ABSTRACT

Title of dissertation: MULTI-SENSOR ASSIMILATION OF AMSR-E SPECTRAL DIFFERENCES AND GRACE TERRESTRIAL WATER STORAGE RETRIEVALS TO IMPROVE MODELED SNOW ESTIMATES

Jing Wang, Doctor of Philosophy, 2020

Dissertation directed by: Associate Professor Barton A. Forman
Department of Civl and Environmental Engineering

Snow, a key component of terrestrial water storage (TWS) in many watersheds across the globe, is a significant contributor to the Earth's hydrologic cycle, energy cycle, and climate system. This study explores multi-sensor, multi-variate data assimilation (DA) using synthetic Advanced Microwave Scanning Radiometer for EOS (AMSR-E) passive microwave (PMW) brightness temperature spectral differences (ΔT_b) and synthetic Gravity Recovery and Climate Experiment (GRACE) TWS retrievals in order to improve estimates of snow water equivalent (SWE), subsurface water storage, and TWS over snow-covered terrain. A series of synthetic twin experiments are conducted using NASA Catchment land surface model as the prognostic model. AMSR-E ΔT_b DA using a support vector machine as the observation operator improves SWE estimates, but adds little value to subsurface storage estimates. A physically-informed GRACE TWS DA approach significantly enhances the TWS vertical resolution via discretization into SWE and subsurface components

more accurately. When AMSR-E ΔT_b and GRACE TWS are assimilated simultaneously, dual assimilation significantly improves the SWE estimates with a 14.1% reduction of RMSE (relative to the Open Loop without assimilation) and leads to the largest improvement in TWS estimates (RMSE = 66.4 mm) and most reliable subsurface water storage ensemble spread (spread-error ratio = 1.08) as compared to the single-sensor DA scenarios. However, dual DA does not always yield complementary updates, and can at times, lead to conflictory changes to SWE. That is, the assimilation of ΔT_b often generates positive SWE increments whereas assimilation of TWS often removes SWE in the dual DA system, which can ultimately degrade the posterior SWE estimates. This synthetic experiment provides valuable insight for future DA experiments merging real-world AMRS-E/AMSR-2 ΔT_b and GRACE/GRACE-FO TWS retrievals in order to better characterize terrestrial freshwater storage across regional and continental scales.

MULTI-SENSOR ASSIMILATION OF AMSR-E SPECTRAL DIFFERENCES AND GRACE TERRESTRIAL WATER STORAGE RETRIEVALS TO IMPROVE MODELED SNOW ESTIMATES

by

Jing Wang

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2020

Advisory Committee:
Associate Professor Barton A. Forman, Chair/Advisor
Professor Shunlin Liang, Dean's Representative
Associate Professor Kaye L. Brubaker
Dr. Rolf H. Reichle
Dr. Mohamad Hejazi

Acknowledgments

I owe my gratitude to all the people who have made this dissertation possible and because of whom my graduate experience has been one that I will cherish forever.

First and foremost, I would like to express my deepest gratitude to Dr. Barton Forman, for giving me the invaluable opportunity to work on challenging and extremely interesting projects. Thank you for the incredible support and encouragement. He taught me how to be an independent researcher, how to solve problems scientifically as well as to use a logical approach when asking questions. He is always there to help and encourage me whenever I am confused or frustrated with my research. He is also super supportive when it comes to my life decision. He made it possible for me to work from home and move to Oklahoma to be with my family while still pursing my doctorate at the University of Maryland.

Further, I would like to thank Dr. Reichle, Dr. Brubaker, Dr. Liang, and Dr. Hejazi for agreeing to serve on my dissertation committee and for sparing their invaluable time reviewing my manuscript. Their insightful suggestions and comments on the proposal defense has greatly motivated me to learn more and think more broadly about the project. I also would like thank our research collaborator Dr. Manuela Girotto. She gave me a lot of insightful suggestions.

I would like to acknowledge our fantastic research group members: Gaohong, Jawairia, Jongmin, Meg, Yuan, Lizhao, Lu, Jing Tao, and Yonghwan, for their support and help. The lunchtime chats with all of you are unforgettable memories.

Thanks Gaohong, Jawairia, Jongmin, and Lizhao for being such great friends and supporting me along the way. Thanks Meg for showing me the American life and the baby kit from Meg's mother. Thanks Jing Tao for teaching me the model compilation and showering my son with piles of clothes. Thanks for Yonghwan for guiding me the synthetic AMSR-E assimilation and machine learning. Thanks Yuan for teaching me the LDAS modeling and AMSR-E assimilation.

Finally, I owe my deepest thanks to my family. Thank you for all your unconditional love and support. Special thanks goes to my husband, Dr. Qinggong Tang, for supporting me and always being there for me. Thanks to our son, Yuanjin, for supporting me along the way in his own way with his precious face and soft giggles.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction and Motivation

1.1 Hydrologic Cycle and Terrestrial Water Storage

The global hydrologic cycle describes the amount of movement within the Earth system, including the oceans, atmosphere, cryosphere, hydrosphere, and biosphere (see Figure 1.1). The hydrologic cycle starts with water evaporation from the ocean surface. As the water vapor rises, it expands, cools, and condenses to form clouds. If the cloud droplet can collide and coalesce, the water can return to the surface as precipitation. Once the water reaches the ground, it may evaporate or transpire into the atmosphere or infiltrate into ground as soil moisture or groundwater. Groundwater can be transpired into the atmosphere a second time or exfiltrate into the oceans, rivers, and streams. The runoff can carry the water from the land surface into the oceans where the cycle begins again.

Terrestrial water storage (TWS), as a key component of the terrestrial and global hydrological cycles, refers to all water stored on the land surface such as the groundwater, ice and snow, lakes and wetlands [1]. It can be defined as the residual of precipitation minus the amount of water that either evapotranspired from land surfaces or discharges as surface runoff [2]. TWS plays a major role in Earth's climate system by exerting a first-order control over the water, energy and biogeochemical

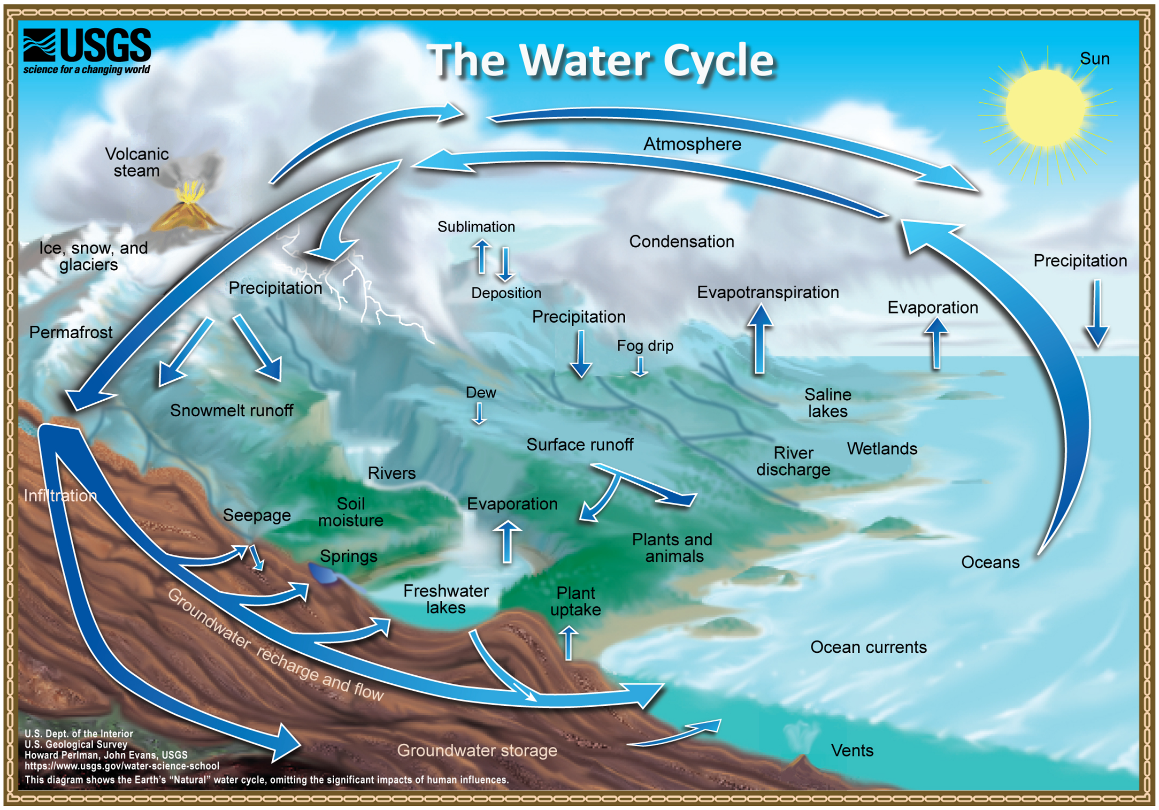

Figure 1.1: Schematic overview of the global nature water cycle. Source: https://www.usgs.gov/media/images/water-cycle-natural-water-cycle

fluxes [3]. As an integrated measure of terrestrial water, TWS has significant implications for water resources management. Since 2002, the Gravity Recovery and Climate Experiment (GRACE) satellites have provided the first global, time-varying estimate of changes in total, column-integrated TWS [4, 5]. However, the application of GRACE TWS retrievals to hydrologic applications is limited by its coarse temporal (∼ monthly) and spatial resolutions (≥150,000 km^2) [6, 7]. In addition, the time latency (∼1 to 2 month) of GRACE TWS retrievals also greatly limits its real-time applicability [8].

1.2 Importance and Challenge of Snow Characterization

Snow, as a key component of terrestrial water storage (TWS) in many basins globally, is a significant contributor to the Earth's hydrologic cycle [9], energy cycle [10], and climate system [11, 12]. Seasonal snow has high spatial and temporal variability and covers between 7% to 40% of the northern hemisphere land surface annually [13]. Further, snow accounts for a large fraction of the available freshwater resources of the northern hemisphere [12]. Due to global warming and an unsteady large-scale atmospheric movement, the snow cover extent in the northern hemisphere has reduced significantly over time during the spring season [14]. In other words, our virtual reservoir of freshwater in the form of snow is disappearing much earlier in the year. Meanwhile, an earlier onset of spring melt results in an earlier ablation season that causes more rapid peak stream flow in mountainous regions, which can increase the occurrence of natural hazards like floods that often adversely impact human life. Therefore, it is important to identify and quantify the snow mass in the natural environment to better understand the hydrologic responses associated with snow as well as to enhance our characterization of TWS.

It remains challenging to obtain high quality snow estimates across time and space. It is difficult to measure snow in conventional acquisition approaches such as point-scale, ground-based techniques due to the spatial and temporal variability inherent to snow processes. In addition, snow measurement collection can be dangerous given the cold weather conditions, limited access, and the everpresent danger of avalanches. As an alternative, snow can be estimated across regional and continental scales with land surface models. However, model-derived SWE estimates contain significant uncertainty due to the model limitations such as initial condition errors, model structure errors, model forcing errors, and model parameterization errors [15–18].

Remotely-sensed measurements from space-borne instrumentations are another way to characterize terrestrial snow mass or snow water equivalent (SWE). Previous studies have shown the potential of GRACE TWS retrievals for snow mass estimation when coupled with a land surface model [8, 19–22]. In addition, remotely-sensed measurements, primarily in the form of passive microwave (PMW) brightness temperature (T_b) measurements (e.g., Advanced Microwave Scanning Radiometer for EOS; AMSR-E) observations) [23–25], contain snow information across regional and continental scales. However, PMW T_b-based SWE retrievals typically contain numerous uncertainties and biases due to a variety of factors. For example, PMW radiation emitted from the underlying snowpack will be attenuated by the overlying vegetation can attenuate. At the same time, overlying vegetation will add on its own contribution to the signal as measured by the radiometer [24]. In addition, the signal saturation typically occurs when SWE is greater than 100 to 200 mm such that the correlation between SWE and T_b (at relative high frequency) can reverse [26–30]. Further, the presence of liquid water within the snowpack alters the electromagnetic response from a dry microwave scatter to a wet microwave emitter [31, 32].

1.3 Snow Assimilation

Fortunately, snow information from both observations and model predictions can be optimally combined through data assimilation (DA) to generate high-quality estimates that are better than the observations or the model alone [17, 33–35]. In addition, DA has the ability to implicitly downscale the integrated, monthly GRACE TWS retrievals to a finer resolution when merged with a land surface model.

A variety of snow-related observations and snow retrieval products can be assimilated into a land surface model. Previous studies have demonstrated that assimilation of point-scale, *in-situ* SWE measurements into a snow model can improve SWE estimation [36–40]. Further, satellite-based snow retrieval products derived from a radiative transfer model or a statistical regression models can also be assimilated into a land surface model [16, 41–46]. In addition to satellite-based snow retrievals, PMW T_b or spectral difference (ΔT_b) observations can be directly assimilated (a.k.a. radiance assimilation) using a physically-based microwave radiative transfer mode (RTM) [46–51] or machine learning in the form of support vector machine (SVM) regression [52–60]. In addition, previous studies have proven that assimilating GRACE TWS anomalies can improve modeled SWE estimates [8,22]. But the application of GRACE TWS retrievals is limited by the temporal ($\sim$monthly) and spatial resolutions ($\geq$150,000 km^2).

Despite the improvements in snow estimation using a single-sensor assimilation approach as highlighted above, it remains a challenge to discriminate only the snow mass-related portion of the PMW T_b or ΔT_b signal given the fact that snow mass estimation using PMW radiometry is fundamentally an ill-posed, underdetermined system [61]. Numerous combinations of snow characteristics can yield the same T_b or ΔT_b observations [61]. In addition, it is known that assimilating one model state may destabilize other model processes [62]. Kumar *et al.* [8] performed GRACE TWS assimilation across North America and the results suggested GRACE single-sensor assimilation may degrade model estimation in some locations such as the eastern United States.

As stated by Pan *et al.* [63] and Rodell *et al.* [64], constraining a model with a broad range of complementary observations is able to provide a more comprehensive understanding of the hydrological cycle. In the context of snow assimilation, Su *et al.* [21] first performed a one-dimensional (1-D) multi-sensor assimilation that merged MODIS snow cover fraction along GRACE TWS information over North America. The study suggested that a multi-sensor approach can significantly improve snow estimation compare to a MODIS-only assimilation approach.

1.4 Goals and Objectives

The limitations and breakthroughs mentioned above motivate this study to further investigate a more comprehensive way to estimate snow using a multi-variate analysis that includes multiple sources of observational information, namely, AMSR-E PMW T_b observations and GRACE TWS retrievals. This goal was achieved by establishing the following objectives:

1). Improve snow estimation using SVM-based PMW ΔT_b assimilation by incorporating a physically-informed assimilation approach along with a "data-thinning" approach

2). Improve snow estimation using GRACE TWS assimilation by including a physically-informed approach

3). Improve snow estimation via multi-sensor, multi-variate assimilation of AMSR-E ΔT_b observations and GRACE TWS retrievals

1.5 Organization of the Thesis

In Chapter 2, the background of the dissertation research and prior research accomplishments are discussed. In Chapter 3, a synthetic identical twin experiment is established over the Volga River basin, which contains a significant amount of snow but where ground-based snow measurements are lacking. In Chapter 4, SVM-based PMW ΔT_b assimilation using a physically-informed approach and a "data-thinning" approach is conducted and then evaluated against the synthetic truth established in Chapter 3. In Chapter 5, multi-sensor, multi-variate assimilation of AMSR-E ΔT_b and GRACE TWS observations is performed and compared against the single-sensor, single-variate assimilation (i.e., AMSR-E-only or GRACE-only assimilation). Chapter 6 summarizes the major findings and future directions for continued research.

Chapter 2: Background and Literature Review

2.1 Components of Terrestrial Water Storage

Terrestrial water storage (TWS) can be defined as the summation of all water components on the land surface and in the subsurface, including biomass interception, snow and ice, soil moisture, groundwater, and surface water impoundments (Figure 1.1) [65]. TWS, as a dynamic part of the hydrological cycle, plays an important role in Earth's climate system by acting as a first-order control on water, energy, and biogeochemical fluxes [1, 3]. It is critical to monitor the TWS in order to characterize changes in water availability and hydrologic extremes, especially in a changing climate [66]. In addition, the accurate estimation of TWS variation can improve our understanding of regional and global water cycles as well as their interactions within the Earth system [3].

The variability of TWS tends to be dominated by snow and ice in polar and alpine regions, by subsurface water (i.e, soil moisture and groundwater) in mid-latitudes, and by surface water in wet and tropical regions (e.g, Amazon) [4]. Wet biomass (i.e., water stored in vegetation) is often considered negligible in TWS budget analyses because of the small weight compared to other TWS components [67].

Surface water impoundments, including rivers, wetlands, inland water bodies, and floodplains, exist at the land-atmosphere interface. In most regions of the tropics, the water from the land surface helps supply agricultural and energy production [65]. The contribution of changes in surface water storage to TWS variability are highly dependent on location. For example, surface water storage change contributes the most in the tropic regions where major rivers flow over arid regions and at high latitudes [68].

Subsurface water storage, including soil moisture and groundwater, is another major component of TWS. Soil moisture refers to the water stored in the unsaturated zone. It plays an important role in partitioning rainfall into surface runoff, infiltration, and evapotranspiration [3]. In addition, soil moisture has a great impact on the occurrences of flood and drought [69, 70]. Soil moisture helps dictate the turbulent fluxes, and hence, it influences global and regional hydrometeorological processes [71, 72]. The high variability of soil moisture is generally the most important component of seasonal changes in TWS in the mid-latitudes areas [65]. Groundwater, on the other hand, provides domestic water for agriculture and energy production [73, 74]. Groundwater responds mores slowly to meteorological conditions compared to soil moisture, and hence, contributes less to seasonal TWS variations but is a significant components to TWS variations at inter-annual to decadal timescales [4, 75, 76].

Snow, as a key component of TWS, is mainly distributed in mountainous regions in the mid-latitude and high latitude regions [4]. It has a significant impact on the Earth's hydrologic cycle [9], energy cycle [10], and climate system [11, 12]. The high variability of snow in space and time induces significant variations in TWS. Due to snow accumulation and ablation, the surface albedo can change and can further affect the surface energy balance [77]. In addition, freshwater from snow (and ice resulting from snow) provides the freshwater supply for more than one billion people globally [78]. However, high quality global snow estimates across time and space are lacking due to the complexity of the terrain when snow if often located and limitations in conventional acquisition approaches such as point-scale, ground-based techniques. For this reason, snow or snow water equivalent (SWE) estimation rely heavily on remote sensing techniques such as microwave or gravimetric measurements [79].

2.2 Catchment Land Surface Model

The NASA Catchment Land Surface Model (Catchment) [80], the land model component of the Goddard Earth Observing System (GEOS-5) modeling and data assimilation framework, is used in this study. Catchment (Figure 2.1) includes an explicit representation of the spatial variation of the soil moisture, water table depth, surface runoff, and evaporation [80]. The hydrological processes for each grid are based on its hydraulic parameters, topographical statistics, and soil texture. Catchment is able to represent shallow groundwater storage changes that is suitable for GRACE TWS data assimilation [22, 81–85].

Snow conditions on the land surface are represented with a three-layer snow model in Catchment, including snowpack consolidation and snow metamorphosis

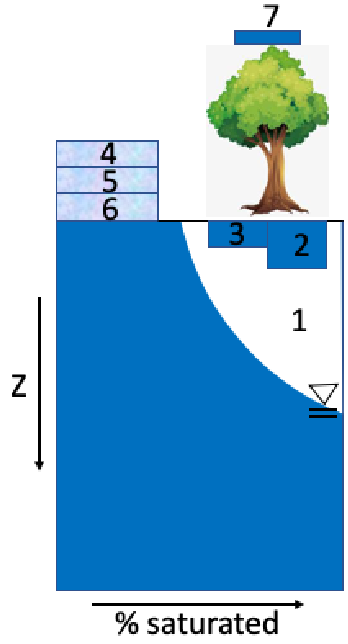

Figure 2.1: Conceptual representation of TWS and its components in the NASA GEOS-5 Catchment Land Surface Model. Number 1 represents the soil moisture deficit, 2 is root-zone layer excess, 3 is surface soil moisture excess, 4-6 represent three-layer snow model, and 7 is canopy interception.

processes [86]. Subsurface water storage is defined as the sum of soil moisture and shallow water storage, which can be calculated by three prognostic variables (i.e., surface excess, root zone excess, and catchment deficit) in Catchment. Catchment deficit is the amount of water that is required to saturate the subsurface, which is related to the unconfined groundwater table depth [87]. Root zone excess and surface excess are the excess or deficit soil moisture relative to equilibrium conditions for the top 100 cm and 5 cm of soils, respectively [87]. Although total groundwater is not explicitly modeled, the vertical distribution of soil moisture has an implicit water table that is located at the depth of the equilibrium saturation. This model feature indicates the model ability to represent time-varying water storage. The Catchment modeled TWS can be calculated as:

$$TWS = \frac{cdcr2}{1 - wpwet} - catdef + rzex + srfexc + capac + SWE \tag{2.1}$$

where cdcr2 is the total water equivalent within the pore space of the subsurface; wpwet is the wilting point wetness; catdef is the catchment deficit; rzex is the root-zone layer excess; srfexc is the surface layer excess; and capac is the canopy interception [22].

One major hydraulic limitation in Catchment is the lack of dynamic surface water impoundments (e.g., reservoirs) and dynamic river routing routines. This deficiency may result in a non-negligible amount in the modeled TWS in regions where surface water storage changes are a significant component of TWS [88].

2.3 Remote Sensing of Terrestrial Water Storage

2.3.1 Overview of GRACE Mission

The Gravity Recovery and Climate Experiment (GRACE) mission is a joint operation between the National Aeronautics and Space Administration (NASA) and the German Aerospace Center. GRACE was launched on March 17, 2002 and monitored the time-variable gravity field of the Earth until 2017 [5]. GRACE Follow-On (GRACE-FO) was launched in 2018 that is a continuation of GRACE's legacy of tracking Earth's water movement across the planet. The Earth's gravity field observations from GRACE can be used to infer global and regional TWS changes and provides a unique view of Earth's climate and hydrologic cycle. GRACE consisted of two almost-identical satellites in tandem formation operating about 200 km apart in one orbital plane with an altitude of 450$\sim$500 km [89].

Figure 2.2 illustrates the basic principle of the gravity measurements by GRACE satellites. As the leading satellite approaches a positive mass or gravity anomaly (e.g., snow-covered mountains), the satellite is pulled toward the mass anomaly, resulting in a larger separation distance between the two satellites. As the trailing satellite approaches the same snow-covered mountain, it is also pulled toward the anomalous mass while the leading satellite is held back by the positive mass anomaly, which results in a decrease in distance between these two satellites. A highly accurate inter-satellite K-Band microwave ranging system was used to measure the distance between the two satellite with a precision better than 1 μm [5].

The precise positions of the twin satellites of GRACE were measured with global positioning system (GPS) and star cameras. By this principle, GRACE was able to detect the mass changes associated with atmospheric circulation, ocean circulation, and redistribution of TWS via the hydrological cycle [65].

2.3.2 Terrestrial Water Storage Retrievals

GRACE TWS retrievals have improved the understanding of hydrologic states and fluxes [90], such as the drought characterization [83,91,92], identification of flood potential [93–95], estimation of streamflow [96], estimation of evapotranspiration from major river basins [97], quantification of snow variations [19–22, 98], quantification of groundwater changes in major aquifer systems [75, 99, 100], estimation of soil moisture [84, 101, 102], tracking of glacier ice mass loss over Greenland and Alaska [103, 104], and hence improved the global water budget estimates [88, 101].

The monthly variations (i.e., anomalies) of TWS are provided by the analysis of the range-rate observations between the twin satellites. In general, there are two methods to produce the monthly TWS anomalies, including the spherical harmonics [105] and the mass concentration (a.k.a., mascon) techniques [6]. TWS, as a part of the gravity field, is often represented using spherical harmonics that is based on the expansion of a set of coefficients (degree and order ≤ 120). Spherical harmonic expansion is used to describe the shape of the geoid with a surface of constant gravitational potential matching the mean sea surface level. Numerical devices such as Gaussian averaging functions can be used with the expansion

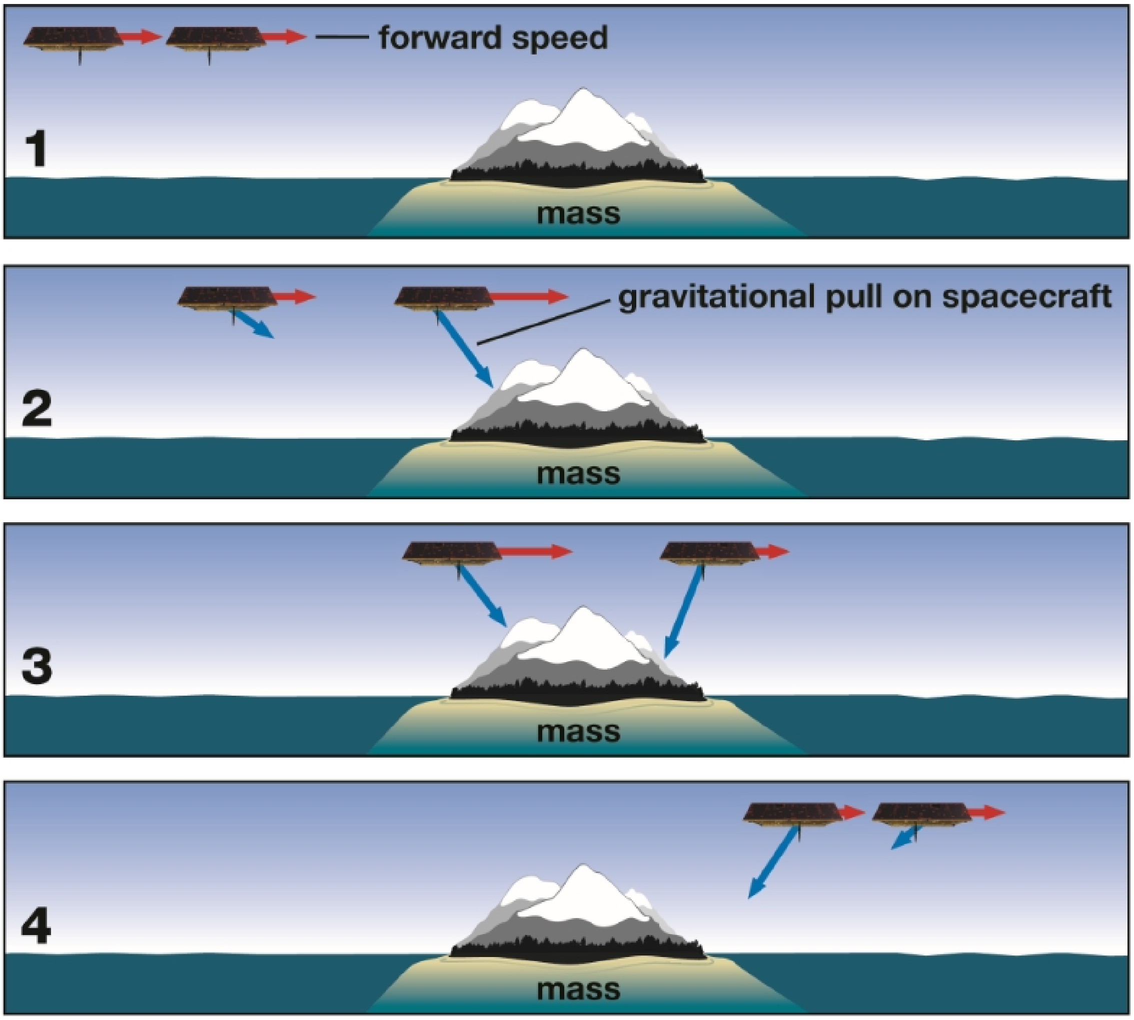

Figure 2.2: Simplified example of how the GRACE satellites detected changes in the gravity field. The distance between the two satellites is constant when there is no mass change (panel 1). As the leading satellite approaches the snow-covered mountain, which represents a positive mass anomaly, the satellite is pulled toward the mountain due to the gravitational force exerted by the mass change (panel 2). Thus, the distance between the two satellites increases. As the second satellite encounters the positive mass anomaly, it is also pulled toward the mass while the lead satellite is held back by the mass, and hence, the distance between them decreases (panel 3). As the satellites move away from the mass anomaly, they return to their original positions of separation (panel 4). Source: https://gracefo.jpl.nasa.gov/resources/50/how-grace-fo-measures-gravity/.

coefficients to isolate mass anomalies from the baseline temporal mean [106]. However, GRACE TWS retrievals derived from spherical harmonic solutions suffer from north-south “stripes” that must be removed with “de-striping” or smoothing algorithms due to the presence of correlated errors between different spherical harmonic coefficients [105]. During the smoothing or “de-striping” processes, a portion of the real, geophysical signal is also removed. In addition, the effectiveness of “de-striping” is strongly affected by the orientation, shape, and size of the signals, and further results in estimation errors that are inversely correlated to the size of the region. In general, the estimation error for continental-scale watershed is 1-2 cm equivalent height of water. For small regions (area $\leq$ 150,000 km^2), the estimation errors are so large that they overwhelm the hydrology signal [4, 105].

Unlike spherical harmonics, the mass concentration (mascons) method uses the inter-satellite ranging observations directly to estimate the gravity solutions [6]. Each mason has a specific known geophysical location that can serve as *a priori* constraints. During the inversion process, the correlated estimation errors in the gravity solution can be removed internally, which can eliminate the need for smoothing or de-striping [65]. In addition, the constrained mascons solutions have a better ability to separate the terrestrial signals from the ocean signals [107].

GRACE-based TWS retrievals suffer from a coarse spatial ($\geq$ 150,000 km^2 at midlatitudes) [6, 7] and coarse temporal ($\sim$ monthly) resolution. In addition, the hydrologic application of GRACE-based TWS retrievals is limited due to the vertical integration of the water storage components. To address these challenges, data assimilation methods can be utilized [81]. Selection of an appropriate observation operator helps partition the vertically integrated TWS into its constituent components and while effectively downscaling TWS information into finer temporal and spatial resolutions [84, 99].

2.4 Remote Sensing of Snow

Remote sensing of snow can be performed using the electromagnetic spectra across a range of wavelengths. Snow cover products have been derived from Moderate Resolution Imaging Spectroradiometer (MODIS) with the infrared and visible bands (0.4-14 μm) [108]. Active remote sensing in the microwave portion of the spectrum based on backscattering effects have been used to study snow cover [109]. Alternatively, NASA Airborne Snow Observatory using a scanning LIDAR and an imaging spectrometer is regularly used to investigate the distribution of snow [110]. Finally, space-borne microwave radiometers are particularly effective for detecting snow and have been utilized in the production of a generation of snow retrieval products [109,111]. This dissertation, in particular, will focus on the passive remote sensing of snow.

2.4.1 Passive Microwave Remote Sensing of Snow

Passive microwave (PMW) radiation is defined as microwave radiation naturally emitted by an object. Compared to optical (visible) radiation, microwave radiation can penetrate through media, and hence, is able to penetrate clouds and be detected by space-borne microwave radiometers. Therefore, PMW sensing of snow can be detected during both day and night under all-weather conditions [112].

In order to quantify emitted microwave radiation associated with snow, brightness temperature (T_b) is often used, which is defined as the equivalent temperature of the microwave radiation thermally emitted by an object [113]. As a fundamental parameter measured by passive microwave radiometers, T_b is calculated based on the Raleigh-Jean approximation:

$$T_b = \varepsilon \times T_{physical} \tag{2.2}$$

where ε is a dimensionless quantity of emissivity ($0 \leq \varepsilon \leq 1$) and $T_{physical}$ is the physical temperature [K] of the emitting surface. Emissivity is wavelength dependent and is largely affected by the dielectric constant of the material.

PMW remote sensing of snow is dependent on preferential scattering of microwave radiation by the snow pack at a relatively low frequency (10.7 GHz or 18.7 GHz) compared to a relatively high frequency (18.7 GHz or 36.5 GHz) (Figure 2.4). This preferential scattering at lower (or higher) frequency increases (or decreases) the emissivity and hence increases (or decreases) the corresponding measured brightness temperature [114]. The difference between T_b's, also known as spectral differences, at lower (10.7 GHz or 18.7 GHz) and higher (18.7 GHz or 36.5 GHz) frequencies is often used in SWE retrievals [111]. The spectral differences, ΔT_b, can be expressed as:

$$\Delta T_{b18V-36V} = T_{b18V} - T_{b36V} \tag{2.3}$$

where T_{b18} represents T_b at 18.7 GHz; T_{b36} represents T_b at 36.5 GHz; and the subscript V represents vertical polarization with a similar equation for horizontal

(H) polarization.

PMW measured brightness temperature is affected by several snow characteristics. For example, a deeper, dry snow pack depth can induce more scattering and less emission of the upward radiation that ultimately results in a lower T_b (Figure 2.4). This effect is more prominent on T_b measured at a relatively high frequency (Figure 2.3) [114]. Snow grain size is generally inversely correlated with measured T_b as the larger mean snow grain size is associated with lower measured T_b for a given amount of SWE (Figure 2.4) [115, 116]. The presence of ice layers inside and on the surface of a snow pack generally leads to overestimation of snow depth and SWE since the ice layers can introduce greater scattering and a lower measured T_b [30]. Similar to an ice layer, the increase in the thickness of depth hoar layer (cup-like large loose snow grains) induces more microwave scattering and less emission, which will ultimately results in lower measured T_b [117]. Liquid water within the snow pack can further complicate the snow estimation. The presence of liquid water within the snowpack alters the electromagnetic response from a dry microwave scatter to a wet microwave emitter due to a larger dielectric constant [31, 32]. In general, T_b has a positive correlation with snow wetness within the snowpack until a saturation threshold (a.k.a. signal saturation) is reached [118]. In addition, It is hard to distinguish wet snow from snow free soil because wet snow behaves like a blackbody [119] which is a perfect emitter for all incident radiation at the physical temperature of the snow layer [120]. Vegetation over snow can attenuate the upwards transport of radiation emitted from the underlying snowpack, and at the same time, it will emit radiation upward microwave radiation as measured by the radiometer (Fig-

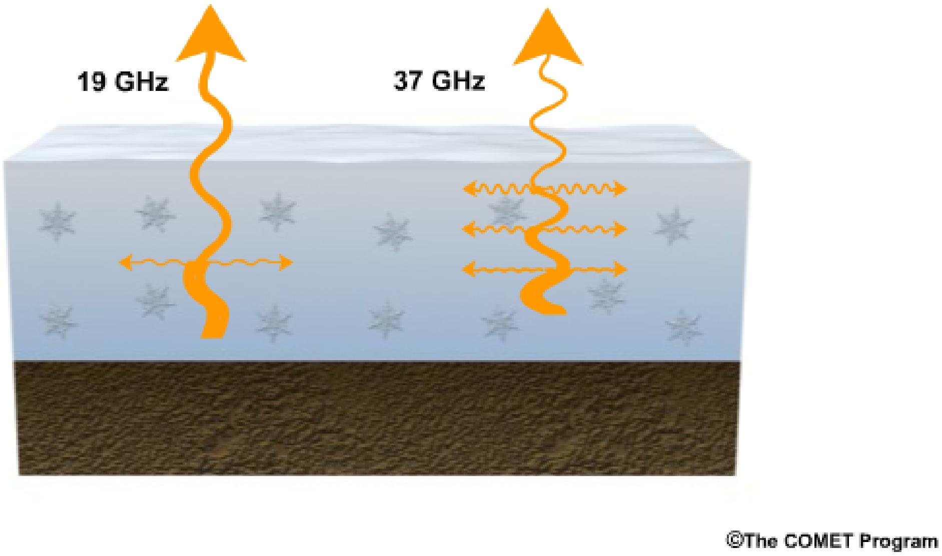

Figure 2.3: Preferential scattering of microwave radiation with frequency 37 GHz compared to microwave radiation with frequency 19 GHz by the snow pack. The orange arrows represent the direction of the wave and the width of the arrow indicates the strength of measured radiation. Source: http://meted.ucar.edu/

ure 2.4) [24]. Clouds and aerosols in the atmosphere are similar to the vegetation in that not only attenuate the microwave radiation emitted by the snowpack but also make its own contribution to the signal [121]. Luckily, the microwave bands (i.e., 10.7 GHz, 18.7 GHz, and 36.5 GHz) utilized for remote sensing of snow are in an atmospheric window that are not significantly affected by the atmosphere [122].

Figure 2.5 shows a comparison of *in-situ* measurements of snow depth from the SNOwpack TELemetry (SNOTEL) network against the AMSR-E PMW spectral difference ($\Delta T_{b18V-36V}$, see notations in Equation 4.1). Although PMW $\Delta T_{b18V-36V}$ captures the snow accumulation and ablation phase, significant high-frequency noise

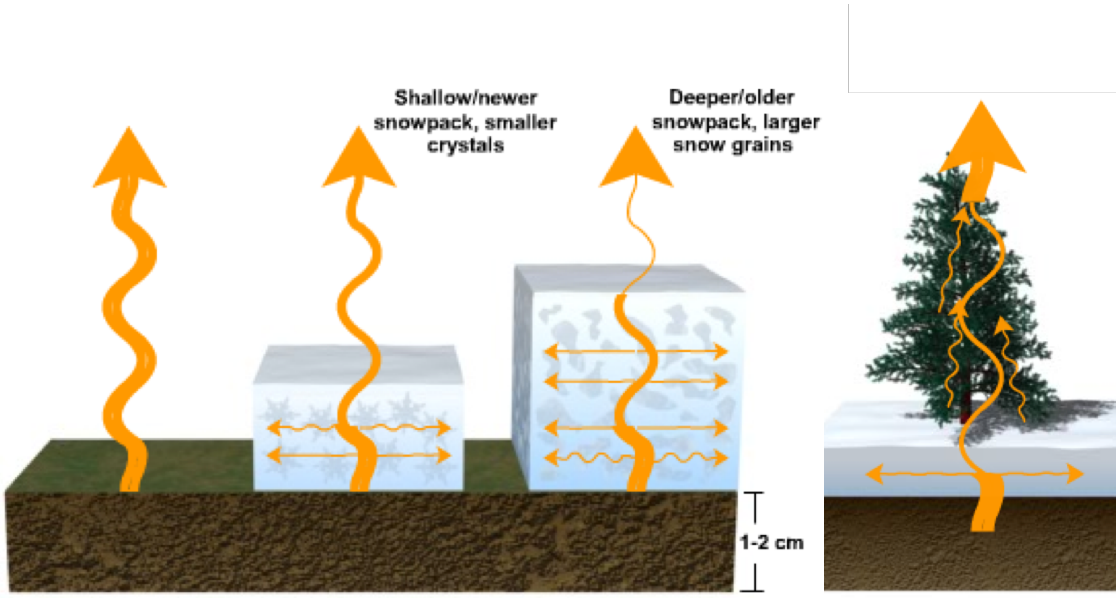

Figure 2.4: Electromagnetic wave emitted by each object (ground, shallow snowpack, deep snowpack, and vegetation) on the surface. The orange arrows represent the direction of the wave and the width of the arrow indicates the strength of measured radiation (reproduced from University Corporation for Atmospheric Research, the COMET ® Program).

still exists and must be carefully considered.

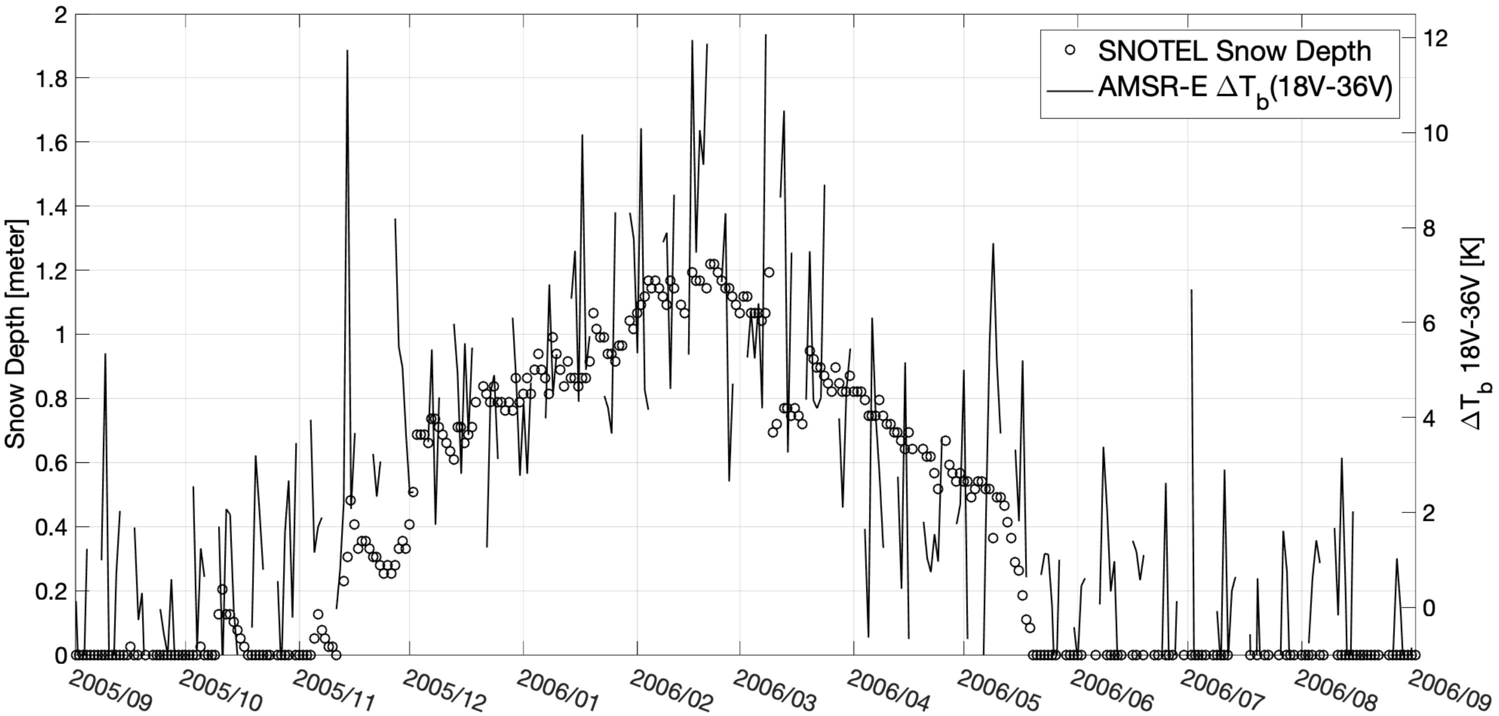

Figure 2.5: Example comparison between AMSR-E $\Delta T_{b18V-36V}$ observations and SNOTEL snow depth measurements for a location in Western Colorado (40.31°N, 105.65°W) from 1 September 2005 to 1 September 2006. ΔT_b (black line) captures the general features of snow depth (black dots), but contains more signals (e.g., snow temperature) *not* related to snow mass as well as the presence of high-frequency noise.

2.4.2 Snow Products

One way to estimate snow-related properties from PMW radiation is to invert snow depth (SD) or SWE from PMW T_b or ΔT_b using regression coefficients. Chang *et al.* [114] first presented the relationship between snow depth and ΔT_b for a dry, uniform snowpack with a constant snow density as:

$$\Delta SD = 1.59 \times (T_{b18H} - T_{b37H}) \tag{2.4}$$

where SD represents the snow depth [cm] and T_{b18H} [K] and T_{b37H} [K] represent T_b at 18 and 37 GHz at horizontal (H) polarization, respectively. Analogously, Goodison *et al.* [123] derived the relationship between SWE and ΔT_b for dry snow as:

$$SWE = a + b \times (T_{b37V} - T_{b19V}) \tag{2.5}$$

where SWE represents the snow water equivalent [cm]; T_{b19V} [K] and T_{b37V} [K] represent T_b at 18 and 37 GHz for vertical (V) polarization, respectively; a $= -20.7$ [m] and b $= -2.74$ [K^{-1}] are fixed parameters.

Kelly *et al.* [124] later employed a radiative transfer model to estimate snow depth by coupling the snow grain size and volumetric fraction. Given the fact that forest cover has an important impact on the snow retrieval algorithm [125], Chang *et al.* [126] considered forest fraction in the algorithm as:

$$SWE = \frac{a \times (T_{b19V} - T_{b37V})}{(1 - FF)} \tag{2.6}$$

where a is a calibrated coefficient [-] and FF ($0 \leq \text{FF} \leq 0.75$) is the forest fraction [-]. Kelly *et al.* [111] calculated snow depth (SD) for both forested and non-forested regions as:

$$SD = FF\times[p1\times\frac{(T_{b18V}-T_{b36V})}{(1-b\times FD)}](1-FF)\times[p1(T_{10V}-T_{36V})+p2(T_{10V}-T_{18V})] \quad (2.7)$$

where SD represents the snow depth [cm]; FF ($0 \leq$ FF ≤ 0.75) is the forest fraction [-]; FD is the forest density; b is a regression coefficient; p1 and p2 are two dynamic coefficients ranging from 1 to 2; and T_{b10V} [K], T_{b18V} [K], and T_{b36V} [K] represent T_b at 10, 18 and 36 GHz at vertical (V) polarization, respectively.

However, there are some certain assumptions, such as constant snow density and uniform snow grain size, have to be made first in order to use above empirical equations. In addition, satellite-based PMW SWE retrievals are affected by signal attenuation in deep snow [127], forest and atmospheric attenuation [24, 128, 129], and the assumed (quasi-) linear relationship between the physical characteristics of SWE and the electromagnetic response of the snowpack [126, 127]. Further, PMW T_b-based SWE retrievals are also affected by an inaccurate model representation of snow morphology [124], stratigraphy [24], ice crusts [130], depth hoar [23], snow grain size [116], and sub-grid scale lake effects [30].

2.4.3 Advanced Microwave Scanning Radiometer – Earth Observing System

During this dissertation, PMW radiation observations from the Advanced Microwave Scanning Radiometer for Earth Observing Systems (AMSR-E) are used for the experiments. AMSR-E was aboard the AQUA polar-orbiting, sun-synchronous satellite from May 4, 2002 to October 4, 2011. AMSR-E instruments are dual-polarized, conical scanning, passive microwave radiometers. They are placed in a near-polar orbit (altitude = 705 km) which allows for up to twice daily sampling of a given location [131]. There are twelve channels and six frequencies, including 6.9 GHz, 10.7 GHz, 18.7 GHz, 23.8 GHz, 36.5 GHz, and 89.0 GHz at both horizontal and vertical polarizations. The spatial resolution of AMSR-E observations is different with different frequencies that higher frequency band is associated with finer resolution. For example, the observations at 36.5 GHz have the spatial resolution of 14×8 km^2 while those at 10.7 GHz and 18.7 GHz have coarser spatial resolution of 51×29 km^2 and 27×16 km^2, respectively [131].

AMSR-E observations are suitable for snow retrievals, especially at the 10.65 GHz, 18.7 GHz and 36.5 GHz frequency bands whose wavelengths are large enough to pass through the atmosphere and clouds with minimal attenuation [111, 115]. AMSR-E observations are also used for retrieving other geophysical fields such as soil moisture, precipitation rate, water vapor, cloud water, sea surface winds, sea surface temperature, and ice [132]. As a successor to AMSR-E, the Advanced Microwave Scanning Radiometer 2 (AMSR2) is onboard the GCOM-W1 satellite that measures weak microwave emission from the surface and the atmosphere of the Earth.

2.5 Data Assimilation

The goal of data assimilation (DA) is to merge complementary information from disparate measurements of any type, including remote sensing observations, into an estimate of one or more geophysical fields from a physically-based model [17].

DA is able to interpolate and, at times, extrapolate remote sensing observations into finer smaller scales in both time and space. An optimal DA system relies on the relative weights of uncertainties that come from the model estimates and the remote sensing observations [17]. In general, assimilated (a.k.a. updated) estimates will be closer to the observations when accurate observations are available and when a reasonable error correlation structure exists.

All DA methods share the same general theory of merging models with observations. However, there are differences between the different methods for a given application. For example, DA systems used for atmospheric and oceanic applications are greatly concerned with the appropriate specification of initial conditions because small errors in the initial conditions can result large differences in the updated results. Alternatively, land surface assimilation is more about estimating errors in boundary conditions (a.k.a. meteorological forcing) and model parameterizations [17].

The skill of a DA framework is largely dependent on the accurate specification of the input error parameters, including the observation error and model background error [133]. Observation errors include instrument errors and the errors of representativeness (such as the errors in the observation operator and the errors associated with the observation interpolation (ϕ in Equation 2.9)) [17]. Model background errors, including uncertainties in boundary conditions, model parameters, and deficiencies in model physics, (a.k.a. model structure error) are usually represented by stochastic perturbations of model prognostics and input forcings [133].

2.5.1 Ensemble Kalman Filter

In this work, the ensemble Kalman filter (EnKF) is used to improve characterization of TWS and its components across regional and continental scales. The standard Kalman filter is the optimal sequential data assimilation method for linear dynamics and measurement processes with Gaussian errors [134]. Given the fact that land surface models are nonlinear, the fully optimal behavior of Kalman filter can not be achieved. Therefore, Evensen [135] developed an ensemble of model trajectories from which the model error covariances are estimated at the time of an update. This Monte-Carlo based approach to filtering problem is known as the EnKF.

EnKF is a reduced-rank approximation that can reduce the number of degrees of freedom to a more manageable level and is suitable for modestly nonlinear problems where a small ensemble of model replicates can capture the relevant parts of the true error structure [17,136]. The EnKF is an attractive option for land surface applications due to four reasons [136]. Firstly, it is relatively easy to implement even with nonlinear land surface models and measurement equations. Secondly, the sequential structure of EnKF is convenient for assimilating measurements in real time. Thirdly, the EnKF provides information on assessing the accuracy of its estimates. Finally, a wide range of possible model errors can be considered in the process [136].

As shown in Figure 2.6, the EnKF alternates between two steps, an ensemble forecast step and an update step. During the forecast step, an ensemble of model state vectors, x, containing the relevant model prognostic variables (i.e., SWE and

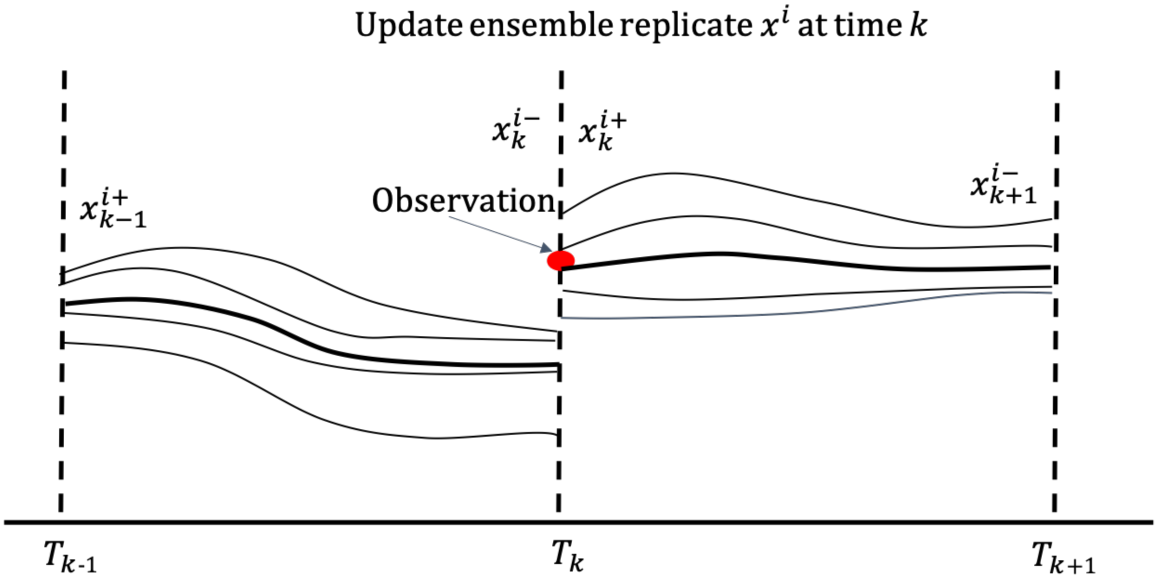

Figure 2.6: Schematic of the ensemble Kalman filter (EnKF) [137].

catdef) are propagated forward in time in the generic form as:

$$x_t^{i-} = f(x_{t-1}^{i+}, u_t^i, \alpha, \omega_t^i), i = 1, ..., N, \tag{2.8}$$

where $f(\cdot)$ represents the nonlinear hydrologic model and u, α, and ω represent the forcing fields, model parameters, and model errors, respectively; i represents a single replicate drawn from an ensemble of size N; and the superscripts $-$ and $+$ refer to the *a priori* state vector and *a posteriori* state vector, respectively. The model error accounts for the uncertainties related to errors in the background model formulation (e.g., model parameters and model structural errors) or the surface meteorological forcing [136]. In the forecast step, the model prognostic variables of interest (e.g., snow water equivalent and catchment deficit described in Section 2.2) are collected into the state vector x_t of dimension M at time t. The EnKF state estimates are computed based on the mean of the ensemble members.

The ensemble of replicates are generated by applying perturbations sampled from randomly generated noise to the boundary condition and model prognostic fields [133]. The typical approach is to employ normally distributed additive or log-normally distributed multiplicative perturbations to model prognostic field errors and meteorological boundary condition errors [133]. The ensemble mean of the perturbations are constrained to zero (additive perturbations) and one (multiplicative perturbations) to avoid introducing systematic biases in the perturbed fields [133].

Using the available observations Z_t at time t (Figure 2.6), the prior state vector, x_t^{i-}, is updated to a new value, x_t^{i+}, based on the relative uncertainties between the state vector and the predicted observation using appropriate weights [8, 81, 138] expressed in the Kalman gain, K_t, via:

$$x_t^{i+} = x_t^{i-} + K_t(Z_t + v^i - \phi_t(x_t^{i-})), i = 1, ..., N, \tag{2.9}$$

where $\phi_t(\cdot)$ is the observation operator that maps the model states (e.g, SWE, catdef) into observation space (e.g., ΔT_b or TWS); i represents a single replicate drawn from an ensemble size N; and v^i represents the observation errors that are assumed here to be Gaussian. Each ensemble member is updated separately. The observation errors include the measurement instrument errors, deficiencies of the observation operators (such as support vector machine regression and radiative transfer models), and representativeness errors from differences in spatial scales [133, 139]. K_t is the Kalman gain matrix computed from the ensemble and can be expressed as:

$$K_t = C^-_{x_t y_t}[C^-_{y_t y_t} + R]^{-1} \tag{2.10}$$

where R is the specified observation error variance; y_t is the observation prediction ($\phi_t(x_t)$ in Equation 2.9); $C^-_{x_t y_t}$ is the error cross-covariance between the modeled state estimates (x^-_t) and the observation prediction (y^-_t) prior to the update; and $C^-_{y_t y_t}$ is the error covariance (a.k.a., sample covariance) of the observation prediction prior to the update. The error covariances (i.e., $C^-_{x_t y_t}$ and $C^-_{y_t y_t}$) can be computed as sample statistics from the ensemble as:

$$C^-_{x_t y_t} \simeq \frac{1}{N-1}[(x^-_t - \overline{x^-_t})(y^-_t - \overline{y^-_t})^T] \tag{2.11}$$

$$C^-_{y_t y_t} \simeq \frac{1}{N-1}[(y^-_t - \overline{y^-_t})(y^-_t - \overline{y^-_t})^T] \tag{2.12}$$

where

$$\overline{x^-_t} = \frac{1}{N}\sum_{i=1}^{N} x^{i-}_t \tag{2.13}$$

$$\overline{y^-_t} = \frac{1}{N}\sum_{i=1}^{N} y^{i-}_t \tag{2.14}$$

where N is the number of ensemble size. The linear update step is, in general, suboptimal. But the degree of suboptimality is relatively small such that the estimates from the EnKF are satisfactory even for moderate ensemble sizes [136]. The reduction of the uncertainty is represented by the reduction of the ensemble spread [137].

As mentioned above, the merging of the observations and model forecasts is based on weighting their respsective sources of errors. Therefore, the appropriate specification of error characteristics, include model errors and observation errors, is critical for the successful implementation of EnKF framework [17, 139–142]. However, the specification of input error covariances is still challenging and remains a subjective process in current land data assimilation systems [17, 140, 142].

2.5.2 Snow Assimilation

There are a variety of observations and snow retrieval products, including SWE, snow depth, snow-covered area, and ΔT_b, that could be assimilated with a land surface model. Previous studies has demonstrated that assimilation of point-scale, *in-situ* SWE measurements into a snow model can improve SWE estimation [36–39]. Point-scale SWE assimilation needs an interpolation algorithm in order to convert point measurements into a spatially continuous observation coverage prior to conducting the DA. Therefore, SWE estimates using DA are relatively good in regions with a sufficiently dense network of observation stations, but remains challenging in places with sparse ground-based measurement networks due to error and uncertainty introduced during interpolation of the the sparse observations [40].

Instead of directly assimilating ground-based measurements, satellite-based snow retrieval products derived from a radiative transfer model or a statistical regression models can be assimilated. For example, Liu *et al.* [45, 46] assimilated bias-adjusted AMSR-E snow depth retrievals in Alaska and Colorado that resulting in improved snow depth estimates. DA results showed a 65% reduction in root-mean-square error (RMSE) versus ground-based snow depth measurements in Alaska as compared to modeled snow depth without assimilation. However, other studies yielded limited improvements in SWE estimates via data assimilation due to the presence of significant negative biases in the assimilated SWE retrievals [16, 41–44].

Satellite-based snow-covered area and snow-covered fraction products derived from visible and thermal imaging have also been assimilated. Preliminary results showed a modest improvement for peak SWE in some study domains such as southern Colorado and the Sierra Nevada mountains [41, 44, 143–145]. However, these studies suggested there were significant uncertainties in the estimation of SWE because of the utilization of a snow depletion curve as a function of fractional snow cover [44]. It has also been shown that AMSR-E-based snow depth assimilation is more effective when using snow-covered area retrievals as an additional data constraint [146]. Beyond satellite-based snow retrieval products, PMW T_b observations can be directly assimilated (a.k.a. radiance assimilation). There are two types of model operators to convert modeled SWE into the PMW T_b observation space: 1) a physically-based microwave radiative transfer mode ls(RTMs), and 2) a machine learning in the form of physically-constrained support vector machine (SVM) regression or an artificial neural network.

With the utilization of a RTM, Pulliainen *et al.* [47] first assimilated spectral differences ΔT_b from AMSR-E and Special Sensor Microwave/Imager (SSM/I) observations. The study showed the combing a snow model with ΔT_b information reduced the systematic errors in SWE associated with signal saturation effects. Durand *et al.* [147] assimilated the ground-based T_b over a relatively small domain and successfully reduced systematic errors in snow depth estimation. Similarly encouraging results were also found in snow estimation via assimilation of AMSR-E T_b observation in Siberia and Nevada [46, 49]. Kwon *et al.* [50, 51] also demonstrated that T_b assimilation can improve continental-scale snow storage estimates [50, 51]. However, the application of a RTM acorss a large region of space is constrained by the nontrivial computational demand [50]. In addition, most global land surface models are unable to accurately represent the snow microstructure such as snow grain size, snow grain shape and internal ice layers to fulfill the RTM requirements [48].

Alternatively, a machine learning technique can be employed within a radiance assimilation framework. A physically-constrained support vector machine (SVM) regression was able to accurately capture the temporal and spatial variability in the modeled T_b or ΔT_b [52–60]. Therefore, SVM is able to be applied across large spatial scales. Xue *et al.* [58] used SVM regression as the PMW ΔT_b observation operator over North America and showed improvements in snow mass estimation under certain conditions such as shallow, dry snow in the absence of forest cover.

However, the SVM-based observation operator has issues related to controllability [59]. Controllability demonstrates the skill of a linear or nonlinear model to guide the model output from any physical plausible initial state towards any physically plausible final state over a finite time period [148]. That is, the system output for a controllable system should be able to be changed by simply changing the system input [134]. For purposes of data assimilation, controllability of the SVM-based observation operator is critical during the analysis update. This is because one of the fundamental assumptions inherent to a DA framework is that model errors are assumed to correlate back to errors in the observation operator predictions [59]. In the context of a SVM-based ΔT_b observation operator, an “uncontrollable” SVM is insensitive to changes of the inputs, which eventually leads to the collapse of the ensemble of SVM-based ΔT_b predictions [59]. This generally happens when the given inputs are outside of the prediction space at which point the SVM is unable to accurately predict snow ΔT_b [60].

2.5.3 GRACE Assimilation

GRACE TWS retrievals have coarse spatial and monthly temporal resolution, and the vertical integration of the water storage components including soil moisture, groundwater, snow, and surface water in rivers, lakes, and wetlands [5]. To realize the full potential of GRACE for hydrology, GRACE TWS retrievals must be disaggregated horizontally, vertically, and in time [81]. One way to vertical disaggregate GRACE TWS retrievals is to use auxiliary information. For example, Yeh *et al.* [149] isolated groundwater storage variations from the GRACE TWS anomalies using ground-based observations of soil moisture.

Alternatively, data assimilation can be used to merge GRACE TWS retrievals with a land surface model. Despite the coarse resolution, GRACE retrievals still have reasonably reliable information about TWS anomalies [150]. Therefore, there is a potential to improve the accuracy of modeled TWS estimates via data assimilation [151]. On the other hand, the hydrological processes captured by the model can be used to enhance the satellite observations. Data assimilation approach provides downscaling and quality control of GRACE TWS retrievals while enabling synthesis of multiple observations in a physically consistent manner [81]. Most importantly, the assimilation of GRACE TWS retrievals has an impact on a number of processes within a land surface model in addition to water storage. Hence, the impacts of GRACE TWS retrievals on model predictions of water and energy fluxes can be quantified [81].

The assimilation method employed for GRACE TWS retrievals is either a “two-step” scheme or a straight application of sequential Kalman filter. Zaitchik *et al.* [81] assimilated monthly GRACE TWS retrievals using an ensemble Kalman smoother approach. That is, the land model runs twice over the course of the same month. The first run is to collect the innovations (i.e., the differences between model predictions and TWS observations). The second run is to update the model states. The TWS retrievals were assimilated on basin-averaged scale with uniformly distributed observation errors [22,82,83,152]. Later research suggests that subbasin scale TWS assimilation is more suitable [21,98,153]. However, the “two-step” scheme is computational expensive since it needs to integrate the model twice. Alternatively, a straight application of sequential Kalman filter can be applied. That is, the model only need to run for one time and the increments are applied at the end of the assimilation window [96]. Instead of assimilation basin or subbasin scale TWS observation, gridded TWS observation at 1° × 1° resolution has been assimilated into a model [8, 84, 85, 96, 153]. But the sequential method fails to capture the sub-seasonal variations in land states. The released daily GRACE TWS generated through the regularized sliding window mascon (RSWM) solution [154] provides another opportunity on TWS applications at a higher temporal resolution. All these studies demonstrate the great potential of GRACE TWS DA to improve TWS and its components.

2.5.4 Multi-sensor Assimilation

Over the past decades, the number of satellites observing the land surface hydrological cycle has increased [155, 156]. Many scientific efforts have focused on multi-sensor, multi-variate data assimilation in order to understand the hydrological cycle in a more comprehensive way. Su *et al.* [21] first performed a 1-D assimilation of GRACE along with MODIS snow cover fraction products over North America from January 2002 to Jun 2007, toward the goal of better estimation of SWE and snow depth. The results showed the multi-sensor approach provided significant improvements over a single-sensor MODIS only approach. Tian *et al.* [101] jointly assimilated GRACE TWS and soil moisture retrievals from Soil Moisture and Ocean Salinity (SMOS, [157]) using a 1-D ensemble Kalman smoother across the Australian continent. The results demonstrated that joint assimilation produced better water balance component estimates compared to GRACE or SMOS single-sensor assimilation.

The studies mentioned above integrated model TWS components over all grids within the GRACE spatial resolution. Alternatively, GRACE TWS retrievals can be interpolated and disaggregated into finer spatial and temporal scales that are similar to those of other satellite missions. For example, Zhao *et al.* [158] combined MODIS, ASMR-E observation, and GRACE daily observations with a deterministic ensemble adjustment Kalman Filter. A daily version of the GRACE TWS retrievals with a finer spatial resolution of 0.5° [154] was assimilated in this study. They found that the ASMR-E assimilation globally improved the soil moisture and snow estimates while the assimilation of GRACE improved snow estimation in high-latitudes. Khaki *et al.* [102] disaggregated GRACE retrievals (5-days) along with assimilating SMOS and AMSR-E observations together to study the water storage changes over South America. The results showed that the sub-surface water storages, especially the variations of groundwater and soil moisture have been significantly improved. Unlike the above studies, Girotto *et al.* [85] applied a spatially distributed (3-D) sequential assimilation scheme without pre-processing of the satellite data over the contiguous US. In this study, SMOS passive T_b observations and GRACE TWS retrievals were assimilated simultaneously. The results suggested that multi-sensor assimilation resulted in the better and more consistent soil moisture and groundwater estimates. They found GRACE assimilation increments can be removed by SMOS assimilation at some times. That is, the SMOS assimilation tried to undo the increments from the GRACE DA.

Chapter 3: Synthetic Identical Twin Experiment Setup

3.1 Study Domain

The study domain for this dissertation is the Volga River basin located in the Eastern European Plain (or Russian Plain) spanning from 45°N to 64°N and from 30°E to 62°E (Figure 3.1). The total area of the Volga River basin is about 1,390,000 km^2 occupying about one-third of the East European Plain. The length of the Volga River is about 3,700 km and ultimately discharges into the Caspian Sea. The main parts of the basin as delineated in Figure 3.1a are the Moskva Oka River basin (237,000 km^2), the upper Volga basin (430,000 km^2), the Kama River basin (500,000 km^2), and the lower Volga River basin (223,000 km^2) [159, 160].

The Volga river basin is located within the Atlantic-continental European climatic region (Myachkova, 1983) that is primarily influenced by the dominant eastward drift of air masses from the northern Atlantic Ocean. As solar radiation increases southward, the recurrence frequency of anticyclonic weather increases, which leads to an increase of continentality of climatic conditions and a decrease in annual precipitation [159]. As shown in Figure 3.2, the Volga River basin experiences minimal postglacial rebound, and hence, most of the gravitational signal changes in the Volga basin are ultimately caused by changes in regional water storage [162].

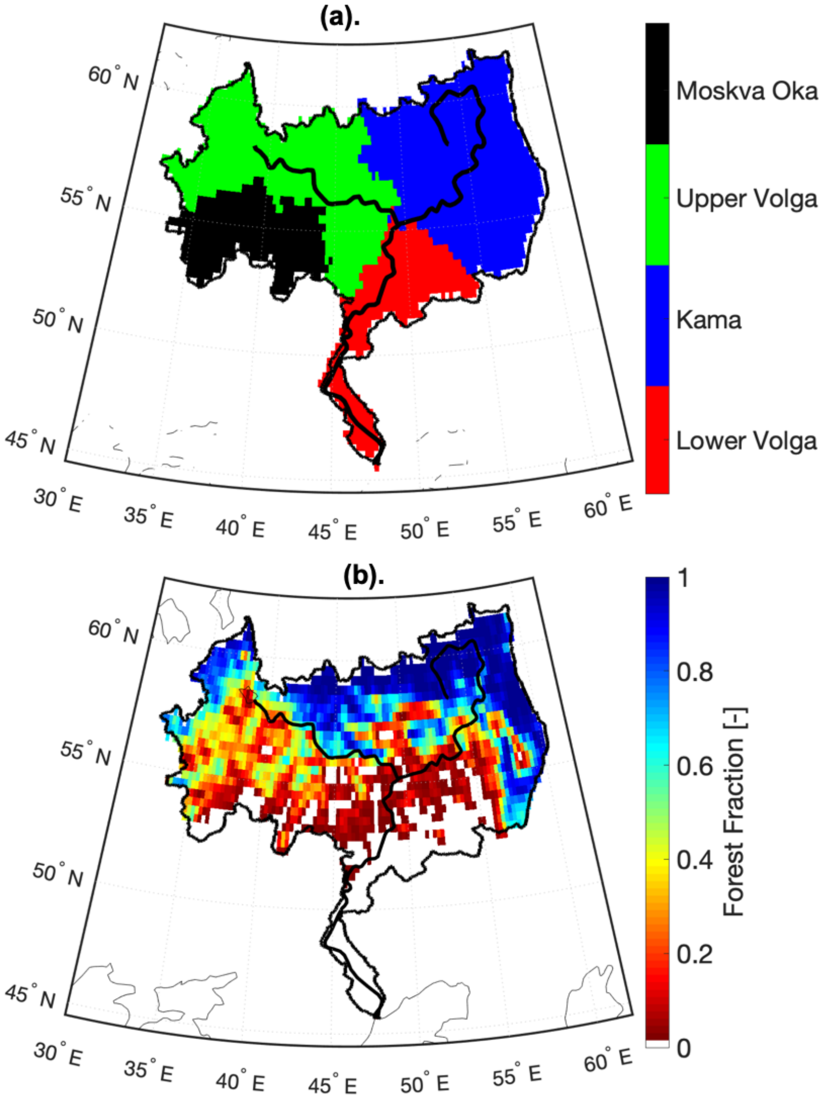

Figure 3.1: (a) The Volga River basin compasses four sub-basins, including the Moskva Oka river basin (black), upper Volga river basin (green), lower Volga river basin (red), and Kama river basin (blue). The black line is the Volga river. (b) Forest cover fraction as derived from the Moderate Resolution Imaging Spectroradiometer [161].

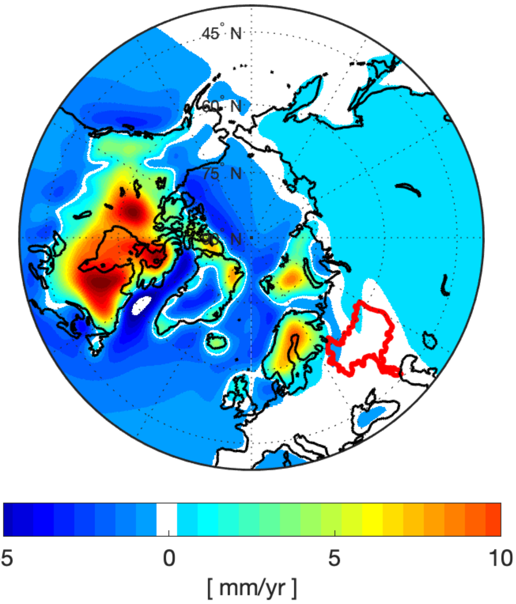

Figure 3.2: Rates of present day global post glacial rebound based on the ICE-6G Model [162]. The Volga River basin demarcated by the red outline experiences relatively little postglacial rebound.

The Volga river basin is dominated by plain landscapes such that mountainous terrain occupies less than 5% of the total basin area [159]. Approximately 80% of the total basin area has an elevation less than 200 m above the sea level [159]. The Ural Mountains are located along the eastern edge of the Volga basin and are mainly located in the Kama subbasin. The average air temperature increases from 3°C in the north to 9°C in the south. The annual precipitation decreases from 750 mm to 150 mm (southward). As a result, the snow depth decreases from 60 cm (north) to 3 cm (south) with a duration of persistence of 240 (north) to 30 days (south) [159].

In general, the Volga River basin can be divided in to three zones based on the spatial pattern of vegetation cover. Coniferous forest occupies the northern part of the Volga River basin. Steppe vegetation is dominant in the middle portion of the basin. The southern part of the basin is mainly semi-desert landscape [159]. As shown in Figure 3.1b), the forest cover fraction of the Volga River basin decreases from the north to the south where the northern area is mainly covered by coniferous forests (taiga) [159].

The Volga river has the largest annual discharge in Europe. The observed annual discharge near the river outlet is 8,380 m^3/s near Volgograd City. The main tributary of the Volga River is the Kama (Figure 3.1a) with an observed annual discharge of 4,100 m^3/s [159]. The primary source of Volga River discharge is snowmelt [159]. In addition, groundwater contributes about 30% of annual flow while rainfall provides about 10% of annual flow [159]. As a result, the Volga River regime can be characterized by a sharp and high spring snowmelt flood wave followed by low-rate flow in the summer and winter. The annual water level ranges

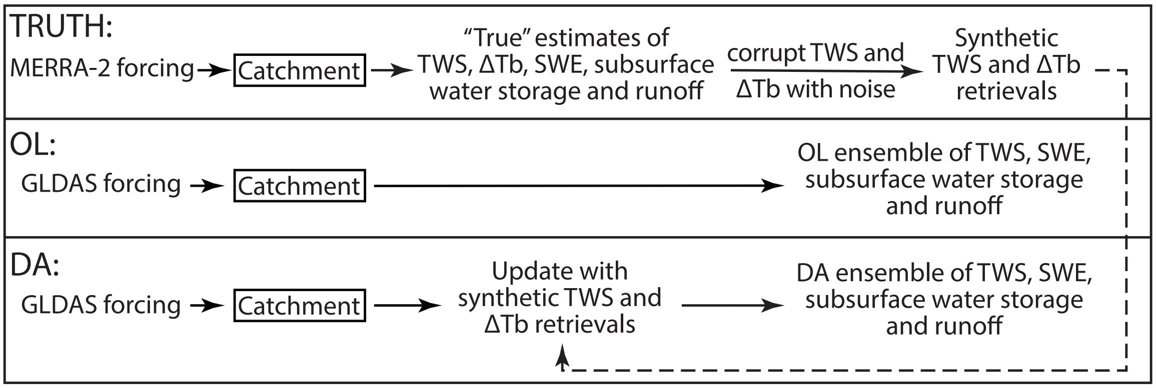

Figure 3.3: Conceptual framework for the synthetic twin experiment, including forcings (a.k.a., boundary conditions), Catchment Land Surface Model (CLSM), synthetic TWS and T_b retrieval generation, and data assimilation.

in the upper, middle, and lower Volga is 4 to 8 m, 10 to 11 m, and up to 5 m, respectively [159].

3.2 Synthetic Identical Twin Experiment

The Volga River basin study domain has few, if any, publicly-available *in-situ* measurements, such as snow, soil moisture, groundwater, or runoff observations, which motivates the application of synthetic identical twin experiment (Figure 3.3).

A synthetic experiment starts with the generation of a synthetic "truth" run to generate "true" model states. In this work, the synthetic truth run is forced by meteorological fields from the Modern-Era Retrospective analysis for Research and Applications, version 2 (MERRA-2) product [163]. More specifically, the precipitation product from MERRA-2 with corrections made by ground-based observations is used [164]. The "true" model states derived from the synthetic "truth" run serve as

Table 3.1: Forcing inputs to the synthetic truth, open loop, and data assimilation

Model run	Synthetic truth	Open loop and Assimilation
Forcing name	MERRA-2	GLDAS
Temporal resolution	1 hour	3 hour
Spatial resolution (latitude × longitude)	$0.5^\circ \times 0.625^\circ$	$2.0^\circ \times 2.5^\circ$

a reasonable proxy for the real-world system variability. Using the nominal output (geophysical) from the synthetic truth, synthetic ΔT_b and synthetic TWS observations are generated using the observation operators discussed in Section 4.4.1. The next step is to assimilate the synthetic observations into a degraded version of the same modeling system that is forced by a different ("erroneous") set of meteorological boundary conditions from the Global Land Data Assimilation System (GLDAS) product [165] (Figure 3.3). The difference between MERRA-2 and GLDAS serves as a reasonable proxy for the boundary condition errors encountered by a real (operational) DA system. The details of the meteorological boundary conditions are summarized in Table 3.1.

This type of synthetic experiment is often referred to as an "identical" twin in that the same land surface model is used in all aspects of the experiment. The major assumption behind an identical twin experiment setup is that the majority of errors encountered by a real-world system originate from the boundary conditions instead of the initial condition errors, model structure errors, or parameter errors [166]. In the context of snow and terrestrial water storage modeling, this assumption is reasonable given the large degree of precipitation error often found in remote, mountainous terrain such as that as the Volga River basin [167]. Alternatively, a “fraternal” twin experiment could be employed in which different land surface models are used during different phases of the experiment. However, this experiment setup is avoided in this current study because we want to focus on model improvements to terrestrial water storage and its components (e.g., snow) associated with erroneous forcings in the mountainous terrain.

In the present study, the spatial resolution of the model grid is the 25 km × 25 km on the Equal Area Scalable Earth (EASE version 2) computational grid [168]. The model initial conditions are generated by looping the model five times over the same 10-year period from 1 September 1992 to 1 September 2002. The model is then initialized and propagated forward from 1 September 2002 in order to initialize the model with a seasonal minimum of snowpack and runoff errors. The experiment period covers 1 September 2002 to 1 September 2011, which coincides with both the majority of the AMSR-E observations and GRACE TWS retrievals.

3.3 Boundary Condition Bias Corrections and Amendments

In the context of a synthetic identical twin experiment, it is critical to first characterize the boundary condition (e.g., precipitation) errors. Boundary condition errors often result in bias or random errors that can be considered representative of the “real-world” errors that could be encountered in an operational assimilation system [166]. In windy conditions, precipitation (snow) measurement errors can range from 20% to 50% [127]. Therefore, an error characterization strategy is employed here such that the difference between MERRA-2 forcings and GLDAS forcings can serve as a reasonable proxy for a range of plausible precipitation error scenarios. That is, the GLDAS precipitation over the study period is rescaled so that its cumulative, domain-averaged total matches 50% (negatively-biased), 100% (neutral), and 150% (positively-biased) of the corresponding total MERRA-2 precipitation by multiplying a fixed factor (γ, Table 3.2) computed as:

$$\gamma = \alpha \times \frac{\sum MERRA2}{\sum GLDAS} \tag{3.1}$$

where $\sum MERRA2$ and $\sum GLDAS$ are the cumulative, domain-averaged MERRA-2 and GLDAS boundary conditions, including snowfall and rainfall over the course of the entire study period, respectively, where α is set to 50%, 100% , and 150% to yield the rescaled GLDAS scenarios for negatively-biased, neutral, and positively-biased boundary conditions relative to MERRA-2 (synthetic truth), respectively. The EnKF assumes unbiased models which is included in the neutral scenario. However, the negatively-biased and positively-biased scenarios are included too, as to explore a range of feasible boundary conditions that could be encountered in a real-world DA system.

These three different scenarios will help explore how data assimilation can improve TWS and snow estimation where the total amount of precipitation in the study domain is under-, well-, or over-estimated. In addition, the downwelling shortwave radiation boundary conditions are also rescaled proportionally. Given the fact the correlation between shortwave radiation and precipitation is negative, the shortwave radiation is decreased as the rainfall is increased due to the present of more cloud (and vice versa). Therefore, the fixed factor (γ, Table 3.2) for shortwave radiation was computed as:

$$\gamma = \alpha \times \frac{\sum GLDAS}{\sum MERRA2} \tag{3.2}$$

where $\sum MERRA2$ and $\sum GLDAS$ are the cumulative, domain-averaged MERRA-2 and GLDAS shortwave radiation over the course of the entire study period, where α is set to 50%, 100% , and 150% to yield the rescaled GLDAS scenarios for negatively-biased, neutral, and positively-biased boundary conditions relative to MERRA-2 (synthetic truth), respectively. This last step is conducted in order to more carefully focus on the first-order control of precipitation (and its error) on TWS and snow mass assimilation. Further, simulations were conducted without the shortwave radiation correction factor and it was shown that the impact of shortwave radiation changes were secondary (or negligible) relative to that of the precipitation correction factor.

Table 3.2: Summary of GLDAS forcing correction factor γ

γ	Rainfall	Snowfall	Shortwave Radiation
Neutral precipitation	1.17	1.95	0.79
positively-biased precipitation	1.75	2.93	0.53
negatively-biased precipitation	0.58	0.97	1.58

As shown in Figures 3.4a and 3.4b), there is a strong precipitation gradient from the north to south across the study domain where the highest precipitation is in the northwest of the domain for both MERRA-2 and GLDAS. Under the positively-biased scenario (Figures 3.4c and d), the “true” precipitation (i.e., MERRA-2) is less than the precipitation forcing field used in both the OL and DA (i.e., positively-biased precipitation) with a gradient from the southeast to northwest. A similar pattern is seen for the negatively-biased scenario (Figures 3.4g and h) but with more "true" precipitation relative to the OL and DA precipitation. Even though the 9-year cumulative amount of precipitation across the domain is identical between MERRA-2 and the neutral precipitation scenarios, differences still exist at different locations in time and space between these two data sets as shown in Figures 3.4e) and 3.4f). As a result, the amount of TWS and SWE could be significantly different at different locations due to the nonlinear hydrologic response of forcing even though the domain-averaged precipitation is identical between the two.

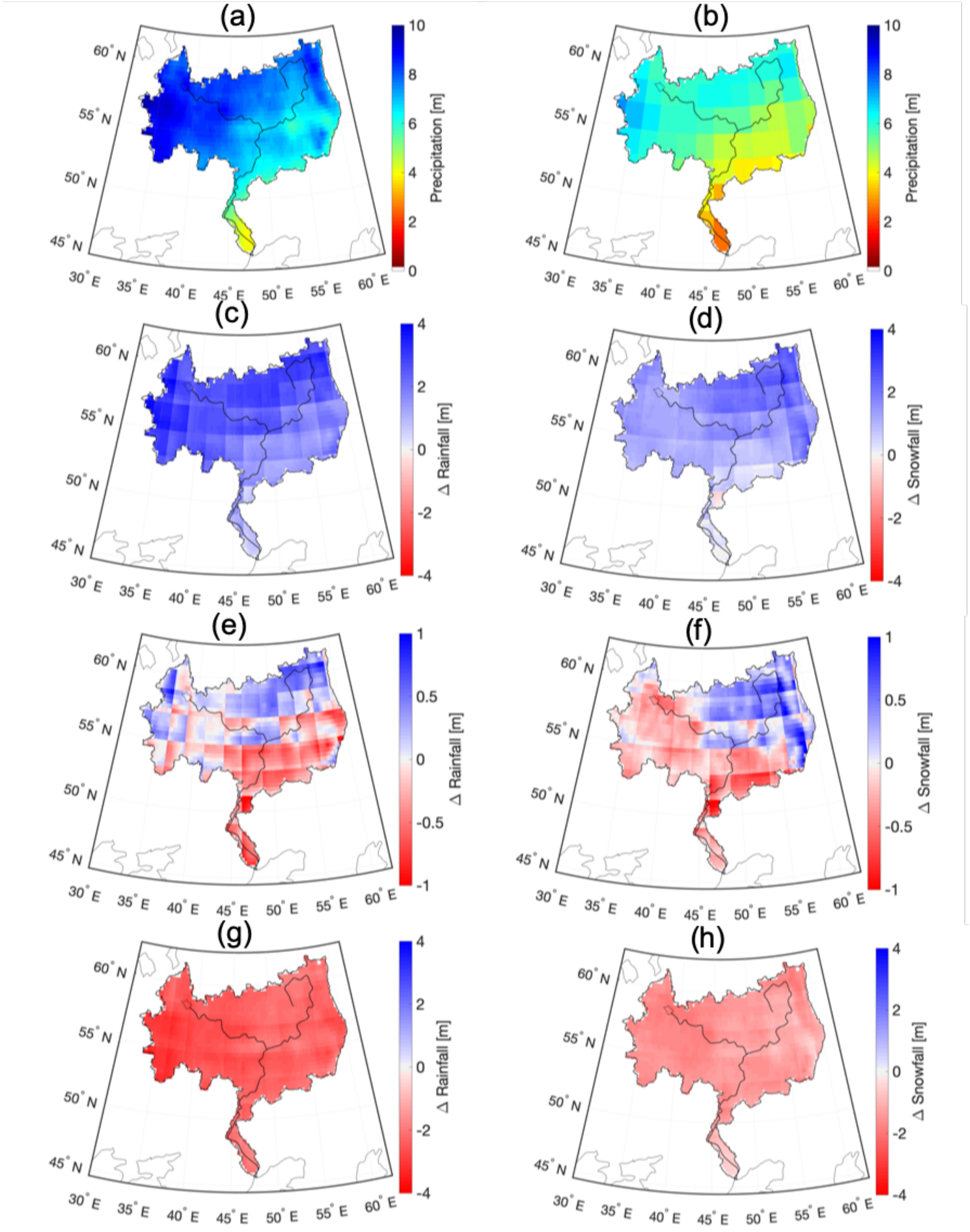

Figure 3.4: Cumulative precipitation (from years 2002 to 2011): (a) MERRA-2 precipitation, (b) GLDAS precipitation, (c) positively-biased minus MERRA-2 rainfall, (d) positively-biased minus MERRA-2 snowfall, (e) neutral minus MERRA-2 rainfall, (f) neutral minus MERRA-2 snowfall, (g) negatively-biased minus MERRA-2 rainfall, and (f) negatively-biased minus MERRA-2 snowfall.

Chapter 4: Exploration of synthetic terrestrial snow mass estimation via assimilation of AMSR-E brightness temperature spectral differences using the Catchment land surface model and support vector machine regression

4.1 Motivation and Objective

Despite the improvements in snow estimation via radiance assimilation, there are still many deficiencies that must be overcome in order to better optimize its use. For example, snow mass estimation using PMW radiometry is fundamentally an ill-posed, underdetermined system [61]. That is, there are numerous combinations of snow depth, snow density, snow temperature, snow grain size, and other snow characteristics that collectively yield the same ΔT_b observation [61]. Therefore, the task of assimilating only the snow mass-related portion of the PMW ΔT_b signal from all of the other signals inherent therein (e.g., vegetation, atmosphere, snow temperature, and snow liquid water content) is a challenge. In addition, the efficacy of SVM-based PMW ΔT_b assimilation is often limited by the controllability and reachability of SVM regression [59]. All of these issues motivate this study and help answer the question: How can we further improve snow estimation with SVM-based

PMW ΔT_b assimilation using a physically-informed approach? To this end, synthetic AMSR-E PMW ΔT_b observations are assimilated into the Catchment Land Surface Model (i.e., Catchment) [80] using SVM regression as the observation operator over snow-covered terrain in Russia.

Unlike previous works that simultaneously assimilating a fixed number of ΔT_b channels [58, 59], *a priori* modeled SWE is used as an indicator to determine which ΔT_b channels should be assimilated into the model. In addition, a simple "data-thinning" strategy is also explored to help mitigate high-frequency error (e.g., changes in snow temperature *not* related to snow mass) embedded in the synthetic AMSR-E PMW ΔT_b observations.

4.2 Synthetic AMSR-E Observations

The synthetic "truth" ΔT_bs are generated using a well-trained SVM (see section 4.4.1) that maps relevant model states (e.g., SWE, snow temperature) derived from the synthetic truth into the corresponding observation space (i.e., ΔT_b). The synthetic ΔT_b truth, including spectral difference between 10.65 and 36.5 GHz, 10.65 and 18.7 GHz, and 18.7 and 36.5 GHz at both horizontal and vertical polarizations [54, 56–58] compactly easily expressed as:

$$\Delta T_{b18V-36V} = T_{b18V} - T_{b36V} \tag{4.1}$$

where T_{b18} represents T_b at 18.7 GHz; T_{b36} represents T_b at 36.5 GHz; and the subscript V represents vertical polarization with a similar equation for horizontal (H) polarization.

The synthetic ΔT_b observations, including $\Delta T_{b10H-36H}$, $\Delta T_{b10V-36V}$, $\Delta T_{b10H-18H}$, $\Delta T_{b10V-18V}$, $\Delta T_{b18H-36H}$, and $\Delta T_{b18V-36V}$, are generated by corrupting the synthetic ΔT_b truth through the inclusion of additive, Gaussian observation noise that is temporally and spatially uncorrelated and is presented in more details in Section 4.3.

4.3 Ensemble Open Loop

As mentioned in Chapter 3, an ensemble open loop (OL) simulation is conducted without the assimilation of observations. The "imperfect" boundary conditions are established using the Global Land Data Assimilation System (GLDAS) product with a 3-hourly temporal resolution and $2.0^\circ \times 2.5^\circ$ (latitude/longitude) spatial resolution [165]. In this study, the difference between the "truth" and the OL ensemble mean is used as a representative proxy for the true system errors.

One key to the success of data assimilation is the appropriate characterization of both model and observation errors [17, 133]. An ensemble of perturbations is generally applied to the forcing variables (e.g., precipitation) as a low-rank approximation of the true system errors. The ensemble mean of the model states is typically used as the expected model estimate and the ensemble spread is used as a proxy for the model error variance [169, 170]. In line with previous works [22, 58, 83, 84, 171], the perturbation settings used in this study are summarized in Table 4.1. The additive observation error is assumed to be Gaussian-distributed with zero mean and a standard deviation of 3 K [50, 58, 170, 172].

With a finite ensemble size, only a small subset of the error space is sampled,

Table 4.1: Parameters for meteorological forcing perturbations used in the assimilation experiments.

Model Variables	Type	Standard Deviation	x, y_{corr}	t_{corr}(day)	Cross-correlation		
					pcp	sw	lw
pcp	M[a]	0.5	2°	3	NA	-0.8	0.5
sw	M	0.3	2°	3	-0.8	NA	-0.5
lw	A[b]	20 W m^{-2}	2°	3	0.5	-0.5	NA

[a] Multiultiplicative (M) or [b] Additive (A) perturbations are applied to precipitation (pcp), downwelling shortwave radiation (sw), downwelling longwave radiation (lw).

and thus, the statistical (or sampling) error is non-negligible [136, 173, 174]. The ensemble size, in part, dictates whether or not the relevant part of the error structure (e.g., error variance) can be reasonably captured by a finite ensemble size of model trajectories, and in turn, provide a reasonable low-rank approximation of the true errors. In this study, a range of ensemble sizes from $N = 14$ to $N = 74$ was tested. An ensemble size of $N = 24$ is ultimately chosen because $N > 24$ show no significant change in the ensemble spread (i.e., the ensemble SWE standard deviation over the study domain) compared to $N = 24$. Therefore, it is assumed an ensemble size of 24 replicates could reasonably represent the low-rank approximation of the true error probability distribution.

4.4 Data Assimilation

An existing, one-dimensional (1-D) ensemble Kalman filter (EnKF) framework [175,176] is employed for daily, synthetic AMSR-E ΔT_b assimilation in this study. In a 1-D EnKF, the computational units are processed independently from one another, which suggests that spatial error correlations between different catchments within the study domain are negligible [175].

During the forecast step, an ensemble of model state vectors containing the relevant model prognostic variables (e.g., SWE) are propagated forward in time by Catchment. For each domain catchment or model grid (j), the prior state vector, $x_{j,t}^{i-}$ at time t:

$$x_{j,t}^{i-} = [SWE]_{j,t}^{i-} \tag{4.2}$$

where i represents a single replicate drawn from an ensemble of size N, in this case, $N = 24$. SWE is the only model state that updated using Equation 2.9. Other snow-related states such as snow depth and snow-specific heat content are recomputed based on the updated SWE while assuming the snow density and temperature in each of the modeled snow layers remained constant before and after the update [22]. The observation predictions, $y_{j,t}^{i-}$, can be expressed as:

$$y_{j,t}^{i-} = \phi_t(SWE, slwc, Tp, Ts)_{j,t}^{i-} \tag{4.3}$$

where $slwc$ is the snow liquid water content; Tp is the top-layer soil temperature; Ts is the skin temperature; and $\phi_t(\cdot)$ is the observation operator that maps the model states into observation space (i.e., ΔT_b). The prior state vector, $x_{j,t}^{i-}$, is then updated (Equation 2.9) to a new value, $x_{j,t}^{i+}$, using the available synthetic ΔT_b observations Z_t at time t:

$$Z_{j,t} = [\Delta T_{b10H-18H}, \Delta T_{b10V-18V}, \Delta T_{b10H-36H}, \Delta T_{b10V-36V}, \Delta T_{b18H-36H}, \Delta T_{b18V-36V}]_{j,t}^{T} \tag{4.4}$$

where the "^{T}" denotes the vector transpose. As stated by Burgers *et al.* [170] , the observations should be treated as random variables at the update step (Equation 2.9). That is, random perturbations with the correct statistics must be added into the observations and generate an ensemble of observations as:

$$Z^i_{j,t} = Z_{j,t} + v^i, v \sim \mathcal{N}(0, 3^2) \tag{4.5}$$

where v^i represents the observation errors that are assumed here to be Gaussian with zero mean and a spatially- and temporally-uncorrelated covariance of 3^2 K^2 [50, 58, 170, 172].

It is worth noting that there is no DA performed around water bodies. This was due to that fact that grid cells with more than 5% coverage by water (ocean or inland water bodies) are excluded from the data assimilation analysis because surface water impoundments are not explicitly accounted for in the Catchment model. Therefore, grid cells containing surface impoundments are not included in the EnKF update.

4.4.1 Observation Operator and SVM Controllability

SVM regression served as the observation operator for mapping the geophysical states (e.g., SWE, snow temperature) into observational (i.e., PMW spectral difference) space. Following Forman *et al.* [52] and Forman *et al.* [53], SVM training used a split-sample, jackknifing procedure where observations used for validation were excluded from the training dataset. The training period was from 1 September 2002 to 1 September 2011. A fortnightly (two weeks) training period was selected to best capture seasonal variability while still providing a sufficiently large enough set for training. The inputs to SVM training were four Catchment model states relevant to PMW remote sensing of snow: (1) SWE, (2) snow liquid water content, (3) top-layer soil temperature, and (4) skin temperature. They were selected based on the results of an extensive sensitivity analysis [57]. The SVM outputs were the synthetic ΔT_b truth, including $\Delta T_{b10H-36H}$, $\Delta T_{b10V-36V}$, $\Delta T_{b10H-18H}$, $\Delta T_{b10V-18V}$, $\Delta T_{b18H-36H}$, and $\Delta T_{b18V-36V}$.

It is important to highlight the issue of controllability with the SVM-based observation operator [59]. As an important factor in optimal control theory, controllability demonstrates the skill of a linear or nonlinear model to guide the model output from any physical plausible initial state towards any physically plausible final state over a finite time period [148]. Controllability is related to the set of training data and the inability of the SVM to accurately predict snow ΔT_b when the given inputs that are outside of the prediction space implicit in the training data [60]. As a result, the model error would no longer correlate with the corresponding error in the SVM-based observation operator, which can lead to spurious error correlations that ultimately degrade the model estimate [59]. To avoid this, prior SWE was updated only when the standard deviation of the prior SVM-predicted ΔT_b was greater than 0.05 K based on heuristics outlined in [59].

4.4.2 Shallow-to-Medium versus Medium-to-Deep Snow Algorithm

Following the methods of [58] and [59], the DA experiments start with simultaneous assimilation of six different ΔT_bs, including $\Delta T_{b10H-18H}$, $\Delta T_{b10V-18V}$,

$\Delta T_{b10H-36H}$, $\Delta T_{b10V-36V}$, $\Delta T_{b18H-36H}$, and $\Delta T_{b18V-36V}$ (baseline DA; Table 4.2) in this study. However, these studies demonstrated that DA, at times, could degrade the snow estimation (also see section 4.7.1). To prevent DA degradation, a physically-informed DA approach (section 4.4.2) and a data thinning DA approach (section 4.4.3) are introduced.

An important assumption behind spectral difference assimilation is that ΔT_b is positively correlated with SWE and that the T_b at the highest frequency (i.e., 36.5 GHz) decreases as SWE increases while the lower frequency (i.e., 10.65 GHz or 18.7 GHz) T_b is relatively insensitive to increasing snow mass [30,177]. However, the correlation between SWE and T_b at 36.5 GHz can reverse once SWE is greater than 100 to 200 mm [26–30,59]. This occurrence is often referred to as signal saturation. Furthermore, ΔT_{b10-18} may introduce errors during shallow snow conditions because the signal is more representative of the soil moisture rather than the snow mass [178]. In other words, the observation error covariance of ΔT_{b18-36} during deep snow conditions or ΔT_{b10-18} during shallow snow conditions may not be adequately represented by the prescribed error parameters, and hence, may introduce spurious errors during the EnKF update. If an observational data set contains data whose errors are not well represented by the prescribed error parameters, the EnKF will not be able to accurately estimate the true fields [17]. Therefore, one hypothesis for DA degradation in the snow estimates is due to the simultaneous assimilation of all six ΔT_b when many of the ΔT_bs are not truly representative of snow mass. Further, the presence of signal saturation during deep snow conditions also result in a degraded estimate when the general assumption that ΔT_b is positively correlated with SWE is no longer the case.

Instead of simultaneously assimilating all available multifrequency and polarization spectral differences into a land surface model as done in the works of [58] and [59], a new assimilation strategy based on prior SWE information is explored here such that ΔT_b is assimilated more selectively. In this new approach, the ensemble mean of the prior SWE is used as an indicator to determine which ΔT_b should be assimilated. That is, shallow-to-medium snow conditions now only utilize $\Delta T_{b18H-36H}$, $\Delta T_{b18V-36V}$, $\Delta T_{b10H-36H}$, and $\Delta T_{b10V-36V}$ whereas medium-to-deep snow conditions now only simultaneously utilize $\Delta T_{b10H-18H}$, $\Delta T_{b10V-18V}$, $\Delta T_{b10H-36H}$, and $\Delta T_{b10V-36V}$ (physically-informed DA; Table 4.2). The SWE threshold used to differentiate between the shallow-to-medium and medium-to-deep snow is somewhat subjective. In this study, the shallow-to-medium snow refers to SWE$\leq$120 [mm] while SWE>120 [mm] is considered as the medium-to-deep snow based on peer-reviewed literature [27, 28, 179].

4.4.3 Data Thinning

Another common issue with snow ΔT_b assimilation is ensemble collapse (i.e., little or no ensemble spread) (Figure 4.1). Ensemble collapse results in an under-representation of the true model uncertainty. Given the fact the optimal combination of the observations with the model is predicated on the consideration of the respective uncertainties of each [142], a poor representation of model uncertainty will often lead to degraded snow estimation.

Ensemble collapse, in part, is exacerbated by the multi-observation nature of the assimilation approach (i.e., multiple observations assimilated daily). In addition, the high-frequency errors embedded in the AMSR-E observations (Figure 2.5) often overwhelm the snow-related information, and thus, can further degrade DA performance. In an attempt to mitigate such high-frequency noise, the synthetic ΔT_b observations (all six channels) are assimilated every 3-, 5-, 7-, 10-, and 15-day intervals rather than daily in order to explore the impact of using fewer observations on DA performance (Table 4.2).

Table 4.2: Descriptions of different ΔT_b data assimilation strategies

Name	Description
Baseline DA	Simultaneous assimilation of six ΔT_b channels including: $\Delta T_{b10H-18H}$, $\Delta T_{b10V-18V}$, $\Delta T_{b10H-36H}$, $\Delta T_{b10V-36V}$, $\Delta T_{b18H-36H}$, and $\Delta T_{b18V-36V}$
Physically-informed DA	Update SWE based on ***prior* SWE ensemble mean**: If SWE$\leq$120 [mm], use $\Delta T_{b10H-36H}$, $\Delta T_{b10V-36V}$, $\Delta T_{b18H-36H}$, and $\Delta T_{b18V-36V}$ If SWE$>$120 [mm], use $\Delta T_{b10H-36H}$, $\Delta T_{b10V-36V}$, $\Delta T_{b10H-18H}$, and $\Delta T_{b10V-18V}$
3-day Thinning DA	Simultaneous assimilation of six ΔT_b channels **every 3 days** including: $\Delta T_{b10H-18H}$, $\Delta T_{b10V-18V}$, $\Delta T_{b10H-36H}$, $\Delta T_{b10V-36V}$, $\Delta T_{b18H-36H}$, and $\Delta T_{b18V-36V}$

4.5 Normalized Innovation and Filter Optimality Assessments

The optimal operation of the Kalman filter is closely related to the statistical properties of the innovation sequence, which is the difference between the observation and model forecast [137]. In theory, the information exchange during the filter update is optimal when the normalized innovation sequence appears as white noise (i.e., mean zero with unit variance and temporally uncorrelated). If the models are unbiased and linear (both the land surface model and the observation operator) and all errors are uncorrelated and Gaussian (and correctly specified), then the normalized innovation sequence, NI, should appear similar in form to a standard normal distribution $\mathcal{N}(0,1)$ [137]. Although both the land surface model and observation operator used here are nonlinear, the investigation of the normalized innovation sequence can still provide useful information as to the performance of the DA procedure.

The normalized innovation (NI) at time t can be written as:

$$NI_t = \frac{y_t - \phi_t(x_t^-)}{\sqrt{C_{y_t y_t} + R}} \tag{4.6}$$

where the numerator is the difference (or innovation) between the synthetic ΔT_b observation (y_t) and SVM-based predicted (prior) ΔT_b observation ($\phi_t(x_t^-)$), and the denominator is the square root of the sum of the background error covariances ($C_{y_t y_t}$) and the observation error covariance (R). The normalized innovation sequence is merely the vector concatenation of all NI_t across the duration of the assimilation experiment that is explored for mean zero, unit variance, temporally-uncorrelated Gaussian-like features that can be used as a proxy for filter optimality.

4.6 Validation Approach

A synthetic, identical twin experiment is designed such that the "true" values of hydrologic states and fluxes are known. Therefore, the validation is performed against the true states (e.g., SWE) derived from the synthetic truth run. Several goodness-of-fit statistics are used for the validation activities: (1) bias, (2) root mean squared error (RMSE), (3) unbiased root mean squared error (ubRMSE), (4) correlation coefficient (R), (5) Nash-Sutcliffe efficiency (NSE), and (6) containing ratio ($CR_{2\sigma}$). In addition, normalized information contribution (NIC) is used to quantify the DA improvement (or degradation) relative to the OL [180,181]. Details about all of these calculations can be found in Appendix A.

4.7 Results and Discussion

4.7.1 SWE Estimates

The DA experiments started with simultaneous assimilation of six different ΔT_bs, including $\Delta T_{b10H-18H}$, $\Delta T_{b10V-18V}$, $\Delta T_{b10H-36H}$, $\Delta T_{b10V-36V}$, $\Delta T_{b18H-36H}$, and $\Delta T_{b18V-36V}$. For illustrative purposes, two relatively ideal locations, Grid #1 and Grid #2, (i.e., long snow season; relatively dry snow conditions; no forest cover; and relatively shallow snow such that $SWE_{max} < 200$ mm) are selected. Given the fundamental physics of PMW T_b remote sensing of snow, if assimilation does not work at these idealized locations (assuming appropriate specification of input error parameters), then assimilation will likely not work at other locations in the Volga basin. Therefore, we chose to present these locations prior to discussing results across the remainder of the basin.

Figure 4.1 highlights the performance of DA (denoted as baseline DA as shown in blue) at these idealized locations under the neutral forcing conditions. As shown in Figure 4.1a) and 4.1b), simultaneous assimilation of all six ΔT_b channels actually degraded model SWE estimates. Starting in late-January 2010 (Figure 4.1a), the DA

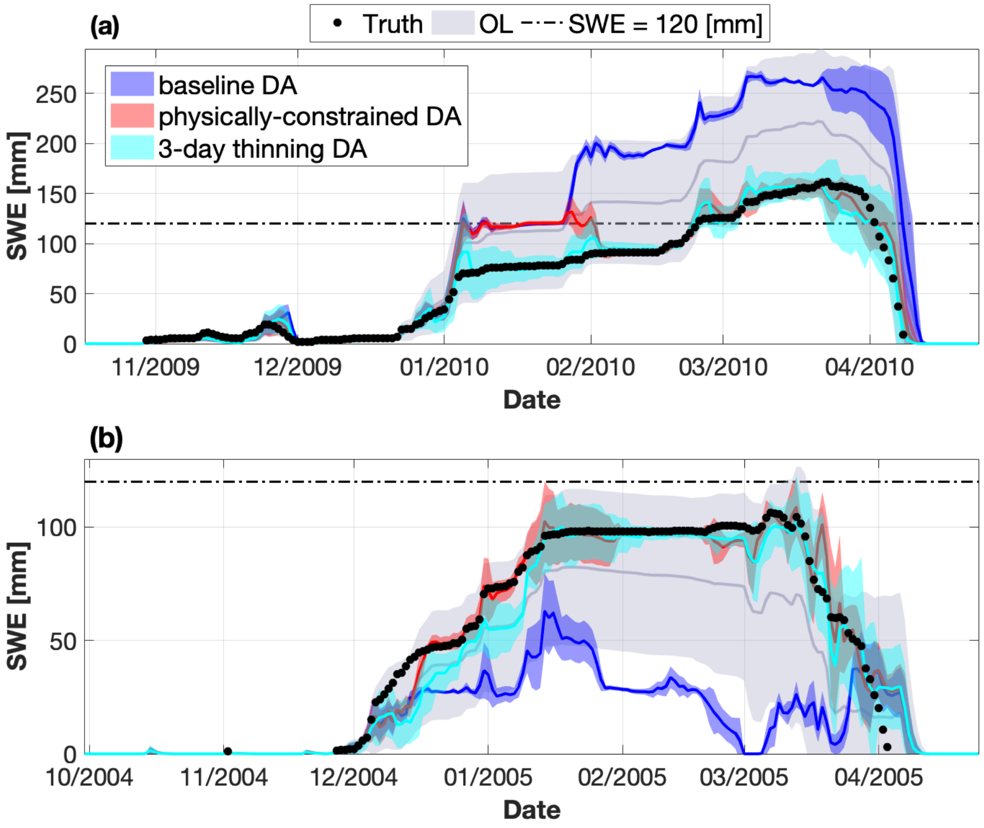

Figure 4.1: Example time series of snow water equivalent (SWE) for (a) Grid #1 (54.1685°N, 47.3343°E) from October 2009 to May 2010 and (b) Grid #2 (49.1489°N, 54.0778°E) from October 2004 to May 2005. Physically-informed DA and data thinning (3-day) improve model results whereas baseline DA (no physical constraint) actually degrades model results relative to the open loop.

SWE estimates diverged from the OL (gray color) and the synthetic truth (black dots). This divergence resulted in degraded SWE estimates with approximately 82%, 80%, and 85% increases in RMSE, bias, and ubRMSE, respectively, relative to the OL. Similarly, the baseline DA SWE estimates at Grid #2 also diverged from the synthetic truth early in the snow season (Figure 4.1b) such that the DA routine was unable to recover.

The following sections will discuss whether the physically-informed DA or the data thinning strategy could serve as a feasible solution in preventing filter divergence, and hence, DA degradation while assimilating PMW ΔT_bs.

4.7.1.1 Shallow-to-Medium versus Medium-to-Deep Snow Algorithm

As shown in Figure 4.1a), the new assimilation strategy (denoted as physically-informed DA, red color) improved the Grid #1 SWE estimates with a 58%, 80%, and 41% reduction in RMSE, bias, and ubRMSE, respectively, relative to the OL. Starting in late-January 2010 (SWE $\approx$ 120 mm), the physically-informed DA converged toward the synthetic truth (black dots) and was able to encapsulate more of the synthetic truth resulting in a larger containing ratio, $\text{CR}_{2\sigma}$ (0.26) than the baseline DA ($\text{CR}_{2\sigma} = 0.10$) (see Appendix A). These results suggest that the physically-based shallow-to-medium versus medium-to-deep snow algorithm effectively mitigated much of the negative influence of spurious correlations between SWE and ΔT_{b18-36} during deep snow conditions. Figure 4.1b) further illustrates the benefits of the physically-informed shallow-to-medium versus medium-to-deep snow algorithm during the shallow snow conditions.

As shown in Figure 4.2a) and 4.2b), the correlation coefficient (R) between the synthetic truth SWE and SVM-based synthetic ΔT_{b10-18} observations for Grid #2 was 0, whereas the correlation coefficients between the synthetic truth SWE and SVM-based synthetic ΔT_{b10-36} (Figure 4.2c, 4.2d) and ΔT_{b18-36} (Figure 4.2e, 4.2f) observations were greater than 0. This implies that ΔT_{b10-18} contained little or no information about shallow snow conditions (i.e., SWE $\leq$ 120 mm) given the fact that both 10 GHz and 18 GHz undergo little or no scattering across such a shallow snow pack [61]. That is, these T_b frequencies (and ΔT_b by construct) are effectively transparent through such shallow snow. After removing ΔT_{b10-18} from the observation vector, the physically-informed DA (red color) was able to correct the model SWE estimates towards the synthetic truth (black dots) during the middle of December 2004, which resulted in a 24%, 92%, and 24% reduction in RMSE, bias, and ubRMSE, respectively, relative to the OL. Further, the $\text{CR}_{2\sigma}$ was increased from 0.04 (baseline DA) to 0.29 (physically-informed DA), which suggested the physically-informed DA was superior to the more uniformed, baseline approach.

However, such a strategy is far from a panacea and is only effective at some locations. One possible reason is that if the prior SWE estimate is incorrect, it is possible the spectral differences used in the update are not the most appropriate. As an alternative, information from the observations can be used directly to help guide which of the ΔT_bs to assimilate into model at a given point in time and space. For example, if ΔT_{b18-36} observation suggests the spectral difference is nearing saturation, then one can assimilate the longer wavelengths as ΔT_{b10-18}. This approach will

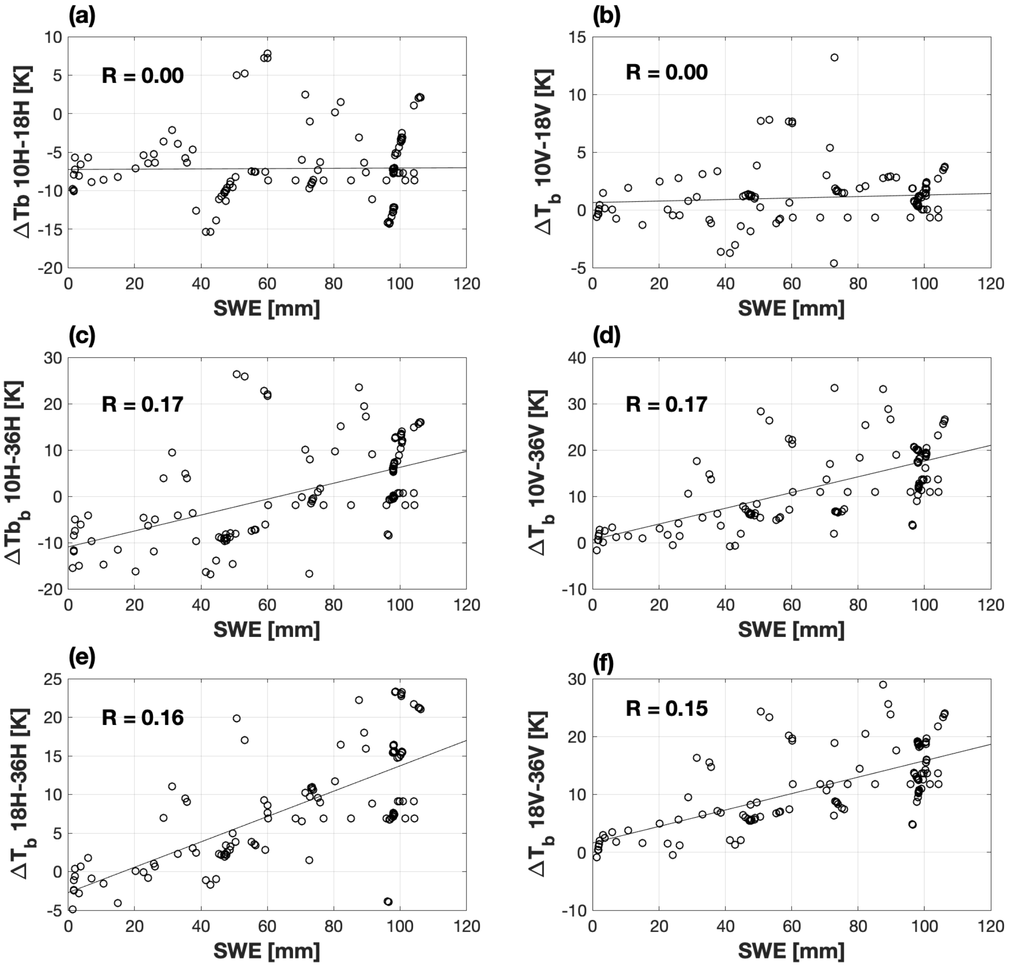

Figure 4.2: Scatter plots (with correlation in upper-left corner) between the synthetic truth SWE and the SVM-based brightness temperature spectral difference (ΔT_b) for (a) $10H - 18H$, (b) $10V - 18V$, (c) $10H - 36H$, (d) $10V - 36V$, (e) $18H - 36H$, and (f) $18V - 36V$ estimates for Grid #2 (49.1489°N, 54.0778°E) from 1 September 2002 to 1 September 2011.

be explored in a future study and is considered beyond the current project scope.

As shown in Figure 4.3, the Volga basin-averaged results suggested the physically-informed DA (bias = -5.1 mm; R = 0.818) showed limited improvements in SWE estimation over the baseline DA (bias = -5.4 mm; R = 0.817) under the neutral forcing conditions. Similarly, the $\text{CR}_{2\sigma}$ increased from 0.24 (baseline DA) to 0.25 (physically-informed DA). Over 57% of the basin grids had improved $\text{CR}_{2\sigma}$. The SWE ensemble spread (σ), which was defined as the long-term time-average of the instantaneous ensemble standard deviation, was also changed from 3.33 mm (baseline DA) to 3.56 mm (physically-informed DA), which suggested the physically-informed DA effectively inflated the ensemble spread, and had better ability to capture more of the synthetic truth. All these results suggest that the physically-informed DA marginally improved the accuracy of SWE estimation relative to the baseline DA under the neutral forcing conditions.

4.7.1.2 Data Thinning

An example of data thinning to once every three days (cyan color) for Grid #1 under the neutral forcing conditions is shown in Figure 4.1a). SWE estimates were improved with a 80%, 98%, and 70% reduction in RMSE, bias, and ubRMSE, respectively, relative to the OL. $\text{CR}_{2\sigma}$ increased from 0.055 (OL) to 0.25 implying that using fewer observations in time during the DA update can help better capture the synthetic truth. In addition, the SWE ensemble spread increased from 1.86 mm (baseline DA) to 3.02 mm (3-day thinning DA), but was significantly less than

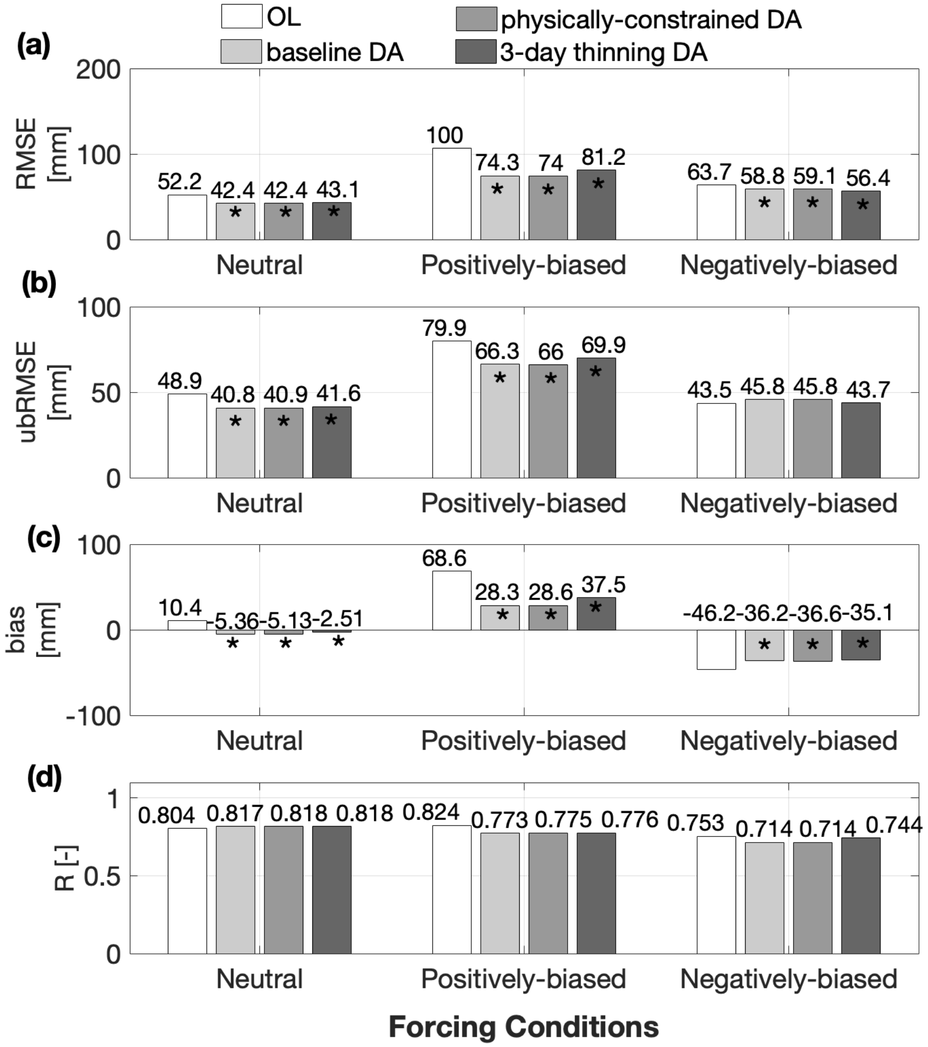

Figure 4.3: Histograms of Volga basin-averaged SWE statistics showing (a) RMSE, (b) ubRMSE, (c) bias, and (d) R under the neutral (first set), positively-biased (second set), and negatively-biased (third set) forcing conditions. The white bar is for the Open Loop (OL). The light gray bar is for baseline DA. The medium gray bar is for physically-informed DA and the dark gray bar is for DA 3-day thinning as listed in Table 4.2. Bars marked with $*$ indicate DA yields statistically significant statistics with a level of significance of 5%.

the OL (9.79 mm). The bigger SWE ensemble spread indicated the 3-day thinning DA effectively prevented ensemble collapse from January to February. The 3-day thinning strategy also significantly improved SWE estimates for Grid #2 as shown in Figure 4.1b). The 3-day thinning strategy helped prevent SWE divergence in December 2004, and as a result, yielded a 50%, 72%, and 25% reduction in RMSE, bias, and ubRMSE, respectively, relative to the OL. As a measure of the standard deviation of the errors, the decrease in ubRMSE suggested that the data thinning strategy effectively mitigated some of the introduction of high-frequency noise (random error) at both locations.

The Volga basin-averaged results under the neutral forcings are summarized in Table 4.3. SWE RMSE increased from 43 mm to 47 mm as fewer observations were assimilated into the model from once a day to every 15 days. This corroborated the earlier results that synthetic assimilation indeed added utility to the model; in the absence of assimilation (i.e., if data thinning approached an infinite amount of time) the results would revert back to the original OL results. Further, there was no statistically significant difference between the baseline DA RMSE and 3-day thinning DA RMSE. In aggregate, these results suggest assimilating the synthetic ΔT_b every 3 days yielded the same amount of SWE errors with daily assimilation.

As fewer observations were assimilated beyond every 3 days, ubRMSE increased from 41 mm (baseline DA) to 45 mm (15-day thinning DA) indicating that assimilating the noisy observations every few days did not help mitigate the random noise embedded in the synthetic ΔT_b observations. The bias, however, were statistically significant (at a level of significance of 5%) and decreased from 10 mm (OL)

Table 4.3: Domain-averaged SWE Statistics for DA thinning experiments from 1 September 2002 to 1 September 2011 under the neutral forcing conditions

Statistics	OL	DA baseline	DA thinning 3-day	5-day	7-day	10-day	15-day
RMSE [mm]	52	42	43	44	45	46	47
ubRMSE [mm]	49	41	42	43	43	44	45
bias [mm]	10	-5.4	**-2.5^a**	**-0.73**	0.16	1.3	3.6
$CR_{2\sigma}$	0.30	0.25	**0.26**	**0.27**	**0.28**	**0.28**	**0.28**

Bold^a number indicates which experiment yields statistically significant statistics relative to the Open Loop with a level of significance of 5%.

to -2.5 mm and -0.73 mm when the model simultaneously assimilated with all six channels every three and five days, respectively, rather than once a day. These results imply that daily assimilation using all six channels tended to underestimate SWE (in part due to filter divergence) and may have overconstrained the model, and hence, often resulted in degraded SWE estimation in terms of bias.

Compared to the baseline DA, all DA thinning strategies enhanced the ability to better capture the synthetic truth (i.e., larger $CR_{2\sigma}$) in part by preventing ensemble collapse, which was also proven by the bigger SWE ensemble spread of 4.6 mm (3-day thinning DA) relative to 3.3 mm (baseline DA). It can be reasonably argued that the 3-day thinning data assimilation strategy was better for SWE estimation under the neutral forcing conditions given the statistical results along with the benefit of a reduction in computational demand.

4.7.1.3 Effects of Precipitation Bias

Figure 4.3 shows the spatially-averaged statistics for the Volga river basin using three different sets of boundary (forcing) conditions. All the DA strategies (including the baseline DA, physically-informed DA, and 3-day thinning DA) had the best performance in terms of SWE estimation under the positively-biased forcing conditions in terms of the greatest reduction of RMSE relative to the OL. Compared to the OL, the RMSE was reduced by approximately 30%, 31%, and 24% with the baseline DA, physically-informed DA, and 3-day thinning DA, respectively. On the contrary, DA with the negatively-biased forcing conditions had relatively smaller improvements with approximately 7.6% (baseline DA), 7.2% (physically-informed DA), and 11% (3-day thinning DA) reduction in RMSE relative to the OL. Under the negatively-biased forcing conditions, all the DA strategies even degraded the SWE estimation in terms of more ubRMSE relative to the OL. The same results were found for bias and ubRMSE.

These results highlights a unique facet of snow assimilation – it is easier for the DA system to remove excess mass than to add missing mass. That is, in part, because the SVM can only make a prediction when snow exists, and hence, can only update the land surface model when snow is present in the model. This behavior is not unique to the SVM, but could also be said when a radiative transfer model is used as the observation operator as part of an ensemble-based DA approach for snow.

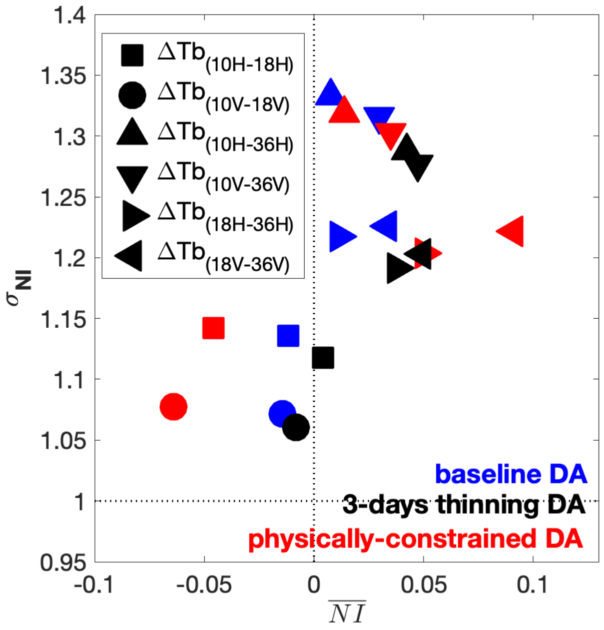

Figure 4.4: Innovation statistics for $\Delta T_{b10H-18H}$, $\Delta T_{b10V-18V}$, $\Delta T_{b10H-36H}$, $\Delta T_{b10V-36V}$, $\Delta T_{b18H-36H}$, and $\Delta T_{b18V-36V}$ shown as different marker shapes. The different marker colors represent different DA strategies as listed in Table 4.2.

4.7.2 Filter Diagnostics

Figure 4.4 shows the temporal mean ($\overline{NI}$) and standard deviation (σ_{NI}) of the normalized innovation sequence (NI) over the study domain under the neutral forcing conditions. In general, the negative $\overline{NI}$s computed at $\Delta T_{b10H-18H}$ and $\Delta T_{b10V-18V}$ suggest the SVM-based ΔT_b forecasts had a small negative bias relative to the synthetic ΔT_b observations. On the contrary, SVM-based $\Delta T_{b10H-36H}$, $T_{b10V-36V}$, $\Delta T_{b18H-36H}$ and $\Delta T_{b18V-36V}$ forecasts had positive biases relative to the synthetic ΔT_b observations.

The σ_{NI} (i.e., the standard deviation of the NI) values computed from horizontally polarized spectral differences were greater than for the vertically polarized spectral differences. In addition, the spatially-averaged σ_{NI} were greater than 1, implying that all DA strategies underestimated the observation and/or forecast errors for each frequency and polarization combination under the neutral forcing conditions. Such underestimation could be corrected using a fraternal twin experiment (rather than an identical twin experiment), but is considered well beyond the scope of this study. The σ_{NI} computed at $\Delta T_{b10H-18H}$ and $T_{b10V-18V}$ were the smallest, which can be explained by the fact $\Delta T_{b10H-18H}$, $T_{b10V-18V}$ are not as sensitive to shallow snow conditions, and hence, typically exhibit smaller variability during the entire study period.

Compared to the baseline DA (blue color), the physically-informed DA (red color) and 3-day thinning DA (black color) had a smaller σ_{NI} for each frequency and polarization combination. It suggested that the prescribed observation error characteristics are more optimal for the physically-informed DA and 3-day thinning DA compared to the baseline DA. However, the SVM-based $\Delta T_{b10H-36H}$, $T_{b10V-36V}$, $\Delta T_{b18H-36H}$ and $\Delta T_{b18V-36V}$ forecasts within the physically-informed DA and 3-day thinning DA had a relatively larger bias relative to the synthetic ΔT_b observations. The $\overline{NI}$ and σ_{NI} computed at $\Delta T_{b10H-18H}$ and $\Delta T_{b10V-18V}$ for the 3-day thinning DA were the closest to 0 and 1, respectively, relative to baseline DA and physically-informed DA. These results suggest that the observation error characteristics for ΔT_b assimilation in this study may be too simplistic. Observation error standard deviations as a function of frequency, polarization, and land cover type (i.e., forested versus non-forested) should be explored in the future.

4.7.3 Seasonality

To further investigate DA performance over the Volga Basin, the basin-averaged bias and RMSE as a function of season for positively-biased, negatively-biased, and neutral forcing conditions are presented in Figure 4.5. Similar patterns were found for ubRMSE (not shown). As expected, the baseline DA performance showed a strong seasonal component under all three forcing conditions. During the snow accumulation period, generally from September to March, DA SWE estimates outperformed OL in terms of smaller bias and RMSE.

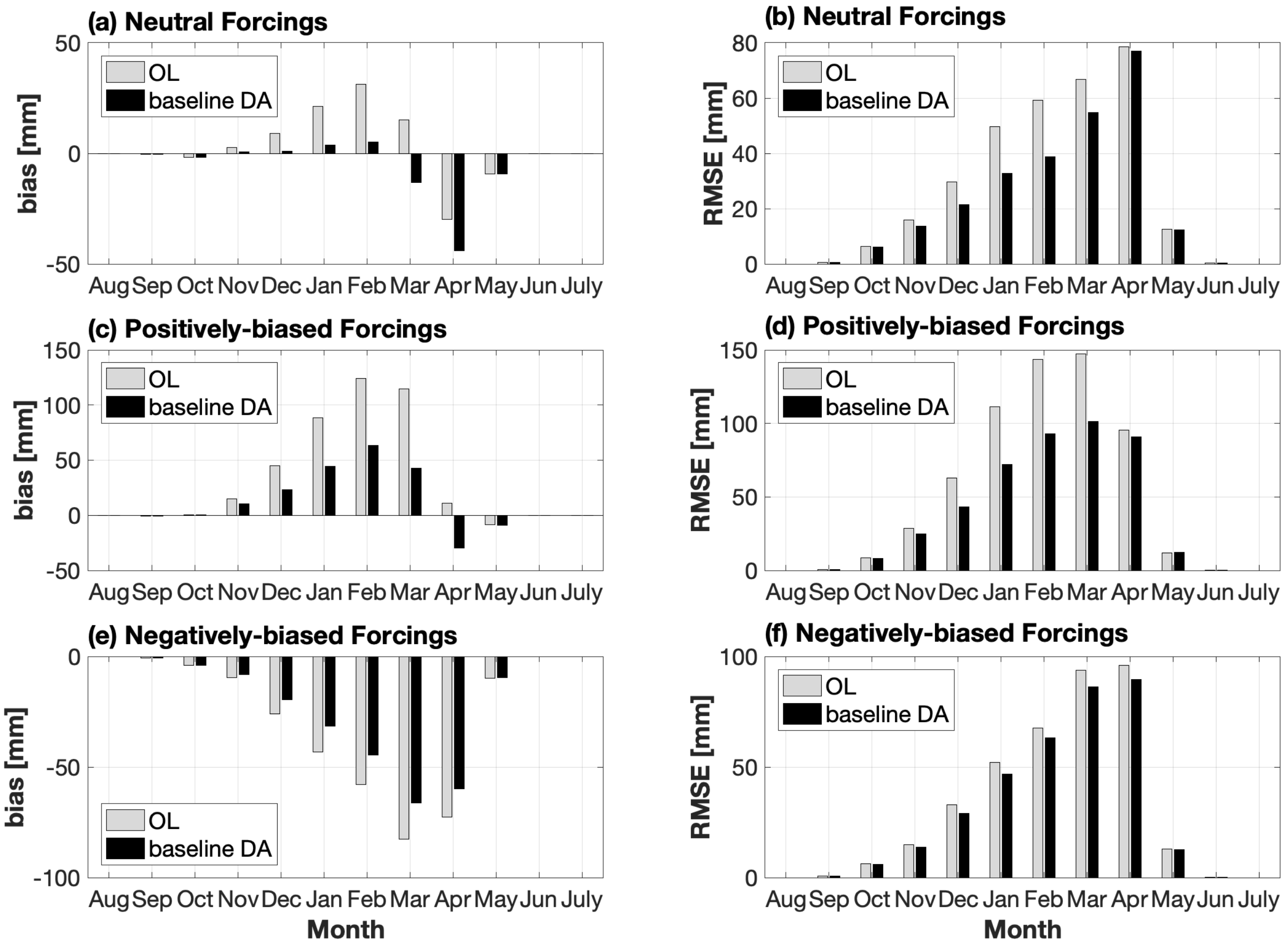

Figure 4.5: Histograms of monthly Volga basin-averaged SWE bias (first column) and RMSE (second column) under the neutral (first row), positively-biased (second row), and negatively-biased (third row) forcing conditions. Bias and root mean squared error (RMSE) were computed by comparing OL or DA SWE ensemble mean against the synthetic truth. The light gray bar is for the Open Loop (OL) and the black bar is for the baseline DA as listed in Table 4.2.

Staring in April, DA performance waned in terms of SWE estimation due to deep snow conditions and/or wet snow conditions given the limited skill of PMW remote sensing of snow [127]. As a measure of the presence of random error, ubRMSE had the largest value during April for DA SWE (not shown). The increase in ubRMSE can be explained, in large part, by the introduction of high-frequency errors originating from the synthetic ΔT_b observations along with the fact that PMW remote sensing skill is least when the snow is deep and/or wet [127]. One main reason for the degradation via DA during April was that snow liquid water (i.e., liquid water coating the snow grains) was commonplace during the snow ablation period.

It has been shown that wet snow introduces additional uncertainties in the estimation of SWE [31, 127]. The presence of liquid water within the snowpack alters the electromagnetic response from a dry microwave scatter to a wet microwave emitter [31, 32]. When the snow is wet, the general assumptions implicit in ΔT_b-based remote sensing of snow are violated [31], and hence, the information content in the ΔT_b observations need not be related to snow mass. As an example shown in Figure 4.6, the correlations between dry (gray plus signs) or wet snow (black dots) and the SVM-based $\Delta T_{b18V-36V}$ synthetic observations changed dramatically. Namely, $\Delta T_{b18V-36V}$ increased as SWE increased for dry snow. Alternatively, $\Delta T_{b18V-36V}$ transitioned to a zero (Figure 4.6a) or negative (Figure 4.6b) correlation with SWE when the snow pack ripens.

It is worth noting that DA had the worst performance in terms of SWE estimation under the negatively-biased forcing conditions in terms of smallest reduction of

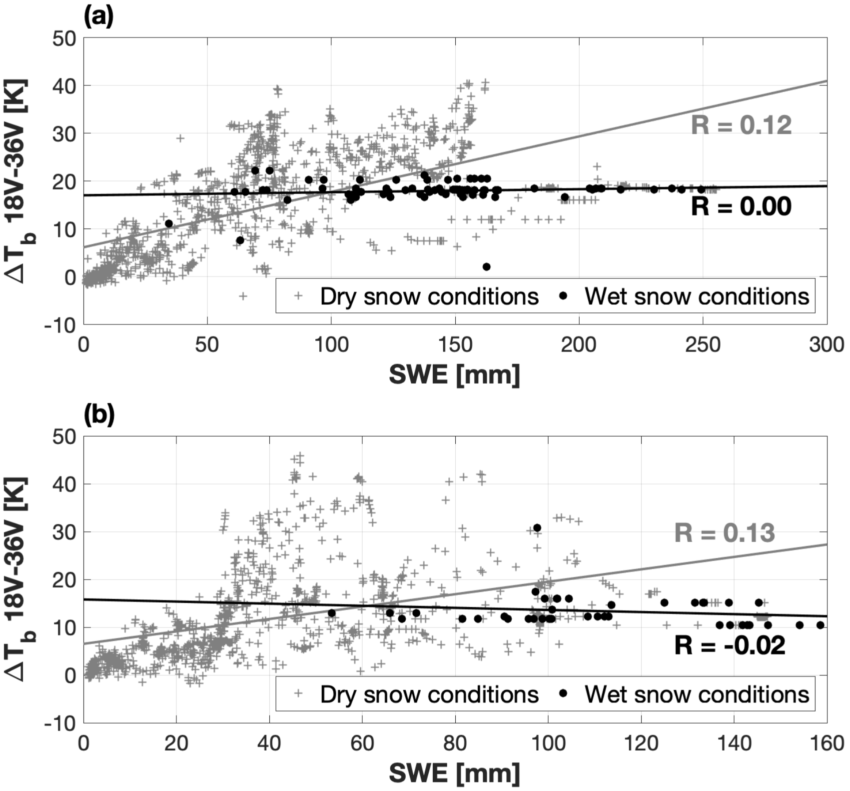

Figure 4.6: Scatter plots (with correlations) between the model dry snow (gray plus signs) and wet snow (black dots) along with SVM-based brightness temperature spectral difference $\Delta T_{b18V-36V}$ estimates for (a) Grid #1 (54.1685°N, 47.3343°E) and (b) Grid #2 (49.1489°N, 54.0778°E) from 1 September 2002 to 1 September 2011.

RMSE relative to the OL. The relatively small change in RMSE between the OL and DA suggested that DA could not significantly improve SWE estimates. In addition, DA had a larger ubRMSE than OL across the entire snow season. It suggested that ΔT_b assimilation under the negatively-biased forcing conditions was suboptimal. This latter point highlights the fact that assimilation works better at ameliorating a positive bias (positively-biased forcings) more so than a negative bias.

4.7.4 Effects of Forest Attenuation

The performance of the snow DA framework in forested regions is explored here in more detail because the presence of forest canopy can significantly alter the PMW ΔT_b signal as measured at the top of the atmosphere. More specifically, a low sensitivity of PMW ΔT_b from terrestrial snow is often observed in densely-forested areas. Overlying vegetation attenuates the PMW radiation emitted from the underlying snowpack while simultaneously adding its own contribution to the signal that is measured by the radiometer [24]. Among all these three frequency channels (i.e., 10.65 GHz, 18.6 GHz, and 36.5 GHz), microwave emission at 36.5 GHz is most strongly absorbed by standing vegetation [29]. Consequently, the scattering signal from the underlying snowpack can be overwhelmed by upwelling microwave radiation from the canopy [29].

Figure 4.7 shows NIC_{RMSE} as a function of forest fraction under the neutral forcing conditions for baseline DA. Similar results were found for other DA strategies under both the positively-biased and negatively-biased forcing conditions

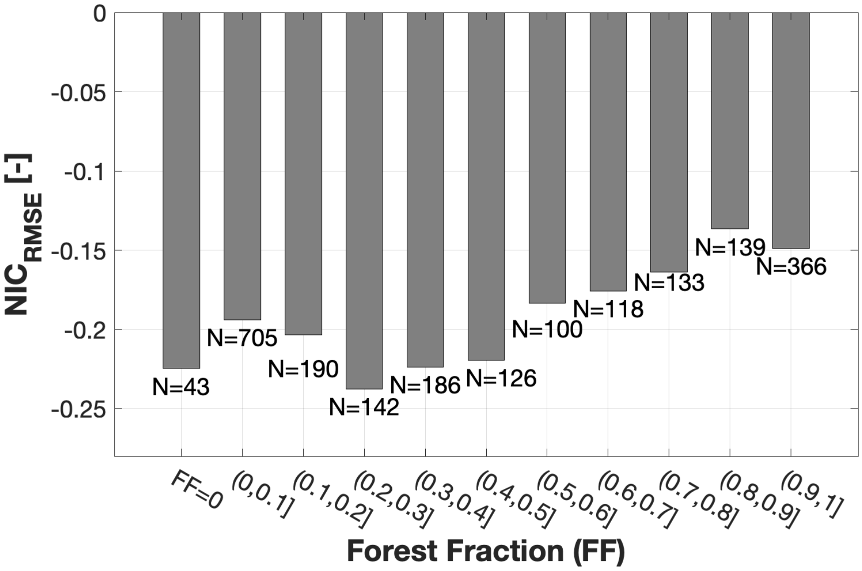

Figure 4.7: Histograms of the domain-averaged SWE NIC_{RMSE} as a function of forest fraction under the neutral forcing condition across the study domain. N is the number of model grid cells. A negative value of NIC_{RMSE} indicates data assimilation (DA) improves SWE estimates relative to the open loop (OL). Note that the largest improvements occur in the relatively sparsely-forested region where PMW attenuation is less pronounced.

(not shown). Overall, DA improved the SWE estimates relative to OL in the most sparsely-forested regions (i.e., forest fraction$\leq$0.4). A hypothesis test at a level of significance of 5% was conducted to investigate whether the forest cover had a significant effect on the DA performance. The null hypothesis was that the mean of NIC_{RMSE} for sparsely-forested areas (FF$\leq$0.4) was significantly smaller than the mean of NIC_{RMSE} for densely forested areas (FF>0.4) (i.e., the forest cover has a negative impact on DA performance). The results suggested the negative effect of forest was statistically significant for the DA algorithm.

4.7.5 Runoff Estimates

Monthly domain-averaged runoff estimates from the OL and DA were compared against true (synthetic) runoff from September 2002 to August 2011. It is encouraging to see that all basins improved runoff estimation skill with the baseline DA, physically-informed DA, and 3-day thinning DA relative to the OL under the neutral, positively-biased, and negatively-biased forcing conditions.

In general, monthly runoff in the Moskva Oka (OL bias = 0.46 mm) and lower Volga (OL bias = 0.12 mm) basins were overestimated whereas runoff in the upper Volga (OL bias = -3.8 mm) and Kama basins (OL bias = -5.5 mm) were underestimated under the neutral forcing conditions. This behavior can be explained by the spatial pattern of precipitation as shown in Figure 3.4f). MERRA-2 (synthetic truth) precipitation was greater than the neutral scenario for the OL run precipitation in the Kama and upper Volga basins, and hence, the runoff from the synthetic truth run was greater than the OL run in the Kama and upper Volga basins.

As a measure of overall hydrograph fit, Nash-Sutcliffe efficiency (NSE) was calculated for all monthly instances when either the synthetic truth or OL/DA runoff estimation was nonzero [182]. All three DA strategies had greater NSE ($NSE > 0.84$) than the OL ($NSE = 0.82$) for all four sub-basins thereby highlighting the DA skill in runoff estimation beyond simply estimating the mean of the synthetic truth under the neutral forcing conditions. In addition, DA (baseline) had better performance in the Moskva Oka (RMSE = 8.43 mm) and lower Volga (RMSE = 3.55 mm) than the upper Volga (RMSE = 15.6 mm) and Kama basins (RMSE = 15.4 mm).

For the Volga basin runoff estimation, the physically-informed DA had the best performance in terms of the greatest reduction of RMSE (relative to the OL, 30.3%) compared to the baseline DA (30.2%) and 3-day thinning DA (23.7%) under the positively-biased forcing conditions. This result further illustrated the fact that assimilation worked better when forced with a positive precipitation bias more so than a negative precipitation bias.

It is worth noting that monthly runoff estimation showed a strong seasonality effect. The spring season had the largest magnitude of runoff among the four different seasons due to the snow melt. All three DA strategies yielded better performance in the runoff estimation during the spring season compared to the OL in terms of bigger NSE and smaller RMSE as shown in Figure 4.8a) and 4.8b), respectively. Most notably during the positively-biased forcing conditions, DA strategies

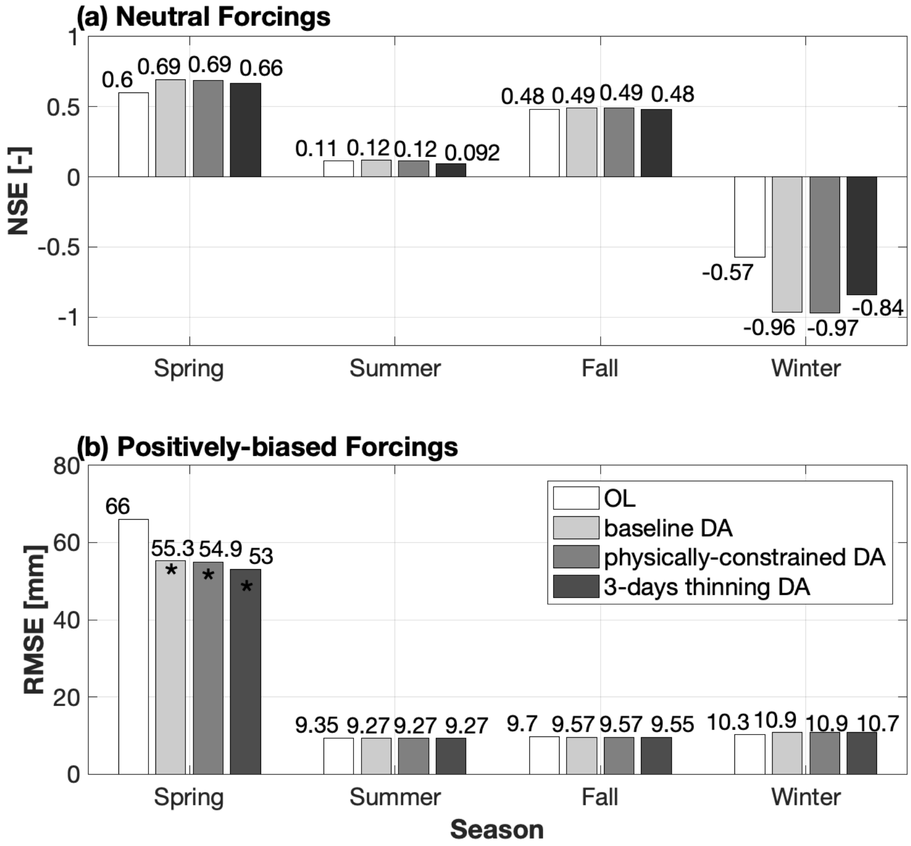

Figure 4.8: Histogram of the Volga basin monthly runoff Nash-Sutcliffe efficiency (NSE) under the neutral forcing conditions and RMSE under the positively-biased forcing conditions. (a) histogram of the Volga basin monthly runoff Nash-Sutcliffe efficiency (NSE) under the neutral forcing conditions and (b) RMSE under the positively-biased forcing conditions. Bars marked with $*$ indicate which experiment yields statistically significant statistics with a level of significance of 5%.

showed significant improvement over the OL (Figure 4.8b). These results suggest DA effectively improved the model performance in capturing relatively high runoff.

4.8 Conclusions

A series of synthetic twin experiments were conducted to explore improvements in the estimation of SWE in the Volga basin based on prescribed precipitation errors. An ensemble Kalman filter (EnKF) was used to merge synthetic PMW brightness temperature spectral differences (ΔT_b) into the NASA Catchment land surface model where well-trained support vector machines served as the observation operator.

The results suggested that simultaneous assimilation of $\Delta T_{b10H-36H}$, $\Delta T_{b10V-36V}$, $\Delta T_{b10H-18H}$, $\Delta T_{b10V-18V}$, $\Delta T_{b18H-36H}$, and $\Delta T_{b18V-36V}$ could degrade SWE estimation due to divergence from the synthetic truth at some experimental locations. One reason for DA degradation was due to simultaneous assimilation of all six ΔT_b channels and the presence of signal saturation during deep snow conditions. To help mitigate this degradation, a physically-informed approach that used the prior SWE ensemble mean as an indicator was explored. That is, $\Delta T_{b18H-36H}$, $\Delta T_{b18V-36V}$, $\Delta T_{b10H-36H}$, and $\Delta T_{b10V-36V}$ were assimilated during shallow-to-medium snow conditions (i.e., SWE $\leq$ 120 mm), while simultaneously assimilating $\Delta T_{b10H-18H}$, $\Delta T_{b10V-18V}$, $\Delta T_{b10H-36H}$, and $\Delta T_{b10V-36V}$ during medium-to-deep snow conditions (i.e., SWE $>$ 120 mm). The physically-informed assimilation approach helped improve SWE estimation at some locations but not all.

In addition, a simple data thinning assimilation strategy was explored to further mitigate the high-frequency noise embedded in synthetic AMSR-E ΔT_b observations. That is, the ΔT_b channels were assimilated every 3-, 5-, 7-, 10-, and 15 days rather than daily. The results suggested DA with 3-day data thinning modestly reduced Volga basin averaged bias from -5.5 mm to -2.5 mm under the neutral forcing conditions. $\mathrm{CR}_{2\sigma}$ was slightly increased from 0.25 (baseline DA) to 0.26 (3-day thinning DA).

DA performance under the neutral, positively-biased, and negatively-biased forcing conditions were investigated. The results suggest AMSR-E ΔT_b DA performed the best under the positively-biased conditions in terms of SWE estimation (relative to the OL). This highlights a unique facet of snow assimilation that it is easier for the DA system to remove excess mass than to add missing mass. This is, in part, due to the fact that the snow-centric DA update can only happen when snow exists in the land surface model.

The investigation in forested regions highlighted the significant negative impact of dense forest on SWE estimation. This is due to the fact that the presence of forest canopy can further alter the PMW ΔT_b signal as measured at the top of the atmosphere. Given the physical limitations of coarse-scale PMW radiometry of snow in forested areas, such scenarios should likely be excluded from the snow DA update in densely-forested areas.

SWE estimation demonstrated a strong seasonality. That is, DA SWE estimates outperformed OL in terms of smaller RMSE, bias, and ubRMSE during the snow accumulation period. However, DA SWE estimates were often degraded during the ablation period due to the presence of liquid water coating the snow grains. The reason for this is that the presence of liquid water within the snowpack elicits a shift in the electromagnetic response from a dry microwave scatter to a wet microwave emitter, and hence, the assumptions implicit in ΔT_b-based remote sensing of snow are regularly violated. The results of runoff estimation also showed a seasonal pattern. Among all four seasons, DA runoff estimates had the best performance relative to the OL during the spring season. This was consistent with the fact that DA SWE estimates were the best during the winter season, and therefore, the runoff derived from snowmelt was vastly improved during the spring season.

In summary, this study investigated new and novel AMSR-E ΔT_b assimilation approaches to further improve SWE estimation. Although the physically-informed and data thinning approach showed no significant improvement in the ensemble mean estimate of SWE, the ensemble spread increased, and as such, was able to encapsulate the synthetic SWE truth. Further, the inflated ensemble spread generally provided more reasonable representation of model SWE uncertainty.

Chapter 5: Estimating Terrestrial Snow Mass via Multi-sensor Assimilation of Synthetic AMSR-E Brightness Temperature Spectral Differences and Synthetic GRACE Terrestrial Water Storage Retrievals

5.1 Motivation and Objective

In this chapter, a multi-variate, multi-sensor assimilation framework is proposed to enhance model-based terrestrial snow estimates across a large watershed in Russia. The feasibility and benefits of jointly assimilating synthetic AMSR-E ΔT_b and GRACE TWS retrievals into the Catchment land surface model (CLSM; Section 2.2) [80] in Volga basin (Section 3) are investigated using an identical twin synthetic experiment [175]. To assess the performance of multi-sensor assimilation, the Open Loop model (OL; model without assimilation), assimilation of synthetic AMSR-E ΔT_b only, assimilation of synthetic GRACE TWS only, and joint assimilation of synthetic AMSR-E ΔT_b and GRACE TWS are systematically evaluated. The study domain lacks sufficient ground-based snow measurements to evaluate the hydrologic cycle estimates produced by the land surface model. Therefore, a synthetic, identical twin experiment [175] was employed in this study (discussed in Section 3).

The major motivation and rationale for the synthetic identical twin experiment is that "true" model states can be generated and subsequently harnessed as a systematic means of evaluating land surface model improvements with and without assimilation (Figure 3.3).

5.2 Synthetic Truth and Synthetic Observations

The synthetic "truth" is generated from the best representation of the land surface model. Afterwards, state variables from this model representation, including SWE, soil moisture, runoff, and surface energy fluxes, are perturbed with a realistic amount of observation error (details provided below) as to generate a synthetic set of observations (retrievals) for later use during assimilation that represents a reasonable and realistic representation of those encountered in the real-world.

5.2.1 Synthetic AMSR-E ΔT_b Observations

As discussed in Section 4.2, the synthetic ΔT_b "truth" is generated using a well-trained SVM that maps the relevant model states (i.e, SWE, snow liquid water content, top-layer soil temperature, and skin temperature) derived from the synthetic truth into the corresponding observation space (i.e., ΔT_b) at 25-km spatial resolution and daily temporal resolution. The synthetic ΔT_b observations include $\Delta T_{b10H-36H}$, $\Delta T_{b10V-36V}$, $\Delta T_{b10H-18H}$, $\Delta T_{b10V-18V}$, $\Delta T_{b18H-36H}$, and $\Delta T_{b18V-36V}$.

5.2.2 Synthetic GRACE TWS Retrievals

The synthetic GRACE TWS retrievals are generated to mimic the level-3, monthly GRACE TWS (unscaled) product based on the RL05 spherical harmonics [183]. The generation of GRACE TWS “truth” starts with a linear observation operator (Equation 2.1) to vertically aggregate the monthly model states from a 25-km model grid. Next, a Gaussian spatial averaging function with a 300-km full width at half maximum is used to spatially aggregate the modeled TWS on the 25-km model grid up to a spatial resolution of 1° × 1°, which corresponds to the GRACE TWS retrieval spatial resolution [85]. The last step is to corrupt the synthetic GRACE TWS “truth” with zero-mean, additive Gaussian noise that is assumed to be temporally uncorrelated [21, 22, 81, 98]. Given the fact that the effective spatial resolution of the GRACE TWS retrievals are >150,000 km^2, the errors in neighboring GRACE TWS retrieval grid cells on the 1° × 1° grid (∼10,000 km^2) are spatially correlated [7, 105]. Therefore, the prescribed Gaussian errors are generated with a spatial correlation length (radius) of 300 km [98]. The horizontal error correlation is directly computed with the function:

$$\rho(b_i, b_j) = \frac{exp(-d(b_i, b_j))}{\lambda} \tag{5.1}$$

where $d(b_i, b_j)$ is the distance between grids b_i and b_j; and λ is the error correlation length (i.e., 300 km).

5.3 Model Perturbation Setup

One important factor in the success of data assimilation is the appropriate characterization of the model forecast errors, which are normally modeled by perturbing the forcing variables (e.g., precipitation) or/and model prognostic variables of each ensemble member [17, 133]. Such perturbations can be temporally and/or spatially correlated [137, 184, 185].

In line with previous work [22, 27, 58, 81, 83–85, 171], horizontal correlation lengths of the forcing fields, including the precipitation, shortwave radiation, and longwave radiation, are set to be isotropic at 2° while the temporal correlations are chosen as three days in order to represent the error scale of precipitation dynamics [175]. Error cross correlations are also imposed between precipitation and shortwave radiation (-0.8), precipitation and longwave radiation (0.5), and shortwave and longwave radiation (-0.5). The model prognostic state perturbations are chosen relative to the particular characteristics of the observations [84, 85, 137, 184, 185]. This work adopted the model state perturbation settings from previous GRACE only assimilation studies [84] but with a revised SWE error perturbation spatial length of 0.25° to correspond with the spatial resolution of the synthetic AMSR-E ΔT_b observations (i.e., 25-km). The perturbation settings are summarized in Table 5.1. An ensemble size of $N = 24$ replicates was chosen as a suitable representation of the low-rank approximation of the true error probability distribution [22, 58, 84, 85].

Table 5.1: Parameters for meteorological forcing perturbations in the assimilation experiments

					Cross-correlation		
Model Variable	Type	Standard Deviation	x, y_{corr}	t_{corr}(day)	pcp	sw	lw
pcp	M[a]	0.5	2°	3	NA	−0.8	0.5
sw	M	0.3	2°	3	−0.8	NA	−0.5
lw	A[b]	20 W m^{-2}	2°	3	0.5	−0.5	NA
catdef	A	0.15 kg m^{-2} h^{-1}	2°	1			
srfexc	A	0.06 kg m^{-2} h^{-1}	2°	1			
swe	M	0.0012	0.25°	1			

[a]Multiplicative (M) or [b]additive (A) perturbations are applied to precipitation (pcp), downwelling shortwave radiation (sw), longwave radiation (lw), catchment deficit (catdef), surface excess (srfexc), and snow water equivalent (swe). Spatial correlations are indicated as x, y_{corr} and temporal correlations as t_{corr}.

5.4 Data Assimilation Methods

In this study, three sets of assimilation experiments were conducted:1) synthetic AMSR-E ΔT_b only assimilation, 2) synthetic GRACE TWS only assimilation, and 3) joint (a.k.a. dual) assimilation of synthetic AMSR-E ΔT_b and GRACE TWS retrievals. The synthetic single-sensor assimilation experiments (i.e., AMSR-E only or GRACE only) were conducted first as comparisons of marginal gain relative to the synthetic, joint assimilation experiment. All assimilation experiments were performed using the ensemble Kalman filter (EnKF; Section 2.5), which is a Monte Carlo implementation of a nonlinear filtering problem [135].

The calculation of increments in a one-dimensional (1-D) EnKF (section 4.4) is different from that of a three-dimensional (3-D) EnKF (section 5.4.2). In the 1-D EnKF, the computational units (i.e., model grids) are processed independently from one another, which suggests only error correlations between model state variables from the same grid are taken into account [175]. On the contrary, the 3-D EnKF takes into account horizontal error correlations of model states from different (adjacent) model grids.

5.4.1 Synthetic AMSR-E ΔT_b Assimilation

As discussed in Section 4.4, the synthetic, daily AMSR-E ΔT_b assimilation framework employed the 1-D EnKF [58]. Only SWE was updated in the model via simultaneous assimilation of $\Delta T_{b10H-36H}$, $\Delta T_{b10V-36V}$, $\Delta T_{b10H-18H}$, $\Delta T_{b10V-18V}$, $\Delta T_{b18H-36H}$, and $\Delta T_{b18V-36V}$ observations (Table 5.2). Other snow-related states (i.e., snow depth and snow-specific heat content) were recomputed based on the updated SWE assuming the snow density and snow temperature in each of the modeled snow layers remained constant before and after the update [22].

Table 5.2: Parameters for the single-sensor AMSR-E ΔT_b and GRACE TWS assimilation frameworks. The same observation types and parameters are also used during dual assimilation (AMSR-E + GRACE DA).

Framework Details	AMSR-E ΔT_b DA	GRACE TWS DA
DA type	one-dimensional (1-D) EnKF	three-dimensional (3-D) EnKF
State vector ($x^j_{i,t}$)	SWE	catdef and SWE
[a]Synthetic Obs.	$\Delta T_{b10H-36H}$, $\Delta T_{b10V-36V}$, $\Delta T_{b10H-18H}$ $\Delta T_{b10V-18V}$, $\Delta T_{b18H-36H}$, $\Delta T_{b18V-36V}$	TWS
[b]Obs. Spatial Resolution	EASEv2 25 km	3° × 3°
[c]Obs. Temporal Resolution	Daily	Monthly
[d]Obs. σ_{error}	3 K	15 mm
[e]Obs. ρ_{error}	–	2°
[f]r_I	–	6°

[a] Synthetic Obs. is synthetic observations. [b] Obs. Spatial Resolution is observation spatial resolution. [c] Obs. Temporal Resolution is observation temporal resolution. [d] Obs. σ_{error} is observation error standard deviation. [e] Obs. ρ_{error} is observation error spatial correlation. [f] r_I is localization radius.

5.4.2 Synthetic GRACE TWS Assimilation

Synthetic monthly GRACE TWS assimilation employed a 3-D EnKF framework that can effectively distribute information horizontally as well as vertically in the TWS space [84,85]. Two model states, including catchment deficit (catdef) and SWE, in each 25-km model grid were updated by assimilating synthetic GRACE TWS retrievals each month within a prespecified radius of influence (see $H(x^{j-})|_{I,T}$ in Equation 5.2; Table 5.2) . The update equation can be expressed as:

$$x_{i,t}^{j+} = x_{i,t}^{j-} + K_{I \to i, T \to t}(Y_{I,T} - H(x^{j-})|_{I,T}) \tag{5.2}$$

where $x_{i,t}^{j-}$ and $x_{i,t}^{j+}$ represent the j^{th} replicate of model states (i.e., catdef and SWE) before and after the update on day t within the month T at a 25-km (i) model grid, respectively. $Y_{I,T}$ is the perturbed, synthetic GRACE TWS retrievals for the month T. Given the fact that the effective spatial resolution of the GRACE TWS retrievals (and the synthetic GRACE TWS retrievals by construct) is around 300-400 km [183], only one 1° × 1° synthetic TWS observation out of every three in the latitude and longitudinal directions were used in the update routine [85]. This was done so that the spatial resolution of assimilated synthetic TWS retrievals was effectively 3° × 3° (I) and served as a means of mitigating the presence of spatially-correlated errors with the 1° × 1° TWS retrievals. $H(\cdot)|_{I,T}$ is the linear observation operator (Equation 2.1) that maps the model states into TWS observation predictions within the monthly time window (T).

The model TWS estimates at 25-km resolution were aggregated spatially up to

$1^\circ \times 1^\circ$ resolution (to match with the assimilated synthetic TWS retrievals) by using a Gaussian filter with a 300 km full width at half maximum. Next, a covariance localization technique was applied in order to collect all of the observation predictions within a prespecified influence radius ($r_I = 6^\circ$; Table 5.2) centered around the given model grid (i) [84, 85, 175, 186]. The daily model TWS estimates were converted to monthly TWS estimates using a simple aggregation operator calculated as:

$$H(x^{j-})|_{I,T} = \frac{1}{N_{days}} \frac{1}{N_{grid}} \sum_{t=t_1}^{N_{days}} \sum_{k=1}^{N_{grid}} [TWS_{i,j,t} \times F(dist, area)] \tag{5.3}$$

where N_{days} is the number of days in the month T and N_{grid} is the number of 25-km model grid cells within the influence radius ($r_I = 6^\circ$; Table 5.2) around the given model grid (i). $TWS_{i,j,t}$ is the j^{th} replicate of the model grid (i) terrestrial water storage for a given day (t) in the month T. $F(dist, area)$ is a weighting factor that was calculated based on the model grid area ($area$) and distance ($dist$) between the model grid (i) and the synthetic TWS observation (I).

$K_{I\to i,T\to t}$ is the Kalman gain matrix weighing the relative uncertainties of the model forecast and the observations. The gain transforms the TWS innovations (i.e., $Z_{I,T} - H(x^{j-})|_{I,T}$) in observation space into the analysis increments (i.e., $\Delta x^j_{i,t} = x^{j+}_{i,t} - x^{j-}_{i,t}$) in model grid space, and can be expressed as:

$$K_{I\to i,T\to t} = C^-_{x_{i,t}H(x^j)|_{I,T}} [C^-_{H(x^j)H(x^j)|_{I,T}} + R_{I,T}]^{-1} \tag{5.4}$$

where $C^-_{x_{i,t}H(x^j)}$ is the error cross-covariance between the modeled states, including catdef and SWE, and the observation prediction (TWS) prior to the update; $C^-_{H(x^j)H(x^j)|_{I,T}}$ is the sample error covariance of the j^{th} replicate of monthly (T) TWS observation predictions prior to the update; and $R_{I,T}$ is the observation error covariance that is a temporally-uncorrelated Gaussian error (zero mean; standard deviation equal to 15 *mm*; see Table 5.2). The spatial error correlation length is set to be 2° as noted in Table 5.2 [84, 85].

The analysis increments (i.e., $K_{I\to i,T\to t}(Y_{I,T} - H(x^{j-})|_{I,T})$ in Equation 5.2) are first calculated for each day (t) within the month T. Next, the average of the daily increments are applied to the initial model states on the first day of the month T and the model is then restarted with the updated initial conditions.

5.4.2.1 Physically-Informed TWS Update

SWE is expected to be the dominant component of TWS that changes during the snow season in high latitude basins [20] such as the Volga basin explored in this study. During the snow accumulation season in these cold, high latitude basins, there is relatively little (or no) snowmelt that can infiltrate into the subsurface. Hence, the subsurface water storage remains relatively unchanged during the snow accumulation season as compared to other times of the year when snowmelt or rainfall can easily infiltrate into the subsurface, and hence, alter the subsurface storage. Therefore, snow mass change is the main contributor to the variability of TWS (instead of subsurface water storage change) during the snow accumulation season. When the snow starts to melt and infiltration increases, both the subsurface water storage changes and the snow mass changes exert a large influence on TWS change.

In consideration of these first-order physics as applied in snow-dominated, high latitude basins, this study modified the original application of the analysis increments (denoted as baseline GRACE DA; see section 5.4.2). That is, a heuristic is applied to each ensemble replicate such that when snow accumulates on the ground without any snowmelt (i.e., SWE $>$ 0 and snmelt $=$ 0), the modified SWE increments are equal to the sum of SWE and the negative value of catdef increments calculated based on Equation 5.2. In other words, the computed increments via Equation 5.2 are assumed to operate only on SWE because it is assumed that SWE is the only component of TWS that is changing. During the snow ablation phase (i.e, SWE $>$ 0 and snmelt $>$ 0), the modified SWE increments are equal to the half of the sum of SWE and the negative value of catdef increments while the other half goes to the modified catdef increments (Table 5.3). In other words, the computed increments via Equation 5.2 are assumed and operate equally between SWE and subsurface storage. It is worth reiterating here that a negative catdef yields a positive change in subsurface storage (Equation 2.1). This new, modified GRACE DA is denoted as Physically-informed GRACE DA hereafter.

5.4.3 Synthetic Dual Assimilation

The dual assimilation (AMSR-E + GRACE DA) merges the daily AMSR-E ΔT_b DA with the monthly GRACE TWS DA [85]. It is a “two-pass” procedure that combines the concepts of the single-sensor AMSR-E DA and GRACE DA as illustrated in Figure 5.1. During the first pass for a given month, GRACE TWS assimilation is conducted provisionally in conjunction with assimilation of the AMSR-

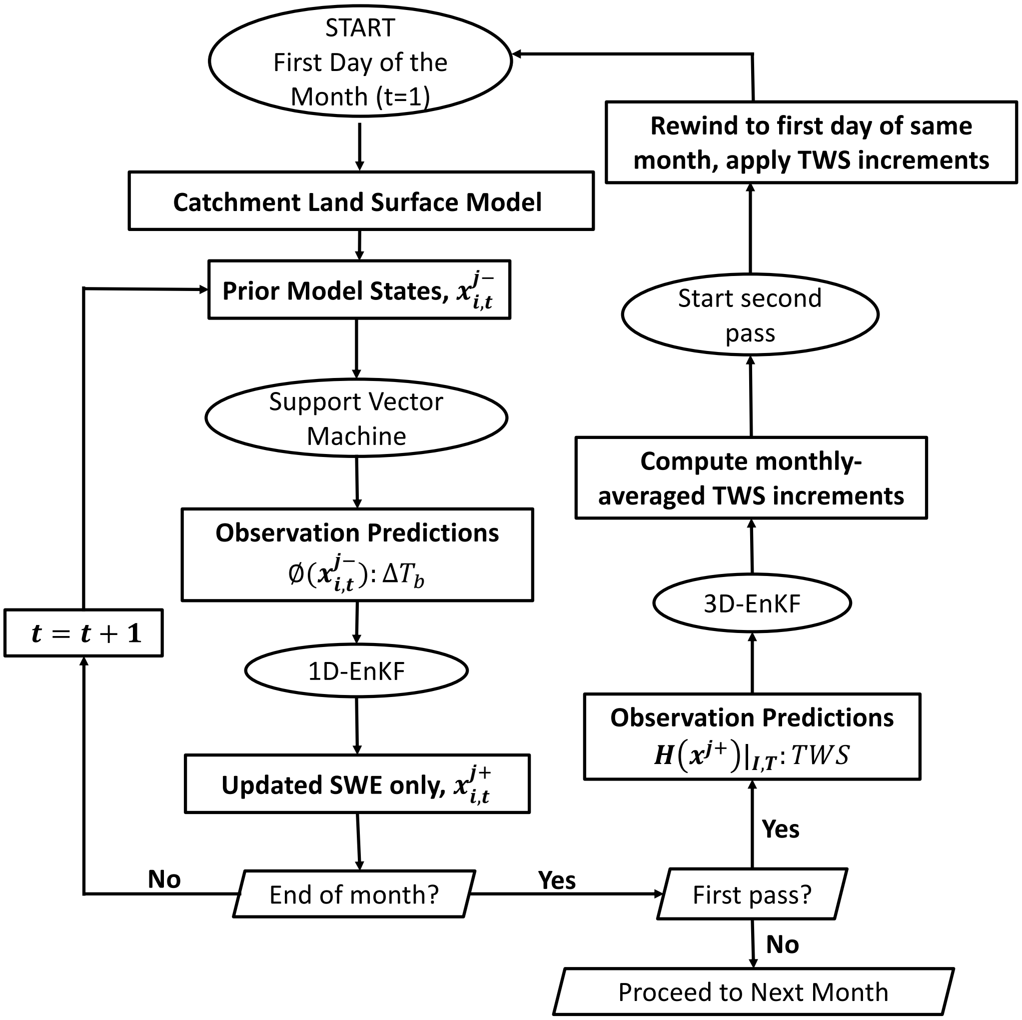

Figure 5.1: Simplified flowchart of dual assimilation for AMSR-E ΔT_b and GRACE TWS retrievals (AMSR-E + GRACE DA).

Table 5.3: Physically-informed, heuristic TWS update strategy appliable to both GRACE-only assimilation and dual assimilation (AMSR-E + GRACE DA).

	Application of Analysis Increments	
Model Scenario	SWE Increments	catdef Increments
SWE $>$ 0 and snmelt $=$ 0	Δx_{SWE} $-\Delta x_{catdef}$	0
SWE $>$ 0 and snmelt $>$ 0	0.5 $\times(\Delta x_{SWE}$ $-\Delta x_{catdef})$	0.5 $\times(\Delta x_{SWE}$ $-\Delta x_{catdef})$
SWE $=$ 0	Δx_{SWE}	Δx_{catdef}

SWE is snow water equivalent; catdef is catchment deficient; snmelt is snowmelt flux. Δx_{SWE} and Δx_{catdef} is SWE and catdef increments, respectively, calculated from Equation 5.2.

E ΔT_b observations. That is, the TWS observation predictions (i.e., $H(x^{j-})|_{I,T}$ in Equation 5.2) is calculated based on the updated model states *after* undergoing single-sensor AMSR-E ΔT_b DA. The provisional AMSR-E ΔT_b assimilation in the first run ensures that the GRACE TWS DA only corrects modeled TWS errors that are not eventually corrected by the AMSR-E ΔT_b DA. At the end of the same month, the increments are calculated in the same manner as the single-sensor GRACE TWS DA outlined in section 5.4.2. Afterwards, the model is then rewound to the beginning of the month at which point the TWS increments are applied to the initial conditions. The second run of the month is to conduct the AMSR-E ΔT_b DA as stated in section 5.4.1 with the updated initial conditions. The same "two-pass" procedure is then repeated for the next month until the end of the simulation. Dual assimilation employs the same observation types and parameters as the single-sensor AMSR-E ΔT_b DA and GRACE TWS DA (Table 5.2).

5.5 Validation Approach

The main advantage of the synthetic, identical twin experiment is the generation of the synthetic "truth" for use during OL and DA evaluation. The synthetic truth provides not only SWE and TWS estimates for using during comparison, but all other states and fluxes associated with the hydrologic cycle. In addition, several different metrics are used for the validation activities: (1) root mean squared error (RMSE), (2) unbiased root mean squared error (ubRMSE) [187], (3) bias, and (4) anomaly correlation coefficient (anomaly R). The anomaly R is computed by first determining the mean seasonal (climatological) cycle over the course of the study period. The second step is to remove the mean seasonal cycle from the original time series. After that, the anomaly R is computed as the Pearson correlation coefficient (R; AppendixA) between the synthetic truth ensemble mean anomalies and the corresponding OL or DA ensemble mean anomalies. The advantage of anomaly R is that it focuses on the skill of the model at estimating the inter-annual variability [188]. In addition, the Nash-Sutcliffe efficiency (NSE; [182]) is selected for use during the runoff evaluation. All statistics are calculated for instances only when either the synthetic SWE truth and OL/DA SWE estimates are nonzero, which yields a more rigorous comparison. Additional details about these calculations are provided in AppendixA.

5.6 Results and Discussion

In this section, the hydrological estimates, including SWE, subsurface water storage, TWS, and runoff, and their subsequent changes via AMSR-E ΔT_b DA, baseline and physically-informed GRACE TWS DA, and dual DA are quantified by validating against the synthetic truth. Figure 5.2a), 5.2b), and 5.2c) show the Volga basin-averaged bias, RMSE, and ubRMSE, respectively, with 99% confidence intervals (CI) calculated based on daily SWE, subsurface water storage (subsrf), and TWS. The anomaly R (Figure 5.2d) is calculated based on monthly SWE, subsrf, and TWS. Runoff comparison (Figure 5.6) is based on monthly basin-averaged runoff, particularly during the snow ablation period (from the March to May) since this study focuses mostly on the change in snow mass.

5.6.1 AMSR-E ΔT_b Assimilation (AMSR-E DA)

AMSR-E ΔT_b DA yields marginal improvements in snow mass assimilation that are consistent with the work by [58] and [59]. The SWE estimates are improved with approximately 8.43%, 13.1%, and 4.45% reduction of RMSE, bias, and ubRMSE relative to OL, respectively (Figure 5.2). However, the interannual variability (anomaly R) of SWE is degraded by AMSR-E DA relative to OL (Figure 5.2d). SWE degradation often occurs in high latitude areas (e.g., northern parts of the upper Volga and Kama basin; Figure 5.3g) that are colocated with deep snow and dense forest fraction (>0.5; Figure 3.1b). Under such conditions, AMSR-E PMW ΔT_b synthetic observations typically suffer from signal saturation due to the deep snow [26–29] and signal attenuation due to the forest cover [24]. Furthermore, the simultaneous assimilation of six different ΔT_b's may also result in degraded performance by the EnKF. For example, during shallow snow conditions, ΔT_{b10-18} is more informative of soil moisture rather than snow mass [178], hence, the assimilation of ΔT_{b10-18} may introduce spurious errors into the SWE estimates during the EnKF update. Finally, the presence of high-frequency noise and non-SWE related signals embedded in the AMSR-E ΔT_b synthetic observations further degrade the SWE estimates by introducing more random errors during the update.

The SWE improvements via AMSR-E DA further influence the subsurface water storage estimates. AMSR-E DA significantly improved subsurface water storage estimates with a RMSE reduction of 7.67% relative to the OL (Figure 5.2b). Similar to SWE estimation, AMSR-E DA improves subsurface water storage estimates the most in Moskva Oka and lower Volga basins with approximately 11.8% and 7.66% reduction of RMSE (relative to OL), respectively. The interannual variability of subsurface water storage is also improved via AMSR-E DA (anomaly R = 0.14) compared to OL (anomaly R = 0.10).

Given the fact that snow is a major component of TWS, the improvement in SWE estimates indicates the improvement of TWS estimates via AMSR-E DA. The Volga-basin averaged RMSE is reduced from 76.2 mm (OL) to 67.8 mm. Consistent with SWE estimates, the Moskva Oka and lower Volga basins have better performance in TWS estimation compared to other two subbasins.

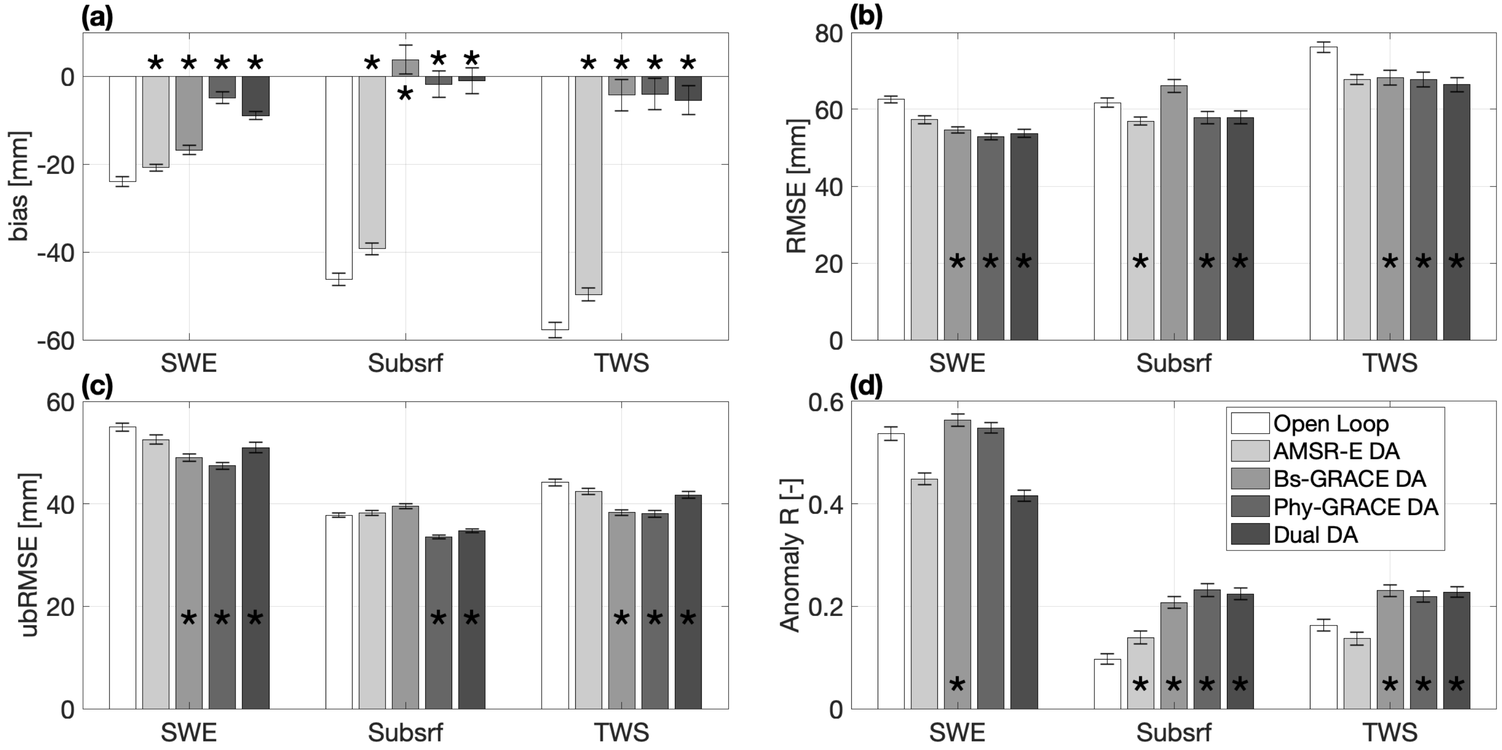

Figure 5.2: Histograms of Volga basin-averaged statistics including (a) bias, (b) root-mean-square-error (RMSE), (c) unbiased-root-mean-square-error (ubRMSE), and (d) anomaly correlation coefficient (R) of snow water equivalent (SWE), subsurface water storage (subsrf), and terrestrial water storage (TWS) for the open loop (OL; no assimilation), AMSR-E DA, baseline GRACE (Bs-GRACE) DA, Physically-informed GRACE (Phy-GRACE) DA, and Dual (AMSR-E + Physically-informed GRACE) DA. All histograms are supplemented with 99% confidence intervals. Bars marked with $*$ indicate DA yields statistically significant improvements with a level of significance of 5% compared to the OL.

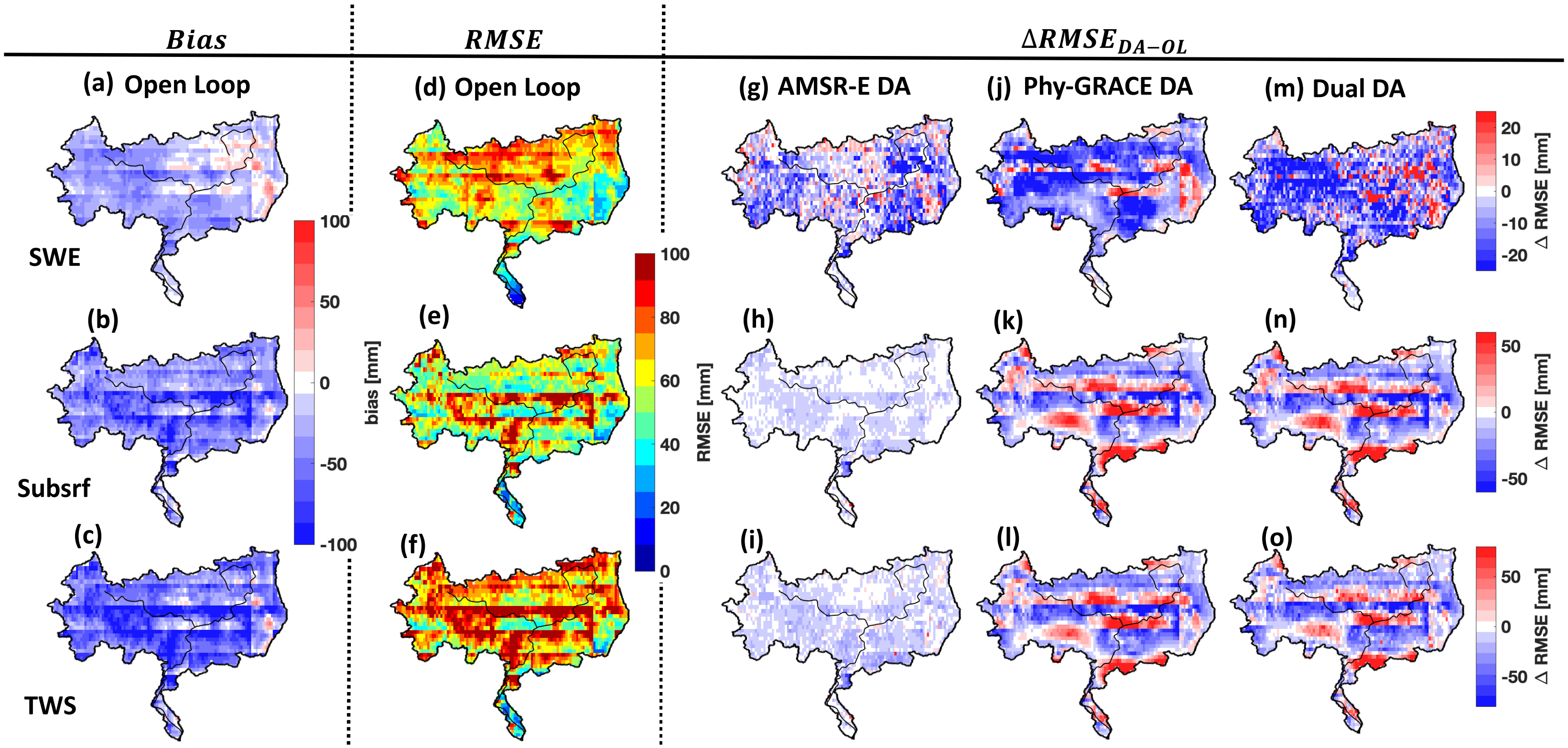

Figure 5.3: Spatial maps of bias [column 1] and root-mean-square-error (RMSE) of open loop [column 2] and [columns 3-5] changes in RMSE ($\Delta RMSE_{DA-OL}$) due to [column 3] AMSR-E data assimilation (DA); [column 4] Physically-informed GRACE DA; [column 5] Dual (AMSR-E + physically-informed GRACE) DA experiments for snow water equivalent (SWE; first row), subsurface water storage (subsrf; second row), and terrestrial water storage (TWS; third row) across the time period from 1 September 2002 to 31 August 2011.

5.6.2 GRACE TWS Assimilation (GRACE DA)

GRACE TWS DA, including both baseline and physically-informed versions of the DA algorithm, significantly improve SWE estimates compared to OL. Compared to baseline GRACE DA, physically-informed GRACE DA yields better SWE estimates with approximately 80.0% and 15.5% reduction of bias and RMSE (relative to OL), respectively (Figure 5.2). SWE degradation using physically-informed GRACE DA approach mainly occurs in the areas where the OL has relatively good performance (e.g., the middle parts of Volga basin; Figure 5.3j). Under such condition, the assimilation of GRACE TWS retrievals introduces unwanted errors that ultimately degrade the SWE estimates. In addition, the OL has more SWE but less TWS than the synthetic truth in the northern parts of the Volga basin (Figures 5.3a and 5.3c). Therefore, the assimilation of synthetic TWS retrievals tends to add snow mass, which results in further SWE overestimation relative to the synthetic truth. It is worth noting the clear edge of the RMSE pattern (e.g., the middle part of the Volga basin) is due to the forcing (i.e., GLDAS) resolution, which also reflects the importance of accurate forcings in the representation of TWS and its constituent components.

Compared to the OL, baseline GRACE DA tends to overestimate subsurface water storage over the Volga basin (Figure 5.2a), whereas the physically-informed GRACE DA using the modified update strategy (Table 5.3) has statistically significant improvements in subsurface water storage estimates. The RMSE and bias of subsurface water storage decrease from 66.1 mm and 3.81 mm (baseline GRACE

DA) to 57.8 and −1.77 mm (physically-informed GRACE DA), respectively. The modified update strategy distributes the analysis increments to SWE and subsurface water storage more accurately. The interannual variability of subsurface water storage is also significantly improved via physically-informed GRACE DA compared to the OL at the level 5% (Figure 5.2d). Similar to GRACE DA SWE estimates, subsurface water storage degradation mainly occurs in the areas where the OL subsurface water storage estimates are close to synthetic truth, and hence, there is relatively little room for improvement.

On the topic of TWS estimation, physically-informed GRACE DA slightly outperforms baseline GRACE DA in terms of smaller RMSE. More importantly, physically-informed GRACE DA enhances the accuracy of TWS discretization into its constituent components. As shown in Figure 5.4c, the example time series of Volga basin-averaged TWS ensemble mean for the baseline (green color) and Physically-informed GRACE DA (blue color) are almost identical in the significant improvements over the OL. Compared to the baseline GRACE DA, the time series of Volga basin-averaged SWE (Figure 5.4a) and subsurface water storage (Figure 5.4b) from the physically-informed GRACE DA more closely match the synthetic truth (black circles). The RMSE calculated based on the Volga basin-averaged SWE and subsurface water storage (from 1 September 2002 to 31 August 2011) is reduced from 27.5 mm and 22.9 mm (baseline GRACE DA) to 19.2 mm and 17.1 mm (physically-informed GRACE DA), respectively. During the snow accumulation season from the December to March, the physically-informed GRACE DA adds more mass to SWE and less mass to subsurface water storage that results in more accurate SWE and subsurface water storage estimates relative to the baseline GRACE DA (smaller RMSE).

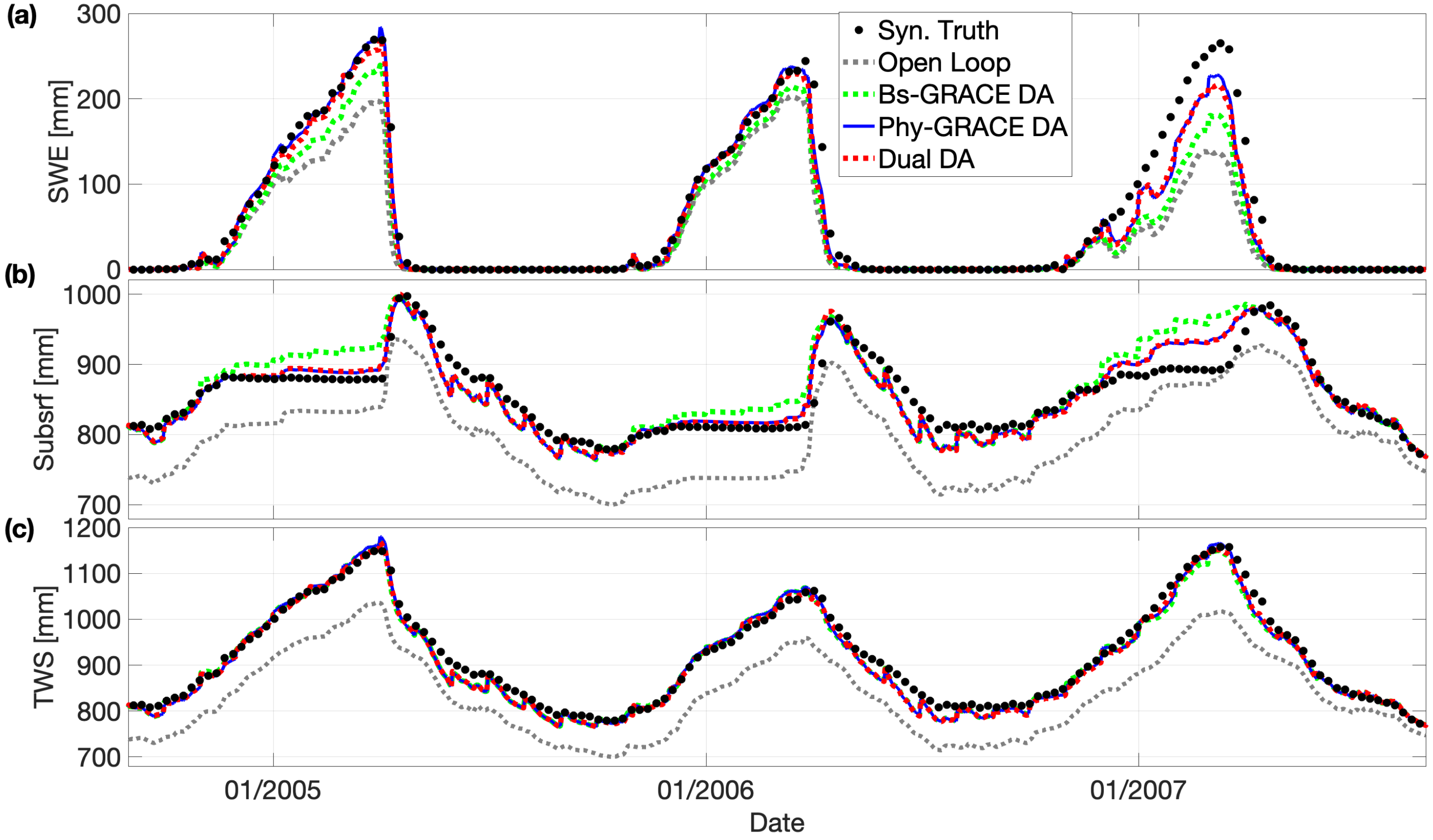

Figure 5.4: Time series of Volga basin-averaged (a) snow water equivalent (SWE), (b) subsurface water storage (subsrf), and (c) terrestrial water storage (TWS) ensemble mean derived from the synthetic truth (Syn. Truth), Open Loop, baseline GRACE (Bs-GRACE) DA, physically-informed GRACE (Phy-GRACE) DA, and dual (AMSR-E + phy-GRACE) DA for 1 September 2004 to 1 September 2007.

5.6.3 Dual Data Assimilation (AMSR-E ΔT_b + GRACE TWS DA)

Dual DA merges the AMSR-E ΔT_b DA and physically-informed GRACE TWS DA procedures. As expected, the results suggest that dual assimilation combines the characteristics of each of the two single-sensor assimilation routines. Dual DA reduces the RMSE (Figure 5.2b) of SWE estimates by 14.1% relative to the OL, which falls between the physically-informed GRACE DA (15.5%) and AMSR-E DA (12.7%) results. Similar results are found for SWE ubRMSE and bias (Figure 5.2a and 5.2b). However, the SWE interannual variability is degraded.

Dual DA has the best performance in SWE estimation in the Moskva Oka and upper Volga basins (relative to OL and single-sensor DAs) in terms of the smallest RMSE (Figure 5.3m). Similar results are found for lower Volga basins. However, dual DA tends to degrade SWE estimates in some parts of Kama basin (Figure 5.3m). The assimilation of both synthetic ΔT_b and TWS retrievals introduce more errors into the system that results in SWE degradation (RMSE = 58.2 mm) compared to single-sensor AMSR-E DA (RMSE = 55.2 mm) and physically-informed GRACE DA (RMSE = 54.8 mm) in the Kama basin where the OL has relatively good performance.

As mentioned earlier, OL has more SWE (Figure 5.3a) but less TWS (Figure 5.3c) than the synthetic truth in the northern parts of the Kama basin due to the differences between the GLDAS and MERRA-2 forcings (Figures 3.4c and 3.4d). Therefore, the assimilation of synthetic ΔT_b observations removes snow mass whereas the assimilation of synthetic TWS retrievals adds snow mass to the system. The

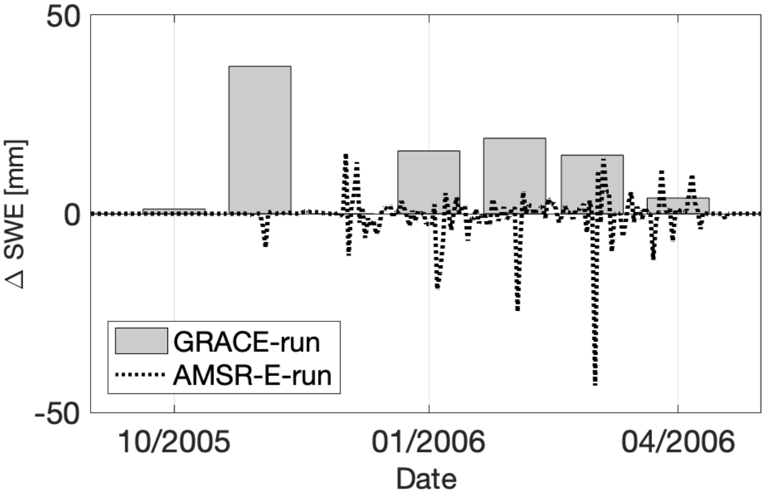

Figure 5.5: Example time series (from 1 September 2005 to 30 April 2006) of snow water equivalent (SWE) increments for one grid cell (55.5191°N, 57.7089°E) in the Kama basin via dual DA (AMSR-E + physically-informed GRACE DA). The increments calculated based on AMSR-E ΔT_b DA are shown as black dash lines while the increments introduced by the Physically-informed GRACE DA are shown as gray bars.

juxtaposed SWE increments introduced by assimilation of synthetic ΔT_b and TWS retrievals degrades the SWE estimates in the Kama basin. Figure 5.5 shows SWE increments (ΔSWE) in one grid cell in the Kama basin that were introduced by the AMSR-E-run (i.e., the assimilation of ΔT_b synthetic observations) and GRACE-run (i.e., the assimilation of TWS synthetic retrievals) during dual DA. As expected, SWE increments introduced by the AMSR-E-run are negative whereas those introduced by the GRACE-run are positive. Such a juxtaposed relationship of increments was also reported by Tian *et al.* [101] and Girotto *et al.* [85].

Similar to dual DA SWE estimation, the statistics for subsurface water storage via dual assimilation fall between those of the AMSR-E DA and physically-informed

GRACE DA results. Compared to the assimilation of synthetic ΔT_b observations, the assimilation of GRACE synthetic TWS retrievals has a larger impact on the subsurface water storage change (Figure 5.3n). Dual DA has the best estimation of subsurface water storage in terms of bias (-0.98 mm; Figure 5.2a) compared to that of the OL (-46.1 mm), AMSR-E DA (-39.3 mm), and physically-informed GRACE DA (-1.77 mm). Similar to the single-sensor, physically-informed GRACE DA, subsurface water storage degradation via dual DA mainly occurs the areas where the OL has relatively good performance such as the middle and northern parts of the Volga basin (Figure 5.3n).

Dual DA has the best TWS estimation in terms of smallest RMSE (66.4 mmm) compared to the OL (76.2 mm), AMSR-E DA (67.8 mm), and physically-informed GRACE DA (67.8 mm). In addition, dual DA maintains the improvements in TWS discretization introduced by the physically-informed GRACE DA. As shown in Figure 5.4, dual DA (red color) is equally adapt at capturing the SWE, subsurface water storage, and TWS synthetic truth comparable to the physically-informed GRACE DA. The change in subsurface water storage and TWS estimates is mainly caused by the assimilation of synthetic TWS retrievals, which is consistent with the work by Girotto *et al.* [84] and Girotto *et al.* [85]. The pattern of TWS degradation is consistent with that of SWE and subsurface water storage (Figure 5.3o).

5.6.4 Runoff Estimates

In order to investigate the runoff change due to snow mass, runoff estimates during the snow ablation period (March, April, and May) are investigated. In general, the runoff estimates are consistent with SWE estimates via single-sensor or dual DA (Figure 5.6). The single-sensor AMSR-E DA, baseline and physically-informed GRACE DA, and dual DA procedures improved runoff estimates relative to the OL for all basins during the snow ablation period expect the Kama basin. Dual DA tends to overestimate the runoff in the Kama basin, which is consistent with the SWE degradation associated with the juxtaposed increments cased by simultaneous assimilation of synthetic ΔT_b and TWS retrievals as discussed in Section 5.6.3.

5.6.5 Ensemble Spread Diagnostics

The ensemble spread is typically used as a proxy for the model error covariance [169, 170]. The reduction of ensemble spread does not necessarily mean better estimates compared to the synthetic truth. In this study, the ensemble spread is defined as the time-averaged ensemble standard deviation (σ) across the study period (from 1 September 2002 to 31 August 2011) [85]. In addition, the spread-error ratio, SR, is also used to assess the ensemble reliability [189, 190]. SR is a comparison of the ensemble spread (σ) and the errors (ubRMSE) that is calculated as $SR_i = \frac{\sigma_i}{ubRMSE_i}$ where i is the model grid within the Volga basin. The ensemble is overdispersed (underdispersed) if SR is greater (smaller) than one [189]. Figure 5.7 maps the ensemble spread of the OL and DA experiments for the SWE and catdef

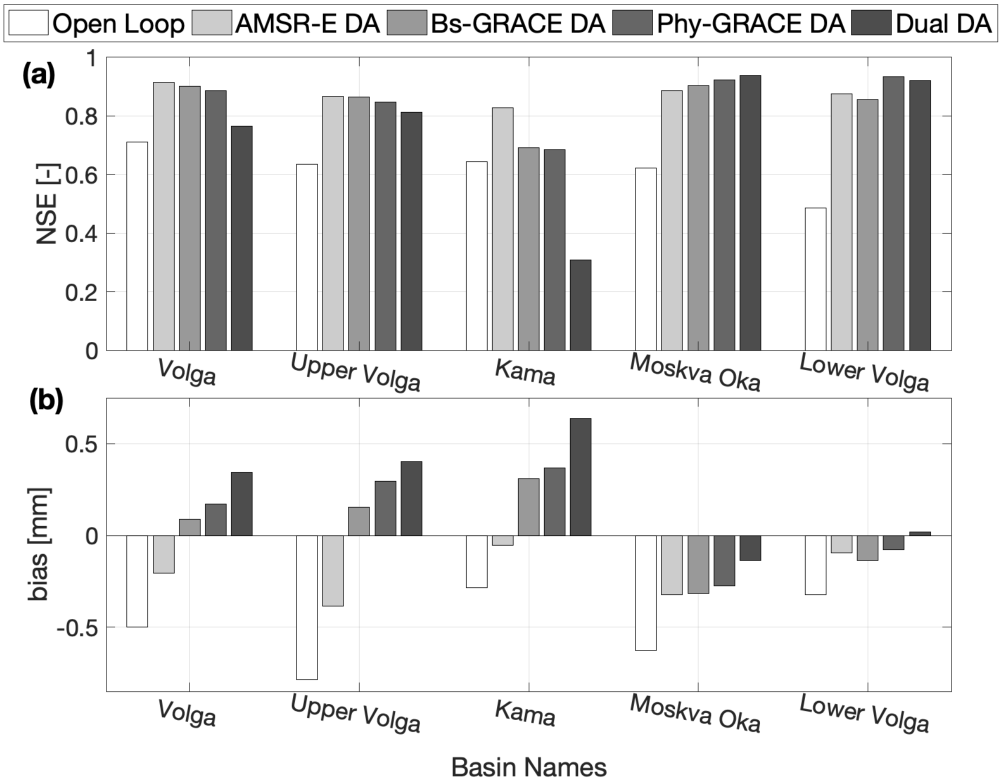

Figure 5.6: Histograms of runoff showing (a) Nash-Sutcliffe efficiency (NSE) and (b) bias of the Volga basin and four sub-basins for the Open Loop (OL; no assimilation), AMSR-E DA, baseline GRACE (Bs-GRACE) DA, physically-informed GRACE (Phy-GRACE) DA, and dual (AMSR-E + physically-informed GRACE) DA. The statistics are calculated by comparing OL or DA basin-averaged monthly runoff ensemble mean against the synthetic truth during the snow ablation period (March, April, and May) from 2002 to 2011.

(i.e., subsurface water storage) state variables. The Volga basin-averaged $\overline{SR}$ is listed in Table 5.4.

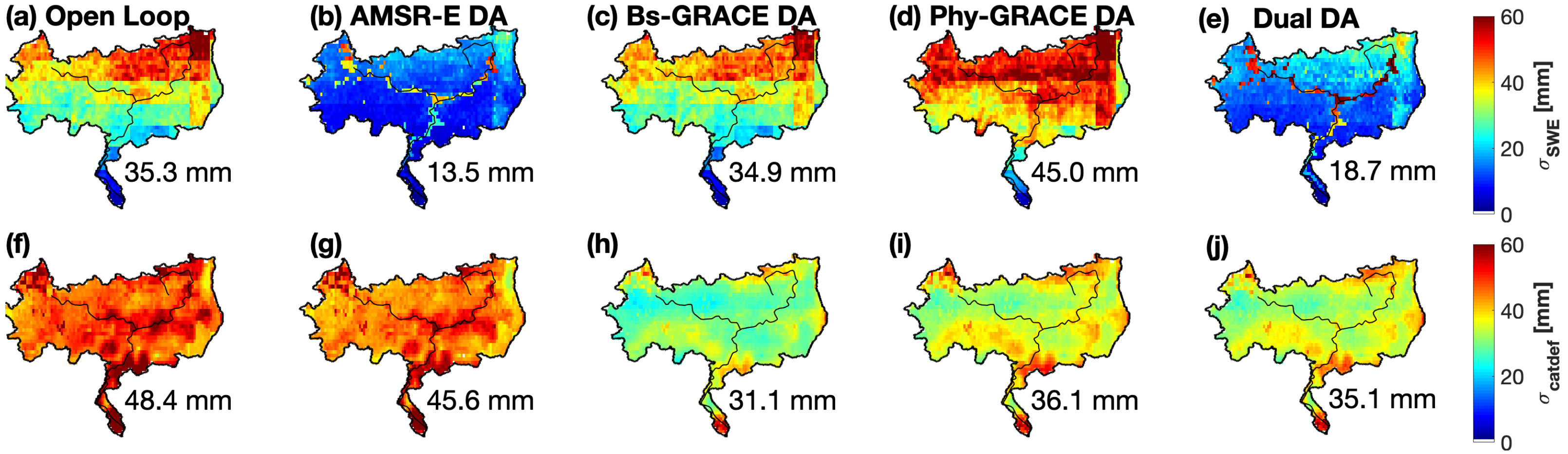

Figure 5.7: Time-averaged (1 September 2002 to 31 August 2011) ensemble standard deviation (σ) of the Open Loop (column 1), AMSR-E DA (column 2), baseline GRACE (bs-GRACE) DA (column 3), physically-informed GRACE (phy-GRACE) DA (column 4), and dual DA (AMSR-E + Physically-informed GRACE DA; column 5) for the snow water equivalent (SWE; first row) and catchment deficient (catdef; second row) state variables. The Volga basin-averaged value is shown in the text in each subplot.

Table 5.4: The Volga basin-averaged spread-error ratio ($\overline{SR}$) for the snow water equivalent (SWE) and catchment deficit (catdef) resulting from the Open Loop, AMSR-E DA, baseline GRACE DA, physically-informed GRACE DA, and dual DA (AMSR-E + physically-informed GRACE DA) from 1 September 2002 to 31 August 2011.

$\overline{SR}$	Open Loop	AMSR-E DA	[a]Bs-GRACE DA	[b]Phy-GRACE DA	Dual DA
SWE	0.67	0.27	0.74	0.98	0.38
catdef	1.36	1.26	0.85	1.15	1.08

[a]Bs-GRACE DA is the baseline GRACE DA. [b]Phy-GRACE DA is the physically-informed GRACE DA.

In general, AMSR-E DA generates the most underdispersed SWE ensemble spread due to the multi-observation nature of the daily assimilation approach. Physically-informed GRACE DA has a larger SWE ensemble spread (45.0 mm; Figure 5.7d) than baseline GRACE DA (34.9 mm; Figure 5.7c). As a combination of AMSR-E DA and physically-informed GRACE DA, dual DA has a moderate SWE ensemble spread (18.7 mm; Figure 5.7e). The $\overline{SR}$ results (Table 5.4) are consistent with the ensemble spread in that AMSR-E DA has the smallest value (0.27) while the physically-informed GRACE DA has the most reasonable ensemble spread (0.98). For OL and all DA scenarios, the spatial distribution of SWE ensemble spread has a strong gradient from the northeast to the southwest. It is worth noting that large rivers also have a large SWE ensemble spread due to the fact that AMSR-E DA is excluded from areas with significant surface water coverage.

In general, OL has the largest Volga basin-averaged catdef ensemble spread

(48.4 mm; Figure 5.7f). AMSR-E DA slightly reduce the catdef ensemble spread to 45.6 mm (Figure 5.7g). Similar to the SWE ensemble spread, physically-informed GRACE DA (31.1 mm; Figure 5.7g) has a larger catdef ensemble spread than baseline GRACE DA (36.1 mm; Figure 5.7i). The reduction in catdef ensemble spread via dual DA is the most pronounced as results in the most reasonable catdef ensemble spread compared to the OL and single DA scenarios (Table 5.4).

5.7 Conclusion

This study investigates the performance of joint assimilation using AMSR-E ΔT_b observations and GRACE TWS retrievals through a series of synthetic twin experiments. In order to better assess the performance of dual assimilation, the Open Loop and single-sensor DA experiments (AMSR-E ΔT_b DA and GRACE TWS DA) are also performed. A well-trained support vector machine is used as the observation operator during the assimilation of AMSR-E ΔT_b observations. The baseline assimilation of GRACE TWS retrievals is further modified using a physically-informed approach during the application of the analysis increments.

Results suggests that single-sensor AMSR-E ΔT_b DA yields improvements in SWE, subsurface water storage, runoff, and TWS estimation. The efficiency of SVM-based PMW ΔT_b DA system is limited by the ill-posed nature of SWE remote sensing using PMW radiometry coupled with limited controllability of the SVM-based observation operator during deep, wet snow conditions. Furthermore, the assimilation approach (i.e., multiple observations assimilated daily) can lead to SWE ensemble collapse, which can ultimately degrade the SWE estimates.

The single-sensor GRACE TWS DA experiment using the physically-informed update approach leads to statistically significant improvements in SWE, subsurface water storage, and TWS estimation. The application of increments based on the presence (or absence) of snowmelt further discretizes TWS into SWE and subsurface water storage more accurately, and hence, effectively enhances the TWS vertical resolution.

Dual assimilation, in general, maintains the benefits introduced by the single-sensor assimilation of AMSR-E ΔT_b observations and GRACE TWS retrievals. Dual DA yields the best TWS estimates (in terms of smallest RMSE) and the most reasonable ensemble spread of subsurface water storage compared to the OL and single-sensor DA. The assimilation of AMSR-E ΔT_b observations significantly reduces the SWE ensemble spread while the assimilation of TWS retrievals reduces the ensemble spread of subsurface water storage. The assimilation of TWS helps mitigate the SWE ensemble collapse often caused by daily assimilation of ΔT_b's, and hence, improves the SWE ensemble reliability. The AMSR-E DA, in general, removes snow mass whereas GRACE DA, in general, adds snow mass to the system, and sometimes results in SWE degradation given this juxtaposed, contradictory behavior.

In summary, this study investigates a new and novel joint assimilation of AMSR-E ΔT_b observations and GRACE TWS retrievals to improve estimation of TWS and its constitute components. In general, the statistics of SWE, subsurface water storage, and TWS using dual assimilation falls between the single-sensor assimilation of AMSR-E ΔT_b observations and GRACE TWS retrievals. The advantages of dual DA is that dual DA is able to generate more reasonable ensemble spread that can better represents the implicit model uncertainty. For example, dual DA inflates model SWE ensemble spread relative to the single-sensor AMSR-E ΔT_b assimilation that mitigates the ensemble collapse. The ensemble spread of subsurface water storage via dual DA is the most reasonable when compared to the single-sensor assimilation scenarios.

Chapter 6: Conclusions and Future Work

6.1 Conclusions

This dissertation explored a multi-variate, multi-sensor assimilation framework in order to enhance model-based terrestrial snow estimates across a large watershed in Russia. The science question addressed in this thesis is: Can the predictability of snow water equivalent (SWE) and terrestrial water storage (TWS) at regional and continental scales be improved via multi-sensor, multi-variate assimilation of AMSR-E spectral differences (ΔT_b) and GRACE TWS retrievals?

In Chapter 4, the NASA Catchment land surface model is used as the prognostic model in the assimilation of AMSR-E passive microwave (PMW) brightness temperature spectral differences (ΔT_b) where support vector machine (SVM) regression is employed as the observation operator. A series of synthetic twin experiments conducted using different precipitation forcing. The results show that, at times, DA degrades modeled SWE estimates (compared to the land surface model without assimilation) over snow-covered terrain. To mitigate this degradation, a physically-informed approach using different ΔT_b for shallow-to-medium or medium-to-deep snow conditions along with a "data-thinning" strategy are explored. Overall, both strategies improve the model ability to encapsulate more of the evaluation data and mitigate model ensemble collapse. The physically-informed DA and 3-day data thinning DA strategies showed marginal improvements in basin-averaged SWE in terms of a reduction in bias from 10 mm (baseline DA) to -5.2 mm and -2.5 mm, respectively. When the estimated forcings are greater than the truth (i.e., positively biased), the baseline DA, physically-informed DA, and 3-day data thinning DA improved SWE the most with approximately 30%, 31%, and 24% reduction of RMSE (relative to OL), respectively. Overall, these results highlight the limited utility of PMW ΔT_b observations in the estimation of snow in the study terrain, but did demonstrate that a physically-based constraint approach and data thinning strategy can add more utility to the ΔT_b observations in the estimation of SWE.

In Chapter 5, multi-sensor data assimilation using synthetic AMSR-E passive microwave (PMW) brightness temperature spectral differences (ΔT_b) and synthetic GRACE terrestrial water storage (TWS) retrievals were used in order to improve estimates of snow water equivalent (SWE), subsurface water storage, and TWS over the Volga basin. Single-sensor AMSR-E ΔT_b DA using a support vector machine as the observation operator improves SWE estimates, but adds little value to subsurface storage estimates. Single-sensor GRACE TWS DA is first modified with a physically-informed increments application. That is, the calculated increments are applied only to the SWE when the snow experiences no melt whereas the increments are divided into two parts equally and applied to the SWE and subsurface water storage, respectively, when snowmelt occurs. Results suggest that a physically-informed GRACE TWS DA approach significantly enhances the TWS vertical resolution via discretization into SWE and subsurface components more accurately.

AMSR-E ΔT_b and GRACE TWS are assimilated simultaneously in the multi-sensor, multi-variate (i.e., dual) assimilation framework. Results suggest significant improvements in SWE estimates via dual assimilation with a 14.1% reduction in RMSE (relative to the Open Loop without assimilation) and leads to the best TWS estimates (smallest RMSE) and most reliable subsurface water storage ensemble spread (spread-error ratio = 1.08) as compared to the single-sensor DA scenarios. However, dual DA does not always yield complementary updates, and can at times, lead to conflicting SWE increments. That is, the assimilation of ΔT_b can generate positive SWE increments whereas assimilation of TWS removes SWE in the dual DA system, which can ultimately degrade the posterior SWE estimates.

In summary, this study supports the conclusion that a multi-sensor, multi-variate data assimilation framework using AMSR-E PMW ΔT_b and GRACE TWS retrievals can improve the predictability of SWE as well as TWS and its constituent components relative to the Open Loop without assimilation.

6.2 Original Contributions

This study conducts the synthetic experiments over the Volga basin in Russia, which is a study domain that receives less attention in the literature as compared to other parts of the globe. Volga basin is dominated by plain landscapes where less than 5% of the total basin area is occupied by complex, mountainous terrain. The relatively mild topography makes AMSR-E ΔT_b snow estimation more applicable than in other areas of the globe such as Himalayas or Colorado Rockies. In addition, most of the gravitational signal changes in the Volga basin are caused by changes in regional water storage due to the presence of minimal postglacial rebound. The study results found in the Volga basin are assumed to be applicable to other basins with mild topographic relief (e.g., northern Canada) that possess long, cold winters.

This study is the first to explore improvements in modeled SWE estimation using either a physically-informed or data thinning approach for AMSR-E ΔT_b assimilation. Results suggest limited improvements in the ensemble mean in term of SWE estimation. However, the advantages of these assimilation approaches include improvements in the posterior ensemble spread (i.e., better representation of the implicit model uncertainty) as well as enhanced controllability of the SVM-based observation operator. In addition, the joint assimilation of AMSR-E ΔT_b observations with GRACE TWS retrievals is a new and novel approach. Dual DA exhibits benefits in terms of the modeled SWE uncertainty as well as TWS and its constituent components relative to the single-sensor assimilation approaches.

6.3 Future Work

6.3.1 AMSR-E ΔT_b Observation Error Characterization

The synthetic experiments assumed that the observation error in AMSR-E ΔT_b is spatially uncorrelated. However, the effective field-of-view (FOV) for each ΔT_b channel changes as a function of frequency. For example, the T_b at 36.5 GHz has a FOV of 14×8 km^2 whereas those at 18.7 GHz have a coarser FOV of 27×16 km^2. The exploration of observation error spatial correlation is expected to further enhance the accuracy of SWE estimates via AMSR-E ΔT_b DA.

6.3.2 Multi-sensor and Multi-variate Data Assimilation

Overall, these synthetic twin experiments demonstrate the potential benefits and limitations of joint assimilation using AMSR-E ΔT_b observations and GRACE TWS retrievals. A future study should conduct experiments using real-world observations including AMSR-E/AMSR-2 ΔT_b's and GRACE/GRACE-FO TWS retrievals. Joint assimilation using real-world data is hypothesized to produce more accurate hydrological estimates and effectively enhance the horizontal and vertical resolution of GRACE/GRACE-FO TWS retrievals.

6.3.3 Implementation of an Ensemble Kalman Smoother

This study demonstrated that the multi-sensor, multi-variate ensemble Kalman filter framework using AMSR-E PMW ΔT_b and GRACE TWS retrievals can lead to improvements in SWE and TWS. However, the multi-sensor, multi-variate ensemble Kalman filter can lead to conflicting SWE increments. That is, the increments generated by assimilating ΔT_b and TWS are often opposite in sign and may ultimately degrade the posterior SWE estimates. A revised assimilation scheme based on the ensemble Kalman smoother could potentially mitigate this degradation due to the fact that the assimilation state vector in the ensemble smoother could be augmented by including observations from both PMW ΔT_b and TWS retrievals and assimilated simultaneously (rather than sequentially) within a given assimilation temporal window.

Appendix A: Goodness-of-Fit Statistics

Goodness-of-fit statistics used in this study include bias, root mean squared error (RMSE), unbiased root mean squared error (ubRMSE), correlation coefficient (R), Nash-Sutcliffe efficiency (NSE), normalized information contribution (NIC), and containing ration ($\mathrm{CR}_{2\sigma}$). The symbol x_{est} denotes the OL or DA ensemble mean and the symbol x_{truth} denotes the synthetic truth. Bias was computed as:

$$bias = \frac{1}{N_t} \sum_{i=1}^{N_t} (x_{est,i} - x_{truth,i}), \tag{A.1}$$

where x_i is the state variable (e.g., SWE) at time i and N_t is the sample size over the time period t. RMSE was computed as:

$$RMSE = \sqrt{\frac{1}{N_t} \sum_{i=1}^{N_t} (x_{est,i} - x_{truth,i})^2}, \tag{A.2}$$

where x_i is the state variable (e.g., SWE) at time i and N_t is the sample size over the time period t. ubRMSE was computed as:

$$ubRMSE = \sqrt{\frac{1}{N_t} \sum_{i=1}^{N_t} (x_{est,i} - x_{truth,i})^2 - (\overline{x}_{est} - \overline{x}_{truth})^2}, \tag{A.3}$$

where $\overline{x}_{est}$ is the time-averaged estimate of the model state variable (e.g., SWE) and $\overline{x}_{truth}$ is the time-averaged synthetic truth. R was computed as:

$$R = \frac{\sum_{i=1}^{N_t} (x_{est,i} - \overline{x}_{est})(x_{truth,i} - \overline{x}_{truth})}{\sqrt{\sum_{i=1}^{N_t} (x_{est,i} - \overline{x}_{est})^2} \sqrt{\sum_{i=1}^{N_t} (x_{truth,i} - \overline{x}_{truth})^2}} \tag{A.4}$$

NSE was computed as:

$$NSE = 1 - \frac{\sum_{i=1}^{N_t}(x_{truth,i} - x_{est,i})^2}{\sum_{i=1}^{N_t}(x_{truth} - \overline{x}_{est,i})^2} \tag{A.5}$$

NIC for RMSE, NIC_{RMSE}, was computed as

$$NIC_{RMSE} = \frac{RMSE_{OL} - RMSE_{DA}}{RMSE_{OL}} \tag{A.6}$$

where the $RMSE_{OL}$ is the OL-based RMSE and $RMSE_{DA}$ is the DA-based RMSE. The containing ratio, $\text{CR}_{2\sigma}$, is the number of synthetic truth that fall within the ensemble mean ± 2 times the ensemble standard deviation normalized by the total number of synthetic truth (N_t), and was computed as

$$CR_{2\sigma} = \frac{\sum_{i=1}^{N_t} I[O(x,i)]}{N_t} \tag{A.7}$$

where $I[O(x,i)] = 1$ if $x_{min,i} \leq x_{truth,i} \leq x_{max,i}$. In other words, if the synthetic truth at time i, $x_{truth,i}$, is equal to or greater than the minimum of OL or DA ensemble estimates, $x_{min,i}$, and also is less than or equal to the maximum of OL or DA ensemble estimates, $x_{max,i}$, the $I[O(x,i)] = 1$. Otherwise, $I[O(x,i)] = 0$.

Bibliography

[1] Tajdarul H Syed, James S Famiglietti, Matthew Rodell, Jianli Chen, and Clark R Wilson. Analysis of terrestrial water storage changes from GRACE and GLDAS. *Water Resources Research*, 44(2), 2008.

[2] Kazuyoshi Suzuki, Koji Matsuo, and Tetsuya Hiyama. Satellite gravimetry-based analysis of terrestrial water storage and its relationship with run-off from the Lena River in eastern Siberia. *International Journal of Remote Sensing*, 37(10):2198–2210, 2016.

[3] James S Famiglietti. Remote sensing of terrestrial water storage, soil moisture and surface waters. *GMS*, 150:197–207, 2004.

[4] Matthew Rodell and James S Famiglietti. An analysis of terrestrial water storage variations in Illinois with implications for the Gravity Recovery and Climate Experiment (GRACE). *Water Resources Research*, 37(5):1327–1339, 2001.

[5] Byron D Tapley, Srinivas Bettadpur, John C Ries, Paul F Thompson, and Michael M Watkins. GRACE measurements of mass variability in the Earth system. *Science*, 305(5683):503–505, 2004.

[6] David D Rowlands, S B Luthcke, S M Klosko, Frank G R Lemoine, D S Chinn, J J McCarthy, C M Cox, and O B Anderson. Resolving mass flux at high spatial and temporal resolution using GRACE intersatellite measurements. *Geophysical Research Letters*, 32(4), 2005.

[7] Sean Swenson and John Wahr. Post-processing removal of correlated errors in GRACE data. *Geophysical Research Letters*, 33(8), 2006.

[8] Sujay V Kumar, Benjamin F Zaitchik, Christa D Peters-Lidard, Matthew Rodell, Rolf Reichle, Bailing Li, Michael Jasinski, David Mocko, Augusto Getirana, Gabrielle De Lannoy, and Others. Assimilation of gridded GRACE terrestrial water storage estimates in the North American Land Data Assimilation System. *Journal of Hydrometeorology*, 17(7):1951–1972, 2016.

[9] Glen E Liston. Interrelationships among snow distribution, snowmelt, and snow cover depletion: Implications for atmospheric, hydrologic, and ecologic modeling. *Journal of Applied Meteorology*, 38(10):1474–1487, 1999.

[10] Richard Fernandes, Hongxu Zhao, Xuanji Wang, Jeff Key, Xin Qu, and Alex Hall. Controls on Northern Hemisphere snow albedo feedback quantified using satellite Earth observations. *Geophysical Research Letters*, 36(21), 2009.

[11] Judith A Curry, Julie L Schramm, and Elizabeth E Ebert. Sea ice-albedo climate feedback mechanism. *Journal of Climate*, 8(2):240–247, 1995.

[12] Tim P Barnett, Jennifer C Adam, and Dennis P Lettenmaier. Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, 438(7066):303–309, 2005.

[13] Dorothy Hall. *Remote Sensing of Ice and Snow*. Springer Science & Business Media, 2012.

[14] Ross D Brown and David A Robinson. Northern Hemisphere spring snow cover variability and change over 1922-2010 including an assessment of uncertainty. *The Cryosphere*, 5(1):219, 2011.

[15] Marc Lynch-Stieglitz. The development and validation of a simple snow model for the GISS GCM. *Journal of Climate*, 7(12):1842–1855, 1994.

[16] Jiarui Dong, Jeffrey P Walker, Paul R Houser, and Chaojiao Sun. Scanning multichannel microwave radiometer snow water equivalent assimilation. *Journal of Geophysical Research: Atmospheres*, 112(D7), 2007.

[17] Rolf H Reichle. Data assimilation methods in the Earth sciences. *Advances in Water Resources*, 31(11):1411–1418, 2008.

[18] Rolf H Reichle, Clara S Draper, Qing Liu, Manuela Girotto, Sarith P P Mahanama, Randal D Koster, and Gabrielle J M De Lannoy. Assessment of MERRA-2 land surface hydrology estimates. *Journal of Climate*, 30(8):2937–2960, 2017.

[19] Frédéric Frappart, Guillaume Ramillien, Sylvain Biancamaria, Nelly M Mognard, and Anny Cazenave. Evolution of high-latitude snow mass derived from the GRACE gravimetry mission (2002–2004). *Geophysical Research Letters*, 33(2), 2006.

[20] Guo-Yue Niu, Ki-Weon Seo, Zong-Liang Yang, Clark Wilson, Hua Su, Jianli Chen, and Matthew Rodell. Retrieving snow mass from GRACE terrestrial water storage change with a land surface model. *Geophysical Research Letters*, 34(15), 2007.

[21] Hua Su, Zong-Liang Yang, Robert E Dickinson, Clark R Wilson, and Guo-Yue Niu. Multisensor snow data assimilation at the continental scale: The value of Gravity Recovery and Climate Experiment terrestrial water storage information. *Journal of Geophysical Research: Atmospheres*, 115(D10), 2010.

[22] Barton A Forman, R H Reichle, and Matthew Rodell. Assimilation of terrestrial water storage from GRACE in a snow-dominated basin. *Water Resources Research*, 48(1), 2012.

[23] Ludovic Brucker, Alain Royer, Ghislain Picard, Alexandre Langlois, and Michel Fily. Hourly simulations of the microwave brightness temperature of seasonal snow in Quebec, Canada, using a coupled snow evolution–emission model. *Remote Sensing of Environment*, 115(8):1966–1977, 2011.

[24] Chris Derksen, Anne Walker, and Barry Goodison. Evaluation of passive microwave snow water equivalent retrievals across the boreal forest/tundra transition of western Canada. *Remote Sensing of Environment*, 96(3-4):315–327, 2005.

[25] Jiarui Dong, Jeffrey P Walker, and Paul R Houser. Factors affecting remotely sensed snow water equivalent uncertainty. *Remote Sensing of Environment*, 97(1):68–82, 2005.

[26] Erwin Schanda, Christian Matzler, and Klaus Kunzi. Microwave remote sensing of snow cover. *International Journal of Remote Sensing*, 4(1):149–158, 1983.

[27] Danielle De Sève, Monique Bernier, Jean-Pierre Fortin, and Anne Walker. Preliminary analysis of snow microwave radiometry using the SSM/I passive-microwave data: the case of La Grande River watershed (Quebec). *Annals of Glaciology*, 25:353–361, 1997.

[28] Christian Mätzler. Passive microwave signatures of landscapes in winter. *Meteorology and Atmospheric Physics*, 54(1-4):241–260, 1994.

[29] Chris Derksen. The contribution of AMSR-E 18.7 and 10.7 GHz measurements to improved boreal forest snow water equivalent retrievals. *Remote Sensing of Environment*, 112(5):2701–2710, 2008.

[30] Chris Derksen, Peter Toose, Andrew Rees, Libo Wang, Michael C. English, Anne E. Walker, and Matthew Sturm. Development of a tundra-specific snow water equivalent retrieval algorithm for satellite passive microwave data. *Remote Sensing of Environment*, 114(8):1699–1709, 2010.

[31] Anne E Walker and Barry E Goodison. Discrimination of a wet snow cover using passive microwave satellite data. *Annals of Glaciology*, 17:307–311, 1993.

[32] Richard L Armstrong and Mary J Brodzik. Recent Northern Hemisphere snow extent: A comparison of data derived from visible and microwave satellite sensors. *Geophysical Research Letters*, 28(19):3673–3676, 2001.

[33] Dennis McLaughlin. An integrated approach to hydrologic data assimilation: interpolation, smoothing, and filtering. *Advances in Water Resources*, 25(8-12):1275–1286, 2002.

[34] Barton A Forman and Steven A Margulis. Assimilation of multiresolution radiation products into a downwelling surface radiation model: 2. Posterior ensemble implementation. *Journal of Geophysical Research: Atmospheres*, 115(D22), 2010.

[35] Clara Draper and Rolf H Reichle. The impact of near-surface soil moisture assimilation at subseasonal, seasonal, and inter-annual timescales. *Hydrol. Earth Syst. Sci*, 19(12):4831, 2015.

[36] Matthias Drusch, Drasko Vasiljevic, and Pedro Viterbo. ECMWF's global snow analysis: Assessment and revision based on satellite observations. *Journal of Applied Meteorology*, 43(9):1282–1294, 2004.

[37] Hsin-Cheng Huang and Noel Cressie. Spatio-temporal prediction of snow water equivalent using the Kalman filter. *Computational Statistics & Data Analysis*, 22(2):159–175, 1996.

[38] Andrew G Slater and Martyn P Clark. Snow data assimilation via an ensemble Kalman filter. *Journal of Hydrometeorology*, 7(3):478–493, 2006.

[39] Glen E Liston and Christopher A Hiemstra. A simple data assimilation system for complex snow distributions (SnowAssim). *Journal of Hydrometeorology*, 9(5):989–1004, 2008.

[40] Jan Magnusson, David Gustafsson, Fabia Hüsler, and Tobias Jonas. Assimilation of point SWE data into a distributed snow cover model comparing two contrasting methods. *Water Resources Research*, 50(10):7816–7835, 2014.

[41] Konstantinos M Andreadis and Dennis P Lettenmaier. Assimilating remotely sensed snow observations into a macroscale hydrology model. *Advances in Water Resources*, 29(6):872–886, 2006.

[42] Hua Su, Zong-Liang Yang, Guo-Yue Niu, and Robert E Dickinson. Enhancing the estimation of continental-scale snow water equivalent by assimilating MODIS snow cover with the ensemble Kalman filter. *Journal of Geophysical Research: Atmospheres*, 113(D8), 2008.

[43] Michael Durand, Edward J Kim, and Steven A Margulis. Quantifying uncertainty in modeling snow microwave radiance for a mountain snowpack at the point-scale, including stratigraphic effects. *IEEE Transactions on Geoscience and Remote Sensing*, 46(6):1753–1767, 2008.

[44] Gabriëlle J M De Lannoy, Rolf H Reichle, Kristi R Arsenault, Paul R Houser, Sujay Kumar, Niko E C Verhoest, and Valentijn R N Pauwels. Multiscale assimilation of Advanced Microwave Scanning Radiometer–EOS snow water equivalent and Moderate Resolution Imaging Spectroradiometer snow cover fraction observations in northern Colorado. *Water Resources Research*, 48(1), 2012.

[45] Yuqiong Liu, Christa D Peters-Lidard, Sujay Kumar, James L Foster, Michael Shaw, Yudong Tian, and Gregory M Fall. Assimilating satellite-based snow depth and snow cover products for improving snow predictions in Alaska. *Advances in Water Resources*, 54:208–227, 2013.

[46] Dongyue Li, Michael Durand, and Steven A Margulis. Estimating snow water equivalent in a Sierra Nevada watershed via spaceborne radiance data assimilation. *Water Resources Research*, 53(1):647–671, 2017.

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

[48] Jaakko Kukkonen, Taru Olsson, David M Schultz, A Baklanov, Thomas Klein, A I Miranda, A Monteiro, M Hirtl, V Tarvainen, M Boy, and Others. A review of operational, regional-scale, chemical weather forecasting models in Europe. *Atmospheric Chemistry and Physics*, 2012.

[49] Tao Che, Xin Li, Rui Jin, and Chunlin Huang. Assimilating passive microwave remote sensing data into a land surface model to improve the estimation of snow depth. *Remote Sensing of Environment*, 143:54–63, 2014.

[50] Yonghwan Kwon, Zong-Liang Yang, Long Zhao, Timothy J Hoar, Ally M Toure, and Matthew Rodell. Estimating snow water storage in North America using CLM4, DART, and snow radiance data assimilation. *Journal of Hydrometeorology*, 17(11):2853–2874, 2016.

[51] Yonghwan Kwon, Zong-Liang Yang, Timothy J Hoar, and Ally M Toure. Improving the radiance assimilation performance in estimating snow water storage across snow and land-cover types in North America. *Journal of Hydrometeorology*, 18(3):651–668, 2017.

[52] Barton A Forman, Rolf H Reichle, and Chris Derksen. Estimating passive microwave brightness temperature over snow-covered land in North America using a land surface model and an artificial neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 52(1):235–248, 2013.

[53] Barton A Forman and Rolf H Reichle. Using a support vector machine and a land surface model to estimate large-scale passive microwave brightness temperatures over snow-covered land in North America. *IEEE Journal of Selected*

Topics in Applied Earth Observations and Remote Sensing, 8(9):4431–4441, 2014.

[54] Yuan Xue and Barton A Forman. Comparison of passive microwave brightness temperature prediction sensitivities over snow-covered land in North America using machine learning algorithms and the Advanced Microwave Scanning Radiometer. *Remote Sensing of Environment*, 170:153–165, 2015.

[55] Barton A Forman and Yuan Xue. Machine learning predictions of passive microwave brightness temperature over snow-covered land using the special sensor microwave imager (SSM/I). *Physical Geography*, 38(2):176–196, 2017.

[56] Yuan Xue and Barton A Forman. Atmospheric and forest decoupling of passive microwave brightness temperature observations over snow-covered terrain in North America. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(7):3172–3189, 2016.

[57] Yuan Xue and Barton A Forman. Integration of satellite-based passive microwave brightness temperature observations and an ensemble-based land data assimilation framework to improve snow estimation in forested regions. In *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 311–314. IEEE, 2017.

[58] Yuan Xue, Barton A Forman, and Rolf H Reichle. Estimating snow mass in North America through assimilation of AMSR-E brightness temperature observations using the Catchment land surface model and support vector machines. *Water Resources Research*, 54(9):6488, 2018.

[59] Yonghwan Kwon, Barton A Forman, Jawairia A Ahmad, Sujay V Kumar, and Yeosang Yoon. Exploring the utility of machine learning-based passive microwave brightness temperature data assimilation over terrestrial snow in high Mountain Asia. *Remote Sensing*, 11(19):2265, 2019.

[60] Jawairia Ashfaq Ahmad, Barton A Forman, and Yonghwan Kwon. Analyzing machine learning predictions of passive microwave brightness temperature spectral difference over snow-covered terrain in High Mountain Asia. *Frontiers in Earth Science*, 7:212, 2019.

[61] Michael Durand and Steven A Margulis. Feasibility test of multifrequency radiometric data assimilation to estimate snow water equivalent. *Journal of Hydrometeorology*, 7(3):443–457, 2006.

[62] Jeffrey P Walker and Paul R Houser. Hydrologic data assimilation. In *Advances in Water Science Methodologies*, pages 45–68. CRC Press, 2005.

[63] Ming Pan, Alok K Sahoo, Tara J Troy, Raghuveer K Vinukollu, Justin Sheffield, and Eric F Wood. Multisource estimation of long-term terrestrial water budget for major global river basins. *Journal of Climate*, 25(9):3191–3206, 2012.

[64] Matthew Rodell, Hiroko Kato Beaudoing, T S L'Ecuyer, William S Olson, James Stephen Famiglietti, Paul Raymond Houser, Robert Adler, Michael G Bosilovich, Carol Anne Clayson, D Chambers, and Others. The observed state of the water cycle in the early twenty-first century. *Journal of Climate*, 28(21):8289–8318, 2015.

[65] Viviana Maggioni and Christian Massari. *Extreme Hydroclimatic Events and Multivariate Hazards in a Changing Environment: A Remote Sensing Approach*. Elsevier, 2019.

[66] Jagadish Shukla and Yale Mintz. Influence of land-surface evapotranspiration on the earth's climate. *Science*, 215(4539):1498–1501, 1982.

[67] Matthew Rodell, Benjamin F. Chao, Andrew Y. Au, John S Kimball, and Kyle C. McDonald. Global biomass variation and its geodynamic effects: 1982–98. *Earth Interactions*, 9(2):1–19, 2005.

[68] Augusto Getirana, Sujay Kumar, Manuela Girotto, and Matthew Rodell. Rivers and floodplains as key components of global terrestrial water storage variability. *Geophysical Research Letters*, 44(20):10–359, 2017.

[69] Paul C. D. Milly. Sensitivity of the global water cycle to the water-holding capacity of soils. *IAHS PUBLICATION*, page 495, 1993.

[70] Wade T Crow, Rajat Bindlish, and Thomas J. Jackson. The added value of spaceborne passive microwave soil moisture retrievals for forecasting rainfall-runoff partitioning. *Geophysical Research Letters*, 32(18), 2005.

[71] Dara Entekhabi, Ignacio Rodriguez-Iturbe, and Fabio Castelli. Mutual interaction of soil moisture state and atmospheric processes. *Journal of Hydrology*, 184(1-2):3–17, 1996.

[72] Sonia I Seneviratne, Thierry Corti, Edouard L Davin, Martin Hirschi, Eric B Jaeger, Irene Lehner, Boris Orlowsky, and Adriaan J Teuling. Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Reviews*, 99(3-4):125–161, 2010.

[73] Tom Gleeson, Yoshihide Wada, Marc F P Bierkens, and Ludovicus P H Van Beek. Water balance of global aquifers revealed by groundwater footprint. *Nature*, 488(7410):197–200, 2012.

[74] James S Famiglietti. The global groundwater crisis. *Nature Climate Change*, 4(11):945–948, 2014.

[75] Matthew Rodell, Isabella Velicogna, and James S Famiglietti. Satellite-based estimates of groundwater depletion in India. *Nature*, 460(7258):999–1002, 2009.

[76] Bailing Li, Matthew Rodell, and James S Famiglietti. Groundwater variability across temporal and spatial scales in the central and northeastern US. *Journal of Hydrology*, 525:769–780, 2015.

[77] Dorothy K Hall, Alfred T C Chang, and Honnappa Siddalingaiah. Reflectances of glaciers as calculated using Landsat-5 Thematic Mapper data. *Remote Sensing of Environment*, 25(3):311–321, 1988.

[78] James L Foster, Dorothy K Hall, John B Eylander, George A Riggs, Son V Nghiem, Marco Tedesco, Edward Kim, Paul M Montesano, Richard E J Kelly, Kimberly A Casey, and Others. A blended global snow product using visible, passive microwave and scatterometer satellite data. *International Journal of Remote Sensing*, 32(5):1371–1395, 2011.

[79] Thomas J Schmugge, William P Kustas, Jerry C Ritchie, Thomas J Jackson, and Al Rango. Remote sensing in hydrology. *Advances in Water Resources*, 25(8-12):1367–1385, 2002.

[80] Randal D Koster, Max J Suarez, Agnès Ducharne, Marc Stieglitz, and Praveen Kumar. A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure. *Journal of Geophysical Research: Atmospheres*, 105(D20):24809–24822, 2000.

[81] Benjamin F Zaitchik, Matthew Rodell, and Rolf H Reichle. Assimilation of GRACE terrestrial water storage data into a land surface model: Results for the Mississippi River basin. *Journal of Hydrometeorology*, 9(3):535–548, 2008.

[82] Bailing Li, Matthew Rodell, Benjamin F Zaitchik, Rolf H Reichle, Randal D Koster, and Tonie M van Dam. Assimilation of GRACE terrestrial water storage into a land surface model: Evaluation and potential value for drought monitoring in western and central Europe. *Journal of Hydrology*, 446:103–115, 2012.

[83] Rasmus Houborg, Matthew Rodell, Bailing Li, Rolf Reichle, and Benjamin F Zaitchik. Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations. *Water Resources Research*, 48(7), 2012.

[84] Manuela Girotto, Gabriëlle J M De Lannoy, Rolf H Reichle, and Matthew Rodell. Assimilation of gridded terrestrial water storage observations from GRACE into a land surface model. *Water Resources Research*, 52(5):4164–4183, 2016.

[85] Manuela Girotto, Rolf H Reichle, Matthew Rodell, Qing Liu, Sarith Mahanama, and Gabriëlle J M De Lannoy. Multi-sensor assimilation of SMOS brightness temperature and GRACE terrestrial water storage observations for soil moisture and shallow groundwater estimation. *Remote Sensing of Environment*, 227:12–27, 2019.

[86] Marc Stieglitz, Agnès Ducharne, Randy Koster, and Max Suarez. The impact of detailed snow physics on the simulation of snow cover and subsurface thermodynamics at continental scales. *Journal of Hydrometeorology*, 2(3):228–242, 2001.

[87] Agnès Ducharne, Randal D Koster, Max J Suarez, Marc Stieglitz, and Praveen Kumar. A catchment-based approach to modeling land surface processes in a general circulation model: 2. Parameter estimation and model demonstration. *Journal of Geophysical Research: Atmospheres*, 105(D20):24823–24838, 2000.

[88] AIJM Van Dijk, Luigi Renzullo, Yoshihide Wada, Paul Tregoning, and Others. A global water cycle reanalysis (2003–2012) merging satellite gravimetry and altimetry observations with a hydrological multi-model ensemble. 2014.

[89] Byron D Tapley and Christoph Reigber. The GRACE mission: status and future plans. *AGUFM*, 2001:G41C—-02, 2001.

[90] James S Famiglietti and Matthew Rodell. Water in the balance. *Science*, 340(6138):1300–1301, 2013.

[91] Alys C Thomas, John T Reager, James S Famiglietti, and Matthew Rodell. A GRACE-based water storage deficit approach for hydrological drought characterization. *Geophysical Research Letters*, 41(5):1537–1545, 2014.

[92] Maike Schumacher, Ehsan Forootan, AIJM van Dijk, Hannes Müller Schmied, Russell Crosbie, Jüergen Kusche, and Petra Döll. Improving drought simulations within the Murray-Darling Basin by combined calibration/assimilation of GRACE data into the WaterGAP Global Hydrology Model. *Remote Sensing of Environment*, 204:212–228, 2018.

[93] John T Reager and James S Famiglietti. Global terrestrial water storage capacity and flood potential using GRACE. *Geophysical Research Letters*, 36(23), 2009.

[94] John T Reager, Brian F Thomas, and James S Famiglietti. River basin flood potential inferred using GRACE gravity observations at several months lead time. *Nature Geoscience*, 7(8):588–592, 2014.

[95] John T Reager, Alys C Thomas, Eric A Sproles, Matthew Rodell, Hiroko K Beaudoing, Bailing Li, and James S Famiglietti. Assimilation of GRACE terrestrial water storage observations into a land surface model for the assessment of regional flood potential. *Remote Sensing*, 7(11):14663–14679, 2015.

[96] Natthachet Tangdamrongsub, Susan C Steele-Dunne, Brian C Gunter, Pavel G Ditmar, and Albrecht H Weerts. Data assimilation of GRACE terrestrial water storage estimates into a regional hydrological model of the Rhine River basin. *Hydrology & Earth System Sciences*, 19(4), 2015.

[97] Matthew Rodell, Eric B McWilliams, James S Famiglietti, Hiroko K Beaudoing, and Joseph Nigro. Estimating evapotranspiration using an observation based terrestrial water budget. *Hydrological Processes*, 25(26):4082–4092, 2011.

[98] Barton A Forman and R H Reichle. The spatial scale of model errors and assimilated retrievals in a terrestrial water storage assimilation system. *Water Resources Research*, 49(11):7457–7468, 2013.

[99] Rolf H Reichle, Gabriëlle J M De Lannoy, Barton A Forman, Clara S Draper, and Qing Liu. Connecting satellite observations with water cycle variables through land data assimilation: Examples using the NASA GEOS-5 LDAS. *Surveys in Geophysics*, 35(3):577–606, 2014.

[100] Manuela Girotto, Gabriëlle J M De Lannoy, Rolf H Reichle, Matthew Rodell, Clara Draper, Soumendra N Bhanja, and Abhijit Mukherjee. Benefits and pitfalls of GRACE data assimilation: A case study of terrestrial water storage depletion in India. *Geophysical Research Letters*, 44(9):4107–4115, 2017.

[101] Siyuan Tian, Paul Tregoning, Luigi J Renzullo, Albert I J M van Dijk, Jeffrey P Walker, Valentijn R N Pauwels, and Sébastien Allgeyer. Improved water balance component estimates through joint assimilation of GRACE water storage and SMOS soil moisture retrievals. *Water Resources Research*, 53(3):1820–1840, 2017.

[102] Mehdi Khaki and Joseph Awange. The application of multi-mission satellite data assimilation for studying water storage changes over South America. *Science of the Total Environment*, 647:1557–1572, 2019.

[103] Scott B Luthcke, T J Sabaka, B D Loomis, A A Arendt, J J McCarthy, and J Camp. Antarctica, Greenland and Gulf of Alaska land-ice evolution from an iterated GRACE global mascon solution. *Journal of Glaciology*, 59(216):613–631, 2013.

[104] Isabella Velicogna, Tyler C Sutterley, and Michiel R Van Den Broeke. Regional acceleration in ice mass loss from Greenland and Antarctica using GRACE time-variable gravity data. *Geophysical Research Letters*, 41(22):8130–8137, 2014.

[105] John Wahr, Sean Swenson, and Isabella Velicogna. Accuracy of GRACE mass estimates. *Geophysical Research Letters*, 33(6), 2006.

[106] John Wahr, Mery Molenaar, and Frank Bryan. Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE. *Journal of Geophysical Research: Solid Earth*, 103(B12):30205–30229, 1998.

[107] Himanshu Save, Srinivas Bettadpur, and Byron D Tapley. High-resolution CSR GRACE RL05 mascons. *Journal of Geophysical Research: Solid Earth*, 121(10):7547–7569, 2016.

[108] Dorothy K Hall, George A Riggs, Vincent V Salomonson, Nicolo E DiGirolamo, and Klaus J Bayr. MODIS snow-cover products. *Remote Sensing of Environment*, 83(1-2):181–194, 2002.

[109] Leung Tsang, Jin Pan, Ding Liang, Zhongxin Li, Donald W Cline, and Yunhua Tan. Modeling active microwave remote sensing of snow using dense media radiative transfer (DMRT) theory with multiple-scattering effects. *IEEE Transactions on Geoscience and Remote Sensing*, 45(4):990–1004, 2007.

[110] Thomas H Painter, Daniel F Berisford, Joseph W Boardman, Kathryn J Bormann, Jeffrey S Deems, Frank Gehrke, Andrew Hedrick, Michael Joyce, Ross Laidlaw, Danny Marks, and Others. The Airborne Snow Observatory: Fusion of scanning lidar, imaging spectrometer, and physically-based modeling for mapping snow water equivalent and snow albedo. *Remote Sensing of Environment*, 184:139–152, 2016.

[111] Richard Kelly. The AMSR-E snow depth algorithm: Description and initial results. *Journal of the Remote Sensing Society of Japan*, 29(1):307–317, 2009.

[112] William H Stiles and Fawwaz T Ulaby. The active and passive microwave response to snow parameters: 1. Wetness. *Journal of Geophysical Research: Oceans*, 85(C2):1037–1044, 1980.

[113] Jay H Zwally and Per Gloersen. Passive microwave images of the polar regions and research applications. *Polar Record*, 18(116):431–450, 1977.

[114] Alfred T C Chang, James L Foster, and Dorothy K Hall. Nimbus-7 SMMR derived global snow cover parameters. *Annals of Glaciology*, 9:39–44, 1987.

[115] Alfred T C Chang, James L Foster, Dorothy K Hall, Albert A Rango, and Beverly K Hartline. Snow water equivalent estimation by microwave radiometry. *Cold Regions Science and Technology*, 5(3):259–267, 1982.

[116] Richard L Armstrong, Alfred Chang, Albert Rango, and Edward Josberger. Snow depths and grain-size relationships with relevance for passive microwave studies. *Annals of Glaciology*, 17:171–176, 1993.

[117] Dorothy K Hall. Influence of depth hoar on microwave emission from snow in northern Alaska. *Cold Regions Science and Technology*, 13(3):225–231, 1987.

[118] Marco Tedesco, Edward J Kim, Anthony W England, Roger D De Roo, and Janet P Hardy. Brightness temperatures of snow melting/refreezing cycles: Observations and modeling using a multilayer dense medium theory-based model. *IEEE Transactions on Geoscience and Remote Sensing*, 44(12):3563–3573, 2006.

[119] John R Howell, M Pinar Menguc, and Robert Siegel. *Thermal Radiation Heat Transfer*. CRC press, 2010.

[120] Dieter Scherer, Dorothy K Hall, Volker Hochschild, Max Konig, Jan-gunnar Winther, Claude R Duguay, FréDérique Pivot, Christian Matzler, Frank Rau, Klaus Seidel, and Others. Remote sensing of snow cover. *GEOPHYSICAL MONOGRAPH-AMERICAN GEOPHYSICAL UNION*, 163:7, 2005.

[121] Marco Tedesco and James R Wang. Atmospheric correction of AMSR-E brightness temperatures for dry snow cover mapping. *IEEE Geoscience and Remote Sensing Letters*, 3(3):320–324, 2006.

[122] Ernest K Smith. Centimeter and millimeter wave attenuation and brightness temperature due to atmospheric oxygen and water vapor. *Radio Science*, 17(06):1455–1464, 1982.

[123] Barry E Goodison and Anne E Walker. Canadian development and use of snow cover information from passive microwave satellite data. *Passive microwave remote sensing of land-atmosphere interactions*, pages 245–262, 1995.

[124] Richard E Kelly, Alfred T Chang, Leung Tsang, and James L Foster. A prototype AMSR-E global snow area and snow depth algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, 41(2):230–242, 2003.

[125] Marco Tedesco and Parag S Narvekar. Assessment of the nasa amsr-e swe product. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 3(1):141–159, 2010.

[126] Alfred T C Chang, James L Foster, and Dorothy K Hall. Effects of forest on the snow parameters derived from microwave measurements during the BOREAS winter field campaign. *Hydrological Processes*, 10(12):1565–1574, 1996.

[127] Debbie Clifford. Global estimates of snow water equivalent from passive microwave instruments: history, challenges and future developments. *International Journal of Remote Sensing*, 31(14):3707–3726, 2010.

[128] Thomas Wilheit, Alfred T C Chang, and A S Milman. Atmospheric corrections to passive microwave observations of the ocean. *Boundary-Layer Meteorology*, 18(1):65–77, 1980.

[129] Matthew H Savoie, Richard L Armstrong, Mary J Brodzik, and James R Wang. Atmospheric corrections for improved satellite passive microwave snow cover retrievals over the Tibet Plateau. *Remote Sensing of Environment*, 113(12):2661–2669, 2009.

[130] Andrew Rees, Juha Lemmetyinen, Chris Derksen, Jouni Pulliainen, and Michael English. Observed and modelled effects of ice lens formation on passive microwave brightness temperatures over snow covered tundra. *Remote Sensing of Environment*, 114(1):116–126, 2010.

[131] Frank J Wentz, C Gentemann, and P Ashcroft. On-orbit calibration of AMSR-E and the retrieval of ocean products. In *Preprints, 12th Conf. on Satellite Meteorology and Oceanography, Long Beach, CA, Amer. Meteor. Soc., CD-ROM P*, volume 5, 2003.

[132] Eni G Njoku, Thomas J Jackson, Venkat Lakshmi, Tsz K Chan, and Son V Nghiem. Soil moisture retrieval from AMSR-E. *IEEE Transactions on Geoscience and Remote Sensing*, 41(2):215–229, 2003.

[133] Sujay V Kumar, Jiarui Dong, Christa D Peters-Lidard, David Mocko, and Breogán Gómez. Role of forcing uncertainty and background model error characterization in snow data assimilation. *Hydrology & Earth System Sciences*, 21(6), 2017.

[134] Arthur Gelb. *Applied optimal estimation*. MIT press, 1974.

[135] Geir Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, 99(C5):10143–10162, 1994.

[136] Rolf H Reichle, Dennis B McLaughlin, and Dara Entekhabi. Hydrologic data assimilation with the ensemble Kalman filter. *Monthly Weather Review*, 130(1):103–114, 2002.

[137] Rolf H Reichle and Randal D Koster. Land data assimilation with the ensemble Kalman filter: assessing model error parameters using innovations. *Developments in Water Science*, 47:1387–1394, 2002.

[138] Rolf H Reichle, Dara Entekhabi, and Dennis B McLaughlin. Downscaling of radio brightness measurements for soil moisture estimation: A four-dimensional variational data assimilation approach. *Water Resources Research*, 37(9):2353–2364, 2001.

[139] Sujay V Kumar, Rolf H Reichle, Kenneth W Harrison, Christa D Peters-Lidard, Soni Yatheendradas, and Joseph A Santanello. A comparison of methods for a priori bias correction in soil moisture data assimilation. *Water Resources Research*, 48(3), 2012.

[140] Dick P Dee. On-line estimation of error covariance parameters for atmospheric data assimilation. *Monthly Weather Review*, 123(4):1128–1145, 1995.

[141] John C Derber, David F Parrish, and Stephen J Lord. The new global operational analysis system at the National Meteorological Center. *Weather and Forecasting*, 6(4):538–547, 1991.

[142] Rolf H Reichle, Wade T Crow, and Christian L Keppenne. An adaptive ensemble Kalman filter for soil moisture data assimilation. *Water Resources Research*, 44(3), 2008.

[143] Michael Durand, Noah P Molotch, and Steven A Margulis. Merging complementary remote sensing datasets in the context of snow water equivalent reconstruction. *Remote Sensing of Environment*, 112(3):1212–1225, 2008.

[144] Steven A Margulis, Gonzalo Cortés, Manuela Girotto, and Michael Durand. A Landsat-era Sierra Nevada snow reanalysis (1985–2015). *Journal of Hydrometeorology*, 17(4):1203–1221, 2016.

[145] Ally M Toure, Rolf H Reichle, Barton A Forman, Augusto Getirana, and Gabrielle J M De Lannoy. Assimilation of MODIS snow cover fraction observations into the NASA catchment land surface model. *Remote Sensing*, 10(2):316, 2018.

[146] Sujay V Kumar, Christa D Peters-Lidard, Kristi R Arsenault, Augusto Getirana, David Mocko, and Yuqiong Liu. Quantifying the added value of snow cover area observations in passive microwave snow depth data assimilation. *Journal of Hydrometeorology*, 16(4):1736–1741, 2015.

[147] Michael Durand, Edward J Kim, and Steven A Margulis. Radiance assimilation shows promise for snowpack characterization. *Geophysical Research Letters*, 36(2), 2009.

[148] Katsuhiko Ogata and Yanjuan Yang. *Modern control engineering*, volume 4. Prentice hall India, 2002.

[149] Pat J-F Yeh, Sean C Swenson, James S Famiglietti, and Matthew Rodell. Remote sensing of groundwater storage changes in Illinois using the Gravity Recovery and Climate Experiment (GRACE). *Water Resources Research*, 42(12), 2006.

[150] Sean Swenson and John Wahr. Estimating large-scale precipitation minus evapotranspiration from GRACE satellite gravity measurements. *Journal of Hydrometeorology*, 7(2):252–270, 2006.

[151] Kevin M Ellett, Jeffrey P Walker, Andrew W Western, and Matthew Rodell. A framework for assessing the potential of remote-sensed gravity to provide new insight on the hydrology of the Murray-Darling Basin. *Australasian Journal of Water Resources*, 10(2):125–138, 2006.

[152] Bailing Li and Matthew Rodell. Evaluation of a model-based groundwater drought indicator in the conterminous US. *Journal of Hydrology*, 526:78–88, 2015.

[153] Annette Eicker, Maike Schumacher, Jürgen Kusche, Petra Döll, and Hannes Müller Schmied. Calibration/data assimilation approach for integrating GRACE data into the WaterGAP Global Hydrology Model (WGHM) using an ensemble Kalman filter: First results. *Surveys in Geophysics*, 35(6):1285–1309, 2014.

[154] Carly Sakumura, Srinivas Bettadpur, Himanshu Save, and Christopher McCullough. High-frequency terrestrial water storage signal capture via a regularized sliding window mascon product from GRACE. *Journal of Geophysical Research: Solid Earth*, 121(5):4014–4030, 2016.

[155] Dennis P Lettenmaier, Doug Alsdorf, Jeff Dozier, George J Huffman, Ming Pan, and Eric F Wood. Inroads of remote sensing into hydrologic science during the WRR era. *Water Resources Research*, 51(9):7309–7342, 2015.

[156] Matthew F McCabe, Matthew Rodell, Douglas E Alsdorf, Diego G Miralles, Remko Uijlenhoet, Wolfgang Wagner, Arko Lucieer, Rasmus Houborg, Niko E C Verhoest, Trenton E Franz, and Others. The future of Earth observation in hydrology. *Hydrology and Earth System Sciences*, 21(7):3879, 2017.

[157] Yann H Kerr, Philippe Waldteufel, J-P Wigneron, JAMJ Martinuzzi, Jordi Font, and Michael Berger. Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission. *IEEE Transactions on Geoscience and Remote Sensing*, 39(8):1729–1735, 2001.

[158] Long Zhao and Zong-Liang Yang. Multi-sensor land data assimilation: Toward a robust global soil moisture and snow estimation. *Remote Sensing of Environment*, 216:13–27, 2018.

[159] Valentin Golosov and Vladimir Belyaev. The Volga River Basin Report, 2016.

[160] Aleksey Yu Sidorchuk, Andrei V Panin, and Olga K Borisova. Morphology of river channels and surface runoff in the Volga River basin (East European Plain) during the Late Glacial period. *Geomorphology*, 113(3-4):137–157, 2009.

[161] Mark A Friedl, Douglas K McIver, John C F Hodges, Xiaoyang Y Zhang, D Muchoney, Alan H Strahler, Curtis E Woodcock, Sucharita Gopal, Annemarie Schneider, Amanda Cooper, and Others. Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment*, 83(1-2):287–302, 2002.

[162] Gordan R Stuhne and William R Peltier. Reconciling the ICE-6G_C reconstruction of glacial chronology with ice sheet dynamics: The cases of Greenland and Antarctica. *Journal of Geophysical Research: Earth Surface*, 120(9):1841–1865, 2015.

[163] Ronald Gelaro, Will McCarty, Max J Suárez, Ricardo Todling, Andrea Molod, Lawrence Takacs, Cynthia A Randles, Anton Darmenov, Michael G Bosilovich,

Rolf Reichle, and Others. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate*, 30(14):5419–5454, 2017.

[164] Rolf H Reichle, Q Liu, Randal D Koster, Clara S Draper, Sarith PP Mahanama, and Gary S Partyka. Land surface precipitation in merra-2. *Journal of Climate*, 30(5):1643–1664, 2017.

[165] Matthew Rodell, P R Houser, U E A Jambor, J Gottschalck, K Mitchell, C-J Meng, K Arsenault, B Cosgrove, J Radakovich, M Bosilovich, and Others. The global land data assimilation system. *Bulletin of the American Meteorological Society*, 85(3):381–394, 2004.

[166] Daniel Günther, Thomas Marke, Richard Essery, and Ulrich Strasser. Uncertainties in snowpack simulations—Assessing the impact of model structure, parameter choice, and forcing data error on point-scale energy balance snow model performance. *Water Resources Research*, 55(4):2779–2800, 2019.

[167] Roy Rasmussen, Bruce Baker, John Kochendorfer, Tilden Meyers, Scott Landolt, Alexandre P Fischer, Jenny Black, Julie M Thériault, Paul Kucera, David Gochis, and Others. How well are we measuring snow: The NOAA/FAA/NCAR winter precipitation test bed. *Bulletin of the American Meteorological Society*, 93(6):811–829, 2012.

[168] Mary J Brodzik, Brendan Billingsley, Terry Haran, Bruce Raup, and Matthew H Savoie. EASE-Grid 2.0: Incremental but significant improvements for Earth-gridded data sets. *ISPRS International Journal of Geo-Information*, 1(1):32–45, 2012.

[169] Peter L Houtekamer and Herschel L Mitchell. Data assimilation using an ensemble Kalman filter technique. *Monthly Weather Review*, 126(3):796–811, 1998.

[170] Gerrit Burgers, Peter van Leeuwen, and Geir Evensen. Analysis scheme in the ensemble Kalman filter. *Monthly Weather Review*, 126(6):1719–1724, 1998.

[171] Gabriëlle J M De Lannoy and Rolf H Reichle. Global assimilation of multiangle and multipolarization SMOS brightness temperature observations into the GEOS-5 catchment land surface model for soil moisture estimation. *Journal of Hydrometeorology*, 17(2):669–691, 2016.

[172] Michael Durand and Steven A Margulis. Correcting first-order errors in snow water equivalent estimates using a multifrequency, multiscale radiometric data assimilation scheme. *Journal of Geophysical Research: Atmospheres*, 112(D13), 2007.

[173] Christian L Keppenne. Data assimilation into a primitive-equation model with a parallel ensemble Kalman filter. *Monthly Weather Review*, 128(6):1971–1981, 2000.

[174] Geir Evensen. The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics*, 53(4):343–367, 2003.

[175] Rolf H Reichle and Randal D Koster. Assessing the impact of horizontal error correlations in background fields on soil moisture estimation. *Journal of Hydrometeorology*, 4(6):1229–1242, 2003.

[176] Rolf H Reichle, Sujay V Kumar, Sarith P P Mahanama, Randal D Koster, and Qing Liu. Assimilation of satellite-derived skin temperature observations into land surface models. *Journal of Hydrometeorology*, 11(5):1103–1122, 2010.

[177] Eung J Kim and Anthony W England. A yearlong comparison of plot-scale and satellite footprint-scale 19 and 37 GHz brightness of the Alaskan North Slope. *Journal of Geophysical Research: Atmospheres*, 108(D13), 2003.

[178] Alexandre Roy, Ghislain Picard, Alain Royer, Benoit Montpetit, Florent Dupont, Alexandre Langlois, Chris Derksen, and Nicolas Champollion. Brightness temperature simulations of the Canadian seasonal snowpack driven by measurements of the snow specific surface area. *IEEE Transactions on Geoscience and Remote Sensing*, 51(9):4692–4704, 2013.

[179] Christian Matzler, Erwin Schanda, and Walter Good. Towards the definition of optimum sensor specifications for microwave remote sensing of snow. *IEEE Transactions on Geoscience and Remote Sensing*, (1):57–66, 1982.

[180] Sujay V Kumar, Rolf H Reichle, Randal D Koster, Wade T Crow, and Christa D Peters-Lidard. Role of subsurface physics in the assimilation of surface soil moisture observations. *Journal of Hydrometeorology*, 10(6):1534–1547, 2009.

[181] Sujay V Kumar, Christa D Peters-Lidard, David Mocko, Rolf Reichle, Yuqiong Liu, Kristi R Arsenault, Youlong Xia, Michael Ek, George Riggs, Ben Livneh, and Others. Assimilation of remotely sensed soil moisture and snow depth retrievals for drought estimation. *Journal of Hydrometeorology*, 15(6):2446–2469, 2014.

[182] J Eamonn Nash and Jonh V Sutcliffe. River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, 10(3):282–290, 1970.

[183] Felix W Landerer and S C Swenson. Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resources Research*, 48(4), 2012.

[184] Michael Durand and Steven A Margulis. Effects of uncertainty magnitude and accuracy on assimilation of multiscale measurements for snowpack characterization. *Journal of Geophysical Research: Atmospheres*, 113(D2), 2008.

[185] Rolf H Reichle, Randal D Koster, Ping Liu, Sarith P P Mahanama, Eni G Njoku, and Manfred Owe. Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR). *Journal of Geophysical Research: Atmospheres*, 112(D9), 2007.

[186] Gregory Gaspari and Stephen E Cohn. Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, 125(554):723–757, 1999.

[187] Dara Entekhabi, Eni G Njoku, Peggy E O'Neill, Kent H Kellogg, Wade T Crow, Wendy N Edelstein, Jared K Entin, Shawn D Goodman, Thomas J Jackson, Joel Johnson, and Others. The soil moisture active passive (SMAP) mission. *Proceedings of the IEEE*, 98(5):704–716, 2010.

[188] Daniel S Wilks. *Statistical Methods in the Atmospheric Sciences*, volume 100. Academic press, 2011.

[189] Vincent Fortin, Mabrouk Abaza, Francois Anctil, and Raphael Turcotte. Why should ensemble spread match the RMSE of the ensemble mean? *Journal of Hydrometeorology*, 15(4):1708–1713, 2014.

[190] Matthieu Lafaysse, Bertrand Cluzet, Marie Dumont, Yves Lejeune, Vincent Vionnet, and Samuel Morin. A multiphysical ensemble system of numerical snow modelling. *The Cryosphere*, 11(3):1173, 2017.