ABSTRACT

Title of dissertation: Characterizing a Multi-Sensor System for Terrestrial Freshwater Remote Sensing via an Observing System Simulation Experiment (OSSE)

Lizhao Wang, Doctor of Philosophy, 2022

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

Terrestrial freshwater storage (TWS) is the vertically-integrated sum of snow, ice, soil moisture, vegetation water content, surface water impoundments, and groundwater. Among these components, snow, soil moisture, and vegetation are the most dynamic (i.e., shortest residence time) as well as the most variable across space. However, accurately retrieving estimates of snow, soil moisture, or vegetation using space-borne sensors often requires simultaneous knowledge of one or more of the other components. In other words, reasonably characterizing terrestrial freshwater requires careful consideration of the coupled snow-soil moisture-vegetation response that is implicit in both TWS and the hydrologic cycle.

One challenge is to optimally determine the multi-variate, multi-sensor remote sensing observations needed to best characterize the coupled snow-soil moisture-vegetation system. Different types of sensors each have their own unique strengths and limitations. Meanwhile, remote sensing data is inherently discontinuous across time and space, and that the revisit cycle of remote sensing observations will dictate much of the efficacy in capturing the dynamics of the coupled snow-soil moisture-vegetation response.

This study investigates different snow sensors and simulates the sensor coverage as a function of different orbital configurations and sensor properties in order to quantify the discontinuous nature of remotely-sensed observations across space and time. The information gleaned from this analysis, coupled with a time-varying snow binary map, is used to evaluate the efficacy of a single sensor (or constellation of sensors) to estimate terrestrial snow on a global scale. A suite of different combinations, and permutations, of different sensors, including different orbital characteristics, is explored with respect to 1-day, 3-day, and 30-day repeat intervals. The results show what can, and what cannot, be observed by different sensors. The results suggest that no single sensor can accurately measure all types of snow, but that a constellation composed of different types of sensors could better compensate for the limitations of a single type of sensor. Even though only snow is studied here, a similar procedure could be conducted for soil moisture or vegetation.

To better investigate the *coupled* snow-soil moisture-vegetation system, an observing system simulation experiment (OSSE) is designed in order to explore the value of *coordinated* observations of these three separate, yet mutually dependent, state variables. In the experiment, a "synthetic truth" of snow water equivalent, surface soil moisture, and/or vegetation biomass is generated using the NoahMP-4.0.1 land surface model within the NASA Land Information System (LIS). Afterwards, a series of hypothetical sensors with different orbital configurations is prescribed in order to retrieve snow, soil moisture, and vegetation. The ground track and footprint of each sensor is approximated using the Trade-space Analysis Tool for Constellations (TAT-C) simulator. A space-time subsampler predicated on the output from TAT-C is then applied to the synthetic truth. Furthermore, a hypothesized amount of observation error is injected into the synthetic truth in order to yield a realistic synthetic retrieval for each of the hypothetical sensor configurations

considered as part of this dissertation.

The synthetic retrievals are then assimilated into the NoahMP-4.0.1 land surface model using different boundary conditions from those used to generate the synthetic truth such that the *differences* between the two sets of boundary conditions serve as a realistic proxy for real-world boundary condition *errors*. A baseline Open Loop simulation where no retrievals are assimilated is conducted in order to evaluate the added utility associated with assimilation of one (or more) of the synthetic retrievals. The impact of the assimilation of a given suite of one or more retrievals on land surface model estimates of snow, soil moisture, vegetation, and runoff serve as a numeric laboratory in order to assess which sensor(s), either separate or in a coordinated fashion, yield the most utility in terms of improved model performance.

The results from this OSSE show that the assimilation of a single type of retrieval (i.e., snow or soil moisture or vegetation) may only improve the estimation of a small part of the snow-soil moisture-vegetation system, but may also degrade of other parts of that same system. Alternatively, the assimilation of more than one type of retrieval may yield greater benefits to all the components of the snow-soil moisture-vegetation system, because it yields a more complete, holistic view of the coupled system. This OSSE framework could potentially serve as an aid to mission planners in determining how to get the most observational "bang for the buck" based on the myriad of different sensor types, orbital configurations, and error characteristics available in the selection of a future terrestrial freshwater mission.

Characterizing a Multi-Sensor System for Terrestrial Freshwater Remote Sensing via an Observing System Simulation Experiment (OSSE)

by

Lizhao 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
2022

Advisory Committee:
Associate Professor Barton A. Forman, Chair/Advisor
Professor Kaye Brubaker
Professor Matthew Hansen
Assistant Professor Michelle Bensi
Dr. Sujay Kumar

Preface

Looking back on my five-year Ph.D. career, I am very glad that I have chosen this life path. The scenery of this path ranges from noise to peace, from wide to narrow, and then to wide. My life also flows forward from countless streams into a wide river.

My biggest dream is to accumulate my trivial contribution to the world in my ordinary life, through my work step by step. I want to add meanings to meaningless life from what I am doing. Fortunately, I joined a project about how to apply satellite data in land surface model by data assimilation, which expanded my original knowledge frame about remote sensing. The desire for this research work has originally come from my long-time interest for better applying the satellite information in understanding earth surface dynamics. As satellite data has been explored in hydrology studies from different dimensions, it is also important to better design the satellite characteristics to monitor the earth surface more efficiently. I have a strong belief that remote sensing information could provide better support in decision making. This five years of study and teamwork helped me build a pleasant rhythm and pace of work. Every day I immerse myself in doing my own things, then get a good rest and communicate with team members on time.

Also, this lengthy work was impossible without the teamwork of our project. Their amount of hard work is truly commendable. The most lucky thing is that I met my mentor, Dr. Forman, who had the greatest impact on me through the time. He is excellent, energetic and smart. He not only helped me a lot, but also influence me imperceptibly as a role model. I am also deeply grateful to Professor McCuen. He gave me the most wonderful learning experience of my life. His teaching let me know the best state of a professor. Thank Professor Brubaker for giving me

a lot of guidance and help in my study. Further, I would like to express my deep gratitude to Dr. Brubaker, Dr. Bensi, Dr. Kumar, and Dr. Hansen for agreeing to serve on my dissertation committee and for giving their constructive and insightful comments and suggestions on the proposal and dissertation, which has greatly expanded my vision about the studies.

In my doctoral career, I am glad to study with many excellent and friendly students. Thank Gaohong and Jongmin as we have spent a lot of lunchtime together in interesting talks. Also thank them and Jawairia for their discussion with me in research and inspiration to me. Thank Yonghwan for his help and guidance in my work. Thank Jing Wang for her encouragement and support. Thank Yuan Xue for the experience exchange in learning, and research. Thank Colin and Alireza for our wonderful cooperation in projects. Thanks to all partners in the laboratory, my study life is more wonderful because of all of you. In the project cooperation, thanks to my partners, Dr. Kumar and Dr. Wrzesien, the cooperation with them is both effective and pleasant. Thanks to Dr. Kim, his ideas always enlighten our work. During those work experience, I felt the real team spirit and common interest. I would like to thank the staff, IT team and administrative staff of the department of Civil and Environment Engineering for providing fast and effective support to my research.

Thanks to my family also for supporting me all the time. Behind all the turbulence and trough, I always feel the love and support from my family. I am especially grateful to my daughter Ellen. Your arrival is the most wonderful experience and bloom in my life. Your dependence makes me realize the beauty and enrichment of a deep love. Your growth is the biggest driving force and inspiration in my life. Thank my wife Jingjing, the love of my life and the best partner. Your enthusiasm for life have given me the most gorgeous color in my life. I am willing to work hard for us. Thank my parents for your love and company. Thank my father-in-law and

mother-in-law for your support and encouragement.

The stream of life is still rushing forward, and I will pack my bags and run to the next journey. No matter where my dream continues to shine, I will spend every day with my heart and grow up with my family.

Acknowledgments

Thanks for the financial support from the NASA AIST projects (under contract 80NSSC17K0254 and 80NSSC20K0210), so that I can work on such a meaningful and interesting topic.

Thanks to the TAT-C team, their support is fundamental for my simulation in this work. Thanks to the LIS team, they help me a lot on LIS developments and make this work possible. Thanks to the UMD DIT team, their prompt and strong assistance solved most problems I met in the high performance computing system.

I also want to acknowledge the assistance of UMD EEIGS program, helping me on academic writing.

I would like to express my deepest appreciation to my committees, for their precious suggestions and comments.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AMSR-2	Advanced Microwave Scanning Radiometer 2
AMSR-E	Advanced Microwave Scanning Radiometer - Earth Observing System
CLSM	Catchment Land Surface Model
CSA	Canadian Space Agency
DA	Data Assimilation
EKF	Extended Kalman Filter
EnKF	Ensemble Kalman Filter
GEDI	Global Ecosystem Dynamics Investigation
GRACE	Gravity Recovery And Climate Experiment
HYMAP	Hydrological Modeling and Analysis Platform
ICESat2	Ice, Cloud and land Elevation Satellite-2
IMS	Interactive Multisensor Snow and Ice Mapping System
ISS	International Space Station
LAI	Leaf Area Index
LiDAR	Light Detection And Ranging
LIS	Land Information System
LSM	Land Surface Model
MERRA2	Modern-Era Retrospective Analysis for research and applications (v2)
MODIS	Moderate Resolution Imaging Spectroradiometer
MPDI	Microwave Polarization Difference Index
MW	Microwave
NASA	The National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
NIC	Normalized Information Contribution
NIR	Near Infrared
NLDAS-2	Phase 2 of the North American Land Data Assimilation System
Noah-MP	Noah multi-physics parameterization
NR	Nature Run
OL	Open Loop
OSSE	observing system simulation experiment
PF	Particle Filter
PMW	Passive Mcirowave
RADAR	Radio Detection and Ranging
RMSE	Root Mean Squared Error
SAR	Synthetic Aperture RADAR
SAST	Simple Snow–Atmosphere–Soil Transfer Model
SLWC	Snow Liquid Water Content
SMAP	Soil Moisture Active Passive
SMMR	Scanning Multi-channel Microwave Radiometer
SMOS	Soil Moisture and Ocean Salinity
SNOTEL	SNOwpack TELemetry
SPD	Spectral Polarization Difference
SSM/I	Special Sensor Microwave/Imager

SWE	Snow Water Equivalent
TAT-C	Tradespace Analysis Tool for Constellations
T_b	Brightness Temperature
TIR	Thermal Infrared
TM	Landsat Thematic Mapper
TRMM	Tropical Rainfall Measuring Mission
TWS	Terrestrial Water Storage
ubRMSE	Unbiased Root Mean Squared Error
VIS	Visible
WMO	World Meteorological Organization

Chapter 1: Introduction

1.1 Motivation

Freshwater is a fundamental resource required for human well-being [1–3]. Water scarcity can lead to poor health, and has an especially adverse impact on poor people. More than 1.1 billion people live in areas of physical water scarcity [2]. Demand is increasing for this finite resource as the world's population continues to grow [1]. Furthermore, with existing climate change, it is anticipated that almost half the world's population will be living in areas of high water stress by 2030 [3]. It is crucial for sustainable socio-economic development to monitor and understand global and regional water resources [4].

Natural surface water distributions are dynamic [5]. Achieving a good understanding of water resources dynamics requires knowledge and understanding based on models and observations. In particular, information is needed on hydrological variables that constitute the hydrological cycle on land and how they vary over time and space. Thus water management requires frequent observations to detect its existence, extent, volume, and dynamics.

However, quantifying the amount of freshwater at a global scale is difficult and complex. Terrestrial water storage (TWS) data is limited and often inaccessible in many parts of the world due to geographic remoteness and/or closed data policies [6]. Spaceborne remote sensing is one of the most feasible ways to monitor freshwater resources across the globe. Among TWS

components, snow, soil moisture, and vegetation are among the most visible parts from space [7–11]. Although snow, soil moisture, and vegetation only account for a relatively small part of the TWS [12], the snow-soil moisture-vegetation system plays a key role in the hydrological cycle, and serves as a first-order control on surface runoff, transpiration, and evapotranspiration, and the prediction of weather and climate [13–15].

A perplexing problem while trying to quantify the components of the snow-soil moisture-vegetation system is that they can at times shroud the measurement of one another such that quantifying one of these components individually often requires knowledge of one or more of the other components [7–11]. Many studies have been completed to estimate the individual parts of the snow-soil moisture-vegetation system [16–18], but a true estimation of this system can be made only by studying them as a cohesive whole rather than as the mere sum of its individual parts.

1.2 Literature Review of Terrestrial Freshwater Storage Monitoring Techniques

1.2.1 Terrestrial Freshwater Storage Retrieved via Remote Sensing

Remote sensing is a viable way to monitor freshwater resources across the globe [19]. Many state variables within the hydrologic cycle can be remotely monitored by spaceborne sensors. For example, NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) products provide daily retrievals of vegetation leaf area index and snow cover extent [20, 21]. Imagery from L-band and X-band radiometers such as the Soil Moisture and Ocean Salinity (SMOS) and the Advanced Microwave Scanning Radiometer (AMSR-E) can be used to retrieve soil moisture information [22]. However, it should be noted that these retrievals are not a direct measurement

of the hydrologic variables, but rather a remotely-sensed estimate of these variables [16, 18, 23].

Each retrieval has its own specific footprint size, spatial resolution, inherent spatiotemporal gaps, and unique error characteristics. It is important to consider the specifications of different satellite retrieval methods because hydrologic variables are often coupled to one another [24]. Interaction between different state variables often demands that quantifying one of these variables requires knowledge of one or more of the other variables. **Therefore, the choice of an optimal, concurrent sensor configuration in order to utilize the strengths of different sensors is a key question for the next generation of terrestrial hydrology missions.**

1.2.1.1 Snow Remote Sensing

Different snow variables, such as snow cover extent, snow depth, and snow water equivalent (SWE; equal to the mass of water per unit area), can be obtained using remotely-sensed observations. SWE is the main snow-related variable of interest because it best characterizes the snow mass. Remote sensing of SWE is an important way to obtain global SWE estimates since in-situ measurements are difficult to obtain given the harsh environment and the heterogeneous large-scale snow distribution [9]. Based on the electromagnetic response of the snowpack, SWE can be empirically derived as a function of the spectral difference between the bands of 18 GHz and 36 GHz obtained from a passive microwave (PMW) radiometer, such as the Special Sensor Microwave/Imager (SSM/I) or AMSR-E [25–27]. However, the effect of vegetation on PMW radiometer observations is a factor that cannot be ignored while retrieving SWE. Studies show that the presence of vegetation can lead to an underestimation of SWE of up to 50% [28, 29]. In the AMSR-E SWE product, vegetated and non-vegetated fractions are retrieved separately

[30] and assessed as a function of the forest cover fraction. Furthermore, the spatiotemporal distribution of snow density is a dominant source of errors in the AMSR-E SWE retrievals, and hence, must also be carefully considered [9].

The existing PMW SWE products are too coarse (> 10 km) in spatial resolution to provide accurate information for snow studies in regions of complex terrain [9]. PMW SWE retrievals, in general, also have poor agreement with reference data from ground-based measurements, particularly in deep snow, wet snow, or snow with overlying vegetation [31].

Active microwave (AMW) synthetic aperture radar (SAR) is also a feasible technique for global snow mass detection. Bernier et al. [32] estimated SWE using RADARSAT images collected during the winter of 1998 and 1999 at La Grande River watershed in eastern Canada. Zhu et al. [33] developed a SWE algorithm using SnowSAR (airborne X- and Ku-band) and tested it over a mixed vegetation area as well as an open tundra landscape. These practices showed SAR has the potential to retrieve SWE at a much finer spatial resolution (~ 0.5 km) compared to PMW sensors, but that SWE retrieval accuracy is limited by the difficulty in properly capturing complex scattering processes over snow-covered [34]. The accuracy of AMW SWE retrievals is also degraded over wet snow or snow with overlying forest cover terrain [34]. Ka-band and Ku-band spaceborne SAR sensors are more sensitive to snow signals, and are expected to show better performance in shallow snow [35, 36].

Airborne LiDAR has been successfully applied to retrieving both snow depth and SWE at the watershed scale [37, 38]. Over non-mountainous terrain, airborne LiDAR systems exhibit 15-20 cm vertical accuracy [39, 40]. Spaceborne LiDAR applications to global SWE are limited because the revisit period of existing spaceborne LiDARs is too long (several months) to catch the seasonal variations of snow [41]. However, the fundamental features of LiDAR allows them to

provide key information of snow depth in the presence of overlying vegetation, which is difficult to obtain using microwave sensors [38].

1.2.1.2 Soil Moisture Remote Sensing

Due to the dynamic and heterogeneous nature of soil moisture, remotely-sensed soil moisture products are a good supplement in order to fill in and augment the ground-based observational network across regional and global scales [42]. Soil moisture can be retrieved from land surface reflectance derived from optical remote sensing. The relationship between soil moisture and vegetation (e.g., Normalized Difference Vegetation Index, NDVI) has been investigated in a number of studies [43–45]. However, the accuracy of this method to retrieve soil moisture is limited because soil reflectance observations are also strongly affected by the inherent (physical) properties of the soil itself, which makes reflectance observation highly variable depending on the ecological setting being monitored [46].

Another method to retrieve soil moisture from space involves using thermal infrared (TIR) observations. The basic principle of TIR soil moisture remote sensing is that the temperature of the land surface is sensitive to the soil moisture given that water has a relatively large heat capacity. By measuring the amplitude of the diurnal temperature change, one can develop a relationship between the temperature change and the corresponding soil moisture content. However, this relationship is also a function of soil type. Further, this method is limited in the presence of vegetated soil. As a result, TIR methods are difficult to apply to large-scale soil moisture monitoring projects [46].

Early in the 1970s, researchers recognized that passive microwave sensor measurements are

useful for estimating surface soil moisture content [47]. The brightness temperature (T_b) observed by the radiometer is a function of the soil physical temperature and surface emissivity [48]. In order to control the influence of overlying vegetation, a number of PMW algorithms have been developed that employ a microwave polarization difference index (MPDI) as an indicator for characterizing vegetation conditions. Existing spaceborne remote sensing soil moisture products, such as AMSR-E, SMOS, and Soil Moisture Active Passive (SMAP), are widely used in hydro-meteorological, climate, and agricultural applications [42]. Evaluations have shown that they perform well in capturing the annual cycle and short-term variability of the surface soil moisture. [49, 50]. The advantage of using PMW radiometers is that they are not limited by the presence of clouds, weather, or glint. However, the coarse spatial resolution (on the order of tens of kilometers) makes PMW soil moisture remote sensing products difficult to apply to basin- or watershed-scale applications without first undergoing some form of spatial downscaling.

Backscatter observations from SARs are also sensitive to soil moisture. However, the retrieval of soil moisture from SAR is confounded by the effects of soil surface roughness and overlying vegetation [51]. Furthermore, the geometric properties of the soil surface and vegetation have a significant effect on SAR observations [46]. As a result, there are a limited number of applications using SAR as an operational source for soil moisture as applied to hydrology despite the benefits of a high-resolution, distributed soil moisture product [52].

1.2.1.3 Vegetation Remote Sensing

Leaf Area Index (LAI) is an important land surface variable for monitoring the condition of vegetation productivity, and is used as an input in a number of ecosystem process models [53].

LAI retrievals are often produced using spaceborne optical sensors at a variety of different spatial resolutions. LAI has been retrieved with the Landsat Thematic Mapper (TM) data at a 30-m resolution over the boreal forest zone in Canada [54]. Global LAI estimates have been produced using the MODIS data at 1-km resolution [55]. World Meteorological Organization (WMO) reported a relative error less than 10% is required in the applications on weather prediction, global climate-carbon modeling, and hydrologic modeling [56]. Studies show that MODIS, the most popular global LAI product, is slightly worse than the accuracy requirement by WMO [57, 58] in all species average. Species related bias is a major source of error that impacts LAI retrieval uncertainty.

LiDAR is a relatively new form of remote sensing used in forest applications. Airborne LiDAR has been widely used in retrieving LAI and monitoring forest density [17, 59]. However, there are relatively few applications of spaceborne LiDAR given the limited availability of spaceborne LiDAR systems specifically designed to study vegetation [60]. The Global Ecosystem Dynamics Investigation (GEDI) mission is the first spaceborne LiDAR specifically optimized to measure vegetation structure [61, 62]. It will form the basis of critical reference datasets for vegetation biomass between 52° S and 52° N [61]. The GEDI instrument employs three laser transmitters, beam-splitting optics, and an optical dithering device to produce 14 ground tracks spaced 500 m across-track to yield a total swath width of 6.5 km. The transmitted pulse shape and returned waveform is recorded for each 25 m diameter footprint (Figure 1.1) [63]. The GEDI liner track is then resampled to a $500m \times 500m$ grid [64].

The launch of GEDI raises the prospect of researchers retrieving LAI with lower errors using spaceborne LiDAR. Dubayah et al. [61] reported the accuracy of a single point of GEDI observations should be better than 20%. Qi et al. [65] evaluated the 1-km GEDI biomass retrievals

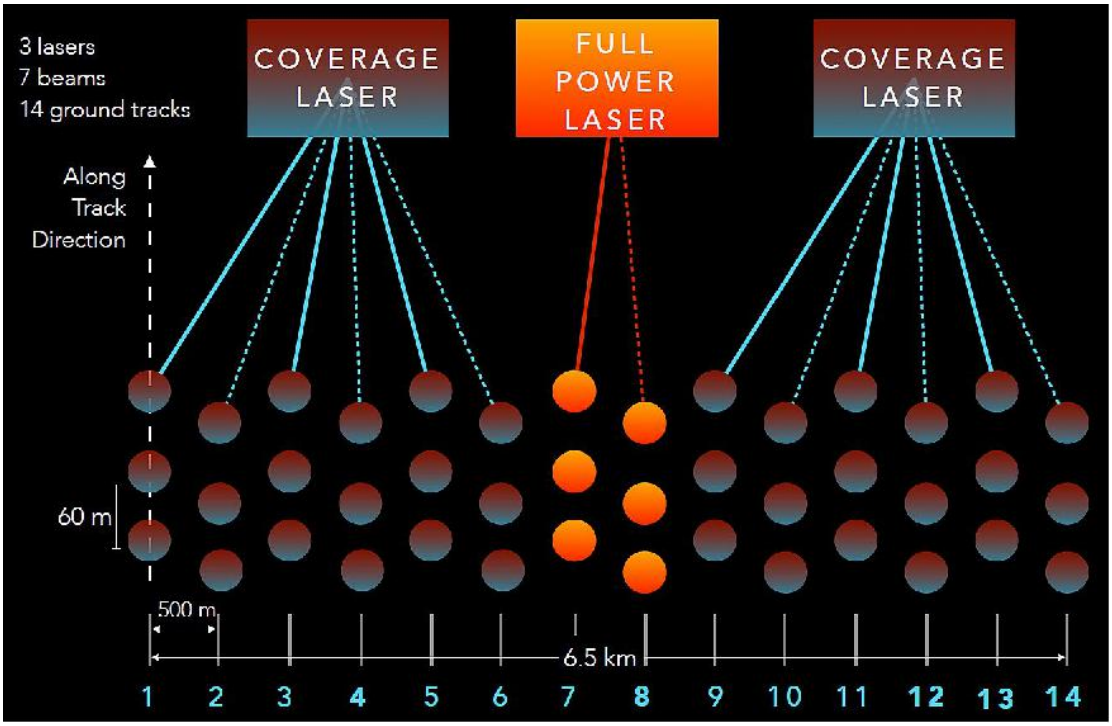

Figure 1.1: GEDI's ground sampling pattern. The red-blue points refer to samples of coverage laser, and the orange points refer to full power laser. (Image credit: https://gedi.umd.edu/instrument/specifications/)

Table 1.1: Orbital configurations of typical sensors used in SWE, soil moisture and LAI retrievals.

Geographical Retrieval	Proxy	Sensor Type	Spatial Resolution	Orbital Height	Inclination Angle
SWE	AMSR-E	PMW	25 km	705 km	97°
	Sentinel-1A	SAR	0.04 km	693 km	98°
	GEDI	LiDAR	1 km	400 km	51°
Soil Moisture	SMOS	PMW	30 km	765 km	98°
	SMAP	SAR	1 km	685 km	98°
LAI	MODIS	Optical	0.25 km	705 km	98°
	GEDI	LiDAR	1 km	400 km	51°

Acronyms:
SWE = Snow Water Equavalent; LAI = Leaf Area Index;
PMW = Passive Microwave; SAR = Synthetic Aperture RADAR;
LiDAR = light detection and ranging

against ground-based measurements, and reported uncertainties ranging from 11%–20% across the different sites.

In summary, remote sensing is an important tool to obtain information related to snow, soil moisture, and vegetation. Various remote sensing sensors, such as PMW radiometers, RADARs, passive optical imagers, and active LiDARs are widely used in hydrological applications. The spatial resolution of each, in general, varies from tens of meters to tens of kilometers. The orbital configurations of these sensors are typically designed to meet specific mission goals (Table 1.1). However, **differences in spatial resolution, revisit frequency, sensor error characteristics, and swath widths complicate the unified coordination of these different sensors** in order to explore the terrestrial hydrologic cycle as a holistic system.

1.2.2 Land Surface Models

To meet the need of water resource monitoring and prediction, researchers have developed land surface models (LSMs) as simplified representations of the conservation of mass and energy

equation (Figure 1.2). It is an important and necessary tool for water and environment resource management [66]. An LSM describes the interaction in the atmosphere-canopy-soil surface system and helps quantify unobserved states related to the movement of heat and water in the natural environment. Most LSMs contain representations of the atmosphere, hydrology, plant physiology, and soil physics, which are interlinked [67].

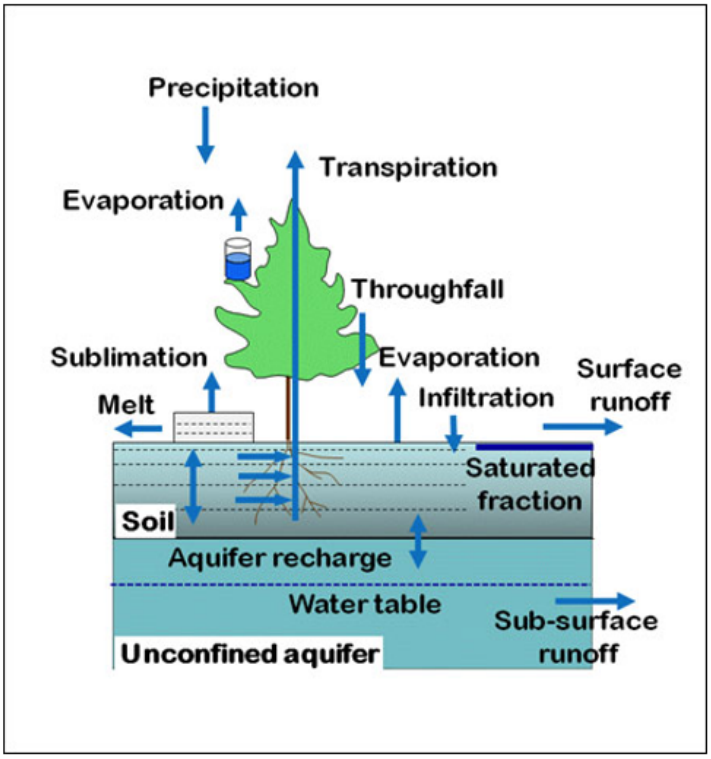

Figure 1.2: Conceptual representative of the state variables and fluxes in an example LSM (Community Earth System Model). Image credit: http://www.cesm.ucar.edu/models/clm/hydrologic.html.

LSMs have become more comprehensive over time as they increasingly represent more physical interactions at the land surface. For example, land-atmosphere models evolved from the simple, single-layer model into two-layer and multi-layer models [68]. In the single-layer (or "big-leaf") model, the land surface is treated as one homogeneous surface [69]. The only

resistance between the land surface and the atmosphere is assumed to be the atmospheric resistance defined as the ability of the air to transport a given quantity away from the land surface. This assumption has since been relaxed to define surface resistance as a function of vegetation and soil moisture [70]. The single-layer model is shown to perform well over dense vegetation [71], however, does not perform as well for the clustered vegetated surface. Furthermore, a single-layer model does not allow for interactions between the soil surface and vegetation, which are significant especially when the vegetation is sparse [72].

In the two-layer models, vegetation is viewed as a single layer located above the soil surface [73]. To better describe the interaction between vegetation and soil, some models extend the one vegetation layer to multiple layers [74]. Although the detailed models such as two-layer and multi-layer models provide more realistic results, they require more detailed information about the vegetation parameters, and therefore, are hard to apply outside the intensely observed experimental areas. Compared to the detailed models, the single-layer models are widely used because of their simplicity. Unfortunately, these simple models make the model parameters hard to infer and highly dependent on calibration [68].

Even though the LSMs are evolving, large uncertainty still remains in our understanding and modeling of the interactions in the atmosphere-canopy-soil surface system. First, most of the processes in LSMs are simplified due to computational cost and input data availability constraints. For example, many of LSMs classify plant species into plant functional types, within which the parameters are undifferentiated. Simulations consisting of a limited number of plant functional types may ignore biodiversity within a simulation grid [75]. Second, the inaccurate initial conditions of state variables can introduce uncertainties into the simulations. It is difficult to provide correct initial fields for state variables that have significant spatial heterogeneity [76].

The third source of uncertainty comes from boundary conditions, such as precipitation and air temperature. The uncertainty from boundary conditions could be considerable in some cases. For example, studies show that large biases and uncertainties exist in precipitation products over mountainous areas, where precipitation is heterogeneously distributed and the ground gauges are relatively sparse [77, 78].

1.2.3 Data Assimilation

There are advantages and disadvantages of LSMs and remote sensing observations [75]. Data assimilation (DA) is a technique to leverage both remote sensing observations and LSMs together to increase model (posterior) skill. The basic idea behind DA is to combine (in a Bayesian sense) complementary information from observations and LSMs, and thus, optimally estimate geophysical state variables.

DA algorithms first emerged from the Kalman filter with Bayesian estimation techniques [79]. Then a series of algorithms were developed based on this theory, such as the [Extended Kalman Filter (EKF) for nonlinear systems [80], and the Ensemble Kalman Filter (EnKF) using a Monte Carlo implementation of Bayesian updating [81]. Another alternative algorithm based on Monte Carlo methods is the Particle Filter (PF) [82]. Compared to classical Kalman Filtering methods, PFs can handle the propagation of non-Gaussian distributions through nonlinear models.

In the area of hydrology, DA has been widely applied as a method of improving the characterization of snow, soil moisture, or runoff estimates. For example, Durand et al. [83] assimilated Special Sensor Microwave Imager (SSM/I) and AMSR-E PMW T_b observations into the Simple Snow–Atmosphere– Soil (SAST) transfer model [84] to improve SWE estimation.

The result showed SWE and snow melting timing improvements via data assimilation. Lievens et al. [85] leveraged the soil moisture retrieved from Sentinel-1A/B into the Catchment Land Surface Model (CLSM) and improved the estimate of soil moisture [85]. Kumar et al. [86] used multisensor, multivariate data assimilation that included a suite of soil moisture, snow depth, and snow cover data from Scanning Multi-channel Microwave Radiometer (SMMR), the Advanced Scatterometer (ASCAT), the Advanced Microwave Scanning Radiometer (AMSR-E/AMSR2), SSM/I, MODIS, SMOS mission, and SMAP mission [86]. The practice proved that the assimilation of terrestrial land remote sensing products provided a high quality estimate of land surface state variables.

1.3 The Problem

The complexity of the land surface system has resulted in complex sources of uncertainty in LSMs. Assimilation of remote sensing observations helps to control the growth of uncertainty in model simulations. However, remote sensing observations collected by spaceborne passive optical (VIS/NIR) radiometers, passive microwave (MW) radiometers, RADAR, and LiDAR are not snow or soil moisture or vegetation observations, per se, but rather information *related* to snow or soil moisture or vegetation [87–89]. Further, these remote sensing data contain their own unique characteristics (e.g., error structure, field of view, overpass time, spatial resolution) that need to be carefully considered when integrating with a LSM. To optimally utilize the spaceborne sensors collecting measurements relevant to snow, soil moisture, and vegetation retrievals from space-based instrumentation, a robust framework is needed to explore the interplay (a.k.a., tradeoff space) between the different sensor types, swath widths, geophysical retrievals, and error characteristics.

Another problem is determining how to optimally utilize the observations from multiple sensors. For a single geophysical state variable assimilation, observations from multiple types of sensors (i.e., passive optical, passive MW, active MW, LIDAR or any combination thereof) can provide complementary information for each other, since measurements from each of these instrument types have their own strengths and weaknesses. For the multiple state variables assimilated into a coupled snow-soil moisture-vegetation system, multiple observations are essential to capture the synergistic interplay among different variables. However, the optimal utilization of multiple observations first requires that one quantify the individual contribution of each single measurement type prior to quantifying the synergistic interplay between a suite of different measurement types.

In order to solve the problems highlighted above, an observing system simulation experiment (OSSE) is developed in this dissertation. In this OSSE, the "synthetic truth" is first generated by the Nature Run (NR) using the NoahMP-4.0.1 land surface model within the NASA Land Information System (LIS). Then, synthetic observations from a series of existing *and* hypothetical sensors are simulated and assimilated into the NASA LIS. An Open Loop (OL) is set with different model parameters and boundary conditions compared to the NR in order to represent the model only performance **without** assimilation. Then, the assimilation results are evaluated relative to the NR and OL, so that the scientific value of each sensor/constellation configuration can be quantified. This OSSE quantifies the added value to the model from the synthetic observations, and their combinations, to help answer the overarching science question: What is the optimal sensor/constellation configuration choice with the goal of quantifying water in a coupled snow-soil moisture-vegetation system?

1.4 Dissertation Structure

In Chapter 2, analysis of the viewing extents of different sensors on-board hypothetical satellites/constellations is conducted in the context of terrestrial snow, in order to evaluate the additive value of new observations in terms of spatial coverage. In Chapter 3, synthetic observations are assimilated into the Noah-MP 4.0.1 land surface model in order to evaluate the value of different observations in terms of model accuracy. In Chapter 4, the findings from Chapter 2 and Chapter 3 are discussed and summarized.

1.5 Original Contributions

1) This study linked the Trade-space Analysis Tool for Constellations (TAT-C) [90] and a time-varying snow binary map to **evaluate the efficiency** of a single sensor (or constellation) to **observe terrestrial snow on a global scale**. The study explores **a range of snow researchers' requirements** (in term of temporal integration periods) for remote sensing snow observations. Therefore, it could assist researchers and mission planners weighing the different trade-offs based on sensor selection and orbital configurations in future snow missions.

2) This study built an integrated OSSE framework coupling TAT-C with LIS to quantify the gains via assimilating synthetic retrievals from a range of feasible mission design options (e.g., orbital configuration, retrieval error characteristics, and mixes of sensor types). This framework could **quantify the scientific gains as a function of mission design**. Therefore, it can serve as **an effective planning tool** for the next generation hydrological missions with particular relevance to snow, soil moisture, and vegetation.

3) The OSSE built in this work is used to assimilate the snow, soil moisture, and vegetation simultaneously. The **synergistic interplay** among components in the snow-soil moisture-vegetation system is explored. It provides a more **cohesive estimation** of terrestrial hydrology, which provides knowledge to the community about what sensors/constellations are needed to better characterize the coupled snow-soil moisture-vegetation system.

4) This study developed a metric, i.e., integrated normalized information contribution (NIC) to evaluate the multi-sensor, multi-variate assimilation cohesively. The integrated NIC can provide an overall assessment of the impact of assimilating synthetic retrievals on their own or in combination with others.

Chapter 2: Exploring the Spatiotemporal Coverage of Terrestrial Snow Mass using a Suite of Satellite Constellation Configurations

2.1 Introduction

2.1.1 Motivation and Objectives

Terrestrial snow is a vital freshwater resource for more than 1 billion people. Remotely-sensed snow observations can be used to retrieve snow mass or integrated into a snow model estimate; however, optimally leveraging remote sensing observations of snow is challenging. One reason is that no single sensor can accurately measure all types of snow because each type of sensor has its own unique limitations. Another reason is that remote sensing data is inherently discontinuous across time and space, and that the revisit cycle of remote sensing observations may not meet the requirements of a given snow applications. In order to quantify the feasible availability of remotely-sensed observations across space and time, this study simulates the sensor coverage for a suite of hypothetical snow sensors as a function of different orbital configurations and sensor properties. The information gleaned from this analysis coupled with a time-varying snow binary map is used to evaluate the efficiency of a single sensor (or constellation) to observe terrestrial snow on a global scale.

2.1.2 Literature Review

Snow is an important component of global freshwater storage. It provides freshwater supply for more than 1 billion people [91–93]. Snow covered terrain serves as a natural reservoir that slowly attenuates freshwater runoff during the snow ablation season [94]. Snow albedo also plays an important role in energy balance and climate change. For example, atmospheric warming could reduce the seasonal snow cover, and hence, increase shortwave absorption at the land surface, which could introduce a positive feedback [95].

Snow storage estimation is increasingly important as the virtual reservoir of snow is threatened by global warming and climate change [13,96,97]. Earlier snowmelt due to global warming could exacerbate severe floods and droughts [31, 98, 99]. As a result, the vulnerability of snow storage has attracted considerable interest from the hydrologic community to monitor the equivalent amount of liquid water contained within the snowpack (a.k.a, snow water equivalent or SWE) so that this vital resource may be better managed and preserved.

2.1.2.1 Limitations of Existing Spaceborne Snow Products

Spaceborne remote sensing is the only viable technique to detect SWE across the globe in a timely manner [9] when considering the large spatial extent of snow and the difficulties in collecting *in-situ* measurements over such large regions. Despite the extensive efforts of researchers to provide accurate SWE retrievals, current SWE products using passive microwave (PMW) remote sensing observations still do not meet the accuracy requirements (±15%) needed to support operational decision-making at a continental or global scale [100, 101]. One main reason is that the knowledge gap in coupling precise physical emission models (i.e., radiative

transfer models [RTMs]) of snowpacks to remotely-sensed observations [9].

A general approach for using PMW observations is the spectral polarization difference (SPD) [102] employing the Chang algorithm [103] and its modification for forested areas [104]; this approach uses PMW spectral difference, i.e., the difference in brightness temperature between two microwave frequency channels. This technique is effective in some regions e.g., dry, shallow snow on flat terrain, but is unable to detect thin snow due to a low signal-to-noise ratio, or accurately retrieve deep snow due to signal saturation, or snow with overlying vegetation due to vegetation attenuation [9, 105]. To help overcome the problems of PMW observations across regional or continental scales, some studies have fused satellite data with ground-based snow measurements to better estimate SWE [106, 107].

There has never been a dedicated satellite mission for snow mass (SWE) detection. Existing spaceborne sensors used to estimate snow mass have typically been designed for a different purpose [108]. With the above limitations, most operational, stand-alone passive microwave SWE products produced using the spectral difference method are far from optimal. Typically, these products are inconsistent with independent reanalysis data and ground-based measurements from meteorological stations and snow courses [31, 109], particularly in deep snow, wet snow, snow in complex terrain, or snow with overlying vegetation [31, 105, 110].

Analogously, the Moderate Resolution Imaging Spectroradiometer (MODIS), a passive visible and thermal infrared radiometer, was designed to view the spatial extent of snow rather than the mass of snow within that snow-covered extent. Thus, MODIS is limited in skill in terms of snow mass estimation although it does a good job of viewing where snow is found on the ground in the absence of dense forest [98, 111]. More recently, active microwave (AMW) synthetic aperture radar (SAR) has been employed for global snow mass detection [112, 113], but

two issues remain unsolved: one is the limited repeat overpass (relative to Advanced Microwave Scanning Radiometer (AMSR) or MODIS, e.g.), the other is that C-band microwave radiation on Sentinel-1, e.g., is less sensitive to snow volume scattering than X-band and less skilled than Ku-band in the detection of shallow snow [35, 36], although recent studies showed its potential at mapping snow mass in mountainous regions [114]. Suffice it to say that the development of such sensors for purposes other than terrestrial snow mass limits the skill of these sensors in the application to global SWE estimation. As a result, the snow science community is discussing the prospect of a future, dedicated spaceborne snow mission which would be the first of its kind [115, 116].

2.1.2.2 Limitations of Remote Sensing Snow Techniques

The complexity of snow further confounds the retrieval of snow properties using remotely-sensed observations such that no single technique can work for all types of snow. One significant source of uncertainty in retrieving snow mass comes from the sensitivity of remotely-sensed observations to other variables such as snow microstructure (e.g., grain size), snow stratigraphy, the amount of liquid water content coating the snow grains, overlying vegetation, complex topography, and atmospheric and cloud conditions [105]. Microwave radiation, in general, is sensitive to the snow liquid water content (SLWC) such that even a small amount of liquid water in the snowpack greatly alters the dielectric constant, and hence the emissivity and absorptivity, of the snowpack. SLWC, which is relatively large during the melting season, often introduces large uncertainties into SWE retrievals from both PMW and AMW retrievals [117]. Snow density, snow grain size, and snow grain shape are other variables that influence the snow emissivity and scattering

characteristics, and as a result, the corresponding electromagnetic response of the snowpack [9]. The variability in these snow characteristics can result in a strong correlation between the PMW signals and snow mass for some years, but not for other years. Additionally, the complex microstructure of a snowpack due to variations in depth hoar, internal ice layering, and vertical heterogeneity increase the spatial and temporal variability of the snowpack [9]. Overlying vegetation further complicates snow remote sensing by attenuating microwave emission from the snowpack while simultaneously contributing its own signal as measured by the spaceborne radiometer. Findings have shown that PMW SWE retrievals tend to underestimate SWE in forested areas [29]. Complex terrain such as in mountains, where much of the global snow is located, also reduces the efficacy of coarse-resolution sensors such as PMW radiometers or AMW scatterometers [105, 118, 119]. All of these uncertainties are exacerbated by the coarse-scale resolution of these measurements that cannot adequately capture the true spatial variability of SWE [120, 121]. To overcome this problem, recent studies improved large-scale patterns of snow mass estimation by merging PMW observations with ground-based measurements [107, 114].

In the context of spaceborne LiDAR, which is another option for retrieving snow depth and snow mass, a major limitation is the relatively narrow swath width of LiDAR [41] ($\sim$ 10 km) compared to SAR ($\sim$ 100 km) or PMW radiometry ($\sim$ 1000 km). The individual LiDAR beams typically obtain tracks with widths of 100 meters or less. Using multiple LiDAR beams as part of a sampling strategy, the swath width of a spaceborne LiDAR retrieval is typically around 6 km [61, 122]. Furthermore, the optical signal used by a snow LiDAR cannot penetrate optically thick clouds [123], hence, the snow under the clouds remains unobserved. LiDAR is an effective tool for retrieving snow depth, but that effectiveness is severely curtailed when considering swath width limitations and cloud attenuation.

In short, no single spaceborne sensor will adequately measure all types of snow under all conditions required for global snow monitoring. Rather, a mixture of observations from different sensors, each with its own strengths and weaknesses, is needed to yield the best estimate of global snow mass [41].

2.1.2.3 Snow Mass Mission Trade-offs

To make global snow mass estimation even more complicated, a future snow mission will face a trade-off between sensor design, spatial resolution, and revisit frequency. For example, a different orbital configuration (largely as a function of inclination angle and satellite altitude) changes the nadir track, which directly influences which portions of the globe are, or are not, observed. Similarly, a wider swath width for a given sensor increases the revisit frequency but likely results in larger errors along the swath periphery due to slant range geometry effects or significant reductions in backscatter, e.g., associated with changes in forward scattering characteristics as the sensor looks increasingly off-nadir [124]. An increase in satellite altitude or changes in orbital parameters could impact the spatial resolution as well as the frequency with which the globe is viewed. However, both the fine spatial resolution and short revisit interval are of interest in SWE estimation considering the strong spatiotemporal dynamics of the snow [125, 126]. Additional concerns about this trade-off include non-uniform distribution of snow cover and diversity of snow types or snow features. The majority of snow occurrence is distributed in the high-latitude or high-altitude regions. Consideration of the different suitable remote sensing techniques to capture the different types of snow such as tundra, taiga, or ephemeral snow, adds even more complexity to the task of global snow mass estimation.

2.1.3 Objectives

Given the difficulties listed above, along with the general lack of uptake of PMW estimates of SWE into hydro-meteorology and hydro-climatology applications [120], the research presented here aims to study the efficacy of different orbital configurations and sensor characteristics on snow mass detection from the perspective of maximizing the global snow coverage to be viewed. The goal of this exercise is to facilitate the mission planning process and enhance the future potential of snow remote sensing, e.g., PMW, LiDAR, and SAR.

The first challenge is to link and combine the prediction ability of different satellite orbits and estimation of snow dynamic extent. The sensor's viewing extent is estimated under various orbital configurations coupled with a snow cover climatology as a function of different snow classes. The goal is not to quantify global snow mass (which is to be pursued in a follow-on study), but rather to explore the different options to quantify snow mass, and how best to maximize global coverage (in space and time) en route to estimating global snow mass.

2.2 Methodology

The snow mass detection capability of sensors (and sensor constellations by construct) is limited by the sensors' viewing extent, the distribution of snow in space and time, and the sensors' efficacy to specific snow conditions (e.g., ability to retrieve deep snow, wet snow, snow overlain by vegetation). This section introduces the methods to analyze these three factors, including the methodology used to simulate the sensor viewing extent, different scenarios for simulations, application of a time-varying snow mask, and metrics for use in evaluating observation efficacy.

2.2.1 Simulation of Sensor Viewing Extent

In order to simulate the viewing extent of the single sensor, the Trade-space Analysis Tool for Constellations (TAT-C) simulator is used to explore the ground track of the sensor orbit under different orbital configurations. The module for simulating the sensor orbits in TAT-C has been employed to investigate the nadir position track of a variety of different satellite sensors [127]. The second step in simulating the viewing extent of a single sensor is to adjust the sensor swath width to enhance the hypothetical sensor coverage for this study. Specifically, the satellite viewing extent is generated by extending the ground track in the cross-track direction to a given swath width of interest. The viewing extent simulation is ultimately expressed as a binary map marking the global surface as viewed (or not) in the absence of clouds. The viewing extent simulation is conducted at a 0.01° spatial resolution in this study, and subsequently aggregated in space to match relatively coarse-scale geophysical retrievals in the following analysis. Finally, a realistic cloud mask is convolved with the viewing extent to explore the effects of cloud attenuation on terrestrial snow observability.

2.2.2 Orbital Configuration and Sensor Type

In short, the orbital configuration mainly depends on satellite altitude and inclination angle. Orbital configuration and swath width determine the repeat cycle of the sensor viewing extent. To represent typical (i.e., polar-orbiting, sun-synchronous) configurations, six different sensors are selected here to represent a range of hypothetical instruments including PMW radiometers, SARs, and LiDARs.

2.2.2.1 Passive Microwave Radiometer

The first evaluated sensor is an AMSR2-like PMW radiometer [128]. It has a wide swath and relatively high revisit frequency, but typically has a coarse spatial resolution ($\sim$10 km). PMW radiometers have served as stalwarts for snow mass estimation over the last 30+ years [9] and have demonstrated considerable skill in estimating relatively dry, shallow snow mass in relatively flat terrain and in the absence of dense vegetation [31].

2.2.2.2 Synthetic Aparture RADAR

The second hypothetical sensor is a Ku/Ka dual-band SAR similar to the Terrestrial Snow Mass Mission (TSMM) that is currently under consideration by the Canadian Space Agency (CSA) [108]. A Ku/Ka dual-band SAR is expected to have a better response to snow mass than other existing spaceborne SARs. The third evaluated sensor is a Sentinel1-like C-band SAR [129]. It represents an existing SAR instrument that is currently used for snow mass detection even though the scattering characteristics of C-band radiation in dry or shallow snow can be of limited value [35]. However, given that Sentinel-1 is operational now and into the future, it is considered here as a viable snow information source that should be included in this study. Two C-band SAR instruments are included here to mimic the Sentinel-1 A/B constellation. The two C-band SARs share the same orbital plane, but with a 180° phase difference.

2.2.2.3 LiDAR Altimetry

The fourth hypothetical sensor is a wide-swath (imaging) LiDAR with a 20 km swath width. This specific configuration (and the assumed instrument errors) for use in space may

Table 2.1: Orbital configurations of the tested sensors.

ID	Sensor Type	Orbit Altitude [km]	Inclination Angle [°]	Swath Width [km]	Prototype (status)
1	PMW Radiometer	510	97	1450	AMSR-2 (Existing)
2	Ku-band SAR	705	98	500	TSMM (Hypothetical)
3	C-band SAR	705	98	250	Sentinel-1 A/B (Existing)
4	Wide LiDAR	481	92	20	ICESat-2 (Hypothetical)
5	Narrow LiDAR	481	92	6	ICESat-2 (Existing)
6	Low-inclination LiDAR	415	51.6	6.5	GEDI (Existing)

PMW = passive microwave; SAR = synthetic aperture RADAR;
LiDAR = light detection and ranging;
AMSR-2 = Advanced Microwave Scanning Radiometer 2 (AMSR2);
TSMM = Terrestrial Snow Mass Mission;
ICESat-2 = Ice, Cloud and land Elevation Satellite-2;
GEDI = Global Ecosystem Dynamics Investigation

not be achievable given the engineering requirements of today, but this aspirational sensor is considered here as a feasible part of a future, hypothetical snow constellation configuration, and as such, is explored in this study. The fifth evaluated sensor is an ICESat-2-like narrow-swath LiDAR [130]. It has a similar orbit as ICESat-2, but with better spatial coverage assuming the use of a hypothetical 6 km continuous swath width to replace the original ICESat-2 observations that are sampled by six laser beams each with a 10-meter footprint across a 6 km field of view. The sixth evaluated sensor is a GEDI-like (Global Ecosystem Dynamics Investigation onboard the International Space Station) low-inclination angle LiDAR [61]. As a low-inclination angle platform, it only views regions within $\pm 51.6°$ latitude. While losing the ability to monitor snow over high latitudes, it yields a higher revisit frequency in low latitude areas of snow. The orbital configurations for each of these sensors are provided in Table 2.1.

Among the existing spaceborne altimetry and LiDAR instruments, ICESat-2 has a 6-km

total swath width [131] and GEDI has a total swath width of 6.5 km [132]. Therefore, the assigned swath width of the hypothetical LiDAR is set to 6 km in order to appropriately represent the current, state-of-the-art technology. The swath width of the hypothetical wide-swath LiDAR is assigned as 20 km. Even though such a spaceborne LiDAR does not currently exist, it is worth conducting this experiment to consider the added value associated with a hypothetical increase in swath width relative to that which is currently operational.

2.2.2.4 Orbital Parameters

Sensor IDs 1-3 (Table 2.1) have orbital configurations that are polar and sun-synchronous, which means the local overpass time at the equator will be similar from one day to the next. A consistent overpass time is critical for microwave-based remote sensing, be it active or passive in nature, because the presence (or absence) of liquid water can drastically change the snow dielectric constant that could in turn affect the observed signals, and therefore introduce additional errors or uncertainties in the snow retrievals. To further reduce the diurnal variations of the observations in sensor IDs 1-3, only one overpass direction representing the nighttime overpass is used here in order to minimize wet snow effects. By contrast, both ascending and descending observations are used for sensor IDs 4, 5, and 6 since the snow depth measured by LiDAR is less impacted (related to microwave sensors) by snow wetness or snow temperature.

2.2.2.5 Constellations

In addition, the performance of four hypothetical constellations (i.e., mixtures of different sensors) is also considered. Four specific constellations were selected from a near-infinite number

Table 2.2: Sensor makeup of hypothetical snow constellation configurations.

Constellation ID	Sensor Marks	Sensor Mixture
C1	✠★★\	PMW & two C-band SARs & narrow LiDAR
C2	✠✯\	PMW & Ku-band SAR & narrow LiDAR
C3	✠✯▮	PMW & Ku-band SAR & wide LiDAR
C4	✠★★✯\▮	PMW & two C-band SARs & Ku-band SAR & narrow LiDAR & wide LiDAR

✠ = PMW sensor; ★ = C-band SAR; ✯ = Ku-band SAR;
\ = narrow LiDAR; ▮ = wide LiDAR.

of possible configurations as the focus of the paper to make the analysis tractable. The combinations explored here include (Table 2.2):

(a) PMW radiometer, two C-band SARs, Narrow LiDAR;

(b) PMW radiometer, Ku-band SAR, narrow-swath LiDAR;

(c) PMW radiometer, Ku-band SAR, wide-swath LiDAR;

and (d) PMW radiometer, Ku-band SAR, two C-band SARs, wide-swath LiDAR, narrow-swath LiDAR.

The different constellations represent: (a) currently available techniques for snow remote sensing [C1], e.g., the two C-band SARs represents the Sentinel-1 A/B constellation; (b) proxies of sensors feasibly applied in the near future [C2]; (c) incorporate the additive value of a wide swath LiDAR [C3]; and (d) represent what could be achieved if all sensors are simultaneously spaceborne [C4]. The selection of these different sensors within each constellation does not consider a likely cost cap to sensor deployment. Rather, the selection of these different sensors aims to explore what could be viewed assuming the financial resources were available to deploy such a configuration.

2.2.3 Time-varying Snow Mask

To help investigate the space-time coverage of terrestrial snow, the Interactive Multisensor Snow and Ice Mapping System (IMS) snow cover [133] is used for the years 2001 - 2020 to serve as a reasonable proxy for binary (yes or no) snow coverage, and a determinant for a time-varying snow mask. Since the vast majority of the terrestrial snow cover is located in the northern hemisphere – about 40 million km^2 [134] compared to less than 1 million km^2 in the southern hemisphere [135] – only terrestrial snow over the northern hemisphere is explored here in order to minimize computational expense.

The IMS snow mask is leveraged here to empirically describe the snow coverage extent. The daily snow-covered probability through statistics of snow occurrence is computed within each 0.04-degree grid from historical data to estimate the probability of snow cover. A pixel with a probability larger than 0.5 is marked as snow. Otherwise, it is considered as snow-free. Then the snow mask is interpolated to 0.01-degree grid using the nearest-neighbor algorithm. The binary snow map is computed as,

$$\begin{cases} \text{snow-covered pixel} & \text{if } \frac{\sum_{i=2001}^{2020} IMS(\vec{x},doy,i)}{N} \geq 0.5 \\ \text{snow free pixel} & \text{if } \frac{\sum_{i=2001}^{2020} IMS(\vec{x},doy,i)}{N} < 0.5 \end{cases} \tag{2.1}$$

where $\vec{x}$ refers to the pixel location in space, doy refers to the day of year, and $N = 20$ represents the total number of years used during the analysis.

This snow cover extent approximates the climatological space-time occurrence of snow as prior knowledge for simulation. It is eventually convolved with the sensor viewing area

(see 2.2.1) to estimate the viewed snow cover extent, which is a necessary precursor to study remotely-sensed snow mass. This chapter focuses on the viewed snow extent as a means of further exploring the snow mass in the next chapter.

2.2.4 Time-varying Cloud Mask

In order to consider the impacts of clouds on the snow retrievals observations obtained via LiDAR, a daily cloud mask is employed to simulate the cloud cover distribution. The cloud mask is extracted from the quality flag (i.e., Coarse Resolution Internal Cloud Mask) of the 0.05-degree MODIS Aqua daily reflectance collection 6 product (MYD09CMG, `https://ladsweb.modaps.eosdis.nasa.gov/filespec/MODIS/6/MYD09CMG`) [136]. The 0.05-degree cloud mask is then interpolated to 0.01-degree grid using the nearest-neighbor algorithm. The cloud mask is not available during the nighttime-only polar winter given the passive (optical) nature of the MODIS sensor. Therefore, a gap filling strategy is adopted to simulate the cloud distribution when the data is missing. A set of gap-free cloud masks collected during the polar summer (i.e., from March 22 to September 20) is employed. When cloud retrievals are missing during the polar winter (i.e., from September 22 to March 21) the gap-free cloud masks from the summer are used as a reasonable surrogate to estimate the impact of cloud attenuation on optical sensor retrievals collected from space. Even though this method does not exactly reproduce the cloud conditions that existed, the filled cloud map serves as a reasonable proxy to represent the true cloud variability across space and time.

2.2.5 Evaluation Metrics

This section defines and describes three different metrics employed to evaluate the sensors' viewability of terrestrial snow: 1) viewed snow coverage percentage, 2) viewed snow classification coverage percentage, and 3) viewing repeat interval for each terrestrial snow class. These metrics help to quantitatively assess the sensor/constellation efficacy in observing snow across space and time while also considering differences in regional snow climatology.

2.2.5.1 Viewed Snow Coverage Percent

The first metric is the snow coverage percent within a certain interval of time, i.e. 1-day, 3-day, and 30-day periods. The three different periods represent time to respond to the daily, synoptic scale, and seasonal variations, respectively [137]. To investigate the viewing effectiveness, as a single sensor or part of a sensor constellation, the normalized snow coverage percentage that is viewed (P_{snow}) over the northern hemisphere terrestrial environment is calculated as:

$$P_{snow} = \frac{\bigcup_{i=1}^{n}[A(\vec{x},t)_{viewed,i} \cap A_{snow}(\vec{x},t)]}{\int A_{snow}(\vec{x},t)d\vec{x}} \times 100\% \tag{2.2}$$

where $A_{snow}(\vec{x},t)$ refers to the dynamic snow-covered terrestrial area during the study period (defined by IMS), $A(\vec{x},t)_{viewed,i}$ is the terrestrial area that is viewed by the sensor i, and $\int A_{snow}(\vec{x},t)d\vec{x}$ is the total snow-covered area for a given day in space ($\vec{x}$) and time (t). The symbol $\bigcup_{i=1}^{n}$ denotes the union of areas viewed by satellites from 1 to n that compose a given constellation. The symbol $\cap$ represents the intersection.

2.2.5.2 Viewing of Snow Classification Coverage Percentage

In addition to the total viewed snow coverage, the second metric explores the efficacy of each sensor configuration to view a specific snow classification as:

$$P_j = \frac{\bigcup_{i=1}^{n}(A(\vec{x},t)_{viewed,i} \times W_{i,j}) \cap A_{snow}(\vec{x},t) \cap A_j(\vec{x},t)}{\int A_j(\vec{x},t)d\vec{x}} \times 100\% \tag{2.3}$$

where $A_j(\vec{x},t)$ refers the area of snow class j, and P_j is the percentage of the viewed snow-covered area of the j-th class. $W_{i,j}$ represents the weight of the efficacy of a given sensor i on the snow class j. The snow classification system proposed by [138] is employed here (see Figure 2.1). Snow is categorized into six different classes based on the physical properties: 1) tundra, 2) taiga, 3) alpine, 4) maritime, 5) ephemeral, and 6) prairie classes. Ice is not discussed in this dissertation. Table 2.3 shows the assigned weight matrix $W_{i,j}$ used in this study. It provides a first-order estimate of the sensor efficacy on specific snow classification according to the assumptions as follows: 1) PMW sensors do not work well for snow under dense forest (taiga), deep snow (maritime), and snow over complex terrain (alpine) [9, 105]; 2) SAR sensors do not work effectively for snow under dense forest (taiga); and 3) LiDAR sensors are affected by cloud attenuation [123]. The situation in the real world is complex, hence, the weight applied here is somewhat subjective over large areas for each class. However, these values are useful in specifying a reasonable estimation of each sensor's efficacy and allows for a relatively transparent understanding of the assumptions made.

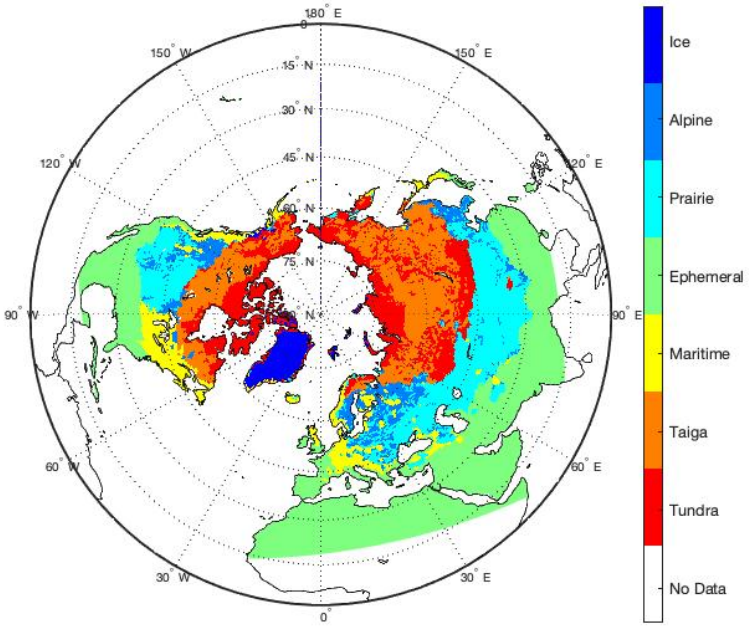

Figure 2.1: Map of snow cover classification based on [138].

Table 2.3: Assumed weight of sensor efficacy ,$W_{i,j}$, for snow mass estimation in each snow class, j. Individual weights are subjective, but serve as an effective skill estimate for each sensor relative to one another.

	Tundra	Taiga	Maritime	Ephemeral	Prairie	Alpine
Radiometer	1	0	0	1	1	0
SAR	1	0	1	1	1	1
LiDAR (cloud-free)	1	1	1	1	1	1
LiDAR (cloud-covered)	0	0	0	0	0	0

0 = contains no skill; 1 = contains skill;

2.2.5.3 Temporal Repeat Interval

The third metric employed here is the temporal repeat interval for each terrestrial snow class. Compared to the snow coverage percentage, the temporal repeat interval better reflects the orbital overlap, which is a strong function of latitude when using a polar-orbiting sensor. The repeat interval, $I(\vec{x})$, is calculated as:

$$I(\vec{x}) = \frac{T}{F(\vec{x})} \tag{2.4}$$

where $\vec{x}$ is space; T refers to a certain period (365 days in this study) in units of days; $F(\vec{x})$ is the number of repeat times during period T considering the sensor efficacy weight shown in Table 2.3; and $I(\vec{x})$ refers to the repeat interval in units of days since last viewed.

2.3 Results

2.3.1 Sensor Simulation of Viewing Extent

Figure 2.2 illustrates the viewing extent simulation steps described in section 2.2.1. Figure 2.2(a) shows the nadir points of a single sensor, e.g., Ku-band SAR in Table 2.1, as simulated by TAT-C tool for a 1-day period of integration. These nadir points are then extended to the swath width coverage as shown in Figure 2.2(b).

Similarly, Figure 2.3 shows the sensor ground track from TAT-C and 1-day set of results of viewing extent for each of the sensors introduced in Section 2.2.1.

The viewed area depends on the sensor's swath width as well as the orbital configuration.

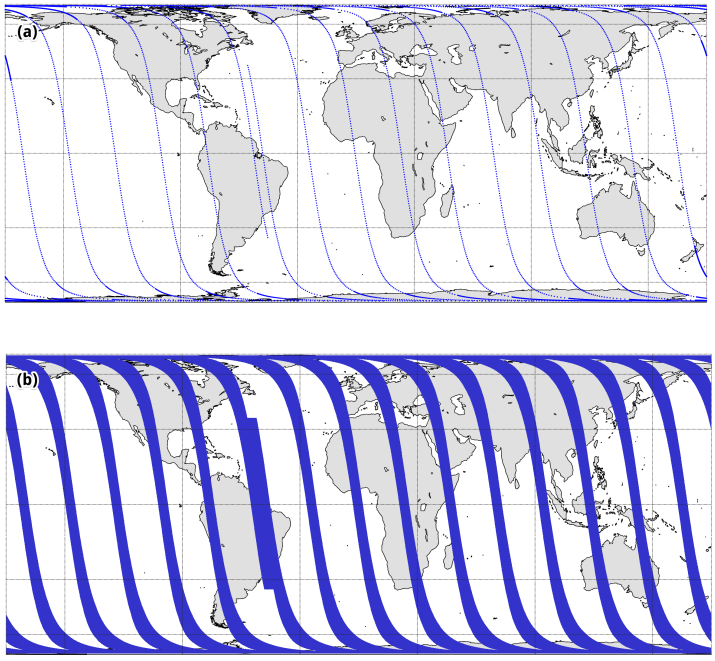

Figure 2.2: Satellite viewing extent simulation of the hypothetical Ku-band SAR in Table 2.1 for (a) nadir points for the ascending pass during a 1-day integration period; and (b) viewing extent for the ascending pass during a 1-day integration period

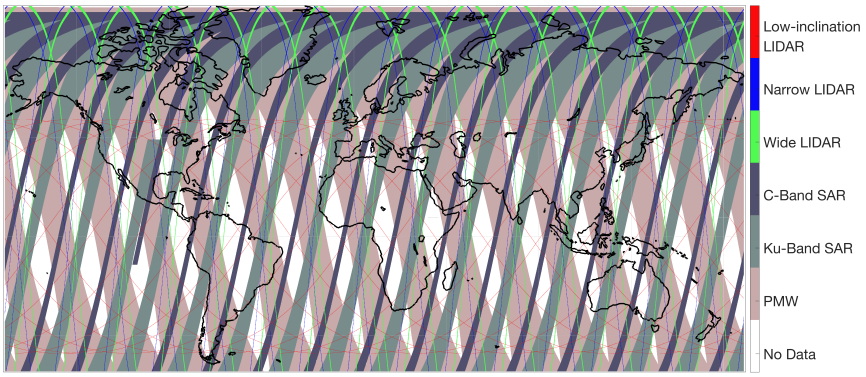

Figure 2.3: Example of daily viewing extent of sensors listed in Table 2.1.

The narrow swath sensors have larger gaps in coverage across space and time, in general, and hence, longer revisit intervals relative to the wide swath sensors.

2.3.2 Time-varying Snow Mask Estimation

Figure 2.4 shows a viewing example for a C-band SAR along with the coincident snow-covered area (based on the IMS snow product) for a single day near peak snow accumulation. The overlap between the blue and the green represents the snow-covered terrain as viewed by the sensor. Since both the snow-covered terrain and viewing area are a function of space and time, the variation of the overlay distribution is complex. This process is repeated over multiple snow seasons for each individual sensor type (section 2.3.3) as well as for a mixture of different sensors (section 2.3.4). The goal of this exercise is to determine the spatiotemporal viewing capability of each sensor on its own as well as in coordination with other sensors. In addition, the use of the weights of each sensor can help discern how best to coordinate these hypothetical sensors in a follow-on study with direct applicability to global snow mass estimation.

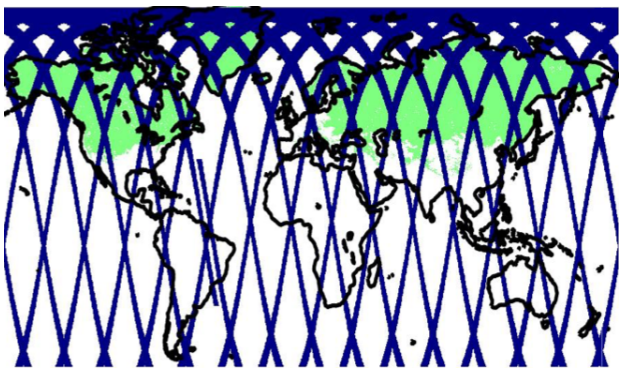

Figure 2.4: Viewing example for a single day using a C-band SAR (No.3 in Table 2.1) overlying the snow-covered area via IMS; blue is sensor coverage; green is snow-covered terrian according to Section 2.2.3.

2.3.3 Evaluation of Single Sensor

2.3.3.1 Viewed snow coverage percentage analysis

To quantitatively assess the seasonal variation of the viewed snow area, the total snow area and viewed snow area by each individual sensor over different periods is illustrated in Figure 2.5. The total snow cover area in the northern hemisphere varies as a function of season and reaches a peak of about $5.4 \times 10^7\ km^2$ during February with a minimum value of about $0.2 \times 10^7\ km^2$ during August. As a result, the spatial coverage for snow varies correspondingly, and yields the largest difference between different sensors during peak snow accumulation. For example, a single PMW sensor could observe more than 80% of the snow-covered terrain in a single day in February while a single SAR sensor could observe between 20% and 40% depending on the swath width. A single LiDAR sensor views less than 5% of the snow-covered terrain area in a single day.

When the integration time increases to three days, a PMW radiometer can view all of the terrestrial snow across the northern hemisphere. However, a single LiDAR sensor coverage is still limited to less than 15%. For a 30-day integration period, the wide-swath LiDAR could cover more than 50% of terrestrial snow in the northern hemisphere while the narrow-swath LiDAR views approximately 38%. In short, a single PMW sensor views most of the northern hemisphere snow-covered terrain in a 1-day period; a single SAR sensor views most of the snow-covered area in a 3-day period; and a 20 km swath width LiDAR sensor cannot view most of the snow-covered terrain even in a 30-day period.

The differences among the sensors' coverage mainly result from swath width configurations.

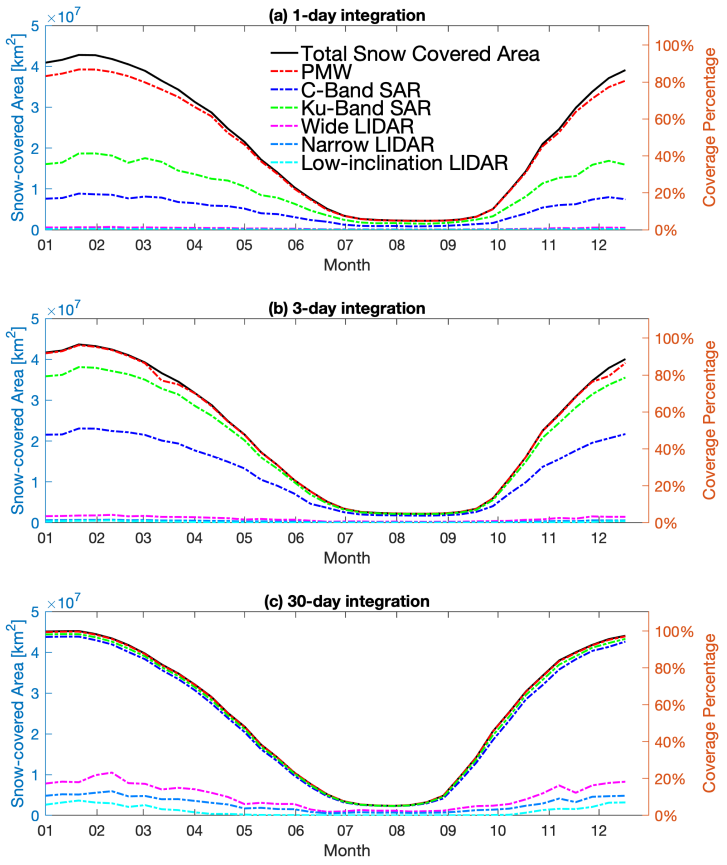

Figure 2.5: Seasonal variation in viewed terrestrial snow area for different types of existing and hypothetical sensors. Subplot (a) shows coverage based on a 1-day integration period; (b) shows a 3-day integration period; (c) shows a 30-day integration period.

Figure 2.6 highlights the viewed snow percentage as a function of swath width. This result is computed as a function of swath widths ranging from 50 km to 1500 km, using an inclination angle of 97° and a 510 km altitude. The increase in viewed snow cover percentage during a 1-day integration period is nearly linear as the swath width increased. The growth rate of percent coverage is asymptotic when the swath width is large for a 3-day or 30-day integration period due to the successive overlaps between different days. When arbitrarily drawing a line of 80% of snow-covered percentage, the swath width required is about 1100 km, 550 km, and 200 km for a 1-day, 3-day, and 30-day integration period, respectively. This result provides a useful benchmark of snow mission demands when considering daily, synoptic, and seasonal variations of snow. This analysis assumes a peak snow period. For other days of the year, the extent of low-latitude snow decreases.

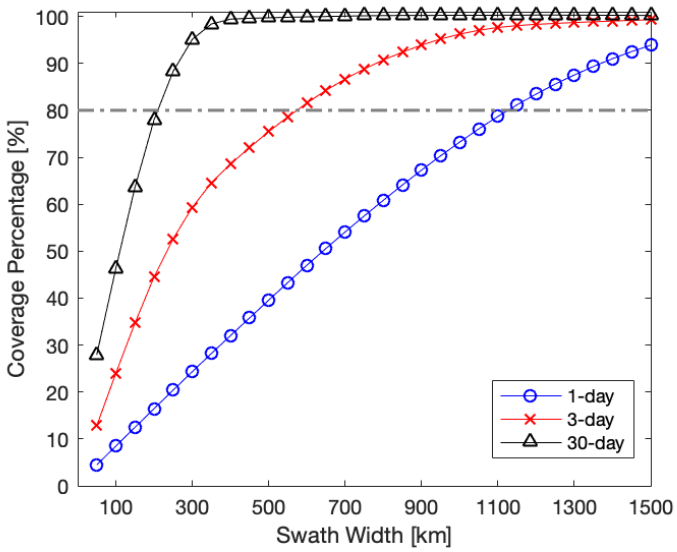

Figure 2.6: Viewed snow coverage percentage as a function of orbit swath width for 1-day, 3-day, and 30-day integration periods during the month of February near peak accumulation of snow in the northern hemisphere. The dot-dashed line represents 80% viewed snow coverage.

The effective coverage as presented in section 2.2.5.2 is employed to reflect the effect of

sensor efficacy on different snow classes along with consideration of cloud attenuation. The effective coverage not only indicates the snow extent viewed by the sensor(s), but also shows the skill of each sensor to detect the snow in different snow classes.

The simulated results for effective snow coverage percentage are shown in Table 2.4. It shows poor percent coverage for all sensors over taiga since only LiDAR sensors, whose swath widths are limited, are assumed to work well for this class. The PMW sensor provides considerable viewing for the tundra, ephemeral, and prairie classes within one day. The SAR sensors cover most areas of these classes when the integration period is increased to three days, which helps mitigate the PMW sensor's limitations in areas such as maritime and alpine snow. The LiDAR sensors represent the smallest viewing percentage. Even over a 30-day period, the viewed percentage can not meet a near-global requirement, per se, but does view relatively large amounts of most snow classes. These results suggest a wider swath width is likely required for LiDAR sensors or that more than one LiDAR will be required in the assessment of northern hemisphere snow.

The assumed weight of sensor efficacy listed in Table 2.3 is to be improved. For example, defining the PMW sensor's efficacy as zero over taiga regions is arguable [9,29,139]. Alternatively, Pulliainen et al. [107] showed there could be a significant correlation between PMW-observed SPD and SWE in a typical boreal forest region. In order to account for a range of feasible efficacies, values ranging from 0.1 to 0.5, with an increment of 0.1, are applied. The results shown in Table 2.5 highlight how the effective coverage of PMW radiometry within taiga snow regions increases with increasing efficacy. During a 1-day integration period, the percentage increases linearly from 9.8% to 49%. During a 3-day or 30-day integration period, the effective coverage is similar to the 1-day integration period because of the relatively large swath width

Table 2.4: Effective coverage of sensors (in units of percent) for different snow classes with integration periods of 1-day, 3-days, and 30-days. Values greater than 80% are in boldface.

		Snow Class					
	Sensor ID	Tundra	Taiga	Maritime	Ephemeral	Prairie	Alpine
1-day	PMW	**98.8**	0.00	0.00	67.1	77.0	0.00
	Ku-band SAR	55.2	0.00	34.5	22.1	25.2	33.5
	C-band SAR	29.3	0.00	17.5	10.2	12.5	17.9
	Wide LiDAR	1.88	1.83	0.795	0.802	0.768	1.61
	Narrow LiDAR	0.686	0.374	0.316	0.344	0.346	0.447
	Low-inclination LiDAR	0.0901	0.105	0.211	0.522	0.158	0.163
3-day	PMW	**100**	0.00	0.00	**93.7**	**95.8**	0.00
	Ku-band SAR	**93.8**	0.00	**80.5**	58.5	63.7	75.8
	C-band SAR	68.6	0.00	47.4	30.6	35.7	46.1
	Wide LiDAR	5.20	3.71	2.08	2.44	2.07	2.73
	Narrow LiDAR	1.92	1.14	0.832	1.04	0.911	0.913
	Low-inclination LiDAR	0.269	0.303	0.597	1.17	1.23	0.519
30-day	PMW	**100**	0.00	0.00	**98.2**	**97.8**	0.00
	Ku-band SAR	**97.7**	0.00	**96.0**	**93.3**	**90.3**	**94.3**
	C-band SAR	**94.9**	0.00	**91.5**	**87.7**	**82.2**	**88.9**
	Wide LiDAR	18.3	13.8	8.84	11.4	9.22	9.70
	Narrow LiDAR	10.6	7.80	5.18	7.00	5.32	5.18
	Low-inclination LiDAR	2.00	2.12	3.60	8.20	7.30	3.01

Table 2.5: Percent effective viewing coverage in taiga snow regions using a PMW radiometers in conjunction with efficacies ranging from 0.1 to 0.5. The different rows represent different integration periods.

Integration period	Efficacy				
	0.1	0.2	0.3	0.4	0.5
1-day	9.8	20	29	39	49
3-day	9.9	20	30	40	50
30-day	10	20	30	40	50

of the PMW radiometer. However, the efficacy of PMW varies in time and space considerably due to the variation of snow properties, such as snow depth, snow wetness, snow density, and snow grain size. This paper only provides a first-order estimate of the approximated efficacy; a dynamic efficacy (in space and time) is likely required in a follow-on study in order to improve model performance.

2.3.3.2 Repeat interval analysis

The results of the repeat interval analysis are presented in Table 2.6. The repeat interval represents an averaged viewing of sensors across space and time as computed from a simulation of an entire year and averaged across the northern hemisphere by each individual snow class. The repeat interval reflects how frequently an observation could be obtained for each sensor. In addition to the taiga snow class, the ephemeral snow class is relatively difficult to view given the lower-latitude position of these snow classes when viewed using polar-orbiting sensor configuration. Even with a 500 km wide swath, the Ku-band SAR does not fully view the ephemeral snow class at a synoptic scale (i.e., approximately 3-day period). The narrow-swath LiDAR takes a long time to revisit a given location, especially for locations at low latitudes. Compared to a polar-orbiting LiDAR, the low inclination LiDAR views the snow classes at low

Table 2.6: Domain-average repeat intervals (in units of days) for different snow sensors as a function of snow class and sensor efficacy (see Table 2.3).

	Snow Class					
Sensor ID	Tundra	Taiga	Maritime	Ephemeral	Prairie	Alpine
PMW	1.03	-	-	1.54	1.31	-
Ku-band SAR	2.12	-	3.40	4.61	3.92	3.07
C-band SAR	4.15	-	6.76	9.11	7.77	6.08
Wide LiDAR	64.1	61.3	179	75.1	93.1	85.0
Narrow LiDAR	172	159	450	172	221	210
Low-inclination LiDAR	390	400	332	136	138	292

latitudes (e.g., ephemeral snow) more frequently, but less so for snow classes at high latitudes (e.g., tundra snow). The wide-swath LiDAR requires over 50 days to revisit the same location depending on the latitude and cloud conditions. This suggests a LiDAR with a swath width larger than 20 km or a constellation with several LiDARs would be required in order to achieve a monthly (or less) repeat interval across the northern hemisphere. All variations of LiDAR explored here have a longer interval to revisit maritime snow compared to other snow classes because cloud attenuation is more prevalent over maritime snow.

2.3.4 Evaluation of Constellations

The results from the individual sensor experiments illustrate how no single sensor can adequately measure all types of snow at all locations across the hemisphere. That is, it is clear that a constellation of different sensors is required to achieve this goal. Table 2.7, therefore, illustrates the effective coverage of the tested constellation cases.

The shortcomings of any single sensor are compensated for by the other sensors in the constellation. For example, the PMW sensor obtains a regular, short duration repeat of observations of the tundra, prairie, and ephemeral snow classes, while the SAR sensors can collect information

Table 2.7: Effective viewing coverage (in units of percent) of constellations for different snow classes with 1-day, 3-day, and 30-day integration periods. Values greater than 80% are in boldface.

		Snow Class					
	Constellation ID	Tundra	Taiga	Maritime	Ephemeral	Prairie	Alpine
1-day	(a)	**97.8**	0.710	34.5	77.1	**90.9**	25.6
	(b)	**98.8**	0.710	29.0	72.7	**91.4**	32.1
	(c)	**98.7**	1.83	29.6	72.9	**91.4**	33.8
	(d)	**98.7**	2.54	47.6	78.3	**93.4**	42.0
3-day	(a)	**100**	2.00	68.9	**98.0**	**99.0**	67.4
	(b)	**100**	2.00	74.2	**98.0**	**99.2**	**81.4**
	(c)	**100**	5.19	74.7	**98.1**	**99.2**	**81.8**
	(d)	**100**	7.14	**87.3**	**98.4**	**99.3**	**87.0**
30-day	(a)	**100**	19.7	**95.9**	**98.9**	**99.0**	**98.4**
	(b)	**100**	19.7	**96.8**	**98.8**	**99.1**	**99.0**
	(c)	**100**	44.5	**96.9**	**98.9**	**99.1**	**99.0**
	(d)	**100**	49.2	**97.8**	**99.2**	**99.2**	**99.2**

regarding maritime and alpine snow. Further, the LiDAR sensors help provide important information about snow with overlying vegetation. Although the viewable area is limited, the LiDAR information could potentially help cross-calibrate other sensor retrievals in other areas. The comparison between constellation (C1) and (C2) shows the impact on viewing coverage by replacing two C-band SARs with one Ku-band SAR. The constellation (C1) achieves more coverage within a 1-day integration period, but less coverage within 3-day and 30-day integration periods compared to the constellation (C2). The difference in viewing coverage between the constellations are not significant, but the higher quality and relatively fine-resolution observations from the Ku-band SAR could potentially improve the snow retrieval quality of the tundra, prairie, and ephemeral snow classes. With a wide-swath LiDAR, constellation (C3) viewed a larger portion of snow in the taiga regions relative to constellation (C2). In constellation (C4), all candidate sensors are introduced, and hence, the viewing coverage is maximized.

2.4 Discussions and Conclusions

What observations are needed to study terrestrial snow across the northern hemisphere? Mission planners want to maximize the scientific value given a fixed budget. However, the exact approach to maximize the benefits to snow science remains an open question. This research is designed to assist researchers and mission planners weighing the different trade-offs based on sensor selection and orbital configurations. In addition, this study explores a range of snow researchers' requirements (in terms of temporal integration periods) for remote sensing snow observations.

This study also provides key information relevant to a future observing system simulation experiments (OSSE) to be conducted in a follow-on study. An OSSE serves to mimic nature and help quantitatively explore the impact of different hypothetical observing systems (such as the snow sensors explored here) on conditional (a.k.a. updated) snow model results. Furthermore, this study provides a slew of sensor coverage simulations with various sensor swath widths and orbital configurations. The results from this study will be applied in a data assimilation experiment in a similar manner as [140], which was only applied to LiDAR remote sensing of snow. For the researchers who are interested in similar topics, the simulated sensor viewing coverage of the hypothetical sensors in this study are published on DRUM at the University of Maryland (`https://drum.lib.umd.edu/handle/1903/27610`) and available for public access.

This study has several aspects that should be refined in the future regarding the simulation method and the metrics for the use of coverage evaluation. First, snow that is viewed does not necessarily mean that it can be retrieved accurately. There is a complex, nonlinear calculation

going from the sensor observations to snow retrievals. The increase of viewing extent only represents the upper-bound on the quantity of observations, not what the retrievals will actually see. Meanwhile, the increase in the number of observations does not guarantee an improvement in snow retrieval quality. To avoid introducing more noise due to information related to something other than snow mass (e.g., snow grain size, snow wetness), efficacy weights are used to reflect the sensor skill as a function of different snow classes. As mentioned in the Results section, the efficacy weights can be improved to reflect the actual performance of sensors under different conditions. Another limitation is that snow is assumed to be uniform within a given snow class. The weight matrix in Table 2.3 is used to describe sensor efficacy as a first-order estimate. However, sensor efficacy depends on the specific snow conditions (e.g., dry versus wet snow, shallow versus deep snow). Besides, the skill of LiDAR to retrieve snow depth is adversely impacted by clouds in a complex, nonlinear manner. When LiDAR's efficacy over cloudy regions is set to zero in Table 2.3, it is a first-order estimate without direct consideration of the cloud optical thickness. That is, even though cloud cover may be present, optically thin clouds (e.g., cirrus clouds) may still allow for snow retrieved using LiDAR.

Additionally, the geophysical retrievals from each sensor are assumed to be mutually unbiased. As a result, constellation case (a) yields an overly optimistic view of snow-covered terrain. This particular constellation scenario represents a constellation composed of existing sensors. However, there are relatively few studies investigating the conjoined use of these sensors because the observations from each sensor are inherently uncertain and posses their own unique error characteristics. It is difficult to first merge these disparate information sources prior to evaluating their additive value relative to one another. Another factor is then needed to carefully consider each sensor's spatial resolution. Before jointly using the different sensor observations, they must

be first integrated into a unified product via data assimilation or some other merging strategy.

This study explores a suite of existing and hypothetical sensors in the viewing of snow-covered terrain as a function of sensor orbital configuration and sensor efficacy. The research explores the viewing of a series of sensors, and constellations of sensors, with distribution of coincident terrestrial snow. The results help quantify the demands on sensor swath width and orbital configuration for the requirement of 1-day, 3-day, and 30-day integration periods.

Viewing extent simulations show that sensors with swath widths of 1100 km, 550 km, and 200 km could meet the demands of 1-day, 3-day, and 30-day repeat intervals, respectively. However, the effective coverage when considering sensor efficacy and cloud attenuation suggests a single sensor cannot observe all snow classes at all locations in the northern hemisphere during all times of the year. Constellations composed of different sensors could better compensate for the shortcomings of a single sensor in a specific snow class. The combination of the PMW and SAR sensors could perform well for all the snow cover classes except taiga, while the LiDAR sensors could potentially provide some key information of snow with overlying vegetation.

To quantify the expected model improvements as a function of orbital configuration and sensor type the results of this chapter are applied in a follow-on snow OSSE study. The results of the snow OSSE combined with a cost analysis could help mission planners decide how to get the most snow-related scientific bang for the scientific buck.

Chapter 3: Impact of multi-variate, multi-sensor assimilation on terrestrial freshwater estimation within a coupled snow-vegetation-soil moisture observing system simulation experiment

3.1 Introduction

Terrestrial freshwater storage (TWS) is the vertically-integrated sum of snow, soil moisture, vegetation canopy storage, surface water impoundments, and groundwater [141]. Among these components, snow, soil moisture, and vegetation are the most temporally-dynamic as well as the most spatially-variable [142, 143]. Although snow, soil moisture, and vegetation only account for a relatively small part of the TWS [12], the snow-soil moisture-vegetation system plays a key role in the hydrological cycle, and serves as a first-order control on surface runoff, transpiration, and evapotranspiration as well as the prediction of weather and climate [13–15]. Previous studies have quantified the snow-soil moisture-vegetation system using land surface models (LSMs) [67, 84, 144], remote sensing retrievals [16–18], or fusing model and remote sensing observations within a data assimilation framework [80, 86, 145–147]. However, the characterization of these highly-dynamic hydrologic state variables in space and time is complex, and inherently uncertain because it is impossible to measure all components of TWS at all times and all locations in space. An additionally perplexing problem is that one component of the snow-soil moisture-vegetation

system can, at times, shroud the measurement of another component such that quantifying one of these components individually requires knowledge of one or more of the other components within the same system [7–11]. Therefore, a true representation of terrestrial freshwater (and its integrated response as it travels through a given watershed) is best studied as a cohesive whole rather than as the mere sum of its individual components.

Observations collected by spaceborne passive optical (visible/near-infrared) radiometers, passive microwave (MW) radiometers, Radio Detection and Ranging (RADAR) sensors, or Light Detection and Ranging (LiDAR) sensors are not snow or soil moisture or vegetation observations, per se, but rather information related to snow or soil moisture or vegetation [87–89]. These disparate data streams contain their own unique characteristics (e.g., error structure, field of view, overpass time) that are often integrated with a land surface model in order to provide estimates of geophysical variables that scientists care most about (e.g., snow mass, soil moisture content, vegetation biomass, runoff) [9, 57, 58, 148]. One major challenge in maximizing the utility of spaceborne sensors collecting measurements relevant to snow, soil moisture, and vegetation retrievals from space-based instrumentation is the lack of a robust framework to explore the interplay between the different sensor types, swath widths, geophysical retrievals, and error characteristics [148,149]. To derive the most utility from a multitude of observations (i.e., passive optical, passive MW, active MW, LiDAR or any combination thereof), one must explore the marginal contribution of each measurement type towards achieving a scientific benchmark while also exploring the synergistic interplay between different measurement types and how they add to, or detract from, the goal of achieving that desired scientific benchmark.

While prior work has demonstrated the utility of each of these measurement types [16–18], there is much debate in the science community as to what the optimal, concurrent sensor

configuration choice should be for a next-generation terrestrial hydrology mission [150–152]. The measurements from each of these instrument types have their own strengths and weaknesses, and they differ in their spatial and temporal coverage, scales, and extents. During a typical mission development process, an assessment of the utility of the measurements is often limited to the immediate physical variables of interest. The impacts on downstream applications is often assumed and not always quantified [153–155]. For example, snow conditions during peak winter significantly impacts the snowmelt processes during the spring time, which are the primary drivers of stream flow in rivers. Similarly, antecedent soil moisture conditions dictate the partitioning of precipitation at the land surface, which in turn affects downstream river discharge. Approaching the debate from the vantage point of a cohesive whole accounts for these nuances whereas considering only one component at a time cannot capture explicitly these coupled nuances.

In order to solve the problems highlighted above, an observing system simulation experiment (OSSE) is developed. In this OSSE, synthetic observations from a series of real-world and hypothetical sensors are simulated and subsequently assimilated into the NASA Land Information System (LIS, version 7.2.0). The OSSE evaluated a suite of alternative observing strategies of snow-soil moisture-vegetation and quantifies the scientific utility, both in terms of a single sensor as well as a coordinated constellation of two or more sensors.

This chapter is structured as follows. In Section 2, the study area is briefly introduced. In Section 3, the methodology is introduced to build the OSSE, setup the experiments, and generate the synthetic retrievals using a variety of different orbital configurations. The evaluation metrics are also introduced in Section 3. Results are presented in Section 4. Sections 5 and 6 provide detailed discussion of the results and a summarized conclusion of the findings, respectively.

3.2 Study Area

The study domain is Western Colorado (37.5°N $\sim$ 40.7°N, 108.6°W $\sim$ 106°W, Figure 3.1) located in the United States. The landscape varies from flat grass-covered plains to forest-covered mountains and semi-arid plateaus [156, 157]. Coniferous forest is the dominant vegetation type across this mountainous region. Annual meteorological characteristics of this area suggest a relatively wet season during the summer with winter storms at high elevations occurring from November to early-May [157]. The complex terrain coupled with a variety of different land cover types make this region an interesting study area with regards to the coupled snow-soil moisture-vegetation system and the remote sensing thereof.

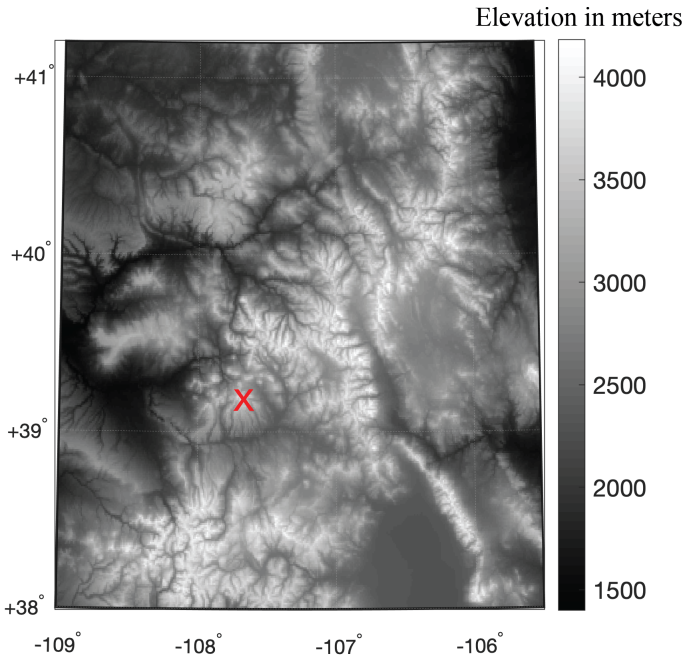

Figure 3.1: Elevation map of study domain. The red "×" marks the location of the example point used in Figure 3.6.

3.3 Methodology

3.3.1 OSSE Framework

The observing system simulation experiment (OSSE) presented here is designed to analyze added values associated with assimilation of remote sensing retrievals into an advanced land surface model. This OSSE includes six distinct parts (Figure 3.2): a) Nature Run (NR), b) Trade-space Analysis Tool for Constellations (TAT-C) space-time subsampler, c) realistic observation error estimate, d) Open Loop (OL) experiment, e) Data assimilation (DA) routines, and f) a method for systematic evaluation of the OL and DA relative to the NR. The state variables and fluxes produced from the NR in Figure 3.2a) are viewed as a proxy for the real-world (a.k.a. synthetic "truth"). Synthetic retrievals are then generated from the NR by applying the TAT-C space-time subsampler (Figure 3.2b) with corresponding error approximation (Figure 3.2c) to the corresponding NR-derived state variables.

The OL experiment illustrated in Figure 3.2d) provides a benchmark for evaluating the added value via assimilation of the synthetic retrievals. The results of the OL and DA experiments are evaluated relative the NR in Figure 3.2e). Each of the framework's six parts is discussed in detail in the following sections.

3.3.2 Experimental Setup

The NASA Land Information System (LIS) is employed during the production of the NR, OL, and DA experiments. LIS is a software framework developed at the NASA Goddard Space Flight Center that integrates a suite of LSMs, satellite observations, ground-based measurements,

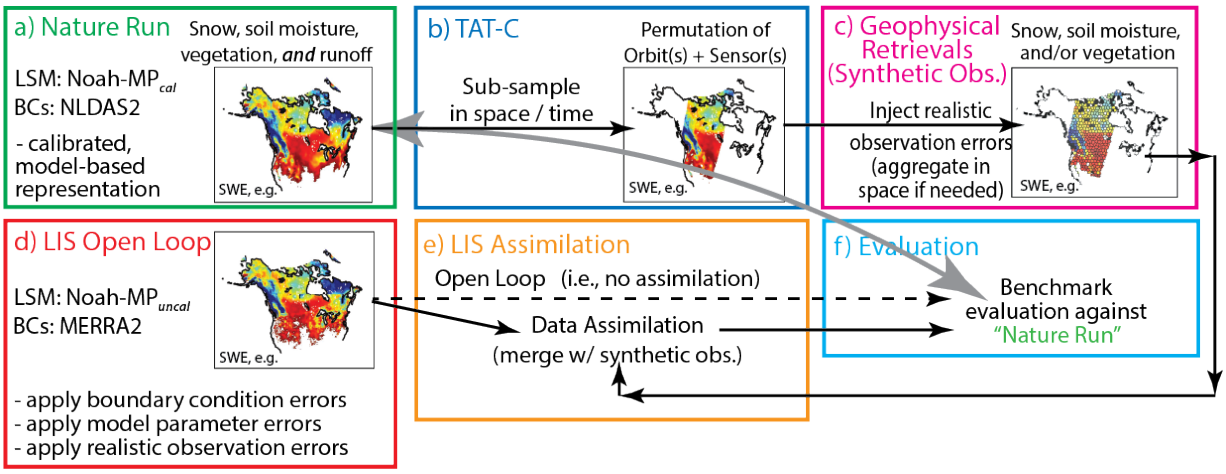

Figure 3.2: Conceptual framework of OSSE. (BCs = boundary conditions; LIS=NASA Land Information System; LSM = land surface model; MERRA2 = Modern-Era Retrospective Analysis for research and applications (v2); NLDAS2 = North American Land Data Assimilation System (v2); Noah-MP = Noah multi-physics parameterization; SWE=snow water equivalent; TAT-C = Tradespace Analysis Tool for Constellations.)

and DA routines in order to obtain a conditioned (posterior) estimate of the hydrologic cycle states and fluxes [86, 158]. The LIS framework is designed with flexibility and high-performance computing capabilities, including advanced software techniques that enable computational tractability in integrating and assimilating observations across regional and global scales [158, 159].

The Noah-MP 4.0.1 land surface model [144] within LIS was employed to generate relevant hydrologic state variables and fluxes such as snow water equivalent (SWE), snow depth, leaf area index (LAI), soil moisture, and runoff. Noah-MP 4.0.1 is a land surface model that includes a one-layer vegetation canopy, a three-layer snowpack, and a four-layer soil moisture profile, thereby allowing for multiple options in terms of key land-atmosphere interaction processes [144]. It is an appropriate model for this study because of its ability to couple snow-soil moisture-vegetation processes in a physically-based manner [160, 161, 161, 162].

A portion of the model parameters in NR were calibrated against the University of Arizona

SWE product [163] using the methodology of [164] in order to reduce the potential bias between the model estimates and the real-world. The NR is able to reflect the true system variability, i.e. hydrologic phenomena during both extreme and non-extreme events, in the real-world [164].

The NR is forced with hourly radiation and precipitation meteorology provided by the North American Land Data Assimilation System (NLDAS2), which is an advanced reanalysis product derived from daily gauge-based precipitation analysis, bias-corrected shortwave radiation, and surface meteorology reanalysis data [165]. Initial conditions were generated via model spinup starting with a uniform soil moisture and groundwater conditions and running.

Compared to the NR, the OL simulation in NoahMP-4.0.1 uses a different set of parameters rather than the calibrated parameters as used in the NR. The differences between the calibrated and uncalibrated versions of the model serves to, in part, represent a reasonable proxy for model structure error. Further, the OL is forced by hourly meteorological data from the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA2) [166], in conjunction with hourly precipitation from the Tropical Rainfall Measuring Mission (TRMM) [167]. To reduce the system bias between the forcing of OL and NR, the TRMM precipitation is corrected using a ratio of monthly domain-averaged NLDAS2 precipitation to monthly domain-averaged TRMM precipitation. The differences between NLDAS2 used as boundary conditions in the NR and MERRA2 + TRMM used as boundary conditions in the OL reflect a reasonable approximation for the boundary condition error encountered by most operational assimilation systems.

Another difference between the OL and the NR is that the OL uses a probabilistic (ensemble) approach whereas the NR is deterministic. In this study, the OL uses an ensemble of 20 members to implicitly represent the system error in the modeled estimate of hydrologic variables [141]. The forcing of precipitation and downward shortwave radiation is perturbed with multiplicative

perturbations with a mean of 1 and standard deviations of 0.5 and 0.5, respectively. Additive perturbations with a standard deviation of 50 Wm^{-2} are applied to the downward longwave radiation fields.

The initial conditions are also different between the NR and OL experiments. The NR initial condition spun up from 1979 to 2009 with single replicate. Model results from 2009 was used to reinitialize the model again from 1979, and then further running the simulation forward in the beginning in September 2009. The OL initial conditions spun up from 1979 to 2008 with single replicate only once. Then model results from September 2008 then continued to run with 20 replicates to allow the ensemble inflation prior to running the OL.

The experimental configurations of the DA experiments are as same as the OL, except that synthetic retrievals are assimilated whereas in the OL no assimilation takes place. The different DA experiments include assimilation of synthetic retrievals for single variables as well as combinations of different variables.

The NR, OL, and DA experiment are then conducted from September 2009 to September 2011 at a 0.01° spatial resolution grid. These two years are typical water year with an average precipitation status.

3.3.3 EnKF Algorithm

3.3.3.1 Univariate Assimilation

A one-dimensional ensemble Kalman filter (EnKF) is used as the assimilation algorithm [168]. The error covariances used in the assimilation updates is estimated from the distributions of the ensemble members [168]. Small perturbations are applied to the assimilated synthetic

observations and state variables at each grid point as part of the stochastic filter in a manner consistent with Brugers et al. [81]. The EnKF algorithm is defined as,

$$y_j^+ = y_j^- + C_{yz}[C_{zz} + C_{vv}]^{-1}[z_j - M(y_j^-)] \tag{3.1}$$

Where, y_j^+ represents the posterior variable of the ensemble j, y_j^- represents the prior variable of the ensemble j. M represents the observation model, which projects the state variable y to the observation z. C represents covariance operator. Among them, C_{yz} represents the cross covariance between the state variable and the observation error. C_{zz} represents the error covariance of the model estimated observation $M(y_j^-)$. C_{vv} represents the covariance of the observation model M. z_j represents the observation of the ensemble j. z_j is perturbed by injecting a random error (see Table 3.2) to the original observation value [81].

In this study, the observation model M is linear, and therefore could be expressed as a linear observation operator H. The EnKF update equation is then simplified as,

$$y_j^+ = y_j^- + C_{yy}^- H^T[H^T C_{yy}^- H^T + C_{vv}]^{-1}[z_j - H y_j^-] \tag{3.2}$$

H represents the observation operator. T represent the transpose operator. C_{yy}^- represents the state error covariance, which could be estimated from the spread of the ensembles of the prior state variables.

3.3.3.2 Multivariate Assimilation

In the multi-variate assimilation, the state variables are updated sequentially, which is the feasible method to update multiple variables in the LIS 7.2.0 version. For example, a triple assimilation is described as,

$$y_{j,1}^{+} = y_{j}^{-} + C_{yy}^{-} H_1^T [H_1^T C_{yy}^{-} H_1^T + C_{vv,1}]^{-1} [z_{j,1} - H_1 y_j^{-}] \tag{3.3}$$

$$y_{j,2}^{+} = y_{j,1}^{+} + C_{yy}^{-} H_2^T [H_2^T C_{yy}^{-} H_2^T + C_{vv,2}]^{-1} [z_{j,2} - H_2 y_{j,1}^{+}] \tag{3.4}$$

$$y_{j,3}^{+} = y_{j,2}^{+} + C_{yy}^{-} H_3^T [H_3^T C_{yy}^{-} H_3^T + C_{vv,3}]^{-1} [z_{j,3} - H_3 y_{j,2}^{+}] \tag{3.5}$$

The $y_{j,1}^{+}$, $y_{j,2}^{+}$, and $y_{j,3}^{+}$ represent the posteriors after the first, second, and third updates. Similarly, H_1, H_2, and H_3 are the observation operators for the first, second, and third observations (z_1, z_2, and z_3). C_{yy}^{-} is calculated from ensemble spread before each updates. For consistency, the assimilation order of the retrievals is decided according to the corresponding spatial resolution and swath width, i.e., the retrievals with a coarser resolution and wider swath width are updated prior to the retrievals with a finer resolution and narrower swath width.

3.3.4 Synthetic Retrievals Generation

3.3.4.1 Work Flow

The synthetic retrievals used in the OSSE are produced in an effort to mimic real-world retrievals as highlighted in Table 3.1. SWE, soil moisture, and LAI retrievals from eight sensors with different orbital configurations are generated in this study. The sensor types include L-band

PMW radiometry (herein referred to as PMW soil moisture), X-/K-/Ka-band PMW radiometry (herein referred to as PMW SWE), C-band SAR (herein referred to as SAR SWE or SAR LAI), and optical LiDAR (herein referred to as LiDAR snow depth or LiDAR LAI).

Figure 3.3 illustrates the work flow to generate a synthetic retrieval. First, state variables and fluxes from the NR are employed as the synthetic truth. Next, a space-time subsampler generated by TAT-C is used as an approximation of what each sensor can "see" at any point of time on Earth given a particular orbital configuration. For the sensor aligned in a sun-synchronous orbit (i.e. PMW and SAR as applied in this experiment), only observations from a single track (i.e. ascending or descending only) are used. For example, only nighttime overpasses are used for PMW snow and SAR snow in order to minimize wet snow effect [29, 169]. Similarly, only a single overpass is used with L-band soil moisture retrievals because soil moisture retrievals from different overpasses (i.e. ascending and descending) show a different error structure [170]. Observations from both ascending and descending tracks of LiDAR are used because LiDAR observations are less sensitive to overpass time than the microwave instruments (i.e., PMW radiometers and SARs). The subsampled image is then upscaled (if needed) into the corresponding resolution of the desired geophysical retrieval. Further, a frozen soil mask (based on NR modeled soil temperature) and cloud attenuation mask (based on optical imagery from MODIS) are applied to the soil moisture or LiDAR retrievals (see Table 3.1) in order to yield a more realistic geophysical retrieval. For the sensors whose proxy has a finer resolution than 1 km, the resolution of their retrievals is assumed to be upscale to 1 km as a means of reducing noise and error within those retrievals. Once the NR has been subsampled and mapped into geophysical space, a realistic amount of error is injected into the retrieval as a proxy of the real-world uncertainty. The processed data then undergo a series of quality control checks. The resulting product is a synthetic

Table 3.1: List of synthetic retrievals explored in this study (blue = snow; orange = soil moisture; green = vegetation).

Geophysical retrieval	symbol	Sensor information				
		Sensor proxy	Sensor type	Orbital type	Sensor swath width [km]	Spatial resolution [km]
SWE or snow depth	○	AMSR2	PMW	PS	1450	25
	×	Sentinel1	SAR	PS	250	1
	+	ICESat2	LiDAR	PS	6	1
	◇	ICESat2	LiDAR	PS	20	1
Soil moisture	○	SMAP	PMW	PS	1000	36
Vegetation LAI	×	Sentinel1	SAR	PS	250	1
	+	GEDI	LiDAR	LI	6.5	1
	◇	GEDI	LiDAR	LI	20	1

PMW = passive microwave; SAR = synthetic aperture RADAR; LiDAR = light detection and ranging.
○ = PMW; × = SAR; + = narrow swath LiDAR; ◇ = wide swath LiDAR.
PS = polar and Sun-synchronous orbit; LI = low inclination orbit;
AMSR2 = Advanced Microwave Scanning Radiometer 2;
GEDI = Global Ecosystem Dynamics Investigation;
ICESat2 = Ice, Cloud and land Elevation Satellite-2;
Sentinel1 = Sentinel-1 A/B by European Space Agency; SMAP = Soil Moisture Active Passive

retrieval (Figure 3.3).

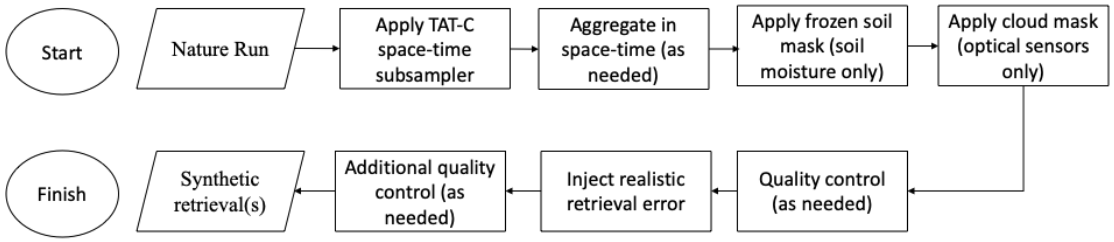

Figure 3.3: Flow chart for synthetic retrieval generation.

3.3.4.2 Space-time Subsampler

TAT-C is employed to simulate the viewing extent of a given sensor [127]. The module for simulating the sensor orbits in TAT-C is used to compute the nadir track of a variety of different satellite sensors (Figure 3.4). Next, the space-time subsampler is generated by extending the ground track in the cross-track direction to yield the desired swath width of interest. The

subsampler is ultimately expressed as a binary map marking the global surface as viewed (or not) in the absence of clouds. The subsampler applied here is produced at a 0.01° spatial resolution grid co-located with the NR. The space-time subsamplers are ultimately upscaled to the sensor resolution if the sensor resolution is coarser than 0.01° (e.g., soil moisture retrievals at approximately 36 km resolution).

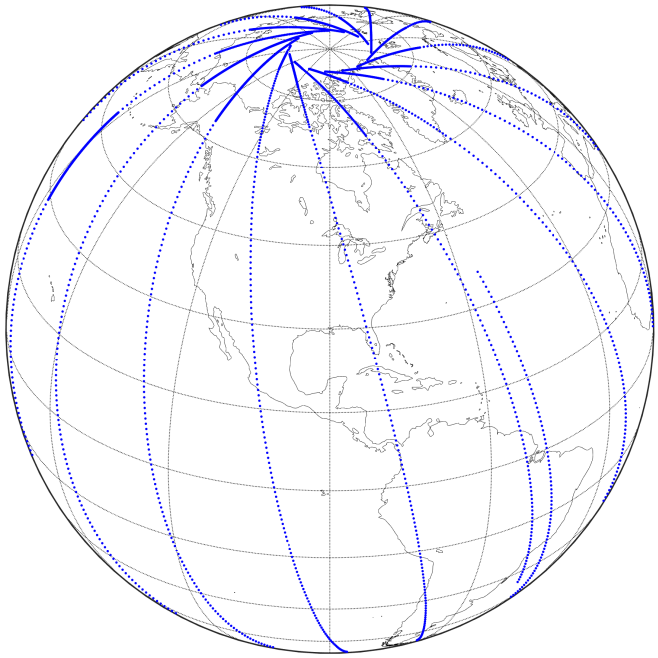

Figure 3.4: Nadir track of ascending overpasses during a 1-day integration period for a satellite platform with an altitude of 693-km and an inclination angle of 98°.

3.3.4.3 Frozen Soil Mask and Cloud Mask

The frozen soil mask is only applied to the L-band synthetic soil moisture retrievals because most L-band PMW radiative transfer models are not applicable to partially-frozen or completely-frozen soils [171]. Frozen soil is identified (via the land surface model) as when the soil temperature

is less than 0°C or when the soil is completely covered by snow (according to the NR simulation). When upscaling the fine-scale Nature Run into the coarse-scale L-band retrieval space, as needed, a coarse pixel with more than 50% frozen soil is demarcated as frozen.

A daily cloud mask is used to approximate cloud attenuation on the LiDAR sensors listed in Table 3.1 because of the large extinction coefficient in the presence of clouds [123]. MODIS cloud mask information [136, 172] is employed in this study. The 0.05-degree cloud mask is then resampled to 0.01-degree grid using the nearest-neighbor algorithm.

3.3.4.4 Retrieval Error

Retrieval error represents the uncertainty in remote sensing retrievals. The random error is assumed to be Gaussian in space. Two types of error models – additive and multiplicative – are adopted for different retrievals, i.e., multiplicative error model is applied for synthetic PMW SWE whereas an additive model is used for SAR SWE, LiDAR snow depth, SAR/LiDAR LAI, and PMW soil moisture (Table 3.2).

The additive error model (mean zero) maintains unbiasedness in the synthetic retrievals, and therefore, does not violate the unbiased assumption of EnKF algorithm. The multiplicative error model (mean one) also maintains unbiasedness in the synthetic retrievals. The multiplicative error model is useful for representing PMW signal saturation errors in deep snow [9] whereas the additive error model is more useful for random errors such as LiDAR-based pointing errors [173].

The additive error model is defined as:

$$\hat{Y}(\vec{x},t) = Y(\vec{x},t) + \varepsilon,\ where\ \varepsilon \sim N(0,\sigma^2) \tag{3.6}$$

Table 3.2: Assumed synthetic retrieval error models and error distribution parameters.

Geophysical retrieval	Symbol	Error model	Error mean	Error standard deviation		
				Small	Medium	Large
SWE or snow depth	○	Multiplicative	1	0.1	0.5	0.9
	×	Additive	0 [m]	0.1 [m]	0.5 [m]	0.9 [m]
	+	Additive	0 [m]	0.1 [m]	0.5 [m]	0.9 [m]
	◇	Additive	0 [m]	0.1 [m]	0.5 [m]	0.9 [m]
Soil moisture	○	Additive	0 $[cm^3cm^{-3}]$	0.02 $[cm^3cm^{-3}]$	-	0.04 $[cm^3cm^{-3}]$
LAI	×	Additive	0 $[m^2m^{-2}]$	0.1 $[m^2m^{-2}]$	-	0.5 $[m^2m^{-2}]$
	+	Additive	0 $[m^2m^{-2}]$	0.1 $[m^2m^{-2}]$	-	0.5 $[m^2m^{-2}]$
	◇	Additive	0 $[m^2m^{-2}]$	0.1 $[m^2m^{-2}]$	-	0.5 $[m^2m^{-2}]$

where $\hat{Y}(\vec{x},t)$ represents a realistic retrieval where $Y(\vec{x},t)$ represents the corresponding subsampled NR in space and time. ε is the error assuming a Gaussian distribution with a zero mean and a variance, σ^2, $\varepsilon \sim N(0,\sigma^2)$. The multiplicative error model is defined in Equation 3.7 as:

$$\hat{Y}(\vec{x},t) = Y(\vec{x},t)\cdot\epsilon,\ \ where\ \epsilon \sim LN(\mu,\varsigma^2) \tag{3.7}$$

where, ε follows a log normal distribution, with parameters of μ, and ς^2. They can be calculated from log normal distribution's mean $m = 1$ and variance v^2 as:

$$\begin{aligned} \mu &= log(m^2/\sqrt{v^2+m^2}) \\ \varsigma^2 &= log(v^2/m^2+1) \end{aligned} \tag{3.8}$$

The exact structure of real-world errors is unknown. Therefore, in an effort to encapsulate the true errors, a series of experiments is conducted using different amounts of σ or v (see Table 3.2), i.e. 0.1, 0.5, 0.9 for SWE and snow depth; 0.1, 0.5 for LAI; and 0.02, 0.04 for soil moisture. Values of assumed observation error are taken from literature where available [31, 50, 174, 175]. Hypothesized errors for non-existent sensors are assumed across a reasonable range in an effort

Table 3.3: Matrix of multi-sensor, multi-variate assimilation experiments (blue = snow; orange = soil moisture; green = vegetation).

Retrieval marks	Description	Motivation
× ○	SAR SWE + PMW soil moisture	Examine synergistic effect on snow and soil moisture assimilation
× ◇	SAR SWE + Wide swath LiDAR LAI	Examine synergistic effect on snow and vegetation assimilation
◇ ○	Wide swath LiDAR LAI + PMW soil moisture	Examine synergistic effect on vegetation and soil moisture assimilation
× ◇ ○	SAR SWE + Wide swath LiDAR LAI + PMW SM	Examine synergistic effect on the snow-soil] moisture-vegetation system

PMW = passive microwave radiometer; SAR = synthetic aperture RADAR; LiDAR = light detection and ranging.
SWE = snow water equivalent; LAI = leaf area index. ○ = PMW; × = SAR; + = narrow Lidar; and ◇ = wide swath LiDAR.

to bound the "true" errors, and thereby help define a necessary threshold error limit for future retrieval developments. The assimilation results are then evaluated to quantify the value of these synthetic retrievals as a function of σ or v.

3.3.5 Evaluation Metrics

In this study, the NR state variables serve as the reference for evaluation of estimated SWE, soil moisture, vegetation, and runoff from the OL and DA experiments. Normalized information contribution (NIC) of root-mean-square error (RMSE) is calculated as the evaluation metric in this study. NIC_{RMSE} is defined in Equation 3.9 as:

$$
\begin{aligned}
NIC_{RMSE} &= \frac{RMSE_{OL} - RMSE_{DA}}{RMSE_{OL}} \\
RMSE_{OL} &= \sqrt{\frac{\Sigma_{i=1}^{n}(y_{OL}(i) - y_{NR}(i))^2}{n}} \\
RMSE_{DA} &= \sqrt{\frac{\Sigma_{i=1}^{n}(y_{DA}(i) - y_{NR}(i))^2}{n}}
\end{aligned}
\tag{3.9}
$$

where, NIC_{RMSE} represents NIC of RMSE. $RMSE_{OL}$, and $RMSE_{DA}$ refer to RMSE of the Open Loop, and data assimilation, respectively. y_{NR}, y_{OL}, and y_{NR} represent the state variable or fluxes from NR, OL, and DA respectively. n refers to the sample size.

Considering that the DA results need to be assessed across multiple states (e.g. SWE, soil moisture, LAI), an integrated NIC is developed as a cohesive evaluation metric. This integrated NIC is a weighted mean of the NIC of individual variables defined in the Equation 3.10) as,

$$NIC_{RMSE,weighted} = \frac{\sum_i^n \alpha(i) \cdot NIC_{RMSE,var(i)}}{\sum_i^n \alpha(i)} \tag{3.10}$$

where, $NIC_{RMSE,weighted}$ represents the weighted NIC; $NIC_{RMSE,var(i)}$ refers to the NIC of each variable; $\alpha(i)$ is the weights for each variable; and n is the number of variables. For this specific study, the NIC of SWE, soil moisture, and LAI is used as the components to be weighted. The weights of each component (i.e., $NIC_{RMSE,SWE}$,$NIC_{RMSE,soil\ moisture}$, $NIC_{RMSE,LAI}$) are set to 1 equally.

3.4 Results

3.4.1 NR, OL, and Synthetic Retrievals

The NR, OL, and synthetic retrievals are presented to demonstrate how they are similar to one another, yet differ in clear and distinct fashions. For example, the NR and OL (Figure 3.5) both capture seasonal variability, but to different magnitudes given differences in boundary conditions and model parameters. Analogously, the NR and synthetic retrievals (Figure 3.3 and Figure 3.6) are initially identical, but the synthetic retrievals progressively deviate from the NR due to subsampling (i.e., space and time) along with the introduction of retrieval errors. In the NR experiment, there is more snow accumulated, and therefore more runoff responding to the snow melting. The NR experiment also exhibits wetter soil and more vigorous vegetation growth

at the start of the spring season. Considerable runoff difference exists between the NR and OL, especially during the snow ablation season.

A domain-averaged comparison between the NR and synthetic retrievals is not appropriate because often the two represent very different regions of space. Therefore, a single location is selected (38.725°N, 107.245°W, Figure 3.1) in the domain to show an example of how synthetic retrievals fluctuate about the NR. Figure 3.6 shows the time series of NR, OL, and the synthetic retrievals from this selected point. Three types of synthetic retrievals, i.e., SAR SWE, PMW soil moisture, and wide swath LiDAR, are shown in Figure 3.6(a), 3.6(b), and 3.6(c) of Figure 3.6 as examples in order to illustrate the features of synthetic retrievals obtained via SAR, PMW, and LiDAR. The results show that sensors with wider swath widths have more frequent observations, which is intuitive. The PMW sensor observes the point nearly daily; the C-band SAR sensor observes approximately every three days. The LiDAR sensor's revisit period is irregular since they are strongly affected by cloud attenuation. There is a gap in the L-band PMW sensor observations of soil moisture during the winter because the L-band PMW sensor retrieval algorithm does not function during frozen soil conditions.

3.4.2 Impact of Univariate Assimilation

The DA performance for each single-sensor, univariate experiment is examined relative to the OL performance using the NR as reference. This examination illustrates how the assimilation of a single variable on the same geophysical variable, along with other relative states and fluxes, in the land surface model. For each type of geophysical retrievals, the DA results of these synthetic retrievals from the most interesting sensor type (i.e., SAR for SWE, PMW for soil moisture, and

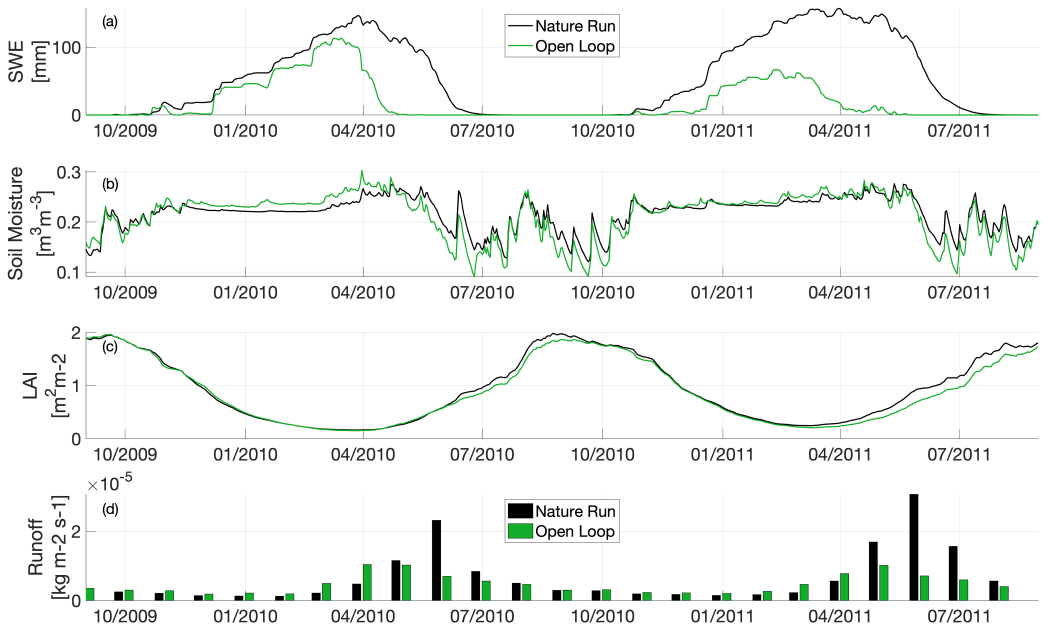

Figure 3.5: Time-series of domain-averaged SWE, soil moisture, and LAI, along with monthly runoff from **the Nature Run(NR) and the Open Loop (OL)**. Black lines/columns represent the NR where the green lines/columns represent the OL.

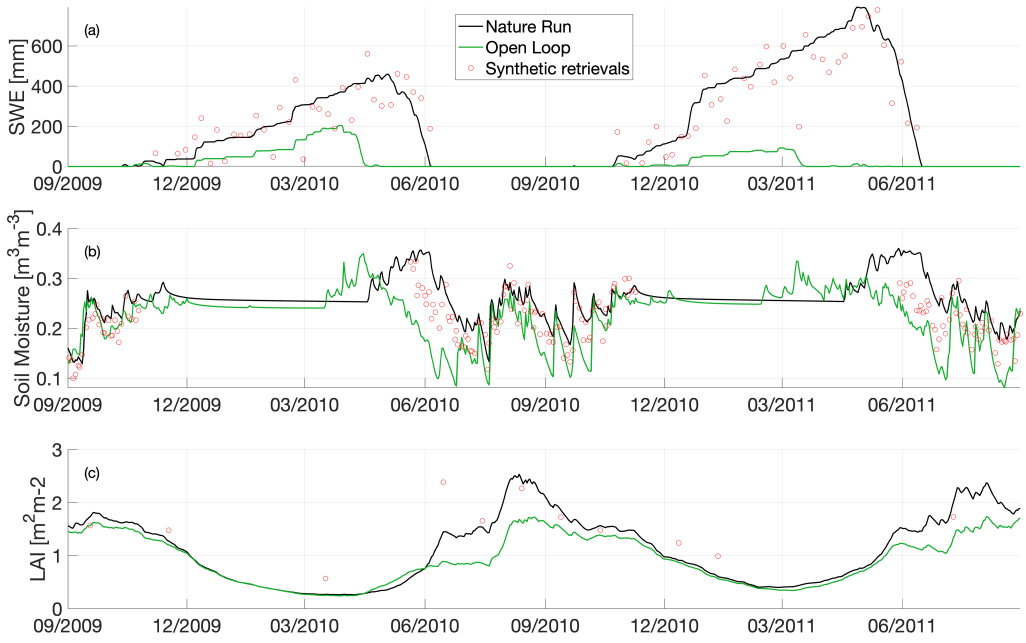

Figure 3.6: Time-series of SWE, soil moisture, and LAI from the **NR, OL, and synthetic retrievals** at a selected point. Black lines represent the NR, green lines/columns represent the OL, and red circles represent synthetic retrievals. Subplots include synthetic retrievals from: (a) C-band SAR SWE (retrieval error: σ=0.1 [m]), (b) soil moisture via PMW radiometry (retrieval error: σ=0.02 [cm^3cm^{-3}]), and (c) LAI via optical LiDAR (retrieval error: σ=0.1 [m^2m^{-2}]).

wide-swath LiDAR for LAI) are illustrated. Only synthetic retrievals with relatively low retrieval errors (i.e., σ=0.1 [m] for SAR SWE, σ=0.02 [cm^3cm^{-3}] for PMW soil moisture, and σ=0.1 [m^2m^{-2}]) for wide-swath LiDAR LAI) are used to better illustrate the impact of assimilation.

Figure 3.7 shows time-series of domain-averaged daily SWE, soil moisture, LAI along with monthly runoff of the NR, OL, and DA using **SAR SWE** during assimilation. The SAR SWE has a finer spatial resolution (relative to the PMW radiometry), thus, provide more information content for assimilation. Figure 3.7 shows that SWE estimation is improved considerably via the DA compared to SWE in the OL, especially in the second year, when the NR and OL SWE yield large differences. In the DA, snow mass is added during snow accumulation season such that SWE is closer to the NR SWE. Snow melt timing also matches better with that in the NR. Compared to SWE, soil moisture does not significantly change. Soil moisture receives only minor improvements between May and July. Wetter soil moisture is associated with increased snowmelt during the ablation season. However, LAI estimation gets worse via the DA compared to that in the OL. Slower vegetation growth in the DA experiment is likely due to the prolonged snow-cover. Monthly runoff is overestimated in the DA in contrast to the underestimation in the OL.

Similarly, Figure 3.8 shows the time-series of daily-averaged, domain-averaged SWE, soil moisture and LAI along with monthly-averaged runoff of the NR, OL, and DA using the synthetic **PMW soil moisture**. Assimilation helps the model better represent the soil moisture variations except during frozen soil conditions. Soil moisture via the DA, in general, increases to better match the NR. Compared to soil moisture, SWE is rarely changed by soil moisture assimilation. The wetter soil moisture leads to denser vegetation, but this trend sometimes makes LAI estimation deviated from the "Truth" (LAI in the NR). Wetter soil moisture also leads to

Figure 3.7: Domain-averaged time-series of SWE, soil moisture, and LAI along with monthly runoff from the **NR, OL, and DA using SAR SWE**. Black lines/columns represent the NR, green lines/columns represent for OL, and red lines/columns represent the DA.

more runoff yield, which causes the overestimation of runoff except in the snow ablation season. In the snow ablation season, the runoff gets less biased due to assimilation but still does not match with the runoff timing in NR.

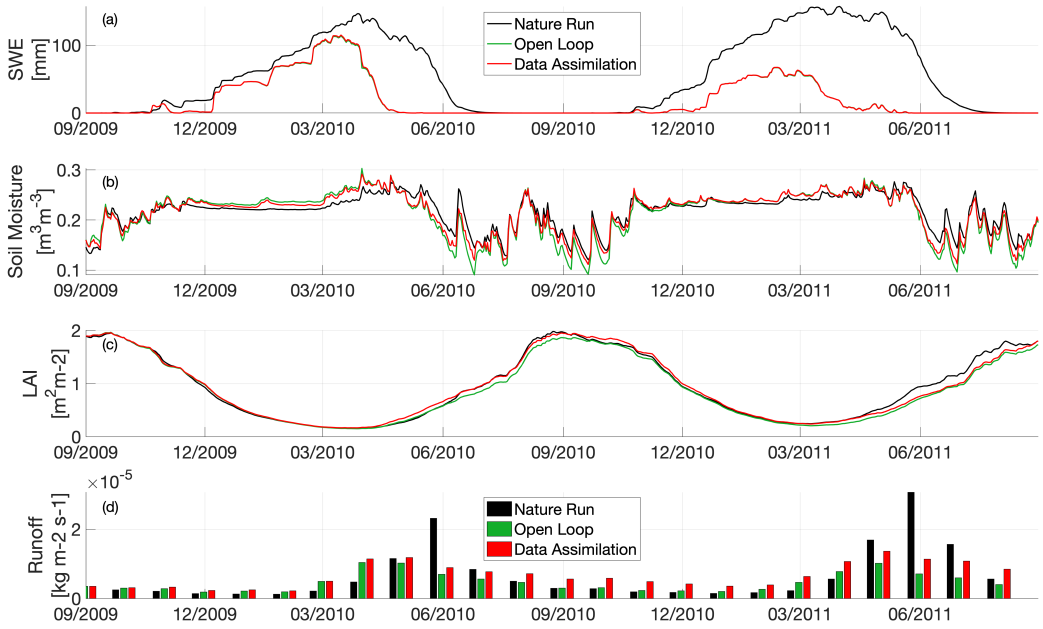

Figure 3.8: Domain-averaged time-series of SWE, soil moisture, and LAI along with monthly runoff from the **NR, OL, and DA using PMW soil moisture**. Black lines/columns represent the NR, green lines/columns represent the OL, and red lines/columns represent DA.

Figure 3.9 shows the time-series of daily-averaged, domain-averaged SWE, soil moisture and LAI along with monthly-averaged runoff of the NR, OL, and DA via **wide-swath (20 km, hypothetical) LiDAR LAI**. The wide-swath LiDAR LAI is observed from a low-inclination orbit, which would provide more information about vegetation that is **not** at high latitudes. The wide-swath LiDAR observes more of the terrestrial environment than the narrow-swath LiDAR (6.5 km). In the Figure 3.9, LAI via the DA is closer to the NR, especially from June to October. The LAI DA impact on SWE, and soil moisture is negligible. The monthly runoff is slightly

reduced via the LAI DA.

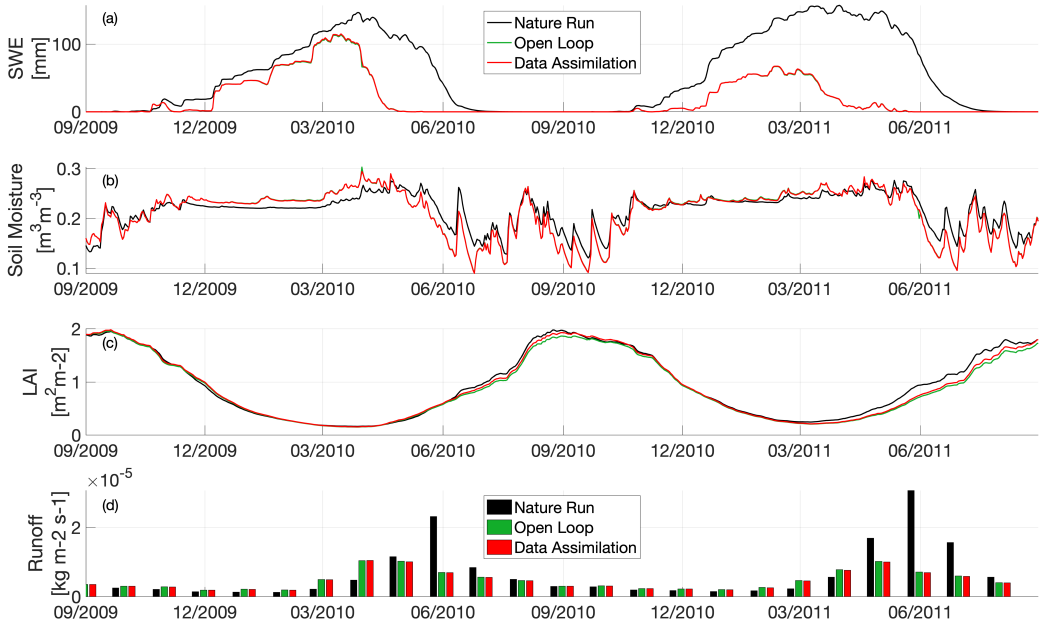

Figure 3.9: Domain-averaged time-series of SWE, soil moisture, and LAI along with monthly runoff from the **NR, OL, and DA using wide-swath LiDAR LAI**. Black lines/columns represent for NR, green lines/columns represent for OL, and red lines/columns represent for DA.

In summary, the univariate assimilation based on relatively low retrieval assumptions improves the model performance of the observed state variable, but does not necessarily improve the estimation of other variables or fluxes. Rather, some state variables and/or fluxes are degraded via assimilation when only one part of the snow-soil moisture-vegetation system is explicitly observed.

3.4.3 Impact of Multi-Variate Assimilation

Univariate assimilation is, in general, insufficient to properly update all the components in the coupled snow-soil moisture-vegetation system. Theoretically, multi-variate assimilation

should improve overall performance. Therefore, different combinations of synthetic retrievals (i.e., SAR SWE, PMW soil moisture, and wide-swath LiDAR LAI) are explored in assimilation experiments in the following section in order to examine the impact of multi-variate versus univariate assimilation, quantify the potential synergistic interplay between variables, and explore an optimal combination of different sensors.

Beginning with the univariate experiments presented above, one additional type observation is added to the univariate assimilation. Compared against the univariate assimilation results, the dual-variate assimilation shows better performance to both observed variables with various degrees of the synergistic interplay. Most notably, the DA using SAR SWE and PMW soil moisture provides apparent improvement to soil moisture estimation (Figure 3.10). The soil moisture output from the univariate soil moisture assimilation experiment [Figure 3.8(b)] can generally recover the NR during most seasons, but is generally underestimated during the snow ablation season. After the SWE retrievals are assimilated, the soil moisture during the snow ablation season increases and moves to the NR [Figure 3.10(b)]. However, there is no positive impact on SWE estimation via soil moisture retrieval assimilation [Figure 3.7(a) and Figure 3.10(a)].

The combination of SAR SWE and wide-swath LiDAR LAI retrievals (Figure 3.11) improves the estimation of LAI. In univariate LAI assimilation experiment, LAI estimation is improved near peak vegetation [from August to September, Figure 3.9(c)]. However, with SWE and LAI dual-variate assimilation, LAI improvements are extended in time [from March to July, Figure 3.11(c)]. The combination of PMW soil moisture and wide-swath LiDAR LAI yields only a small improvement for both state variables (Figure 3.12).

The unobserved variables might be still not improved via the dual-variate assimilation

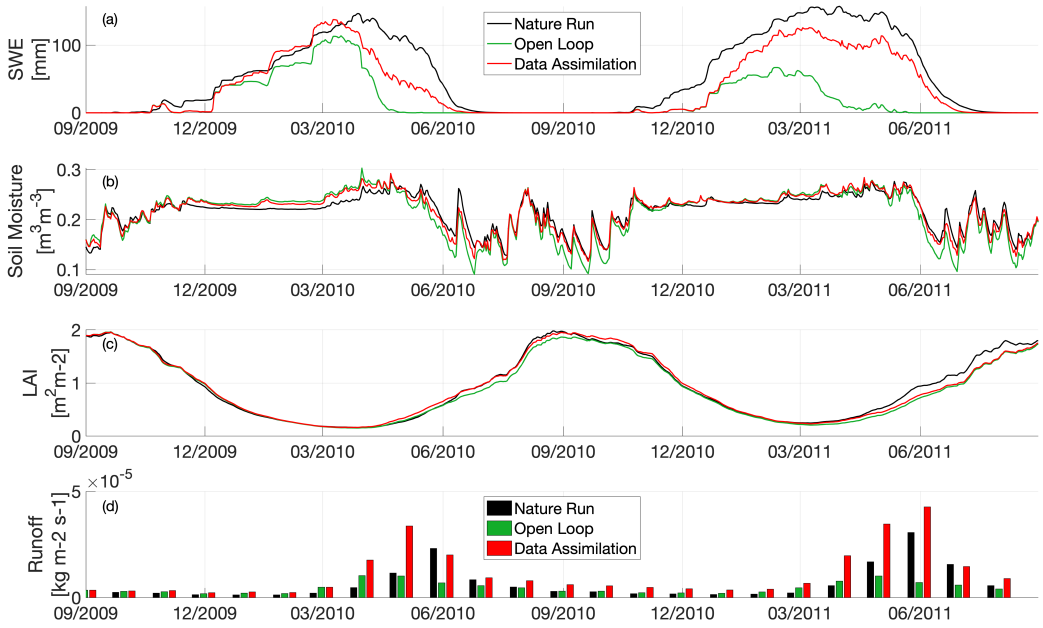

Figure 3.10: Domain-averaged time-series of SWE, soil moisture, and LAI along with monthly runoff from the **NR, OL, and DA using SAR SWE, and PMW soil moisture**. Black lines/columns represent the NR, green lines/columns represent the OL, and red lines/columns represent the DA.

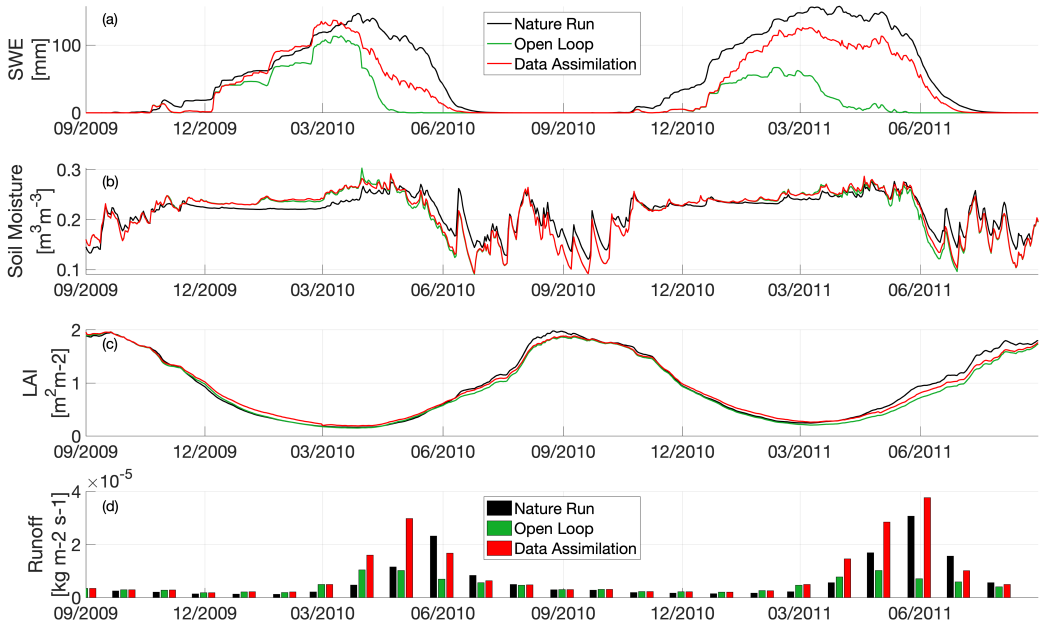

Figure 3.11: Domain-averaged time-series of SWE, soil moisture, and LAI along with monthly runoff from the **NR, OL, and DA using SAR SWE, and wide-swath LiDAR LAI**. Black lines/columns represent the NR, green lines/columns represent the OL, and red lines/columns represent the DA.

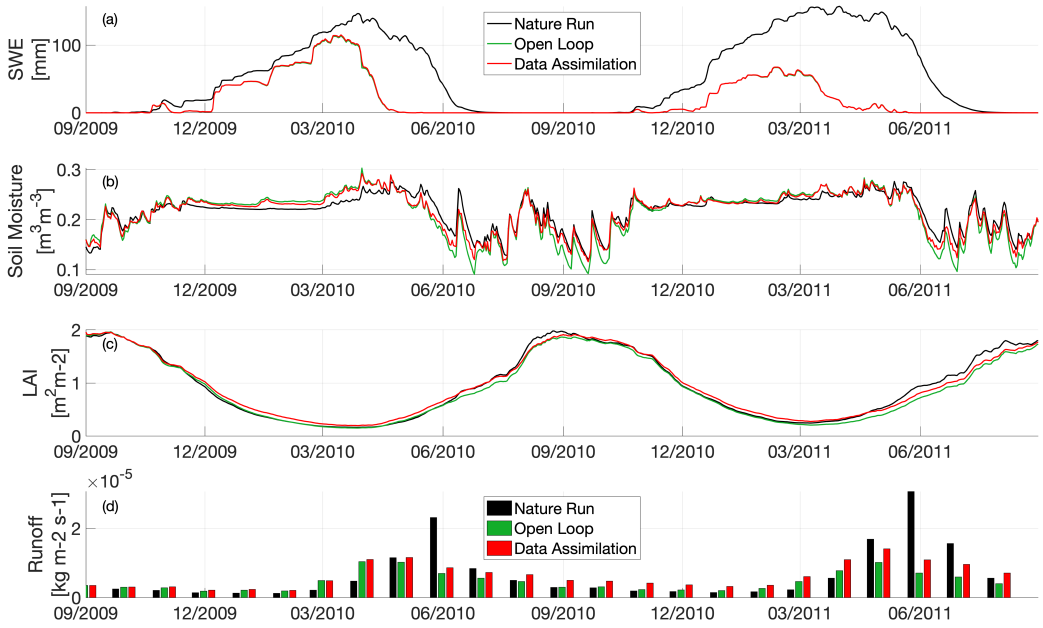

Figure 3.12: Domain-averaged time-series of SWE, soil moisture, and LAI along with runoff from the **NR, OL, and DA using PMW soil moisture, and wide-swath LiDAR LAI**. Black lines/columns represent the NR, green lines/columns represent the OL, and red lines/columns represent the DA.

experiment. For example, Figure 3.10(c) shows the assimilation of SAR SWE and PMW soil moisture avoids the negative impact on LAI witnessed in the univariate experiments. The impact of the SAR SWE plus wide-swath LiDAR LAI DA does not alter the soil moisture estimation compared to the univariate SWE DA. The assimilation using PMW soil moisture plus wide-swath LiDAR LAI similarly does not improve SWE estimation compared to either of the univariate assimilation equivalents.

The triple-variate (i.e. SAR SWE plus PMW soil moisture plus wide-swath LiDAR LAI) assimilation performs an overall best result compared to the dual-variate assimilation and univariate assimilation. All the three variables (i.e, SWE, soil moisture, and LAI) are observed and then improved via the triple-variate assimilation. However, dual-variate assimilation can reach a better performance in some of the observed variables. For instance, the triple-variate DA obtains an equivalent improvement on SWE compared to the dual-variate assimilation using SAR SWE and PMW soil moisture [Figure 3.13(a) and Figure 3.10(a)]. It also obtains an equivalent improvement on soil moisture compared to the dual-variate assimilation using SAR SWE and PMW soil moisture [Figure 3.13(b) and Figure 3.10(b)]. However, a **less** improvement is obtained from triple-variate assimilation on LAI than the dual-variate assimilation using SAR SWE and wide-swath LiDAR LAI [Figure 3.13(c) and Figure 3.12(c)].

It is observed that the phase of the domain-averaged, monthly-averaged runoff for the OL is different from the NR. The OL also yields a lower magnitude runoff than the NR. Via data assimilation, the peak time of runoff is closer to the NR result. However, the runoff is considerably overestimated during the snow ablation season in the univariate snow DA, because the large amount of snow mass added via DA (relative to the OL) ultimately results in the overestimation of runoff from the resulting snow melt. Set the selected point as example, Figure

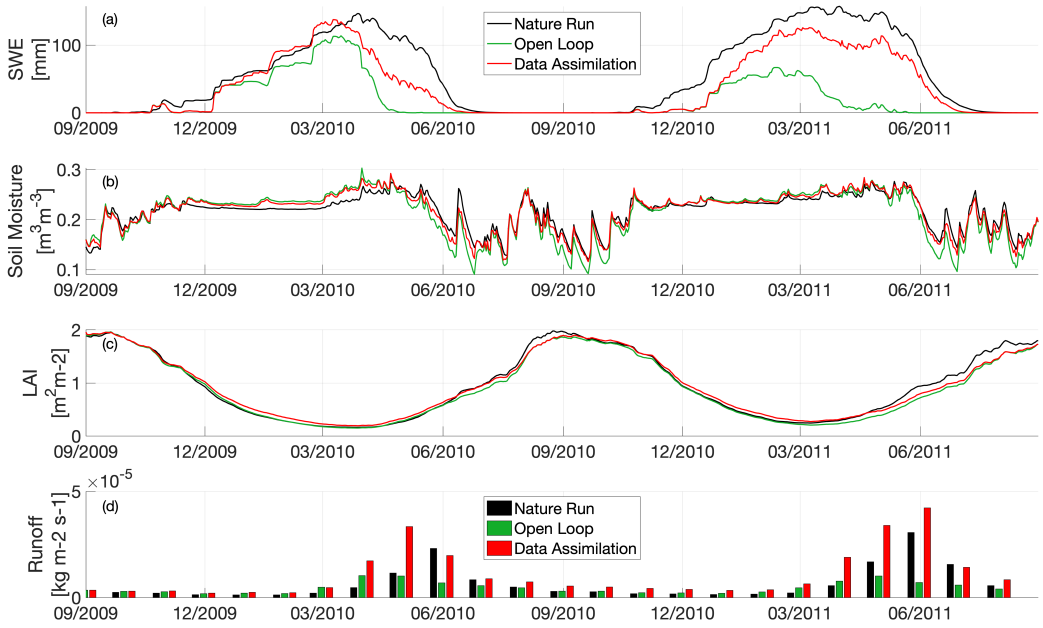

Figure 3.13: Domain-averaged time-series of SWE, soil moisture, and LAI along with monthly runoff from the **NR, OL, and DA using SAR SWE, PMW soil moisture, and wide-swath LiDAR LAI**. Black lines/columns represent the NR, green lines/columns represent the OL, and red lines/columns represent the DA.

3.14(c) shows an overestimation of runoff in DA during May 2010 while significant amounts of snow ablation occur as shown in Figure 3.14(a). The large amount of runoff resulted from snow melt. During May 2010, the daily-averaged temperature of OL and DA is mostly greater than 0 °C. As a result, DA SWE reduced sharply at the same time, which indicates the snow melting due to the warm weather, which is warmer than as modeled in the NR (results not shown). The earlier snowmelt in DA causes an earlier DA runoff peak in May compared with the NR runoff peak in June. Although the update to snow makes DA SWE better match with the NR SWE, the large mass added in DA is ultimately converted to runoff, and therefore, the runoff in the DA simulation yields more runoff earlier in the ablation season as produced by the NR.

In terms of univariate soil moisture DA, the runoff is slightly less underestimated, but the runoff peak timing remains unchanged. The LAI DA has a relatively negligible impact on runoff. In the multi-variate assimilation experiments, updates due to SWE retrieval assimilation dominate changes in model runoff. Secondary to SWE assimilation, updates due to soil moisture retrieval assimilation also alter changes in model runoff.

3.4.4 Integrated Evaluation

Considering the need to cohesively quantify the full suite of experimental scenarios, a simple time-series analysis is necessary but not sufficient. Therefore, an integrated (weighted) NIC is employed here to provide an overall assessment of the impact of assimilating synthetic retrievals on their own or in combination with others (Figure 3.15). The NIC for each single variable (i.e., SWE, soil moisture, or LAI) is also shown to assess the impact of assimilation experiments on the single variable. Positive NIC values indicate improvement via DA, whereas

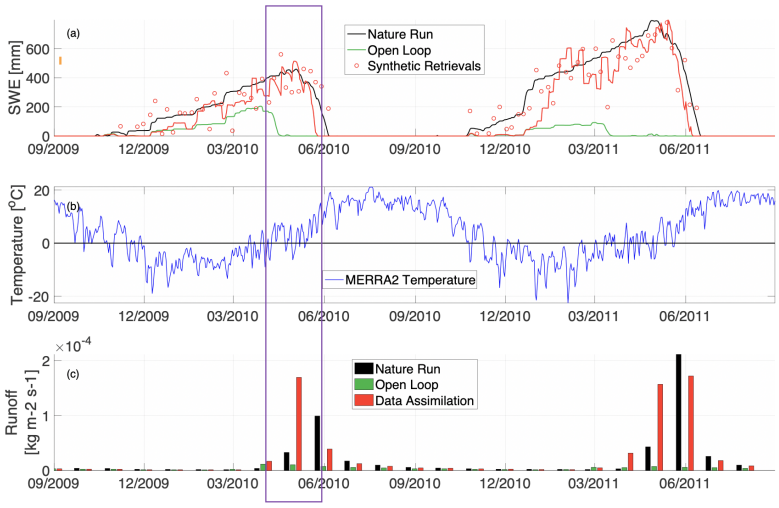

Figure 3.14: Time-series of SWE and runoff from the **NR, OL, and DA**, along with **the MERRA2 (used in OL and DA) daily-averaged temperature** at the selected point. The box highlights when the air temperature start going over the freezing point in 2010.

negative values indicate degradation via DA. A larger NIC indicates better performance on a relative scale. Figure 3.15 shows the NIC values of a suite of scenarios assimilating synthetic retrievals from different sensors/constellations (Table 3.1 and Table 3.3) with different retrieval error assumptions (Table 3.2).

Figure 3.15(a) shows the integrated NIC of 25 DA scenarios. The best integrated NIC scenario is the triple-variate assimilation using **SAR SWE, PMW soil moisture and wide-swath LiDAR LAI** with small retrieval error assumptions. The second and third integrated NIC scenarios are the dual-variate assimilation using **SAR SWE and wide-swath LiDAR LAI**, as well as dual-variate assimilation using **SAR SWE and PMW soil moisture**. The result that the multi-variate assimilation overcomes univariate assimilation is identical with what observed in time-series figures. What comes next in integrated NIC scenarios are the univariate assimilation using SAR SWE and using SAR LAI with small retrieval error assumptions. These two scenarios reach top improvements among univariate assimilation scenarios because the retrievals have

high revisit frequency, fine spatial resolution, and low retrieval errors. Compared to them, the univariate assimilation using PMW soil moisture has a relatively low NIC because the PMW soil moisture retrievals have a coarse spatial resolution (36 km), which makes it difficult to catch the spatial variation of soil moisture in complex terrain. Besides, the existence of frozen soil conditions during winter also hinders the performance of soil moisture assimilation. Compared to the PMW soil moisture, the univariate assimilation using PMW SWE, whose spatial resolution is also coarse (25km), performs better. One obvious reason is that the integrated NIC of SWE assimilation benefits from synergistic interplay, because SWE assimilation has positive effect on soil moisture and negligible effect on LAI, but soil moisture assimilation has negligible effect on SWE and negative effect on LAI (Figure 3.15(b), Figure 3.15(c), and Figure 3.15(d)).

Figure 3.15(a) also shows that model performance may become worse resulting from high retrieval errors. For example, SAR SWE with a 0.9 m (additive) retrieval error has an obvious negative effect on SWE assimilation, but PMW SWE with a 0.9 (multiplicative) retrieval error only has a marginal negative effect. In the study domain, the large additive error (0.9 m) is generally more noisy than the large multiplicative error (0.9), which is smaller than 0.9 m additive error unless the SWE is larger than 1 m. The narrow swath width could be another factor that limits the performance of assimilation. For example, assimilation of LiDAR sensors (for both snow depth and LAI) results in a relatively low NIC since LiDAR sensors have narrow swath widths compared to PMW and SAR sensors. Besides, LiDAR observations are further reduced due to cloud contamination.

Figure 3.15(b), (c), and (d) present the NIC of each single variable (i.e., SWE, soil moisture, and LAI). The NIC indicates the impact of assimilation from each scenario on a specific variable. For example, Figure 3.15(b) shows that the assimilation of soil moisture and/or LAI retrievals

does not affect modelled SWE. Figure 3.15(c) shows that the assimilation of SWE has a positive effect on modelled soil moisture, whereas, the impact of LAI assimilation is negligible. Figure 3.15(d) indicates that the assimilation of synthetic SWE with low retrieval error has a negligible effect on modelled LAI, but produces a negative impact when the retrieval error increases. The soil moisture assimilation has a slightly negative impact on modelled LAI. Figure 3.15(d) exhibits an instance of synergistic interplay between SWE DA and LAI DA. The dual-assimilation using SAR SWE and wide-swath LiDAR has a larger NIC than the summed NIC of univariate assimilation using SAR SWE and using wide-LiDAR LAI. The main difference between LAI time-series shown in Figure 3.9(c) and Figure 3.11(c) is during the growing season in 2011. The wetter soil moisture associated with snow melting in multi-variate assimilation leads to more vigorous vegetation, which is closer to the NR compared to the univariate LAI assimilation.

3.5 Conclusion and Discussion

This study sets up an advanced OSSE to evaluate the impact of assimilation on the coupled snow-soil moisture-vegetation system modelling. 25 scenarios are assimilated including synthetic retrievals for single variables as well as combinations of different variables coupled with different amount of retrieval errors. The results of DA are evaluated against the NR and OL using time-series and NIC.

Multi-variate assimilation better improves the estimation of the coupled snow-soil moisture-vegetation system compared to univariate assimilation. In the univariate assimilation, the improvement of the unobserved variable is not guaranteed even when the synthetic retrieval is of good quality (i.e., high revisit frequency, fine spatial resolution, and low retrieval error, and no cloud attenuation).

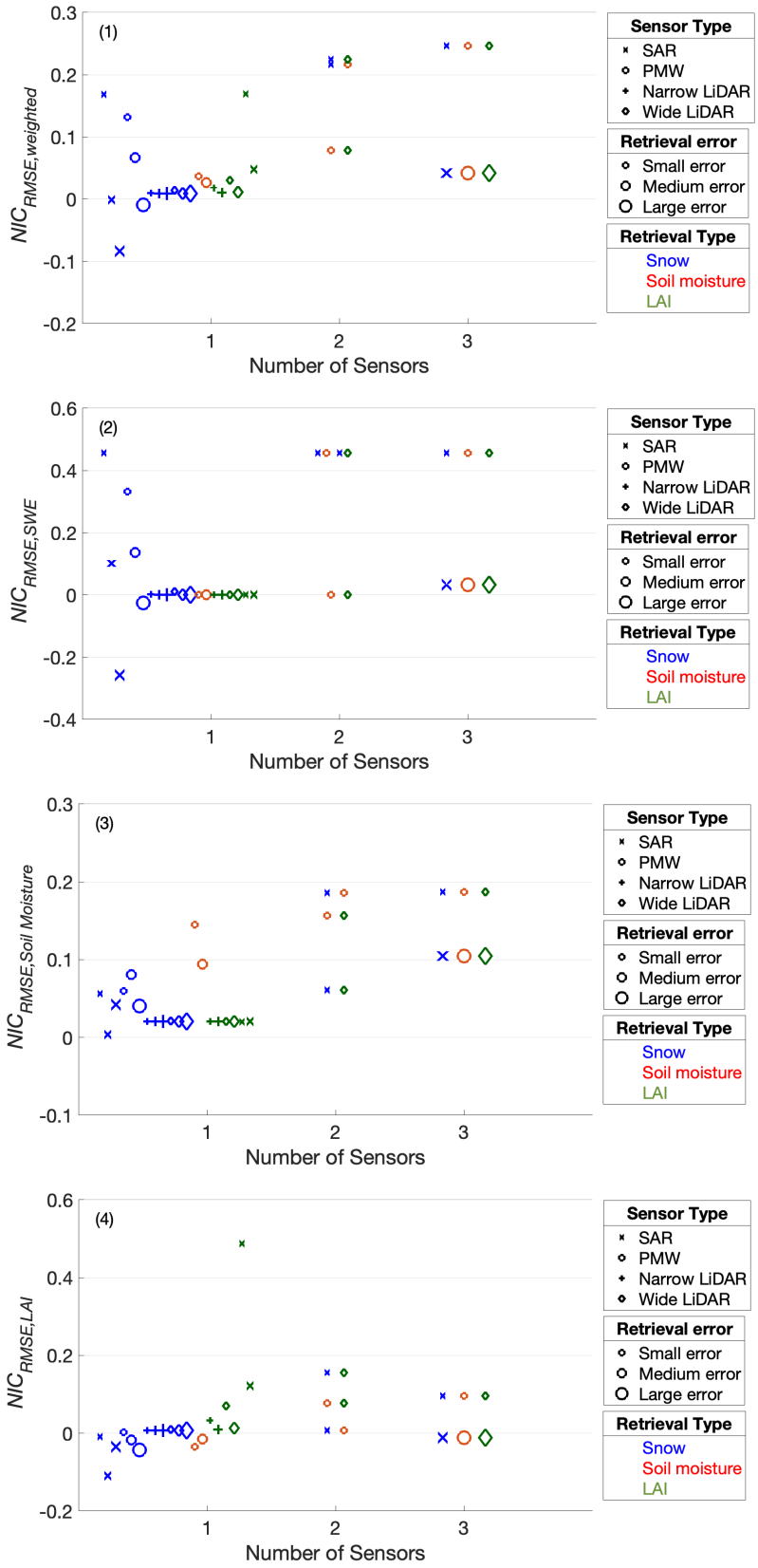

Figure 3.15: NIC_{RMSE} of DA experiments with different kinds of retrievals. Markers represents to sensor types (×=SAR sensors; ○=PMW sensors; +=narrow LiDARs; and ◇=wide LiDARs). The small, medium, and large size of symbols represents the different amount of injected retrieval errors (Table 3.2). The color of symbols refers to synthetic retrievals types (blue=snow; orange=soil moisture; and green=LAI).

For example, the model LAI degrades in the soil moisture assimilation. The multi-variate assimilation assimilated more observed variable(s) and better improved the model in this coupled system, when the synthetic retrievals were of good quality. According to the NIC, the triple-variate assimilation perform best overall.

Synergistic interplay exists in the coupled snow-soil moisture-vegetation system. First, the assimilation of the observed variable(s) can impact the estimation of the unobserved variables (e.g. soil moisture estimation is improved by SAR SWE univariate assimilation). Second, on the observed variable, the multi-assimilation shows better performance compared to the univariate assimilation, e.g., dual-variate assimilation using SAR SWE and PMW soil moisture vs. univariate assimilation using PMW soil moisture on improving soil moisture; dual-variate assimilation using SAR SWE and wide-LiDAR LAI vs. univariate assimilation using wide-LiDAR LAI on improving LAI.

The NIC analysis also indicates the maximum acceptable retrieval error for each type of retrievals in order to get a marginal improvement. The retrieval error larger than the indicated maximum can lead to degradation of model performance.

The impact on runoff is complex. As shown in result, the runoff in the SWE assimilation better matches the peak timing of the NR, but overestimates the runoff magnitude. The soil moisture assimilation reduces the magnitude of underestimation in the OL but does not alter the runoff peak timing.

There are several limitations to this study. First, this study does not include hydraulic routing. Therefore, a monthly-averaged runoff is used in order to reduce the impact of runoff timing errors. Future OSSEs could include a routing model such as HYMAP [176] in order to reduce these errors. Hydraulic routing, however, was avoided to reduce computational costs in

this study. Runoff from a hydraulic model could provide a better understanding of the impacts of assimilation, e.g., understanding the question of how snow mass updates during the snow ablation season impact estimates of peak flow magnitude and timing.

Second, the assumed error characteristics (Table 3.2) have identical independent distribution across time and space. However, the true error varies across space and time due to different land surface conditions. For example, SWE retrieval error from PMW sensors would be considerably affected by the type of overlying vegetation, the relative depth (e.g., shallow vs. deep), and snow wetness (dry vs. wet) [29, 177]. Therefore the observation error of SWE should be modeled as a function of forest density, snow depth, and snow wetness. Such a spatiotemporal variation in the observation error would be more accurate. Improvements to assumed error characteristics would likely lead to a more robust OSSE framework.

In summary, an advanced OSSE is developed in this paper in order to quantify the added value of synthetic retrievals for improving the estimation of the geophysical variables in the coupled snow-soil moisture-vegetation system. The synthetic retrievals of SWE, snow depth, soil moisture, and LAI from a group of sensors and constellations are assimilated and evaluated in the OSSE framework. The results suggest that in univariate assimilation, retrievals with good quality ((i.e., high revisit frequency, fine spatial resolution, and low retrieval error, and no cloud attenuation) better improve the model performance than retrievals with bad quality do. The multi-variate assimilation achieved a better overall performance as additional data containing more information and further the synergistic interplay between different types of observations is contained. The multi-variate assimilation with all the three variables best represents the whole snow-soil moisture-vegetation system.

Chapter 4: Overall Conclusions and Discussions

4.1 Conclusions

This dissertation explores different sensor combinations to observe the snow-soil moisture-vegetation system from space. The science question addressed is which sensor or constellation of sensors is most effective at quantifying and characterizing the hydrologic cycle?

In Chapter 2, this dissertation simulated the snow viewing extent of a set of representative orbit configurations, in order to investigate which kind of sensor or constellation is able to view the most snow covered terrain not only in the terms of total spatial coverage percentage, but also in the terms of coverage percentage as a function of different snow classifications. The results suggest that 1100 km, 550 km, and 200 km are the minimum required swath widths for a polar-orbiting sensor to meet snow-related applications demanding a 1-day, 3-day, and 30-day repeat cycle, without considering cloud attenuation and sensor limitation on detecting specific types of snow. Alternatively, simulations *including* cloud attenuation and sensor efficiency suggests that no single sensor can observe all snow classes at all locations. Instead, a combination of PMW, SAR, and LiDAR sensors works best for snow detection.

Although the viewed coverage percentage simulation in Chapter 2 can evaluate the observation coverage of each sensor/constellation, the simulation cannot tell us the value of the observation coverage in terms of land surface model improvement. In Chapter 3, an advanced OSSE is

developed in order to quantitatively define the value of observations in improving the geophysical variables estimation in the coupled snow-soil moisture-vegetation system. The synthetic retrievals of SWE, snow depth, soil moisture, and LAI from a group of sensor/constellation are assimilated and evaluated in the OSSE framework. The results show that the univariate assimilation for snow/soil moisture/LAI with high quality retrievals is able to improve the model performance of the corresponding observed geophysical variable, however, not necessarily to improve the model estimation of the other geophysical variables. The univariate assimilation does not guarantee model improvement because focusing on one variable may lead to a worse estimation of other variables.

The integrated NIC, which is used to evaluate the synthetic retrievals comprehensively, suggests that the multi-variate assimilation with more assimilated variables better represents the whole system. There is obvious synergistic interplay between snow and soil moisture as well as snow and vegetation observed in dual-variate assimilation. The triple-variate assimilation with high quality retrievals of SWE, soil moisture, LAI performs best overall in estimating the hydrological components of snow-soil moisture-vegetation system. Chapter 3 not only extends the study of Chapter 2 from snow to the coupled snow-soil moisture-vegetation system, but also provides a method to evaluate the synthetic retrievals in terms of improving land surface model performance instead of only reporting the observed coverage percentage. Chapter 3 provides a quantitative solution to the overarching goal: finding the optimal mix of the synthetic retrievals in snow-soil moisture-vegetation hydrologic system.

The OSSE developed in this dissertation could serve as a valuable tool for hydrological mission planners. Synthetic retrievals observed from eight orbital configurations and 25 scenarios with various retrieval errors are simulated in this dissertation, however, there could be an infinite

number of combinations and permutations for the synthetic retrieval simulation for hydrological applications. This OSSE is able to help mission planners to evaluate their ideas and decide where (e.g., polar-orbiting versus low-inclination versus geostationary), when (e.g., single sensor versus constellation), and what (active RADAR versus passive microwave radiometer versus LiDAR) sensors to employ given their specific goals. For example, this OSSE has been applied in another study to evaluate the synthetic snow retrievals from a snow LiDAR with an agile viewing strategy [178]. The agile viewing has the potential to provide 50% more observations of snow during the snow seasons for the sensors with around 20 km swath width, while sacrificing the observations on the snow-free area. According to the evaluation by this OSSE, the extra observations provided by agile viewing strategy helps the model better estimate snow mass with lower RMSE. The NIC score of RMSE for SWE increases from 1.6% (fixed viewing strategy) to 5.0% (agile viewing strategy) when all other sensor and orbit configurations are the same. The evaluation shows that the agile viewing strategy successfully improves the efficiency of narrow-swath-width sensors, such as LiDAR for detecting a specific target. This example shows the OSSE as a "soup to nuts" approach to hydrological mission planning of a satellite (or constellation). Further, the OSSE framework could integrate with the modules of mission cost estimation, and risk analysis, so that it can help mission planners to make a scientific bang for the certain scientific buck.

The advanced OSSE also provides a more cohesive view of the coupled snow-soil moisture-vegetation system. It is proved that synergistic interplay among the different components in this system can be significant. The state variables without assimilation might get improved, or even degraded, from the assimilated components in this system. Further, the multi-variate assimilation approach often yields a better constrained system, and hence, can overcome some of the deficiencies of univariate assimilation. A well-constructed, multi-variate assimilation system

can benefit applications related to the snow-soil moisture-vegetation system, such as agriculture yield production estimates, drought index assessments, and reservoir operations and storage estimation. Applying multi-variate assimilation in such applications would provide better results related to snow, soil moisture, and vegetation.

4.2 Limitations

This dissertation still has many limitations. First, as discussed in Chapter 3, the runoff model used in this study is a rough approximation from the sum of the surface flow and base flow. To reduce the approximation error, the runoff is temporally integrated into a coarse temporal scale, e.g., monthly. To yield a more accurate and physically-realistic daily runoff estimation, a runoff routing model, such as Hydrological Modeling and Analysis Platform (HYMAP) [176] is needed. The runoff from a routing model would allow the OSSE to better analyze the impacts of assimilation on river discharge.

Second, the distribution of error injected in Chapter 3 is assumed identical across time and space. However, the true error varies across space and time due to different land surface conditions. For example, the snow mass retrievals from PMW sensors would be considerably affected by the type of overlying vegetation, the relative depth (e.g., shallow vs. deep), and snow wetness (dry vs. wet), which matches with the knowledge obtained in Chapter 2. Therefore the observation error of SWE should be modeled as a function of dense/sparse forest, deep/shallow snow, dry/wet snow, and topography. The spatiotemporal observation errors would be more realistic. It will also allow the OSSE to ingest the retrievals from sensors and constellations more accurately.

Third, the error used in this study is assumed to be uncorrelated in space in order to reduce computational expense. However, the errors could be spatially correlated in the real world. When the errors are spatially correlated but assumed as independent and identically distributed, the error covariances are generally underestimated. With the inclusion of correlated errors, the error covariances would better represent the real world errors while also inflating the model ensemble to better represent the true magnitude of the model uncertainties. Therefore, The use of spatially-correlated errors is desired in a follow-on study. To apply spatially-correlated error to the model or the observations, a turning bands algorithm [179] could be used to generate the spatially-correlated error fields. Then the one-dimensional EnKF algorithm used in this study could be replaced with a two-dimensional EnKF algorithm in order to allow the updates from the neighboring pixels to harness a spatially-correlated error structure that surrounds those pixels.

4.3 Future Work

The aim of this study is to find the optimal design of sensors and constellations configurations to minimize the cost while providing the most scientific utility. Thus, coupling the OSSE with a tool kit to estimate the cost and risk of designs would be an essential input for mission planners to make better decisions with better science gains and reasonable cost estimates than current solutions.

Bibliography

[1] David Seckler, Randolph Barker, and Upali Amarasinghe. Water scarcity in the twenty-first century. *International Journal of Water Resources Development*, 15(1-2):29–42, 1999.

[2] Guy Howard, Jamie Bartram, Sanitation Water, World Health Organization, et al. Domestic water quantity, service level and health. Technical report, World Health Organization, 2003.

[3] World Health Organization et al. International decade for action water for life, 2005-2015. *Weekly Epidemiological Record*, 80(22):195–200, 2005.

[4] R Warren Flint. The sustainable development of water resources. *Water Resources Update*, 127:48–59, 2004.

[5] Mirela G Tulbure and Mark Broich. Spatiotemporal dynamic of surface water bodies using landsat time-series data from 1999 to 2011. *ISPRS Journal of Photogrammetry and Remote Sensing*, 79:44–52, 2013.

[6] Huilin Gao. Satellite remote sensing of large lakes and reservoirs: From elevation and area to storage. *Wiley Interdisciplinary Reviews: Water*, 2(2):147–157, 2015.

[7] Eni G Njoku and Dara Entekhabi. Passive microwave remote sensing of soil moisture. *Journal of Hydrology*, 184(1-2):101–129, 1996.

[8] Eni G Njoku and Steven K Chan. Vegetation and surface roughness effects on AMSR-E land observations. *Remote Sensing of Environment*, 100(2):190–199, 2006.

[9] 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.

[10] Matthew C Hansen, David P Roy, Erik Lindquist, Bernard Adusei, Christopher O Justice, and Alice Altstatt. A method for integrating MODIS and landsat data for systematic monitoring of forest cover and change in the Congo Basin. *Remote Sensing of Environment*, 112(5):2495–2513, 2008.

[11] Evan Pugh and Eric Gordon. A conceptual model of water yield effects from beetle-induced tree death in snow-dominated lodgepole pine forests. *Hydrological Processes*, 27(14):2048–2060, 2013.

[12] Franklin W Schwartz and Hubao Zhang. *Fundamentals of ground water*. John Wiley & Sons, 2002.

[13] Erika Coppola, Francesca Raffaele, and Filippo Giorgi. Impact of climate change on snow melt driven runoff timing over the Alpine region. *Climate Dynamics*, 51(3):1259–1273, 2018.

[14] 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.

[15] I Rodriguez-Iturbe, P D'odorico, A Porporato, and L Ridolfi. On the spatial and temporal links between vegetation, climate, and soil moisture. *Water Resources Research*, 35(12):3709–3722, 1999.

[16] M Tedesco. Algorithm theoretical basis document NASA AMSR-E SWE product. 2012.

[17] David Moeser, Jiri Roubinek, Patrick Schleppi, Felix Morsdorf, and Tobias Jonas. Canopy closure, LAI and radiation transfer from airborne LiDAR synthetic images. *Agricultural and Forest Meteorology*, 197:158–168, 2014.

[18] Yann H Kerr, Philippe Waldteufel, Philippe Richaume, Jean Pierre Wigneron, Paolo Ferrazzoli, Ali Mahmoodi, Ahmad Al Bitar, François Cabot, Claire Gruhier, Silvia Enache Juglea, et al. The SMOS soil moisture retrieval algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, 50(5):1384–1403, 2012.

[19] 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.

[20] R Ramon Solano, K Didan, A Jacobson, and A Huete. MODIS vegetation index user's guide. *The University of Arizona: Tucson, AZ, USA*, 2010.

[21] George A Riggs, Dorothy K Hall, and Miguel O Román. MODIS snow products collection 6 user guide. *National Snow and Ice Data Center: Boulder, CO, USA*, page 66, 2015.

[22] Xiangjin Meng, Kebiao Mao, Fei Meng, Xinyi Shen, Tongren Xu, and Mengmeng Cao. Long-term spatiotemporal variations in soil moisture in North East China based on 1-km resolution downscaled passive microwave soil moisture products. *Sensors*, 19(16):3527, 2019.

[23] Alan Strahler, Doug Muchoney, Jordan Borak, Mark Friedl, Sucharita Gopal, Eric Lambin, and Aaron Moody. MODIS land cover product algorithm theoretical basis document (ATBD) version 5.0. 1999.

[24] Andrés Varhola, Nicholas C Coops, Markus Weiler, and R Dan Moore. Forest canopy effects on snow accumulation and ablation: An integrative review of empirical results. *Journal of Hydrology*, 392(3-4):219–233, 2010.

[25] 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.

[26] 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.

[27] C Derksen, A Walker, E LeDrew, and B Goodison. Combining SMMR and SSM/I data for time series analysis of central North American snow water equivalent. *Journal of Hydrometeorology*, 4(2):304–316, 2003.

[28] Ross D Brown, Bruce Brasnett, and David Robinson. Gridded North American monthly snow depth and snow water equivalent for GCM evaluation. *Atmosphere-Ocean*, 41(1):1–14, 2003.

[29] James L Foster, Chaojiao Sun, Jeffrey P Walker, Richard Kelly, Alfred Chang, Jiarui Dong, and Hugh Powell. Quantifying the uncertainty in passive microwave snow water equivalent observations. *Remote Sensing of Environment*, 94(2):187–203, 2005.

[30] Marco Tedesco, REJ Kelly, JL Foster, and ATC Chang. AMSR-E/Aqua daily L3 global snow water equivalent EASE-Grids V002. *National Snow and Ice Data Center: Boulder, CO*, 2004.

[31] 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.

[32] Monique Bernier, Jean-Pierre Fortin, Yves Gauthier, Raymond Gauthier, René Roy, and Pierre Vincent. Determination of snow water equivalent using RADARSAT SAR data in Eastern Canada. *Hydrological Processes*, 13(18):3041–3051, 1999.

[33] Jiyue Zhu, Shurun Tan, Joshua King, Chris Derksen, Juha Lemmetyinen, and Leung Tsang. Forward and inverse radar modeling of terrestrial snow using SnowSAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 56(12):7122–7132, 2018.

[34] Hui Li, Zuo Wang, Guangjun He, and Wang Man. Estimating snow depth and snow water equivalence using repeat-pass interferometric SAR in the northern piedmont region of the tianshan mountains. *Journal of Sensors*, 2017, 2017.

[35] Ya-Lun S Tsai, Andreas Dietz, Natascha Oppelt, and Claudia Kuenzer. Remote sensing of snow cover using spaceborne SAR: A review. *Remote Sensing*, 11(12):1456, 2019.

[36] R. Storvold, E. Malnes, Y. Larsen, K. A. Høgda, S. E. Hamran, K. Müller, and K. A. Langley. SAR remote sensing of snow parameters in Norwegian areas - current status and future perspective. *PIERS 2006 Cambridge - Progress in Electromagnetics Research Symposium, Proceedings*, pages 182–186, 2006.

[37] Chris Hopkinson, Mike Sitar, Laura Chasmer, and Paul Treitz. Mapping snowpack depth beneath forest canopies using airborne LiDAR. *Photogrammetric Engineering & Remote Sensing*, 70(3):323–330, 2004.

[38] Jeffrey S Deems and Thomas H Painter. LiDAR measurement of snow depth: accuracy and error sources. In *Proceedings of the 2006 International Snow Science Workshop: Telluride, Colorado, USA, International Snow Science Workshop*, pages 330–338, 2006.

[39] David Mark Cobby. *The use of airborne scanning laser altimetry for improved river flood prediction*. PhD thesis, University of Reading, 2002.

[40] M Nolan, CF Larsen, and M Sturm. Mapping snow-depth from manned-aircraft on landscape scales at centimeter resolution using structure-from-motion photogrammetry. *Cryosphere Discussions*, 9(1), 2015.

[41] Edward Kim, Barton A Forman, Lizhao Wang, Jacqueline Lemoigne-Stewart, Sreeja Nag, Sujay Kumar, Carrie Vuyovich, Bryan Blair, and Michelle Hofton. Space-time coverage scenarios for a global snow satellite constellation. In *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, pages 5614–5616. IEEE, 2019.

[42] L Karthikeyan, Ming Pan, Niko Wanders, D Nagesh Kumar, and Eric F Wood. Four decades of microwave satellite soil moisture observations: Part 1. a review of retrieval algorithms. *Advances in Water Resources*, 109:106–120, 2017.

[43] XIAO Qianguang, CHEN Weiying, SHENG Yongwei, and Li Jing. A study on soil molsture monitoring using NOAA satellite. *Quarterly Journal of Applied Meteorology*, 5(3):312–317, 1994.

[44] L Wang, JJ Qu, S Zhang, X Hao, and S Dasgupta. Soil moisture estimation using MODIS and ground measurements in Eastern China. *International Journal of Remote Sensing*, 28(6):1413–1418, 2007.

[45] Bo-Cai Gao. Ndwi—a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58(3):257–266, 1996.

[46] George P Petropoulos, Gareth Ireland, and Brian Barrett. Surface soil moisture retrievals from remote sensing: Current status, products & future trends. *Physics and Chemistry of the Earth, Parts A/B/C*, 83:36–56, 2015.

[47] William Burke and J. Paris. A radiative transfer model for microwave emissions from bare agricultural soils. 09 1975.

[48] TJ Schmugge, JR Wang, and G Asrar. Results from the push broom microwave radiometer flights over the Konza Prairie in 1985. *IEEE Transactions on Geoscience and Remote Sensing*, 26(5):590–596, 1988.

[49] Hyunglok Kim and Venkat Lakshmi. Use of cyclone global navigation satellite system (CYGNSS) observations for estimation of soil moisture. *Geophysical Research Letters*, 45(16):8272–8282, 2018.

[50] Steven K Chan, Rajat Bindlish, Peggy E O'Neill, Eni Njoku, Tom Jackson, Andreas Colliander, Fan Chen, Mariko Burgin, Scott Dunbar, Jeffrey Piepmeier, et al. Assessment of the SMAP passive soil moisture product. *IEEE Transactions on Geoscience and Remote Sensing*, 54(8):4994–5007, 2016.

[51] M Susan Moran, Christa D Peters-Lidard, Joseph M Watts, and Stephen McElroy. Estimating soil moisture at the watershed scale with satellite-based radar and land surface models. *Canadian Journal of Remote Sensing*, 30(5):805–826, 2004.

[52] Kurt C Kornelsen and Paulin Coulibaly. Advances in soil moisture retrieval from synthetic aperture radar and hydrological applications. *Journal of Hydrology*, 476:460–489, 2013.

[53] SE Plummer. Perspectives on combining ecological process models and remotely sensed data. *Ecological Modelling*, 129(2-3):169–186, 2000.

[54] Jing M Chen and Josef Cihlar. Retrieving leaf area index of boreal conifer forests using landsat tm images. *Remote Sensing of Environment*, 55(2):153–162, 1996.

[55] Ranga B Myneni, S Hoffman, Yuri Knyazikhin, JL Privette, J Glassy, Yuhong Tian, Y Wang, X Song, Y Zhang, GR Smith, et al. Global products of vegetation leaf area and fraction absorbed par from year one of MODIS data. *Remote Sensing of Environment*, 83(1-2):214–231, 2002.

[56] Atmosphere ECVs, Terrestrial ECVs, and Ocean ECVs. Gcos. 2005.

[57] Hongliang Fang, Chongya Jiang, Wenjuan Li, Shanshan Wei, Frédéric Baret, Jing M Chen, Javier Garcia-Haro, Shunlin Liang, Ronggao Liu, Ranga B Myneni, et al. Characterization and intercomparison of global moderate resolution leaf area index (LAI) products: Analysis of climatologies and theoretical uncertainties. *Journal of Geophysical Research: Biogeosciences*, 118(2):529–548, 2013.

[58] Kai Yan, Taejin Park, Guangjian Yan, Zhao Liu, Bin Yang, Chi Chen, Ramakrishna R Nemani, Yuri Knyazikhin, and Ranga B Myneni. Evaluation of MODIS LAI/FPAR product collection 6. Part 2: Validation and intercomparison. *Remote Sensing*, 8(6):460, 2016.

[59] Benjamin Koetz, Felix Morsdorf, Guoqing Sun, K Jon Ranson, K Itten, and B Allgower. Inversion of a LiDAR waveform model for forest biophysical parameter estimation. *IEEE Geoscience and Remote Sensing Letters*, 3(1):49–53, 2006.

[60] Moskal LM Zheng G. Retrieving leaf area index (LAI) using remote sensing: theories, methods and sensors. *Sensors*, 9(4):2719–2745, 2009.

[61] Ralph Dubayah, James Bryan Blair, Scott Goetz, Lola Fatoyinbo, Matthew Hansen, Sean Healey, Michelle Hofton, George Hurtt, James Kellner, Scott Luthcke, et al. The global ecosystem dynamics investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of Remote Sensing*, 1:100002, 2020.

[62] Peter Potapov, Xinyuan Li, Andres Hernandez-Serna, Alexandra Tyukavina, Matthew C Hansen, Anil Kommareddy, Amy Pickens, Svetlana Turubanova, Hao Tang, Carlos Edibaldo Silva, et al. Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote Sensing of Environment*, 253:112165, 2021.

[63] D Barry Coyle, Paul R Stysley, Demetrios Poulios, Greg B Clarke, and Richard B Kay. Laser transmitter development for NASA's global ecosystem dynamics investigation (GEDI) LiDAR. In *LiDAR Remote Sensing for Environmental Monitoring XV*, volume 9612, page 961208. International Society for Optics and Photonics, 2015.

[64] Paul L Patterson and Sean Healey. Global ecosystem dynamics investigation (GEDI) LiDAR sampling strategy. In *In: Stanton, Sharon M.; Christensen, Glenn A., comps. 2015. Pushing boundaries: new directions in inventory techniques and applications: Forest Inventory and Analysis (FIA) symposium 2015. 2015 December 8–10; Portland, Oregon. Gen. Tech. Rep. PNW-GTR-931. Portland, OR: US Department of Agriculture, Forest Service, Pacific Northwest Research Station. p. 245.*, volume 931, 2015.

[65] Wenlu Qi, Svetlana Saarela, John Armston, Göran Ståhl, and Ralph Dubayah. Forest biomass estimation over three distinct forest types using TanDEM-X InSAR data and simulated GEDI LiDAR data. *Remote Sensing of Environment*, 232:111283, 2019.

[66] Gayathri K Devia, BP Ganasri, and GS Dwarakish. A review on hydrological models. *Aquatic Procedia*, 4:1001–1007, 2015.

[67] Rosie A Fisher and Charles D Koven. Perspectives on the future of land surface models and the challenges of representing complex terrestrial systems. *Journal of Advances in Modeling Earth Systems*, 12(4):e2018MS001453, 2020.

[68] J. Overgaard, D. Rosbjerg, and M. B. Butts. Land-surface modelling in hydrological perspective – a review. *Biogeosciences*, 3(2):229–241, 2006.

[69] Syukuro Manabe. Climate and the ocean circulation: I. the atmospheric circulation and the hydrology of the Earth's surface. *Monthly Weather Review*, 97(11):739–774, 1969.

[70] John L Monteith. Evaporation and environment. In *Symposia of the society for experimental biology*, volume 19, pages 205–234. Cambridge University Press (CUP) Cambridge, 1965.

[71] John Monteith and Mike Unsworth. *Principles of environmental physics: plants, animals, and the atmosphere*. Academic Press, 2013.

[72] JL Heilman, KJ McInnes, MJ Savage, RW Gesch, and RJ Lascano. Soil and canopy energy balances in a west Texas vineyard. *Agricultural and Forest Meteorology*, 71(1-2):99–114, 1994.

[73] PJ Sellers, YCSY Mintz, YC e al Sud, and A Dalcher. A simple biosphere model (SiB) for use within general circulation models. *Journal of the Atmospheric Sciences*, 43(6):505–531, 1986.

[74] Lianhong Gu, Herman H Shugart, Jose D Fuentes, TA Black, and Stanley R Shewchuk. Micrometeorology, biophysical exchanges and NEE decomposition in a two-story boreal forest—development and test of an integrated model. *Agricultural and Forest Meteorology*, 94(2):123–148, 1999.

[75] Sujan Pal and Prateek Sharma. A review of machine learning applications in land surface modeling. *Earth*, 2(1):174–190, 2021.

[76] Xin Li. Characterization, controlling, and reduction of uncertainties in the modeling and observation of land-surface systems. *Science China Earth Sciences*, 57(1):80–87, 2014.

[77] Yeosang Yoon, Sujay V Kumar, Barton A Forman, Benjamin F Zaitchik, Yonghwan Kwon, Yun Qian, Summer Rupper, Viviana Maggioni, Paul Houser, Dalia Kirschbaum, et al. Evaluating the uncertainty of terrestrial water budget components over High Mountain Asia. *Frontiers in Earth Science*, 7:120, 2019.

[78] Qinglong You, Jinzhong Min, Wei Zhang, Nick Pepin, and Shichang Kang. Comparison of multiple datasets with gridded precipitation observations over the Tibetan Plateau. *Climate Dynamics*, 45(3):791–806, 2015.

[79] R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45, 03 1960.

[80] Rudi Hoeben and Peter A Troch. Assimilation of active microwave measurements for soil moisture profile retrieval under laboratory conditions. In *IGARSS 2000. IEEE 2000 International Geoscience and Remote Sensing Symposium. Taking the Pulse of the Planet: The Role of Remote Sensing in Managing the Environment. Proceedings (Cat. No. 00CH37120)*, volume 3, pages 1271–1273. IEEE, 2000.

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

[82] Albrecht H Weerts and Ghada YH El Serafy. Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models. *Water Resources Research*, 42(9), 2006.

[83] 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.

[84] Shufen Sun, Jiming Jin, and Yongkang Xue. A simple snow-atmosphere-soil transfer model. *Journal of Geophysical Research: Atmospheres*, 104(D16):19587–19597, 1999.

[85] Hans Lievens, Rolf H Reichle, Qing Liu, Gabrielle JM De Lannoy, R Scott Dunbar, SB Kim, Narendra N Das, M Cosh, Jeffrey P Walker, and Wolfgang Wagner. Joint Sentinel-1 and SMAP data assimilation to improve soil moisture estimates. *Geophysical Research Letters*, 44(12):6145–6153, 2017.

[86] Sujay V Kumar, Michael Jasinski, David M Mocko, Matthew Rodell, Jordan Borak, Bailing Li, Hiroko Kato Beaudoing, and Christa D Peters-Lidard. NCA-LDAS land analysis: Development and performance of a multisensor, multivariate land data assimilation system for the National Climate Assessment. *Journal of Hydrometeorology*, 20(8):1571–1593, 2019.

[87] Giovanni Macelloni, Marco Brogioni, Francesco Montomoli, Juha Lemmetyinen, Jouni Pulliainen, and Helmut Rott. Retrieval of snow water equivalent in forested area using multifrequency SAR data. In *EUSAR 2014; 10th European Conference on Synthetic Aperture Radar*, pages 1–3. VDE, 2014.

[88] Juha Lemmetyinen, Jouni Pulliainen, Anna Kontu, Andreas Wiesmann, Christian Maetzler, Helmut Rott, Karl Voglmeier, Thomas Nagler, Adriano Meta, Alex Coccia, et al. Observations of seasonal snow cover at X and Ku bands during the NoSREx campaign. In *EUSAR 2014; 10th European Conference on Synthetic Aperture Radar*, pages 1–4. VDE, 2014.

[89] Jianxiu Qiu, Wade T Crow, Grey S Nearing, Xingguo Mo, and Suxia Liu. The impact of vertical measurement depth on the information content of soil moisture times series data. *Geophysical Research Letters*, 41(14):4997–5004, 2014.

[90] Jacqueline Le Moigne, Philip Dabney, Olivier de Weck, Veronica Foreman, Paul Grogan, Matthew Holland, Steven Hughes, and Sreeja Nag. Tradespace analysis tool for designing constellations (tat-c). In *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 1181–1184. IEEE, 2017.

[91] James Foster, Son Nghiem Marco Tedesco, George Riggs, Dorothy Hall, and John Eylander. A global snowmelt product using visible, passive microwave and scatterometer satellite data. *The Int. Arch. Photogramm., Remote Sens. and Spatial Inf. Sci., Commission VIII, WG*, 8(8):B8, 2008.

[92] Zbigniew W Kundzewicz, LJ Mata, Nigel William Arnell, P Döll, B Jimenez, K Miller, T Oki, Z Şen, and I Shiklomanov. The implications of projected climate change for freshwater resources and their management. *Hydrological Sciences Journal*, 53(1):3–10, 2008.

[93] Matthew Sturm, Michael A Goldstein, and Charles Parr. Water and life from snow: A trillion dollar science question. *Water Resources Research*, 53(5):3534–3544, 2017.

[94] Tamlin M. Pavelsky, Stefan Sobolowsk, Sarah B. Kapnick, and Jason B. Barnes. Changes in orographic precipitation patterns caused by a shift from snow to rain. *Geophysical Research Letters*, 39(17):1–6, 2012.

[95] Chad W Thackeray and Christopher G Fletcher. Snow albedo feedback: Current knowledge, importance, outstanding issues and future directions. *Progress in Physical Geography*, 40(3):392–408, 2016.

[96] Sonja Wipf, Veronika Stoeckli, and Peter Bebi. Winter climate change in Alpine Tundra: plant responses to changes in snow depth and snowmelt timing. *Climatic Change*, 94(1-2):105–121, 2009.

[97] Iris T Stewart, Daniel R Cayan, and Michael D Dettinger. Changes in snowmelt runoff timing in western North America under a "business as usual" climate change scenario. *Climatic Change*, 62(1-3):217–232, 2004.

[98] Andreas Juergen Dietz, Claudia Kuenzer, and Christopher Conrad. Snow-cover variability in central Asia between 2000 and 2011 derived from improved MODIS daily snow-cover products. *International Journal of Remote Sensing*, 34(11):3879–3902, 2013.

[99] M F Wehner, J R Arnold, T Knutson, K E Kunkel, and A N LeGrande. Droughts, floods, and wildfires. *Climate Science Special Report: Fourth National Climate Assessment*, I, 2017.

[100] Ross Brown, Dominique Tapsoba, and Chris Derksen. Evaluation of snow water equivalent datasets over the Saint-Maurice river basin region of southern Québec. *Hydrological Processes*, 32(17):2748–2764, 2018.

[101] Fanny Larue, Alain Royer, Danielle De Sève, Alexandre Langlois, Alexandre Roy, and Ludovic Brucker. Validation of GlobSnow-2 snow water equivalent over Eastern Canada. *Remote Sensing of Environment*, 194:264–277, 2017.

[102] Peter Ashcroft and Frank J Wentz. AMSR level 2A algorithm. *Rep. 121599B-1, Remote Sensing System*, 2000.

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

[104] JL Foster, ATC Chang, and DK Hall. Comparison of snow mass estimates from a prototype passive microwave snow algorithm, a revised algorithm and a snow depth climatology. *Remote Sensing of Environment*, 62(2):132–142, 1997.

[105] Allan Frei, Marco Tedesco, Shihyan Lee, James Foster, Dorothy K. Hall, Richard Kelly, and David A. Robinson. A review of global satellite-derived snow products. *Advances in Space Research*, 50(8):1007–1029, 2012.

[106] Matias Takala, Jaakko Ikonen, Kari Luojus, Juha Lemmetyinen, Sari Metsämäki, Juval Cohen, Ali Nadir Arslan, and Jouni Pulliainen. New snow water equivalent processing system with improved resolution over Europe and its applications in hydrology. *IEEE*

Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10(2):428–436, 2016.

[107] Jouni Pulliainen, Kari Luojus, Chris Derksen, Lawrence Mudryk, Juha Lemmetyinen, Miia Salminen, Jaakko Ikonen, Matias Takala, Juval Cohen, Tuomo Smolander, et al. Patterns and trends of Northern Hemisphere snow mass from 1980 to 2018. *Nature*, 581(7808):294–298, 2020.

[108] C Derksen, J Lemmetyinen, J King, S Belair, C Garnaud, M Lapointe, Y Crevier, G Burbidge, and P Siqueira. A Dual-Frequency Ku-Band RADAR mission concept for seasonal snow. In *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, pages 5742–5744. IEEE, 2019.

[109] Colleen ., Lawrence Mudryk, Chris Derksen, Kari Luojus, Ross Brown, Richard Kelly, and Marco Tedesco. Evaluation of long-term Northern Hemisphere snow water equivalent products. *The Cryosphere*, 14(5):1579–1594, 2020.

[110] Carrie M Vuyovich, Jennifer M Jacobs, and Steven F Daly. Comparison of passive microwave and modeled estimates of total watershed SWE in the continental United States. *Water Resources Research*, 50(11):9088–9102, 2014.

[111] 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.

[112] Vasco Conde, Giovanni Nico, Joao Catalao, Anna Kontu, and Maria Gritsevich. Wide-area mapping of snow water equivalent by Sentinel-1&2 data. In *EGU General Assembly Conference Abstracts*, page 9580, 2017.

[113] Giovanni Macelloni, Marco Brogioni, Francesco Montomoli, Juha Lemmetyinen, Jouni Pulliainen, Helmut Rott, Karl Voglmeier, Irena Hajnsek, Rolf Scheiber, and Björn Rommen. On the synergic use of Sentinel-1 and CoreH2O SAR data for the retrieval of snow water equivalent on land and glaciers. In *Living Planet Symposium*, September 2013.

[114] Hans Lievens, Matthias Demuzere, Hans-Peter Marshall, Rolf H Reichle, Ludovic Brucker, Isis Brangers, Patricia de Rosnay, Marie Dumont, Manuela Girotto, Walter W Immerzeel, et al. Snow depth variability in the Northern Hemisphere mountains observed from space. *Nature Communications*, 10(1):1–12, 2019.

[115] Edward Kim, Charles Gatebe, Dorothy Hall, Jerry Newlin, Amy Misakonis, Kelly Elder, Hans Peter Marshall, Chris Hiemstra, Ludovic Brucker, Eugenia De Marco, et al. NASA's SnowEx campaign: Observing seasonal snow in a forested environment. In *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 1388–1390. IEEE, 2017.

[116] C Derksen, J King, S Belair, C Garnaud, V Vionnet, V Fortin, J Lemmetyinen, Y Crevier, P Plourde, B Lawrence, et al. Development of the terrestrial snow mass mission. In *2021*

IEEE International Geoscience and Remote Sensing Symposium IGARSS, pages 614–617. IEEE, 2021.

[117] Sheldon D. Drobot and David G. Barber. Towards development of a snow water equivalence (SWE) algorithm using microwave radiometry over snow covered first-year sea ice. *Photogrammetric Engineering and Remote Sensing*, 64(5):415–423, 1998.

[118] Dongyue Li, Michael Durand, and Steven A Margulis. Potential for hydrologic characterization of deep mountain snowpack via passive microwave remote sensing in the Kern River basin, Sierra Nevada, USA. *Remote Sensing of Environment*, 125:34–48, 2012.

[119] Jinjun Tong, Stephen J Dery, Peter L Jackson, and Chris Derksen. Testing snow water equivalent retrieval algorithms for passive microwave remote sensing in an Alpine watershed of Western Canada. *Canadian Journal of Remote Sensing*, 36(sup1):S74–S86, 2010.

[120] Nastaran Saberi, Richard Kelly, Margot Flemming, and Qinghuan Li. Review of snow water equivalent retrieval methods using spaceborne passive microwave radiometry. *International Journal of Remote Sensing*, 41(3):996–1018, 2020.

[121] Matias Takala, Kari Luojus, Jouni Pulliainen, Chris Derksen, Juha Lemmetyinen, Juha-Petri Kärnä, Jarkko Koskinen, and Bojan Bojkov. Estimating northern hemisphere snow water equivalent for climate research through assimilation of space-borne radiometer data and ground-based measurements. *Remote Sensing of Environment*, 115(12):3517–3529, 2011.

[122] R Kwok, G Cunningham, T Markus, D Hancock, JH Morison, SP Palm, SL Farrell, and A Ivanoff. ATLAS/ICESat-2 L3A sea ice freeboard, version 3. *Boulder, Colorado USA: NSIDC: National Snow and Ice Data Center.*, 10:5067, 2020.

[123] Shunlin Liang. *Comprehensive Remote Sensing*. Elsevier, 2017.

[124] AM Sayer, NC Hsu, and C Bettenhausen. Implications of MODIS bow-tie distortion on aerosol optical depth retrievals, and techniques for mitigation. *Atmospheric Measurement Techniques*, 8(12):5277–5288, 2015.

[125] Kathryn J Bormann, Seth Westra, Jason P Evans, and Matthew F McCabe. Spatial and temporal variability in seasonal snow density. *Journal of Hydrology*, 484:63–73, 2013.

[126] Matthew Sturm and Carl Benson. Scales of spatial heterogeneity for perennial and seasonal snow layers. *Annals of Glaciology*, 38:253–260, 2004.

[127] Sreeja Nag, Steven P. Hughes, and Jacqueline Le Moigne. Streamlining the design tradespace for Earth imaging constellations. In *AIAA SPACE 2016*. 2016.

[128] Taikan Oki, Keiji Imaoka, and Misako Kachi. AMSR instruments on GCOM-W1/2: Concepts and applications. In *2010 IEEE International Geoscience and Remote Sensing Symposium*, pages 1363–1366. IEEE, 2010.

[129] Evert Attema, Pierre Bargellini, Peter Edwards, Guido Levrini, Svein Lokas, Ludwig Moeller, Betlem Rosich-Tell, Patrizia Secchi, Ramon Torres, Malcolm Davidson, et al. Sentinel-1 - the RADAR mission for GMES operational land and sea services. *ESA Bulletin*, 131:10–17, 2007.

[130] Waleed Abdalati, H Jay Zwally, Robert Bindschadler, Bea Csatho, Sinead Louise Farrell, Helen Amanda Fricker, David Harding, Ronald Kwok, Michael Lefsky, Thorsten Markus, et al. The ICESat-2 laser altimetry mission. *Proceedings of the IEEE*, 98(5):735–751, 2010.

[131] Kelly M Brunt, Thomas A Neumann, Kaitlin M Walsh, and Thorsten Markus. Determination of local slope on the Greenland Ice Sheet using a multibeam photon-counting LiDAR in preparation for the ICESat-2 mission. *IEEE Geoscience and Remote Sensing Letters*, 11(5):935–939, 2013.

[132] Wenlu Qi and Ralph O Dubayah. Combining TanDEM-X InSAR and simulated GEDI LiDAR observations for forest structure mapping. *Remote sensing of Environment*, 187:253–266, 2016.

[133] Sean R Helfrich, Donna McNamara, Bruce H Ramsay, Thomas Baldwin, and Tim Kasheta. Enhancements to, and forthcoming developments in the Interactive Multisensor Snow and Ice Mapping System (IMS). *Hydrological Processes: An International Journal*, 21(12):1576–1586, 2007.

[134] David A Robinson and Allan Frei. Seasonal variability of Northern Hemisphere snow extent using visible satellite data. *The Professional Geographer*, 52(2):307–315, 2000.

[135] JL Foster, DK Hall, REJ Kelly, and L Chiu. Seasonal snow extent and snow mass in South America using SMMR and SSM/I passive microwave data (1979–2006). *Remote Sensing of Environment*, 113(2):291–305, 2009.

[136] EF Vermote, SY Kotchenova, and JP Ray. MODIS surface reflectance user's guide. *MODIS Land Surface Reflectance Science Computing Facility, version*, 1, 2011.

[137] Ngar-Cheung Lau and Mark W Crane. Comparing satellite and surface observations of cloud patterns in synoptic-scale circulation systems. *Monthly Weather Review*, 125(12):3172–3189, 1997.

[138] Matthew Sturm, Brian Taras, Glen E Liston, Chris Derksen, Tobias Jonas, and Jon Lea. Estimating snow water equivalent using snow depth data and climate classes. *Journal of Hydrometeorology*, 11(6):1380–1394, 2010.

[139] Liu Jiuliang and Li Zhen. Temporal series analysis of snow water equivalent of satellite passive microwave data in northern seasonal snow classes (1978–2010). In *2013 IEEE International Geoscience and Remote Sensing Symposium-IGARSS*, pages 3606–3609. IEEE, 2013.

[140] Yonghwan Kwon, Yeosang Yoon, Barton A Forman, Sujay Kumar, and Lizhao Wang. Synthetic study of spaceborne LiDAR snow depth retrieval assimilation within the nasa land information system. In *AGU Fall Meeting Abstracts*, 2018.

[141] Barton A. Forman and Steven A. Margulis. Assimilation of multiresolution radiation products into a downwelling surface radiation model: 1. prior ensemble implementation. *Journal of Geophysical Research: Atmospheres*, 115(D22), 2010.

[142] R Schmidt, Svetozar Petrovic, Andreas Güntner, Franz Barthelmes, J Wünsch, and Jürgen Kusche. Periodic components of water storage changes from GRACE and global hydrology models. *Journal of Geophysical Research: Solid Earth*, 113(B8), 2008.

[143] Andreas Güntner, Jochen Stuck, Susanna Werth, Petra Döll, Kerstin Verzano, and Bruno Merz. A global analysis of temporal and spatial variations in continental water storage. *Water Resources Research*, 43(5), 2007.

[144] Guo-Yue Niu, Zong-Liang Yang, Kenneth E Mitchell, Fei Chen, Michael B Ek, Michael Barlage, Anil Kumar, Kevin Manning, Dev Niyogi, Enrique Rosero, et al. The community noah land surface model with multiparameterization options (noah-mp): 1. model description and evaluation with local-scale measurements. *Journal of Geophysical Research: Atmospheres*, 116(D12), 2011.

[145] 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.

[146] Hamid Moradkhani. Hydrologic remote sensing and land surface data assimilation. *Sensors*, 8(5):2986–3004, 2008.

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

[148] Ross N Hoffman and Robert Atlas. Future observing system simulation experiments. *Bulletin of the American Meteorological Society*, 97(9):1601–1616, 2016.

[149] Michiko Masutani, John S Woollen, Stephen J Lord, G David Emmitt, Thomas J Kleespies, Sidney A Wood, Steven Greco, Haibing Sun, Joseph Terry, Vaishali Kapoor, et al. Observing system simulation experiments at the national centers for environmental prediction. *Journal of Geophysical Research: Atmospheres*, 115(D7), 2010.

[150] Lajiao Chen and Lizhe Wang. Recent advance in earth observation big data for hydrology. *Big Earth Data*, 2(1):86–107, 2018.

[151] Martin J Unwin, Nazzareno Pierdicca, Estel Cardellach, Kimmo Rautiainen, Giuseppe Foti, Paul Blunt, Leila Guerriero, Emanuele Santi, and Michel Tossaint. An introduction to the HydroGNSS GNSS reflectometry remote sensing mission. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14:6987–6999, 2021.

[152] M Khaki, H-J Hendricks Franssen, and SC Han. Multi-mission satellite remote sensing data for improving land hydrological models via data assimilation. *Scientific Reports*, 10(1):1–23, 2020.

[153] Marc Leisenring and Hamid Moradkhani. Snow water equivalent prediction using Bayesian data assimilation methods. *Stochastic Environmental Research and Risk Assessment*, 25(2):253–270, 2011.

[154] Hongxiang Yan, Mahkameh Zarekarizi, and Hamid Moradkhani. Toward improving drought monitoring using the remotely sensed soil moisture assimilation: A parallel particle filtering framework. *Remote Sensing of Environment*, 216:456–471, 2018.

[155] Anouk I Gevaert, Luigi J Renzullo, Albert IJM Van Dijk, Hans J Van Der Woerd, Albrecht H Weerts, and Richard AM De Jeu. Joint assimilation of soil moisture retrieved from multiple passive microwave frequencies increases robustness of soil moisture state estimation. *Hydrology and Earth System Sciences*, 22(9):4605–4619, 2018.

[156] Mel Griffiths and Lynnell Rubright. *Colorado: a geography*. Routledge, 2019.

[157] Randall S Cerveny. The western san juan mountains: Their geology, ecology, and human history. *American Scientist*, 85(3):281–282, 1997.

[158] Sujay V Kumar, Christa D Peters-Lidard, Yudong Tian, Paul R Houser, James Geiger, S Olden, L Lighty, Joseph L Eastman, B Doty, P Dirmeyer, et al. Land information system: An interoperable framework for high resolution land surface modeling. *Environmental Modelling & Software*, 21(10):1402–1415, 2006.

[159] Karen I Mohr, Wei-Kuo Tao, Jiun-Dar Chern, Sujay V Kumar, and Christa D Peters-Lidard. The NASA-goddard multi-scale modeling framework–land information system: global land/atmosphere interaction with resolved convection. *Environmental Modelling & Software*, 39:103–115, 2013.

[160] Robert E Dickinson. Land surface processes and climate—surface albedos and energy balance. In *Advances in geophysics*, volume 25, pages 305–353. Elsevier, 1983.

[161] Guo-Yue Niu and Zong-Liang Yang. Effects of vegetation canopy processes on snow surface energy and mass balances. *Journal of Geophysical Research: Atmospheres*, 109(D23), 2004.

[162] Zong-Liang Yang and Guo-Yue Niu. The versatile integrator of surface and atmosphere processes: Part 1. model description. *Global and Planetary Change*, 38(1-2):175–189, 2003.

[163] Xubin Zeng, Patrick Broxton, and Nicholas Dawson. Snowpack change from 1982 to 2016 over conterminous united states. *Geophysical Research Letters*, 45(23):12–940, 2018.

[164] Melissa L Wrzesien, Sujay Kumar, Carrie Vuyovich, Ethan D Gutmann, Rhae Sung Kim, Barton A Forman, Michael Durand, Mark S Raleigh, Ryan Webb, and Paul Houser. Development of a "nature run" for observing system simulation experiments (OSSEs) for snow mission development. *Journal of Hydrometeorology*, 2022.

[165] Brian A Cosgrove, Dag Lohmann, Kenneth E Mitchell, Paul R Houser, Eric F Wood, John C Schaake, Alan Robock, Curtis Marshall, Justin Sheffield, Qingyun Duan, et al. Real-time and retrospective forcing in the north american land data assimilation system (NLDAS) project. *Journal of Geophysical Research: Atmospheres*, 108(D22), 2003.

[166] Ronald Gelaro, Will McCarty, Max J Suárez, Ricardo Todling, Andrea Molod, Lawrence Takacs, Cynthia A Randles, Anton Darmenov, Michael G Bosilovich, Rolf Reichle, et al. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate*, 30(14):5419–5454, 2017.

[167] GJ Huffman and A Pendergrass. The climate data guide: TRMM: tropical rainfall measuring mission. *Climate Data Guide [Internet, https://climatedataguide.ucar.edu/climate-data/trmm-tropical-rainfall-measuring-mission]*, 3, 2019.

[168] 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.

[169] Purushottam Raj Singh and Thian Yew Gan. Retrieval of snow water equivalent using passive microwave brightness temperature data. *Remote Sensing of Environment*, 74(2):275–286, 2000.

[170] Thomas W Collow, Alan Robock, Jeffrey B Basara, and Bradley G Illston. Evaluation of SMOS retrievals of soil moisture over the central United States with currently available in situ observations. *Journal of Geophysical Research: Atmospheres*, 117(D9), 2012.

[171] Hylke E Beck, Ming Pan, Diego G Miralles, Rolf H Reichle, Wouter A Dorigo, Sebastian Hahn, Justin Sheffield, Lanka Karthikeyan, Gianpaolo Balsamo, Robert M Parinussa, et al. Evaluation of 18 satellite-and model-based soil moisture products using in situ measurements from 826 sensors. *Hydrology and Earth System Sciences*, 25(1):17–40, 2021.

[172] MODIS Cloud Mask Team, Steve Ackerman, Kathleen Strabala, Paul Menzel, Richard Frey, Chris Moeller, Liam Gumley, Bryan Baum, Crystal Schaaf, and George Riggs. Discriminating clear-sky from cloud with MODIS algorithm theoretical basis document (MOD35), 2010.

[173] Rod Frehlich. Errors for space-based Doppler LiDAR wind measurements: Definition, performance, and verification. *Journal of Atmospheric and Oceanic Technology*, 18(11):1749–1772, 2001.

[174] Hans Lievens, Isis Brangers, Hans-Peter Marshall, Tobias Jonas, Marc Olefs, and Gabriëlle De Lannoy. Sentinel-1 snow depth retrieval at sub-kilometer resolution over the European Alps. *The Cryosphere*, 16(1):159–177, 2022.

[175] Hao Tang, Ralph Dubayah, Anu Swatantran, Michelle Hofton, Sage Sheldon, David B Clark, and Bryan Blair. Retrieval of vertical LAI profiles over tropical rain forests using waveform LiDAR at La Selva, Costa Rica. *Remote Sensing of Environment*, 124:242–250, 2012.

[176] Augusto CV Getirana, Aaron Boone, Dai Yamazaki, Bertrand Decharme, Fabrice Papa, and Nelly Mognard. The hydrological modeling and analysis platform (HyMAP): Evaluation in the amazon basin. *Journal of Hydrometeorology*, 13(6):1641–1665, 2012.

[177] 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.

[178] Colin McLaughlin. Quantifying the added value of agile viewing relative to non-agile viewing to increase the information content of synthetic satellite retrievals. Master's thesis, University of Maryland, College Park, 2022.

[179] Xavier Emery. A turning bands program for conditional co-simulation of cross-correlated Gaussian random fields. *Computers & Geosciences*, 34(12):1850–1862, 2008.