and in groundwater. Land surface models, similar to
snow energy balance models, are very complicated and require extensive computing
time.

1.3.2 Snow Observation

The observation of snow properties is performed in a variety of ways using both
in-situ and remote sensing methods. The purpose of snow observation is to quantify the
amount of water that is held in the snowpack. Estimating the amount of snow in
mountainous areas is necessary for effective water resources management because
snowmelt dominates runoff in these regions [31]. In-situ measurements are performed at
both constructed sites with automatic sensors and snow-course sites with manual
measurements. In the US, automatic snow sensors are located at snow telemetry sites,
called SNOTEL, managed by the Natural Resources Conservation Service (NRCS). At
these sites, SWE, snow depth, precipitation and temperature are all measured on an
hourly or daily basis. Snow course measurements are made throughout the western US
and consist of NRCS employee’s manually gathering snow depth and other properties. At
the point scale, these observations are quite accurate, but due to large spatial heterogeneities in snowpack, especially in high elevation regions, these data are not very representative of the overall snowpack. Remote sensing of snow properties has the potential to overcome issues of spatial representativeness of in-situ measurements. Remote sensing of snow is performed by using data from airborne or satellite based sensors that measure electromagnetic radiation, which is reflected or emitted from the land surface, to infer properties of the snowpack. This can provide spatial datasets of snow cover at about a 500m resolution and snow water equivalent at about a 25 km resolution. Although remote sensing of snow has the ability to create a spatial dataset, problems owing to sensor obstruction, imperfect inversion techniques and sensor saturation make remotely sensed snow properties quite uncertain.

1.3.3 Hydrologic Uncertainty and Data Assimilation

In the realm of hydrologic modeling, accurate prediction of significant hydrologic variables is made difficult due to several sources of uncertainty. Uncertainty in hydrologic predictions is a result of persistent errors in forcing data, initial states, model calibration and model structure. Forcing data, typically precipitation and temperature, are difficult to measure accurately for use in a hydrologic model due to spatially varying values, sensor error, topographic roughness and vegetation, to name a few. Calibration of hydrologic models is difficult because of the large number of model parameters, complex parameter spaces and imperfect input/observation data. The structure of hydrologic models creates uncertainty because a model cannot perfectly model the physics of the basin being modeled. In addition to uncertainties in the prediction of hydrologic
quantities, the observation of such quantities is also difficult. Problems owing to sensor errors, complex topography and imperfect inversion algorithms, among others, create uncertainties in hydrologic observation datasets. Due to the uncertainties of predicted and observational datasets, much hydrologic research has focused on statistically combining predicted predictions and observations in a framework that combines the knowledge of both datasets to produce a single, more accurate dataset, which also quantifies the uncertainty of the prediction. This process is called Data Assimilation.

Data assimilation techniques have been used extensively to predict model and land surface states and are gaining popularity for use in parameter estimation[33, 34]. The land surface states that are most commonly estimated through data assimilation are streamflow, soil moisture and snow water equivalent (SWE). In the prediction of these states, many observational datasets, both in-situ and remotely sensed, have been studied for the purpose of data assimilation including streamflow, SWE, snow cover percentage, soil moisture, evapotranspiration, and observed microwave emission. All of these datasets contain important information about the hydrologic cycle and can be used to add knowledge to a model. For the purposes of this study, streamflow and SWE are the target quantities.

Snow data assimilation can utilize a variety of datasets, both in-situ and remotely sensed, to infer properties about the snowpack. While in-situ datasets of SWE can be used to perform snow data assimilation, this is typically only used for technique development and validation due to the low representativeness of point snow measurements. This has led researchers to use distributed snow datasets. Most recent
research in snow data assimilation has focused on the use of remotely sensed snow properties. Remote sensing of snow utilizes observations of electromagnetic radiation from the earth’s surface to estimate snow properties. Typically observations of electromagnetic radiation used for snow applications are in the visible/infra-red or microwave frequencies. Land snow cover datasets, estimated from measurements in the visible/infra-red frequencies, have been the most widely studied for purposes of snow data assimilation. This is because snow cover datasets can have a very fine resolution, typically about 500m. While the resolution of this dataset makes it very attractive, the loose relationship between snow cover and SWE detracts from its applicability to hydrologic prediction. This has led some researchers to examine the utility of remotely sensed data in the microwave frequencies, which are represented in the form of brightness temperature (TB). Theses frequencies, though at a much coarser resolution (~25km), can penetrate the snowpack and are therefore sensitive to snow depth. By using observations from two frequencies in the microwave range, typically around 18 and 36 GHz, SWE can be estimated. SWE data created in this way have been implemented in a data assimilation study but, as pointed out by [11], errors in this dataset owing to the inversion techniques make this dataset unreliable in deep or wet snowpack. Since the inversion techniques have inherent errors, the most recent research has looked to assimilate TB directly. The feasibility of observed TB assimilation has been tested using the Simple Snow-Atmosphere Soil Transfer model coupled with the Microwave Emission Model for Layered Snow (MEMLS) [16, 19]. The Variable Infiltration Capacity model coupled with the Dense Media Radiative Transfer Model has also been shown to
This study expands on recent research by applying TB assimilation methods to operational models, with the intent of improving both SWE reconstruction and streamflow prediction.

1.3.4 Model Calibration

All hydrologic models, from the simplest to the most complex, contain parameters aimed at correcting model prediction for basins of different physical and climatic structure. In order for a given model to produce an accurate prediction, these parameters must be accurately calibrated. The choice of an applicable calibration technique is necessary for successful model parameter estimation. Techniques for calibrating hydrologic models range in complexity from manual calibration to algorithms like global optimization, Markov Chain Monte Carlo (MCMC) and data assimilation techniques. Manual calibration consists of a user inspecting the performance from a model, compared to the observation, to determine the best parameters. This method, though effective in some situations, is subjective and time-consuming, making automatic calibration techniques much more useful [12, 22, 40]. Global optimization techniques, which attempt to find the most accurate set of parameters with respect to an objective function, can be effective and efficient methods of calibration. This provides accurate, reproducible parameters for a model, during a given time period [12]. Optimization techniques, though useful for finding the best possible parameter sets, lack the ability to provide any information regarding the uncertainty of a model, with respect to parameters. In order to get an accurate estimation of the uncertainty of parameters, one must use Monte Carlo techniques to estimate the posterior parameter distribution. MCMC and data assimilation
methods can both account for uncertainty in the model parameters and provide an accurate parameter distribution. MCMC techniques predict the posterior parameter distribution by building parallel ergodic Markov Chains. The stationary distribution of these chains reflects the posterior density of parameters. In contrast, data assimilation methods create an ensemble of prior parameters, from which a posterior distribution can be calculated at each observation time step. While both MCMC and data assimilation methods have the ability to accurately predict the distribution of parameters, MCMC techniques require a long term dataset and are unable to account for possible temporal variability of parameters [33]. Thus MCMC techniques are restricted to basins with stationary parameters and long terms observation datasets. Up to this point, no study has compared to effectiveness of hydrologic model calibration using Global Optimization, MCMC and data assimilation. This is the focus of the second portion of this thesis and will be discussed further in later sections. In addition to standard filtering implementations, this study runs each filtering technique in a smoothing framework.
2. Hydrologic Datasets

2.2 Basin Descriptions

2.2.1 East River Basin

The East River watershed, a tributary of the Gunnison River in the Colorado River basin, was chosen as the study area for this research (see Figure 1). The river basin is 291 sq. miles (754 sq. km) with the elevation ranges from 8,022 ft to 12,789 ft (2445 to 3898m). The Gunnison River is the fifth largest tributary of the Colorado River basin. This makes estimation of the water volume stored in snow and streamflow in upper tributaries, in the Gunnison River Basin important for water supply forecasting in the Colorado River. Model cell elevations, which are necessary for determining cell elevation bands, were aggregated from the USGS seamless data set (1km resolution) to the modeling resolution (1/8\textsuperscript{th} degree). The modeling is conducted from September 1, 2002 through September 30, 2005. This time period was chosen because of the availability of all desired datasets which encompasses an entire snow accumulation and ablation period. All forcing data including precipitation and temperature at a 6-hourly time-step were provided by the NWS-CBRFC. The data are split into three elevation bands to accommodate running the models in a distributed format and cells are assigned forcing data depending on the average elevation of the cell.
2.1.2 Leaf River Basin

This study takes place over the Leaf River basin in southern Mississippi. The basin is 1944-km$^2$ and is the main tributary of Pascagoula River, which drains into the Gulf of Mexico. Data for this study was obtained from the National Weather Service Hydrology Laboratory, which consists of precipitation (mm/d), potential evapotranspiration (mm/d) and streamflow (cm$^3$/s). The calibration for all methods takes place over a 10 year period from July, 28$^{th}$ 1952 to July 28$^{th}$, 1962 and is validated for the following 3 years.
Figure 2. The Leaf River Basin in Southern Mississippi, USA

2.3 In-Situ Observations

2.3.1 Streamflow

In the U.S., streamflow measurement is provided by streamflow gauges managed by the United States Geological Survey. Streamflow observations are gathered from gauges that measure the depth of a river, at the watershed outlet. This is usually performed automatically but can be performed manually. This is then translated to flow through the stage-discharge relationship, which has been predetermined for a given stream.

2.3.2 SNOTEL

SNOTEL stations are distributed throughout the mountainous areas of the U.S. and are managed by the NRCS. The NRCS currently has over 800 SNOTEL sites in operation. A SNOTEL site typically consists of a precipitation gauge, a temperature gauge, a snow depth gauge and a snow pillow. The snow pillows at these sites are scales
that weigh the snowpack over about three meter square. This weight is then transferred to the volume of water stored in the snowpack by multiplying by the density of water. The amount of water is then reported as a depth of water (in inches) and referred to as snow water equivalent (SWE).

2.4 Remotely Sensed Data

2.4.2 Visible and Infra-red

Visible and infra-red radiation refers to wavelengths from about .4 to 14.4 µm. Currently, NASA provides data retrieved from these wavelengths as observed by the Moderate Resolution Imaging Spectroradiometer (MODIS), which is flown on both the Aqua and Terra satellites. The MODIS instrument makes observations at a 500 m resolution. Observations in the visible and infrared range are useful for snow observation because they are sensitive to both the presence of snow and the albedo of snow because of the high reflectivity of snow. Presence of snow can be inferred for each observation cell based on the Normalize Difference Snow Index, which compares the observations at MODIS bands 4 (545 - 565nm) and 6 (1628 – 1652nm). If the NDSI is greater than .4, then the whole cell is considered snow covered. This is used to create a binary map (snow or no snow) at the 500m resolution. Products at a coarser resolution than this provide percentage of snow cover determined from the 500m map. The albedo product has a separate algorithm that will not be discussed in this thesis. While the datasets provided in this range have the advantage of being at a very fine resolution, the downside of these products is that no data can be gathered in areas that are cloud covered. Another problem
with these products, with respect to this study, is this range of wavelengths cannot penetrate a snowpack and is therefore insensitive to SWE.

2.4.3 Passive Microwave

For the purpose of SWE estimation, passive microwave radiometer data is more useful than radiation in the visible and infrared range as microwave upwelling from the snowpack can originate from below the snowpack. This allows inference about snow properties, including SWE, based on these observations. The newest instrument that NASA has in operational use for passive microwave observation is the Advanced Microwave Scanning Radiometer- Earth Observing System (AMSR-E), which is flown on the Aqua satellite. This instrument measures in six different channels: 6.9, 10.7, 18.7, 23.8, 36.5 and 89 GHz, from which the TB is provided. TB is the temperature of the earth’s surface, calculated from these measurements, assuming the earth is a blackbody (emissivity of 1). For the purpose of SWE estimation, bands between 18 and 37GHz are the most useful as they penetrate the snowpack the furthest and are less sensitive to atmospheric effects or radio interference. The newest remotely sensed SWE datasets provided by NASA are gathered from the AMSR-E instrument. SWE products from the AMSR-E instrument are created through the [7] algorithm, which uses an empirical relationship of the difference of the 18.7 and 36.5GHz channels with SWE. While it is quite useful to have a dataset that provides remotely sensed SWE, datasets inverted from microwave TB are subject to errors from inversion owing to topographic roughness, proximity to water, sensor saturation, air temperature, dense forest, liquid water in the snowpack and changing snowpack structure. Since direct inversion of TB is associated
with many errors, recent research has looked to more complicated methods of SWE reconstruction through a radiative transfer model.

Radiative Transfer Models (RTM) are used to predict the microwave emission from the land surface based on the temperature and physiographic characteristics. Predicted microwave emission from snowpack is a factor of several variables including snow depth, density, grain size, liquid water content and temperature. Two popular techniques used in predicting snow microwave emission are using six-flux theory to describe multiple scattering within the snowpack, as is performed in the Microwave Emission Model For Layered Snowpack (MEMLS), and Dense Media Radiative Transfer theory, as is performed in the Dense Media Radiative Transfer Model [44]. While both of these methods are effective in prediction snow microwave emission, to date they have not been extensively compared and thus neither has been proven more effective. In addition to modeling the emission of microwave radiation from the snowpack, RTMs must also account for vegetation effects and atmospheric attenuation. The output from a RTM is a prediction of the TB above the earth’s atmosphere. This provides validation for a snow prediction model and a framework for data assimilation, which is only recently being studied.

2.4.4 Data Processing

Data from remote sensors come in a wide variety of formats including binary, ASCII, netCDF, Hierarchical Data Format-Earth Observing System (HDF-EOS) among others. For the purpose of this study, only TB data from the AMSR-E instrument were used, which are distributed in HDF-EOS. This is very similar to HDF format and can be
processed by a variety of software. Software for geolocation and transformation of this data into a binary or ASCII format is available from the National Snow and Ice Data Center website. This study used the HDF import tool within MATLAB to load and manipulate the data.
3. Hydrologic Models

3.1 Hydrologic Models

3.1.1 HyMod Model

The first model used in this study is the HyMod model, which has been used previously by several authors in testing of calibration strategies [33, 45]. This is a simple conceptual, lumped model containing 5 calibration parameters. The parameters and the calibration bounds applied to them are shown in Table 1. Inputs to HyMod are precipitation and potential evapotranspiration and the output is streamflow. The model allocates water between a series of three quick-flow tanks and one slow-flow tank. The model also consists of five state parameters. The states are the storage in each of the four tanks and a non-linear storage capacity. Model parameters are summarized in Table 1 and model state variables are summarized in Table 2.

Table 1. Parameters for the HyMod model from

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>R_q</td>
<td>Quick Flow Tank Parameter</td>
</tr>
<tr>
<td>R_s</td>
<td>Slow Flow Tank Parameter</td>
</tr>
<tr>
<td>Alpha</td>
<td>Partitioning Factor</td>
</tr>
<tr>
<td>Beta</td>
<td>Variability of Soil Moisture Capacity</td>
</tr>
<tr>
<td>C_max</td>
<td>Maximum Watershed Storage Capacity</td>
</tr>
</tbody>
</table>

Table 2. HyMod model state variables

<table>
<thead>
<tr>
<th>State Variable</th>
<th>Description</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>X_1</td>
<td>Quick Flow Tank 1 Storage</td>
<td>mm</td>
</tr>
<tr>
<td>X_2</td>
<td>Quick Flow Tank 2 Storage</td>
<td>mm</td>
</tr>
<tr>
<td>X_3</td>
<td>Quick Flow Tank 3 Storage</td>
<td>mm</td>
</tr>
<tr>
<td>X_4</td>
<td>Slow Flow Tank Storage</td>
<td>mm</td>
</tr>
<tr>
<td>S</td>
<td>Total Watershed Storage</td>
<td>mm</td>
</tr>
</tbody>
</table>
3.1.2 Sacramento Soil Moisture Accounting Model

The SAC-SMA model, which was first introduced by Burnash [6], is a conceptual water balance model used operationally at the NWSRFC. The model simulates water storage with two soil moisture zones: an upper and a lower zone. The upper zone accounts for short term storage of water in the soil, while the lower zone models the longer term groundwater storage. Water can move vertically from the upper zone to the lower zone, laterally out of the system depending on the state variables and the parameterization, or vertically out of the system through evapotranspiration. For the snow data assimilation portion of this study, the model is run in three different elevation bands, as is performed by the SNOW-17. The SAC-SMA is run with information from the SNOW-17 model and the potential evapotranspiration (PET), is linearly interpolated from the NWSRFC monthly PET values for the study basin, for each elevation band. The model calculates the water balance for the system and any excess is routed to the basin outlet using the unit hydrograph method. Model parameters are summarized in Table 3. Similarly the model state variables with their descriptions are summarized in Table 4.
Table 3. Parameters for the SAC-SMA model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Units</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Capacity Parameters</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UZTWM</td>
<td>Upper zone tension water maximum</td>
<td>mm</td>
<td>1.0-150</td>
</tr>
<tr>
<td>UZFWM</td>
<td>Upper zone free water maximum</td>
<td>mm</td>
<td>1.0-150</td>
</tr>
<tr>
<td>LZTWM</td>
<td>Lower zone tension water maximum</td>
<td>mm</td>
<td>1.0-500</td>
</tr>
<tr>
<td>LZFPM</td>
<td>Lower zone free primary maximum</td>
<td>mm</td>
<td>1.0-1000</td>
</tr>
<tr>
<td>LZFSM</td>
<td>Lower zone free secondary maximum</td>
<td>mm</td>
<td>1.0-1000</td>
</tr>
<tr>
<td>ADIMP</td>
<td>Additional impervious area</td>
<td>-</td>
<td>0.0-0.4</td>
</tr>
<tr>
<td><strong>Recession Parameters</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UZK</td>
<td>Upper zone depletion parameter</td>
<td>1/day</td>
<td>0.1-0.5</td>
</tr>
<tr>
<td>LZPK</td>
<td>Lower zone primary depletion parameter</td>
<td>1/day</td>
<td>0.0001-0.025</td>
</tr>
<tr>
<td>LZSK</td>
<td>Lower zone secondary depletion parameter</td>
<td>1/day</td>
<td>0.01-0.25</td>
</tr>
<tr>
<td><strong>Percolation and other</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZPERC</td>
<td>Maximum percolation rate</td>
<td>-</td>
<td>1.0-250</td>
</tr>
<tr>
<td>REXP</td>
<td>Percolation equation exponent</td>
<td>-</td>
<td>0.0-5.0</td>
</tr>
<tr>
<td>PCTIM</td>
<td>Impervious area of watershed</td>
<td>-</td>
<td>0.0-1.0</td>
</tr>
<tr>
<td>PFREE</td>
<td>Free water percolation from upper to</td>
<td>-</td>
<td>0.0-0.1</td>
</tr>
<tr>
<td></td>
<td>lower zone</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Routing Parameter</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kq</td>
<td>Nash-Cascade Routing Parameter</td>
<td>1/day</td>
<td>0.01-0.99</td>
</tr>
<tr>
<td><strong>Not Estimated</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RIVA</td>
<td>Riparian vegetated area</td>
<td>-</td>
<td>0.0</td>
</tr>
<tr>
<td>SIDE</td>
<td>Deep recharge to channel base flow</td>
<td>-</td>
<td>0.0</td>
</tr>
<tr>
<td>RSERV</td>
<td>Lower zone free water not transferable to</td>
<td>-</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>tension water</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4. SAC-SMA model state variables

<table>
<thead>
<tr>
<th>State Variable</th>
<th>Description</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>UZTWC</td>
<td>Upper zone temperature water content</td>
<td>mm</td>
</tr>
<tr>
<td>UZFWC</td>
<td>Upper zone free water content</td>
<td>mm</td>
</tr>
<tr>
<td>LZTWC</td>
<td>Lower zone tension water content</td>
<td>mm</td>
</tr>
<tr>
<td>LZFPC</td>
<td>Lower zone free primary water content</td>
<td>mm</td>
</tr>
<tr>
<td>LZFSSC</td>
<td>Lower zone free secondary water content</td>
<td>mm</td>
</tr>
<tr>
<td>ADIMC</td>
<td>Additional impervious area water content</td>
<td>mm</td>
</tr>
</tbody>
</table>
3.1.3 SNOW-17 Model

The hydrologic model used in this study is a distributed version of the NWS’s SNOW-17 model [1, 39]. This model is run at a spatial resolution of 1/8° and at a 6-hourly time step. The SNOW-17 model is a snow hydrology model, which is currently used operationally at the NWSRFC to model snow accumulation and ablation. The main processes simulated by SNOW-17 include: form of precipitation (snow or rain), accumulation of snow cover, energy exchange at the snow-air interface, internal states of snow cover (temperature, liquid/frozen water content, density, etc.), transmission of liquid water through the snowpack, and heat transfer at the soil-air interface.

The model is forced with precipitation and temperature data, and predictions are made for the SWE and snowmelt depth, at each time step, averaged over the modeling domain. Model parameters with their feasible ranges and the state variables are given in Table 5 and Table 6 respectively.

Table 5. Parameters in the SNOW-17 model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Units</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>PXTEMP</td>
<td>Temperature that separates rain/snow</td>
<td>°C</td>
<td>0.5-4</td>
</tr>
<tr>
<td>UADJ</td>
<td>Wind function for rain on snow events</td>
<td>mm/mb</td>
<td>0.02-0.2</td>
</tr>
<tr>
<td>MFMAX</td>
<td>Maximum melt factor without rain</td>
<td>mm/(˚C•6hr)</td>
<td>0.5-2</td>
</tr>
<tr>
<td>MFMIN</td>
<td>Minimum melt factor without rain</td>
<td>mm/(˚C•6hr)</td>
<td>0.05-0.6</td>
</tr>
</tbody>
</table>

Stationary Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Units</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCF</td>
<td>Factor for adjusting gauge catch errors</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SI</td>
<td>SWE when land is fully snow-covered</td>
<td>mm</td>
<td>-</td>
</tr>
<tr>
<td>Areal Depletion Curve</td>
<td>11 points on the snow depletion curve</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NMF</td>
<td>Maximum negative melt factor</td>
<td>mm/(˚C•6hr)</td>
<td>-</td>
</tr>
<tr>
<td>TIPM</td>
<td>Antecedent temperature index</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>MBASE</td>
<td>Base temperature for melt computations</td>
<td>°C</td>
<td>-</td>
</tr>
<tr>
<td>PLWHC</td>
<td>Percent liquid water holding capacity of snowpack</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DAYGM</td>
<td>Amount of melt which occurs daily at snow-soil interface</td>
<td>mm/day</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 6. State variables in the SNOW-17 Model

<table>
<thead>
<tr>
<th>State Variable</th>
<th>Description</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_i$</td>
<td>Frozen water equivalent in pack</td>
<td>mm</td>
</tr>
<tr>
<td>$D$</td>
<td>Heat Deficit</td>
<td>mm</td>
</tr>
<tr>
<td>ATI</td>
<td>Antecedent temperature index</td>
<td>°C</td>
</tr>
<tr>
<td>$W_{\text{max}}$</td>
<td>Maximum water equivalent that has existed during accumulation</td>
<td>mm</td>
</tr>
<tr>
<td>$W_{\text{ns}}$</td>
<td>Water equivalent of new snowfall on bare ground</td>
<td>mm</td>
</tr>
<tr>
<td>$A_{\text{ns}}$</td>
<td>Areal cover when new snow falls on partly bare ground</td>
<td>-</td>
</tr>
<tr>
<td>$W_{100}$</td>
<td>Amount of water equivalent at which areal cover drops below 100%</td>
<td>mm</td>
</tr>
<tr>
<td>$S$</td>
<td>Amount of lagged excess liquid water</td>
<td>mm</td>
</tr>
<tr>
<td>$A_{\text{adj}}$</td>
<td>Value computed for the depletion curve computation</td>
<td>mm</td>
</tr>
<tr>
<td>$E_t$</td>
<td>Hourly average lagged excess water for precipitation time interval</td>
<td>mm</td>
</tr>
<tr>
<td>$H$</td>
<td>Depth of total snow cover</td>
<td>cm</td>
</tr>
<tr>
<td>$T_s$</td>
<td>Average snow cover temperature</td>
<td>°C</td>
</tr>
<tr>
<td>$T_{a,h-M}$</td>
<td>Air temperature for precious time step</td>
<td>°C</td>
</tr>
</tbody>
</table>

3.2 Observation Model

In this study, an observational operator is necessary to transform snow properties into TB. This operator is a model referred to as a RTM. A RTM is a numerical program that translates several land surface variables into TB. TB of the land surface is sensitive to many variables including surface temperature, soil moisture, vegetation, SWE, and snow grain size. Many experiments have been performed to invert TB to SWE, but due to the non-unique relationship between TB and SWE [18], the RTM is strictly used as a forward model for this assimilation experiment. This provides a framework for estimating the possible snowpack states, at each time-step, constrained by the precipitation and temperature inputs to the system. The RTM used in this study is the Microwave Emission Model for Layered Snow (MEMLS) [47]. This model is designed to work in the frequency range of 5 to 100GHz for both polarizations and the correlation length...
(inferred from snow grain size according to [29]) range from 0.01 to 0.3mm. In this experiment, the TB is modeled at 36.5GHz frequency at vertical polarization. This frequency was chosen because of its sensitivity to snow parameters [43].

MEMLS assumes a snowpack with homogeneous horizontal layers of depth, density, correlation length, liquid water content and temperature. The model is based on multiple scattering radiative transfer and internal scattering is based on the six-flux theory, but simplified to upwelling and downwelling radiation. Scattering and absorption coefficients are derived from the frequency of the model and the snowpack temperature, correlation length and density [31]. Since MEMLS is only designed to calculate the microwave emission of the snowpack and does not take into account other land surface characteristics, updates are only made when the ground is snow covered. For comparison with AMSR-E data, the microwave radiation must be predicted through the top of the atmosphere, thus vegetation and atmospheric effects must be accounted for outside the MEMLS. Vegetation and atmospheric characteristics were modeled as implemented in both [37] and [17].

### 3.3 Snow Grain Size Model

In addition to a distributed version of the original SNOW-17 model, a snow grain size calculation algorithm [26] was implemented in the SNOW-17, as it is a necessary quantity for radiative transfer calculations. In this algorithm, snow grains growth is calculated from the diffusive vapor flux in the snowpack and the liquid water content. For dry snow conditions, equation 1 is used; for snow with liquid water content of less than 9%, equation 2 is used and equation 3 is performed for snow with liquid water content
greater than 9%. The snow grain size is then used to calculate correlation length according to [29]. Correlation length is necessary for running the RTM. This methodology is similar to the implementation used by [17].

\[
\frac{\partial d}{\partial t} = \frac{g_1}{d} (U_f) \tag{1}
\]

\[
\frac{\partial d}{\partial t} = \frac{g_2}{d} (\theta + 0.05) \tag{2}
\]

\[
\frac{\partial d}{\partial t} = \frac{g_2}{d} (0.14) \tag{3}
\]

where, \(d\) is the grain diameter in mm, \(t\) is time step, \(g_1\) is the dry snow coefficient \((5 \times 10^{-7} \text{m}^4/\text{kg})\), \(g_2\) is the wet snow coefficient \((4 \times 10^{-12} \text{m}^2/\text{s})\), \(U_f\) is the vapor flux and \(\theta\) is the liquid water content. Since the SNOW-17 model does not predict vapor fluxes within the snow, this calculation is simplified by assuming the maximum vapor flux \((10^{-6} \text{kg/m}^2/\text{s})\). Due to simplistic grain size calculations here, it is assumed that there is large uncertainty in the grain size prediction. Therefore a 15% error is introduced to the grain size prior to all radiative transfer calculations.
4. Data Assimilation and Calibration Techniques

4.1 Global Optimization vs. Monte Carlo Methods

Recently, much interest within the hydrologic, and the wider geophysical, modeling community has been placed on the potential of the use of Monte Carlo Methods for calibration of hydrologic models as opposed to the more widely used Global Optimization schemes [38]. Global optimization techniques attempt to find the exact single set of parameters which minimize a given objective function. These techniques often rely on approximations of the gradient of the given objective function, because this can dramatically increase the efficiency of convergence over random walk algorithms. Due to the extreme non-linearity and the presence of multiple local optima, gradient based approaches are subject to failure because of poor estimation of the gradient, a result of the non-linear nature of the parameter space, and converging in a local as opposed to the global optimum. Since gradient based methods are prone to failure, methods involving a random search of the parameter space (Monte Carlo optimization) gained popularity. This reduces errors due to the gradient misrepresentation and convergence in local optimum. Though Monte Carlo techniques reduce the error, the computational burden of random search algorithms is much higher than gradient methods and lacks stopping criteria. In order to improve the efficiency of Monte Carlo optimization, hybrid methods of Monte Carlo and gradient approximation were created. Genetic algorithms, simulated annealing, and, specifically in hydrologic modeling, Shuffle Complex Evolution-University of Arizona (SCE-UA) are popular hybrid techniques of gradient
based Monte Carlo techniques for optimization. This led to accurate and efficient methods for global optimization [38].

With the success of Monte Carlo techniques for optimization, interest in Monte Carlo methods, for the purpose of parameter distribution estimation, has become an area of active research. Monte Carlo methods, in a Bayesian framework, can be used to provide the distribution of parameters. Since there is uncertainty associated with parameters in hydrologic models, it is advantageous to characterize the uncertainty by finding the distribution of the parameters. An accurate distribution of parameters provides knowledge of the uncertainty with respect to the parameters, and allows for a better account of the total uncertainty [23]. Monte Carlo methods calculate a posterior probability density function of parameter distribution based on a prior model distribution and observed data. Through Bayes Law, the posterior is calculated as the product of the prior distribution and the likelihood. This can be performed in a sequential (filtering) or non-sequential (MCMC) framework. In filtering, a Monte Carlo algorithm estimates the likelihood at each model timestep in which an observation is available and the posterior distribution at that time-step is then estimated. This method requires assumptions about forcing, model and observation error are necessary for accurate prediction of the posterior distribution. This method also relies on a large number of samples to characterize the distribution accurately. In MCMC techniques, a model is run for set period of time and the posterior distribution is iteratively sampled, relaxing the need for error assumptions but also increasing the computational burden.
4.2 Ensemble Kalman Filter

4.2.1 Filter

The Ensemble Kalman Filter (EnKF) is commonly used for hydrologic modeling applications. In hydrologic prediction, the EnKF is often used to update states but in this study the ability of the EnKF to estimate parameters is examined. The EnKF is an ensemble version of the Kalman Filter, performed as a Monte Carlo simulation, to overcome issues with applying the traditional Kalman Filter, including the need to linearize the model and to estimate the prior and posterior model error covariances for updating the model states. Supplementary to the description below, Figure 3 provides a descriptive flowchart of the EnKF. As the model progresses forward in time, the prior distribution of states is

\[ x_t^{i-} = f(x_{t-1}^{i+}, u_t^i, \theta^{(l-1)+}) + \omega_t^i \]  \hspace{1cm} (4)

where \( f \) is the forward operator (hydrologic model), \( x_t^{i-} \) represents the model predicted (prior) states, \( x_{t-1}^{i+} \) represents the updated model states at the previous time-step, \( u_t^i \) represents the meteorological forcing data, \( \theta^{(l-1)+} \) represents the posterior model parameters from the previous time-step, \( \omega_t^i \) is the model error, \( i \) is the ensemble member and \( t \) is the time-step. Prior to update, an observational operator must be applied to the states to bring them into the observation space, as in equation 5.

\[ Q'_t = h(x_t^{i-}) + v_t^i \]  \hspace{1cm} (5)

where \( Q'_t \) is the predicted observation and \( v_t^i \) is the observation error. Then the states and parameters are updated with the Kalman update equation as follows,
where $K_x$ and $K_\theta$ are the Kalman Gains for states and parameters respectively. The Kalman Gain is calculated from:

$$K_t = P_t^{-}H^T (HP_t^{-}H^T + R_t)^{-1} = C_{xy} (C_{yy} + R_t)^{-1}$$  \hspace{1cm} (8)$$

Where $P_t^{-}H^T = C_{xy}$ is the covariance of the states/parameter ensemble with the predicted observation, $H P_t^{-}H^T = C_{yy}$ is the variance of the predicted observations, and $R_t$ is the variance of the observational error in (8). $H$ is the linearized observation operator ($H = \frac{\partial h}{\partial \mathbf{x}}$), which translates prior states from model space to measurement space. The model state error covariance $P_t^{-}$ can now be computed directly from the ensemble deviations ($e_t^{-}$):

$$P_t^{-} = \frac{1}{N_{ens}} e_t^{-} e_t^{-T}$$  \hspace{1cm} (9)$$
Figure 3. Flowchart of the Ensemble Kalman Filter
\[ e_t^i = x_t^i - \frac{1}{N_{\text{ens}}} \sum_{i=1}^{N_{\text{ens}}} x_t^i \]  \tag{10} 

\[ K = C_{xy}(C_{yy} + C_v)^{-1} \]  \tag{11} 

### 4.2.2 Smoother

Each smoothing technique used in this study is implemented in a batch framework as outlined in [15] (see Figure 5). The EnKS performs the batch update the same as the EnKF in this application but instead of a single observation, multiple observations and predictions are applied in a vector in the Kalman gain equation. The observation and prediction vector is shown in 12 and 13.

\[ Q_t = [Q_t \; Q_{t+1} \; \ldots \; Q_{t+10}]^T \]  \tag{12} 

\[ Q''_t = [Q''_t \; Q''_{t+1} \; \ldots \; Q''_{t+10}]^T \]  \tag{13} 

The states and parameters at each time-step are calculated according to Kalman Filter update equation. This calculation provides a state and parameter value for each time-step in the sliding window. While it is clear that the model, beginning at time \( t \), should be propagated forward with \( x_t \), the parameters, \( \theta_t \), must be set equal to the average over the batch window, \( \theta_{t:t+n} \), as it is assumed that parameters are constant in this window. With an updated ensemble of states and parameters, the batch window is moved forward one step and the process continues. This algorithm represented as a flowchart in Figure 4.
Figure 4. Flowchart of the Ensemble Kalman Smoother
4.3 Particle Filter

4.3.1 Filter

The PF, similar to the EnKF, sequentially calculates a posterior distribution of states and parameters. The advantage of the PF is, unlike the EnKF, it does not assume a Gaussian error structure, which allows the PF to more accurately predict the posterior distribution. This is accomplished by resampling sets of state variables and parameters, or “particles”, with higher posterior weights, as opposed to the linear model state updating of the EnKF. Though this method is more accurate, it is more computationally...
demanding than the EnKF [49]. The PF used in this study is the Sequential Importance Resampling (SIR) PF.

Based on the recursive Bayes Law (equation 14), the PF sequentially samples prior states and parameters to create an accurate posterior distribution, at each observation time-step.

\[
p(x_t | Y_t) = p(x_t | y_t, Y_{t-1}) = \frac{p(y_t | x_t)p(x_t | Y_{t-1})}{\int p(y_t | x_t)p(x_t | Y_{t-1})dx}
\]  

(14)

Equation 14 shows mathematically that a posterior conditional probability distribution of model predicted states and parameters, \(x\) given all previous observations \(Y_t\), can be computed sequentially in time. It should be noted that all \(y\) in equation 14 are observations as is signified in other equations by \(Q\). Similar to the MCMC description above, the probability of each particle in the PF is based on the likelihood equation. The sequential likelihood is calculated according to equation 15.

\[
L(Q_t | x_{t+1}^i, \theta_t^i) = \frac{1}{(2\pi)^{1/2}|R_{k+1}|^{1/2}} \exp \left( -\frac{1}{2R_{k+1}} [Q_t - Q_t']^2 \right)
\]  

(15)

The normalized likelihood, \(p(Q_t | x_{t+1}^i, \theta_t^i)\), can easily be calculated by:

\[
p(Q_t | x_{t+1}^i, \theta_t^i) = \frac{L(Q_t | x_{t+1}^i, \theta_t^i)}{\sum_{i=1}^{N_p} L(Q_t | x_{t+1}^i, \theta_t^i)} = p(Q_t - Q_t'R_{k+1})
\]  

(16)
This probability is necessary to transform the prior particle weights into the posterior via equation 15.

\[
    w_t^{l+} = \frac{w_t^{l-} \cdot p(Q_t | x_{t+1}^l, \theta_t^l)}{\sum_{i=1}^{N_p} w_t^{l-} \cdot p(Q_t | x_{t+1}^l, \theta_t^i)}
\]  

In the SIR PF used in this study, prior particle weights, \( w_t^{l-} \), are set equal to \( 1/N_p \) before moving on to the next time-step. This results in a posterior weight, \( w_t^{l+} \), equal to \( p(Q_t | x_{t+1}^l, \theta_t^l) \), the normalized likelihood. The SIR algorithm resamples those states which have a probability greater than the uniform distribution. For further description of the resampling technique used refer to [33]. The PF algorithm is shown in flowchart form in Figure 6.
Figure 6. Flowchart of Particle Filter
4.3.2 Smoother

The PF is extended to the PS by calculating a weight for each time-step in the sliding update window (see Figure 5), then resampling states and parameters at the beginning of the sliding window. This is performed by backward smoothing of particle weights. The formula for the backward smoothing of states is provided by [9, 27] and is presented in equation 18. This is then discretized in equation 19. This allows the weighting to be calculated based on multiple observations.

\[
p_{t|T} = \frac{p(x_{t+1}|x_t)p_{t|t}(x_t)}{\int p(x_{t+1}|x_t)p_{t|t}(x_t)dx_t} \, dx_{t+1} \tag{18}
\]

\[
p^{(n)}_{t|T} = \sum_{m=1}^{N} \frac{p^{(m)}(x_{t+1}^{(m)}|x_t^{(n)})p_{t|t}^{(n)}}{\sum_{i=1}^{N} p^{(m)}(x_{t+1}^{(m)}|x_t^{(i)})p_{t|t}^{(i)}} \tag{19}
\]

After applying equation 19 to calculate the weight at time \( t \), the particles are resampled using the SIR algorithm based on the weights at time \( t \). After resampling, the batch window is moved forward one step and the process repeats. The PS algorithm is shown in flowchart form in Figure 7.
Figure 7. Flowchart of the Particle Smoother algorithm
4.4 Global Optimization via SCE-UA

In this study, Global Optimization is implemented using the SCE-UA algorithm [12-14]. This is a combination of the simplex gradient estimation technique and a Monte Carlo sampling to avoid convergence to a local optimum. This is performed by evolving multiple simplexes, which are referred to as a complex, and shuffling the estimates from separate simplexes within a complex after a certain number iterations. By shuffling the estimates within a complex, the possibility of convergence to a local optimum is reduced while increasing the speed of convergence. This method has been proven to be a robust method and therefore is used to estimate the optimal parameters.

4.5 Markov Chain Monte Carlo Methods

MCMC techniques have become a popular method of hydrologic model calibration due to success in recent studies [3, 28, 30, 45, 46]. These studies showed that MCMC techniques can effectively calibrate hydrologic models and are becoming increasingly efficient. MCMC techniques are general methods for sampling a posterior distribution. Through various sampling techniques, a proposal distribution $\theta_p$ is created based on the prior distribution $\theta_{old}$. The probability $\pi$ of a certain parameter set is calculated according to the likelihood equation (20) assuming a non-informative prior distribution. Equation 20 is provided by [3] assuming homoscedastic and uncorrelated error.

$$
\pi(Q|\theta) = (2\pi\sigma^2)^{-t/2} \prod_t^t \exp \left\{ -\frac{[Q_t - Q^*_t]^2}{2\sigma^2} \right\}
$$  (20)
In this equation, $\sigma^2$ is the variance, $Q_t$ and $Q''_t$ represent the time-series of observed and simulated data and $t$ is the number of observations. Equation 20 can be simplified by integrating out the variance term giving equation 21 as described in [4, 45, 46].

$$\pi(Q|\theta) = \left[ \sum_{t=1}^{t} (Q_t - Q''_t)^2 \right]^{-\frac{t}{2}}$$  \hspace{1cm} (21)

The probability of $\theta_{old}$ is compared to $\theta_p$ in equation 22 to determine the $\alpha$ used in the acceptance/rejection rule.

$$\alpha = \min \left[ \frac{\pi(\theta_p|Q)}{\pi(\theta_{old}|Q)}, 1 \right]$$  \hspace{1cm} (22)

At each Markov Chain iteration, $\theta_p$ is accepted/rejected if $\alpha$ is greater/less than a random draw from the uniform distribution. After a sufficient number of chain iterations the model will converge to a stationary distribution. The assumption in this method, based on information presented in [32], is that this stationary distribution is the true posterior distribution. In order to determine if the model has converged to a stationary distribution, the variance of each Markov Chain is compared to variance of all Markov Chains as described by [21]. This statistic is assumed to show convergence when the value for all parameters is less than 1.2. This is because a value of 1, the perfect value, is very difficult and time consuming to achieve [21].

Currently, most hydrologic implementations of MCMC techniques are variants of the Metropolis Algorithm. The original Metropolis Algorithm is performed as a random walk to converge to the solution, making convergence of the algorithm very slow in complicated parameter spaces. Similar to how Monte Carlo techniques for optimization were improved through the addition of gradient approximations, the Metropolis
Algorithm has been improved with the use of adaptive search techniques. The Adaptive Metropolis, proposed by [23], has utilized the variance of the chains, at each iteration, to adapt the search area for each parameter based on new knowledge at each iteration. With the success of the Adaptive Metropolis, the Delayed Rejection Adaptive Metropolis [24] was presented. In addition to variance estimation of parameter predictions, genetic algorithms, such as Differential Evolution, have also been shown to be effective in minimizing the search area [5]. Separately these algorithms have been shown to be effective but in recent studies [46]. This is referred to as Differential Evolution Adaptive Metropolis (DREAM). The two techniques used in DREAM have been shown to be complimentary in some situations are capable of minimizing the number of chains necessary for convergence. Due to the efficiency and availability of the DREAM program, this algorithm is used to calibrate the hydrologic models in the comparison of global optimization, MCMC and filtering techniques.
5. Experimental Setup

This study is made up of two separate experiments. The first experiment examines the usefulness of remotely sensed microwave TB data assimilation on NWS operational models for the purpose of enhancing streamflow. The second study compares the effectiveness of Global Optimization, MCMC and data assimilation techniques for calibration of hydrologic models.

5.1 Snow Data Assimilation Experiment

In the snow data assimilation portion of this study, the effects of microwave TB assimilation of prediction of SWE from the NWS SNOW-17 model and the streamflow from a the coupled SNOW-17 and SAC-SMA model are analyzed. The SWE estimation is compared in a synthetic experiment. A synthetic experiment was used because of the lack of an accurate validation dataset. This compares the effectiveness of the EnKF and the PF in estimation of SWE, using state only and joint state-parameter estimation. In the synthetic experiment the effectiveness of both methods is compared.

In order to determine the effectiveness of TB assimilation in the SNOW-17 model on streamflow prediction in the SAC-SMA model, six different experiments were run. Experiment 1 is a control run of both models to simulate the prediction in the current NWSRFC system. Experiment 2 is an ensemble run of both models, with the same input noise as the data assimilation experiments, without the implementation of data assimilation. Experiments 3-6 includes TB assimilation in the SNOW-17 model including (3) PF state estimation (PF-SU), (4) PF state-parameter estimation (PF-PU), (5) EnKF state estimation (EnKF-SU) and (6) EnKF state-parameter estimation (EnKF-PU). In
order to ensure consistency in comparing the assimilation techniques, all experiments use 1000 ensemble members, except for the PF-PU experiment which was raised to 3000 ensemble members to reduce the possibility of sample impoverishment. In addition to providing consistency between the methods, using the same ensemble size makes the computational cost of each method nearly exactly the same and thus computational cost will not be discussed in this study.

5.2 Comparison of Calibration Techniques

Global Optimization, MCMC and data assimilation techniques are compared in this study to determine the relative benefits of each technique, in terms of calibration. The accuracy of all techniques will be compared with the Root Mean Square Error (RMSE) in a validation period, based on the average value at each prediction timestep for the ensemble assimilation methods. Each method is used to calibrate the HyMod and SAC-SMA models over a 10 year period. The models are then validated for the following 3 years.

5.3 Performance Metrics

In the TB assimilation experiment, in order to determine the improvement of the assimilation experiments in relation to the control model run, the Benchmark Efficiency (BE) for streamflow prediction from each assimilation experiment is calculated. The BE is a measure of the improvement (if positive) over a simulation run. This is analogous to the Nash-Sutcliffe efficiency as the error is scaled by the control model run.
BE = 1 - \frac{\sum_{t=1}^{T}(AR_t - O_t)^2}{\sum_{t=1}^{T}(CR_t - O_t)^2} \quad (23)

where AR is the assimilation run, O is the observation, CR is the control run and T is the total number of timesteps.

Ranked Probability Score is another widely used measure for evaluating the quality of probabilistic predictions [48]. By definition Rank Probability Score is the sum of squared error of the cumulative probability forecasts averaged over multiple events. In streamflow prediction, the probability forecast is usually expressed using a non-exceedance probability forecast within pre-specified categories (i.e., 5%, 10%, 25%, 50%, 75%, 90%, 95% and 99% non-exceedance). The observed value for a given threshold (forecast category) takes on the value of 1 if the observed flow value is less than the threshold for that category. Otherwise, the observed value is 0. The discrete expression of Rank Probability Score is given as:

\[ RPS_t = \sum_{i=1}^{I} [F^t_i - O^t_i]^2 \quad (24) \]

Where \( F^t_i \) is the forecast probability at time \( t \) given by \( P(\text{forecast}_i < \text{thresh}_i) \) and \( O^t_i \) is the observed probability given by \( P(\text{observed} < \text{thresh}_i) \) where \( i \) is the probability category. The Rank Probability Skill Score (RPSS) is also computed as the percentage improvement over a reference score (e.g. climatology)[48]:

\[ RPSS = \left(1 - \frac{RPS}{RPS_{\text{climatology}}} \right) \times 100 = \left(1 - \frac{RPS_{\text{climatology}}}{RPS_{\text{climatology}}} \right) \times 100 \quad (25) \]

Where \( RPS_{\text{climatology}} \) is the rank probability score for the observation. A positive value shows the percentage of improvement over the reference RPS.
6. Results and Discussion

6.1 Snow Data Assimilation

6.1.1 Synthetic Experiment

The TB assimilation experiments were performed using the EnKF and PF for both state and joint state-parameter estimation. Each experiment assimilates TB produced by running the model with unperturbed forcing data and adding homoscedastic noise with a variance of 1°C. The data is then assimilated into the coupled SNOW-17/MEMLS model for each scenario. This implementation is similar to the synthetic study outlined in [35]. It is important to note that in a synthetic experiment, state estimation is expected to perform better than joint state-parameter estimation because state only estimation is run with predefined parameters. The results from each experiment, averaged through all model cells, are summarized in Figure 8. Figure 8 shows that in state estimation, the PF-SU experiment clearly outperformed its respective EnKF-SU experiment, while in state-parameter estimation, the PF-PU performed slightly better than the EnKF-PU. A lower RMSE and higher RPSS for the PF simulations in comparison to the EnKF simulations suggests that the PF is providing more accurate expected values of SWE, as well as a more accurate uncertainty representation through the SWE ensemble. Figure 9 and 10 are time series plots of cells in the middle and upper elevation band of the SNOW-17 model. Each plot shows all four assimilation techniques. As was suggested in the RMSE values, the expected value lines of the PF experiments appear to match the true value line more accurately than the expected value line from the EnKF experiments. Further, the predictive bounds from the PF experiments, in comparison to the EnKF experiments,
more effectively bracket the true value, which indicates a more accurate representation of uncertainty. This is supported by Figure 8b where the RPSS of both PF experiments was higher than the EnKF experiments. The superior performance of the PF in this experiment is attributed to the method of resampling as opposed to linear updating of states through the EnKF. By updating the SWE states through the EnKF, errors arise through a non-linear relationship between TB emission from a snowpack and the state of the snowpack.

**Figure 8.** (a)RMSE and (b) RPSS of SWE from September 1st 2002 through September 30th 2005 for all grid cells
Figure 9. The SWE posterior 95% predictive bounds (dotted), expected value (solid black), and the true synthetic value (solid gray) in the SNOW-17 model using TB assimilation for the middle elevation band for the filtering procedures: (a) PF-SU (b) EnKF-SU, (c) PF-PU, and (d) EnKF-PU from September 1st, 2002 through June 30th, 2005.
In order to better understand the progression of errors through the snow accumulation and ablation seasons, figures 9 and 10 are presented. It is important to note the seasonal differences (winter and spring) in all four scenarios. The tendency is for the uncertainty bands to grow much wider during the spring for the joint state-parameter estimation experiments. This is important to note because problems arise in prediction of TB emission from snowpack during the spring due to the presence of liquid water in the
This leads to errors in the prediction of microwave emission and ultimately leads the uncertainty in the prediction of all variables in a data assimilation framework to rise. Because the MEMLS model has more variability in TB prediction from wet snowpack, the spread of the ensembles grows in the spring months. The state only estimation experiments are less prone to this wider uncertainty in spring because there is no error added with respect to parameters. Though the errors grow in the melt season, causing a higher overall RMSE for the parameter estimation experiments, the overall RPSS for the parameter estimation experiments is higher than for the EnKF-SU experiment. While the uncertainty is growing, leading to errors in the expected value of parameter estimation experiments, the ensemble of predictions still reflects the uncertainty accurately. This is surprising as both state estimation experiments have predefined parameters for each model run.

In order to compare the effectiveness of the assimilation methods under dry snow conditions, Figure 11 is presented. This figure compares the RMSE and RPSS during the winter months (January, February, and March) of all the experiments. Note that in this figure, for the EnKF-SU experiment, two outlier cells are removed to remove to prevent skewed averages. This figure shows that during the winter months, the EnKF-SU, PF-PU and EnKF-PU experiments all provide similar results, in terms of RMSE. This suggests that the PF-PU with added parameter uncertainty is able to predict SWE as accurately as the EnKF with predefined parameters. In addition, as shown in Figure 8, the RPSS for the PF-PU experiment provides a more accurate probabilistic prediction than both of the EnKF experiments. While similar results are found based on the expected value of the
PF-PU, EnKF-SU and EnKF-PU, the PF-PU is more effective based on predicting an accurate ensemble. After removing the effects of the spring snowmelt, the accuracy of the PF-PU, in terms of RMSE, is improved over the EnKF-SU without diminishing the probabilistic forecast skill. Since the EnKF-SU is run with the correct predefined parameters, this demonstrates that the PF is able to identify parameter and state values from TB more effectively than the EnKF can identify state values alone. When the PF is used for state only estimation, the results are even better.

![Graph](image)

**Figure 11.** (a) RMSE and (b) RPSS of SWE for the winter months (January through March) during the modeling time period for all grid cells for years 2003, 2004, 2005. Note that in order to prevent a skewed mean error EnKF-SU experiment, two outlier cells were removed.

An analysis of the spatial SWE estimation error, over the winter months, is provided in Figure 12. This provides a comparison of the assimilation of TB over different elevation bands, and therefore shows the error relating to different snowpack depths. From Figure 12, it appears that the PF-SU is more consistent in SWE estimation,
through the three elevation bands, than the other experiments. Also, values for the lower elevation band, in all experiments, were the most accurate, which was expected as this elevation band will have the shallowest snowpack and the TB is therefore more sensitive to its depth. It appears that through the winter months, the errors between the middle and upper elevation bands are not significantly different, within each separate experiment, except for the EnKF-SU. In the EnKF-SU experiment, the errors appear to be inconsistent spatially. Since consistent error was placed on the synthetic data, this suggests the EnKF is less stable than the PF in estimating SWE based on TB. Further evidence of this instability is shown in the third year of figures 9 and 10. In these figures, the SWE in the early spring are severely underestimated. This is due to errors accumulating over the winter season, especially as the SWE grows and TB becomes less sensitive to depth and more sensitive to the surface characteristics. Figure 13 shows spatial RPSS of the winter months to supplement the spatial analysis. This figure also shows that the EnKF-SU gives inconsistent results in a TB assimilation experiment. The other experiments show consistent results between the middle and upper elevation band, but the lower elevation band shows the lowest RPSS. This is due to a short snow accumulation and ablation season with a shallow snowpack. Since the snow season is shorter and the snowpack remains shallow, the ensemble of SWE values remains very thin. From this plot, both the PF experiments appear to have performed better than the respective EnKF experiments.
Figure 12. Spatial SWE RMSE (mm) for the winter months (January through March) during the modeling time period for all grid cells for years 2003, 2004, 2005. Note that the lower left cell in each experiment is outside the modeling domain and the elevation bands of all other cells are labeled Up (upper), Mid (middle) and Low (lower)
Comparison of the snow estimates from the EnKF and PF show that the PF is a more effective method than the EnKF in the radiance assimilation, for the SNOW-17 model. The better performance of the PF in comparison to the EnKF is attributed to the assumptions associated with the EnKF updating technique. While the PF resamples state variables that produce a more accurate TB estimate, the EnKF linearly updates state variables based on the covariance of the observations and state variables with the
Gaussian assumption of errors in these variables. In addition to an incorrect Gaussian error assumption, the EnKF adjusts state variable values, changing the relationship between snowpack states. Because TB becomes more sensitive to snowpack variables other than depth and SWE in deep snowpacks, and the model is predicting these snowpack variables as they relate to SWE and depth, keeping the relationship between all snowpack variables through resampling creates a more accurate SWE estimate than can be created by adjustment of each variable based on its covariance with TB. Thus, resampling the more probable sets of state variables is a more effective means of data assimilation than updating individual states via the EnKF, with respect to TB. This is consistent with findings in [10], where it was found that the EnKF is not an effective method for assimilating TB in soil moisture experiments due to the assumption of Gaussian error structure in conjunction with skewed model ensembles and non-Gaussian errors.

Increased uncertainty in the model prediction in the melt season (spring) shows the high variability of the radiative transfer calculations when high water contents are present. Since the accuracy of TB prediction is lower, the uncertainty in model updating/resampling is higher. This puts more weight on the model and less on the observation, as more ensemble members are resampled. Since the model is forced with uncertain input, the uncertainty in prediction is increased significantly. This is consistent with findings in previous studies [2] where radiative transfer calculations are shown to be less accurate in wet snow conditions. This creates a scenario where the assimilation becomes less reliable in terms of predicting an expected value. For this reason, the
predictive bounds grow much wider showing the uncertainty in prediction. This is important in the context of streamflow forecasting as the ensemble of snowmelt from the SNOW17 model should contain the true value leading to accurate input into the SAC-SMA model, and consequently more accurate streamflow prediction.

6.1.2 Real Experiment

Streamflow prediction from the coupled SNOW-17/SAC-SMA was performed for the six experiments outlined above. In order to determine if TB assimilation improves streamflow prediction, all experiments are compared against the control run in terms of BE. This shows the accuracy of the expected values of each method and the relative improvements over the control model run. In addition to comparing the accuracy of expected values, the accuracy of the distribution, in terms of RPSS, is compared for all experiments other than the control run. By comparing the RPSS of the ensemble model run to the assimilation model runs, the merit of each assimilation technique in producing an accurate probabilistic prediction is shown.

Figure 14 compares the benchmark efficiency of each of the six experiments. This figure shows that all data assimilation experiments improved the final streamflow accuracy (positive benchmark efficiency). The most accurate expected values from all the data assimilation techniques were provided by the PF-PU followed by the EnKF-PU, the PF-SU and the EnKF-SU. Both state-parameter estimation experiments performed considerably better than the other experiments. Since the addition of parameter estimation produced more accurate results than the state only estimation experiments, it is likely that the NWS parameters are not perfectly calibrated. Though this makes the calibration of the
NWS parameter suspect, assimilation with these set parameters still had a positive BE. This suggests that, in terms of accuracy of expected value of streamflow, radiance data assimilation has the potential to improve streamflow estimates with the implementation of either the EnKF or the PF, especially when parameter estimation is included.

**Figure 14.** Benchmark Efficiency as compared with the control run for all TB assimilation experiments

In order to determine the effectiveness of each method from a probabilistic forecast standpoint, Figure 15 shows the RPSS of the ensemble model run and all four assimilation experiments. In this figure, it is interesting to note the low RPSS of the EnKF-PU. While the BE from the EnKF-PU was the second best of all assimilation runs, it produced the lowest RPSS and was the only assimilation run that produced a lower RPSS than the ensemble model run. A lower RPSS than the ensemble model run indicates that the skill of the prediction was actually lowered by the EnKF-PU. This figure also shows that the two PF experiments produced the highest RPSS values. While
the PF-SU appears, in terms of RPSS, to give the best probabilistic forecast, the PF-PU had a considerably higher BE value than the PF-SU experiment. Thus, while the skill of the forecast for the PF-SU was better, it was not capable of providing better expected value than the PF-PU experiment. Since the PF experiments had a similar probabilistic forecast, but the PF-PU experiment had significantly higher BE, the PF-PU is considered to have performed the best in terms of streamflow prediction.

Figure 15. RPSS for all ensemble streamflow experiments. Note that ENS is the ensemble run with no assimilation.

Figure 16 shows the predicted streamflow, from both joint state-parameter estimation data assimilation experiments, as compared to the observed streamflow. Since the parameters are not perfect, only the parameter estimation experiments are shown. This figure illustrates the effectiveness of the two methods in creating predictive bounds. Overall the EnKF has produced wider predictive bounds. The wider predictive bounds suggest that the EnKF has a higher uncertainty in the snowmelt prediction. A higher uncertainty is predicted by the EnKF, as compared to the PF, because the EnKF was
unable to narrow the distribution of parameters as effectively as the PF. The PF experiment was able to select parameters in a smaller range than the EnKF without sacrificing accuracy. This creates a more accurate representation of the uncertainty in streamflow. Since the PF is able to more accurately predict the uncertainty in the system, it is preferable for operational prediction.

Figure 16. Streamflow prediction in the East river basin using the coupled SNOW-17/SAC-SMA model for the period of September 1st, 2002 through September 30th, 2005. (a) SNOW-17 state-parameter estimation via the PF, (b) SNOW-17 state-parameter estimation via the EnKF. The shaded region is the 95% predictive bound, the dotted line is the observed flow and the solid line is the expected value.

Overall, results from the streamflow experiment have shown that the PF-PU provides the best prediction of ensemble snow melt data from the SNOW-17 model. This
method produces a positive BE with respect to the control run, and a higher RPSS than the ensemble run, indicating an increase in prediction accuracy over the NWSRFC’s current system. Though both the PF and EnKF improved the streamflow accuracy over the control run, the PF-PU produces the most accurate expected value and the best overall prediction.

6.2 Streamflow Data Assimilation

In the streamflow assimilation experiments, an estimation of input, observation and model error must be used to update the model states and parameters. To keep all experiments consistent, the same value was used in each experiment for the input and observation error, while separate model errors were used for the HyMod and SAC-SMA Models. The errors are summarized in Table 5. Because the SAC-SMA model has the potential to produce a more accurate prediction, it requires less error model error for data assimilation.

Table 7. Data Assimilation Error Assumptions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Error Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input Error</td>
<td>20%</td>
</tr>
<tr>
<td>Observation Error</td>
<td>15%</td>
</tr>
<tr>
<td>HyMod Model Error</td>
<td>30%</td>
</tr>
<tr>
<td>SAC-SMA Model Error</td>
<td>25%</td>
</tr>
</tbody>
</table>

6.2.1 HyMod Model Results

The calibration of the HyMod model is the first experiment in this study. By using the simple (5 parameter) HyMod model, the basic differences in the calibration techniques are explored. The RMSE from the validation period of all calibration techniques is shown in Figure 17. In this figure, the most accurate calibration was
produced by the PS followed by the PF, SCE-UA, EnKF, EnKS and MCMC. Both particle filtering techniques performed slightly better than global optimization while Kalman filtering techniques performed slightly worse. While global optimization and the ensemble of parameters from filtering techniques performed similarly, the ensemble prediction from MCMC techniques performed significantly worse. This is likely due to the small number of chains used in this experiment (50). While the accuracy of the MCMC techniques theoretically will achieve the perfect posterior parameters, due to the computational demand of the MCMC techniques, it is not feasible to allow time for convergence with a large enough ensemble size to achieve this distribution. In comparing the filtering and SCE-UA techniques, it is interesting to note that the ensemble mean from the particle method parameters are capable of achieving the most accurate prediction. Though the SCE-UA found the best (or very close to) possible parameters for a single model run, an ensemble run achieves an accuracy that has surpassed it. This suggests that either the filtering techniques are capable of locating better parameters than global optimization or the mean of the ensemble predictions from particle filtering/smoothing are more accurate than a single parameter set. This has important implications for hydrologic modeling because it appears that a simple hydrologic model can be more accurate, simply by calibrating the model over a distribution, as opposed to the more traditional single parameter set optimization. Figure 18 is presented to compare the RMSE of the expected value of parameters from filtering calibration, with the RMSE from the mean of ensemble prediction values. This shows that for all filtering techniques, the ensemble mean outperformed the prediction by a single (expected value) parameter
set. This is because the filtering techniques are creating a distribution of parameters based on the relationship between the parameter and the output. Since the model is non-linear, the best single parameter set from filtering is not capable of predicting the streamflow as accurately as the ensemble of parameters.

**Figure 17.** The RMSE of each assimilation method, in the HyMod model

**Figure 18.** Comparison of ensemble prediction and single parameter values prediction
The calibration of each of the five HyMod parameters is presented in Figure 19. This figure is a box and whisker plot of the distribution of each of the HyMod model parameters, produced by each of the calibration techniques. Since the MCMC technique performed far worse than the other methods, it is important to understand the cause of the poor calibration. First note that for alpha, beta, Cmax and Rq the method converges to a very small distribution, while the Rs value remains a very wide distribution. This is because the Rs parameter is less identifiable than the other parameters, which is a result of the small contribution it has in producing streamflow. Since the parameter is the least identifiable, it cannot correctly converge until after the other four parameters converge. This is also apparent, though not as severe, in the other calibration techniques. Since the parameters produced by MCMC converge as a result of the scaled variance (adaptive metropolis) and Differential Evolution, sampling of this parameter is much less dependent on other parameters. Particle filtering is more dependent on parameter relationship because parameters are resampled as opposed to being either accepted or rejected. The EnKF adjusts parameters based on the covariance of each parameter and the output, which further increases the speed of convergence for less identifiable parameters, but at the expense of accuracy of the Beta parameter. Thus particle filtering/smoothing proves to be the more effective form of parameter estimation in this model.
Figure 19. HyMod calibration parameters
Figure 20. Calibration of the Alpha parameter in the HyMod model. The dashed lines show the 95% predictive bounds and the solid line shows the median value.
Figure 21. Calibration of the Beta parameter in the HyMod model. The dashed lines show the 95% predictive bounds and the solid line shows the median value.
Figure 22. Calibration of the Cmax parameter in the HyMod model. The dashed lines show the 95% predictive bounds and the solid line shows the median value.
Figure 23. Calibration of the Rs parameter in the HyMod model. The dashed lines show the 95% predictive bounds and the solid line shows the median value.
Figure 24. Calibration of the Rq parameter in the HyMod model. The dashed lines show the 95% predictive bounds and the solid line shows the median value.
In order to understand the differences in calibration of the EnKF and PF, the reader is directed to the calibration of the Beta value in Figure 19. This figure shows that the EnKF and EnKS converge to a similar range, which in comparison to the PF, PS and SCE-UA is underestimated. In order to test the importance of the sensitivity of this parameter in the EnKF and EnKS, the validation run for the parameters of both methods was performed again, but with the beta value set to the SCE-UA predicted value. The results are shown in Figure 25. This figure shows that this parameter significantly contributed to the errors in EnKF and EnKS parameter estimation. It should also be noted that from Figure 21 that this value converges very quickly in the EnKF and EnKS as opposed to particle based methods. The cause of convergence to this value, which is the main factor in its poor performance in comparison to the PF and PS, is due to the relationship between the parameter and model prediction. This value, though very important to prediction from the model, determines the complicated relationship (exponential) between moisture storage and evapotranspiration. Since the EnKF and EnKS rely on a Guassian assumption, and this parameter is exponential, it is incorrectly identified. Because the PF and PS do not rely on a Guassian relationship, they are more capable of identifying this parameter accurately. Though the Guassian assumption was detrimental to the performance of the EnKF and EnKS in the Beta value, the benefit of the EnKF and EnKS is shown in the Rs value. From figure 19, it is clear that the EnKF and EnKS converge to about the same value as the SCE-UA. While the PF and PS had difficulty converging to this value, the EnKF and EnKS appear to converge efficiently and accurately. In addition the EnKS converged very quickly to this value, as is shown in
Figure 23. Note that this parameter has a linear relationship with model prediction. Due to the simplicity of this parameter, the Gaussian assumption is beneficial. This is also true of the Rq value, which is also a parameter for linear reservoirs. While the EnKF and EnKS may incorrectly identify parameters with a highly complicated relationship with model prediction, they are accurate in identifying much simpler parameters. This is addressed further in the analysis of the SAC-SMA model calibration.

![Bar chart showing RMSE (cms) for EnKF and EnKS with Filter Estimated Beta and SCE-UA Estimated Beta]

**Figure 25.** Validation of EnKF and EnKS calibration with Beta value set to SCE-UA estimate.

Another important difference in the PF and the EnKF can be seen in the calibration of Rq, which is the most identifiable parameter. While the EnKF and EnKS were successful in calibrating this parameter, both techniques took nearly the entire 10 year period to reach the correct value. This is a result of a severe miscalculation of the Rq value in the first year. Large miscalculations are made because the outflow is very sensitive to this parameter, as well as the states. This leaves the calibration subject to
errors during early timesteps, especially at low flows. The PF and PS appear to work a little slower but correctly estimate the error, with respect to the parameter at each time step.

### 6.2.2 Sacramento Model Results

A similar analysis of parameter estimation in the SAC-SMA model was performed. The RMSE for all techniques, over the validation period, is presented in Figure 26. From this figure, it is clear that, as was found in the HyMod model experiment, the MCMC technique performed the worst. From the experiments provided, it seems clear that MCMC techniques, in this implementation, are not capable of performing as accurately as global optimization or filtering. The best calibration of the SAC-SMA model, in this experiment, is provided by the EnKF. Although the PF and PS outperformed the EnKF in calibrating the HyMod model, in the SAC-SMA model, the EnKF performed better. In addition, the EnKF was the only method to locate better parameters than SCE-UA. The PF and PS performed similarly to SCE-UA and the EnKF performed worse than all other filtering methods. In order to compare the ensemble parameters with the expected value of parameters, Figure 27 is presented.
Figure 26. RMSE of the validation period of the SAC-SMA model

Figure 27 shows that, similar to the HyMod model calibration, the EnKF, EnKS and PS performed better when run as an ensemble prediction experiment than as a single value experiment. Contrary, to the HyMod model calibration, the PF performed best when the expected value of parameters, as opposed to the distribution of parameters, was used. Due to the number of parameters and states in the SAC-SMA model, as compared to the HyMod model, and the differing performance of the PF in comparison to the EnKF, it is likely that the PF suffered from sample impoverishment. This may also be the cause of the better performance of the single parameter set as opposed to the ensemble of parameters. Because the single parameter set is performing better than the ensemble prediction in the PF, it is likely that the final parameter distribution does not accurately quantify the uncertainty in the prediction. More evidence of sample impoverishment is provided by Figures 42, 43 and 44.
Figure 27. Comparison of ensemble prediction and single parameter values prediction

From these figures, it should be noted that the PF converged to a much smaller parameter range than all other assimilation techniques. The speed of convergence is also shown in Figures 28-. This behavior indicates that the PF was unable to characterize the uncertainty with an ensemble size of 1000 ensemble members. While this was a large enough ensemble size for the HyMod model, the complexity of the SAC-SMA model requires a larger ensemble size to correctly identify the parameters. In addition, the PS appears to be subject the same problem.
Figure 28. Calibration of the UZTWM parameter in the SAC-SMA model. The dashed lines show the 95% predictive bounds and the solid line shows the median value.
Figure 29. Calibration of the UZFWM parameter in the SAC-SMA model. The dashed lines show the 95% predictive bounds and the solid line shows the median value.
Figure 30. Calibration of the UZK parameter in the SAC-SMA model. The dashed lines show the 95% predictive bounds and the solid line shows the median value.
Figure 31. Calibration of the PCTIM parameter in the SAC-SMA model. The dashed lines show the 95% predictive bounds and the solid line shows the median value.
Figure 32. Calibration of the ADIMP parameter in the SAC-SMA model. The dashed lines show the 95% predictive bounds and the solid line shows the median value.
Figure 33. Calibration of the ZPERC parameter in the SAC-SMA model. The dashed lines show the 95% predictive bounds and the solid line shows the median value.
Figure 34. Calibration of the REXP parameter in the SAC-SMA model. The dashed lines show the 95% predictive bounds and the solid line shows the median value.
Figure 35. Calibration of the LZTWM parameter in the SAC-SMA model. The dashed lines show the 95% predictive bounds and the solid line shows the median value.
Figure 36. Calibration of the LZFSM parameter in the SAC-SMA model. The dashed lines show the 95% predictive bounds and the solid line shows the median value.
Figure 37. Calibration of the LZFPM parameter in the SAC-SMA model. The dashed lines show the 95% predictive bounds and the solid line shows the median value.
Figure 38. Calibration of the LZSK parameter in the SAC-SMA model. The dashed lines show the 95% predictive bounds and the solid line shows the median value.
Figure 39. Calibration of the LZPK parameter in the SAC-SMA model. The dashed lines show the 95% predictive bounds and the solid line shows the median value.
Figure 40. Calibration of the PFREE parameter in the SAC-SMA model. The dashed lines show the 95% predictive bounds and the solid line shows the median value.
Figure 41. Calibration of the $K_q$ parameter in the SAC-SMA model. The dashed lines show the 95% predictive bounds and the solid line shows the median value.
In figures 42-24, it should be noted that the only parameter in which the MCMC algorithm converged to a wider range than all other methods is the LZSK. This is the lower zone reservoir parameter, which determines the persistent supplementary flow from groundwater. Similar to the groundwater parameter in the HyMod model, this is the least identifiable parameter because streamflow additions from this parameter are very small and therefore may be masked by the effects of other parameters. Further, the Kq value, the parameter for Nash Cascade reservoir routing, was the most identifiable. Since all water from the SAC-SMA model is affected by the reservoir routing parameter, the outflow is very sensitive to this parameter. Also note the speed of convergence of this parameter in Figure 41 and the relative convergence the other parameters in the PF. While the EnKF, due to independent estimation of each parameter, is capable of minimizing the Kq distribution without speeding convergence of other values, all parameters in the PF converge in conjunction with the Kq value. This is because only a few particles are resampled at this timestep, due sensitivity of this parameter in relation to the range, leading to a poor distribution of other parameters (sample impoverishment). SCE-UA is also incapable of estimating the values as accurately as the EnKF because it converged to a local optimum, which is a result of the streamflow being so sensitive to Kq value.
Figure 42. SAC-SMA parameter distributions
Figure 43. SAC-SMA parameter distributions
Figure 44. SAC-SMA parameter distributions
In order to further explain the differences in the PF and EnKF in calibrating the SAC-SMA model, both methods were used to calibrate the SAC-SMA model with two different scenarios. The first scenario was to raise the ensemble size to 5000 members and the second was to make the range of possible Kq values smaller (0.2 to 0.25). Both of these scenarios reduce the probability that sample impoverishment will occur and allow for a comparison of the EnKF and PF from a more objective standpoint. This analysis is summarized in Figure 45. This figure shows that both scenarios had very positive effects on the PF calibration but the EnKF did not benefit from these scenarios. This suggests that sample impoverishment was avoided in the PF and the EnKF was not subject to these issues. Surprisingly, the EnKF actually performed worse in the two extra studies than in the original study. This is because filtering relies on random numbers and is therefore subject small errors. In addition, the PF appears to perform more consistently than the EnKF, if sample impoverishment does not occur. Not only does this figure suggest that the PF can perform as well as the EnKF, it also suggests that it tends to be more robust if sample impoverishment does not occur.
From the difference in calibration of the HyMod model and the SAC-SMA model, the relative benefits and problems associated with each method are shown. First, MCMC techniques fail to compete with either global optimization or filtering. This is likely a result of the low number of chains. Though problems of MCMC techniques converging to the incorrect solution may be remedied by using a larger number of chains, this would greatly increase the computational burden and was therefore not included in this experiment. On the other hand, global optimization proved to be an efficient and relatively accurate technique. Though the SCE-UA algorithm performed well, it was outperformed by particle filtering for each model, as long as sample impoverishment did not occur. This showed that filtering has potential to be very useful in calibration of hydrologic models. While filtering was shown to be very effective for model calibration, the accuracy of each technique was variable, depending on the model. Conflicting results were found in the calibration of the HyMod model and the SAC-SMA model. While the
PS and PF perform better than the EnKF in calibrating the HyMod model, the EnKF performed better than the PS and PF in calibrating the SAC-SMA model. This is due to the occurrence of sample impoverishment in the PS and PF. Through the test of two scenarios to remove the occurrence of sample impoverishment, it was shown that the PF performs equally as well as the EnKF for calibrating the SAC-SMA model, if sample impoverishment is avoided. Though the EnKF is less prone to sample impoverishment, if this can be avoided in the PF, it will perform at least as well as the EnKF. In some scenarios, as was seen in the HyMod model, if a parameter is present in the model which has a highly non-Gaussian distribution, the EnKF is incapable of locating the correct parameter value.
7. Conclusion

7.1 Snow Data Assimilation Experiment

This research examined the operational potential for TB assimilation in SWE and streamflow prediction at the NWSRFC. The SNOW-17 model was coupled with a radiative transfer model showing potential for radiance prediction from the SNOW-17 model. Radiance assimilation had a positive influence on SWE prediction during the winter, though the spring estimation was associated with large uncertainties. This experiment also showed the ability of the PF to assimilate TB, and in comparison with the EnKF for this application, appears to be a more effective assimilation technique, which is attributed to errors resulting from the Gaussian assumption of the EnKF and the linear update scheme affecting the physical relationship of snow variables.

Snowmelt predictions, from the SWE estimation experiments, were then used to run the SAC-SMA model. By running the SAC-SMA model with snowmelt prediction from the SNOW-17 model, with TB assimilation, significant improvements in streamflow prediction were shown. Results from all assimilation experiments produced more accurate streamflow results than a control run, displaying the usefulness of TB assimilation for operational streamflow forecasting. Best results overall were from state-parameter estimation via the PF. This led to a more accurate expected value than the control run and provided a sufficient predictive bound. The end result for TB assimilation, in conjunction with streamflow assimilation, showed promise for operational streamflow forecasting in the NWSRFC.
7.2 Streamflow Data Assimilation

This study examined the relative effectiveness of global optimization, MCMC and filtering techniques in hydrologic model calibration. This was performed on two different hydrologic models, the HyMod model and the SAC-SMA model. All techniques were performed successfully in calibration of each model. Overall it was found that MCMC methods performed the worst in model calibration. Global optimization was found to perform quite well for both models but was not able to outperform filtering techniques. Results from this analysis show that filtering performs the best in calibrating models, but which filtering technique to use is a function of the model used/allowable computational demand.
References


[38] Sambridge M, K Mosegaard. Monte Carlo methods in geophysical inverse problems. Rev Geophys, 40 (3); 1009.