DISSERTATION

DATA-DRIVEN IMPROVEMENTS TO GPROF-BASED SATELLITE SNOWFALL RETRIEVALS WITH A FOCUS ON MOUNTAIN SNOWFALL

Submitted by

Ryan L. Gonzalez

Department of Atmospheric Science

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2025

Doctoral Committee:

Advisor: Christian Kummerow

Glen Liston
Christine Chiu
Kristen Rasmussen
Branislav Notaros

ABSTRACT

DATA-DRIVEN IMPROVEMENTS TO GPROF-BASED SATELLITE SNOWFALL RETRIEVALS WITH A FOCUS ON MOUNTAIN SNOWFALL

Snowfall is a critical component of Earth's hydrological and climate system despite only 5% of Earth's annual precipitation falling as snow. Satellite-based snowfall estimates, particularly those obtained from the Global Precipitation Measurement (GPM) Microwave Imager (GMI), struggle to accurately estimate the total annual snowfall accumulations, especially in mountainous regions of the world. Part of the challenge is due to the reference precipitation used in the GMI-based algorithms, while radiometers struggle to distinguish between the microwave signatures of surface snowpack and snowfall. The aim of this dissertation is to evaluate the impact machine learning-based GMI retrievals have on snowfall estimates, explore how temperature and climatological adjustments to the reference precipitation can provide additional information to the retrieval, and asses if these changes lead to improved snowfall accumulations required for modeling the life-cycle of snow. A key objective of this study is to improve snowfall accumulation estimates in mountainous areas, where snowpack is a critical component of water storage.

First, snowfall rates estimated from the Goddard Profiling Algorithm (GPROF) for GMI are compared using three types of GPROF algorithms: one Bayesian (GPROF V7) and two neural network versions (GPROF-NN 1D and GPROF-NN 3D). The highest detection and quantitative statistics are observed using GPROF-NN 3D with both neural network retrieval algorithms out-performing the Bayesian version. It is shown that artificial biases in the retrieval statistics can result from the selected threshold for snow/no-snow classification. Coincident in-situ snowfall and radar data are also used to evaluate the temperature dependency of the reflectivity-snowfall (Z-S) relationship and how it impacts the GPROF retrievals.

Second, an evaluation of the three GPROF algorithms is conducted in the mountains of the western United States. Using data from a snow reanalysis dataset, water year snowfall accumulations from the Multi-Radar Multi-Sensor (MRMS) are adjusted to produce more realistic snowfall magnitudes and spatial patterns. These adjustments were found to decrease errors in snowfall accumulation estimates for all three retrieval algorithms, resulting in significant improvements when compared to independent SNOTEL observations.These results provide a positive outlook for snowfall retrievals in mountainous regions by incorporating additional information to the retrieval algorithm.

Finally, a framework for incorporating satellite precipitation estimates into a snow evolution model in the western United States is presented that offers a flexible design to account for different study domains. The objective of this framework is to present an approach for deriving snow water equivalent (SWE) from satellite precipitation estimates given the difficulties of directly measuring SWE from passive microwave sensors. A UNet-based retrieval model is used to estimate precipitation at 30 minute time resolution across the currently available passive microwave and infrared sensors. The initial precipitation estimates were found to have a systematic bias across the study period, which, after correction, produced realistic spatial patterns of snow depth and snow water equivalent, but underestimated the magnitudes compared to two reference snow model simulations.

ACKNOWLEDGMENTS

I am immensely thankful for my advisor, Professor Chris Kummerow, for his mentorship during me through Master's and Ph.D. programs. We first met in 2015 when I was selected as your summer intern by, I think, shear luck and the choice was clear when it came time to select a graduate program that I would return to the Kummerow group and CSU. I think what I appreciate the most about the Kummerow group is the wonderful people that have been a part of it, past and present, and the connection that exists between us. It takes an advisor and person such as yourself to create that kind of a group. Thank you. Thanks also to Glen Liston, Christine Chiu, Kristen Rasmussen, and Branislav Notaros for agreeing to be on my Ph.D. committee.

To my wife, Rosemary, for moving jobs and states in order for us to be closer. You then moved to Fort Collins during the initial two-week shutdown for COVID, which has now turned into almost 5 years with an almost 3 year old son. Your loving spirit has been invaluable during my Ph.D. program as well as during the writing of this dissertation. To my family for always providing me support over the course of my 20+ years of being in school. I promise I won't go back to school after this.

DEDICATION

To my Dad.

May the road rise up to meet you.
May the wind be always at your back.
May the sun shine warm upon your face,
the rains fall soft upon your fields,
and until we meet again,
may God hold you in the palm of his hand.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

Introduction

1.1 Motivation

Precipitation is a key factor in the availability and replenishment of the world's sources of freshwater. It directly feeds rivers, lakes, and reservoirs where it can be extracted for use across many aspects of human society. The availability of water is sensitive to changes in precipitation both in the short and long term with little agreement among climate models on the direction and magnitude of precipitation in the future climate (Ruiz-Barradas and Nigam 2006; Wu et al. 2022). Alternatively, many climate models are in agreement that near surface air temperature will increase in the future climate with exacerbated warming at high latitudes (Rantanen et al. 2022). The consensus on increased warming poses problems for snow dominated regions of the world.

It is estimated about 5% of the world's annual precipitation falls in the form of snow with snow dominating the total precipitation above 65 degrees latitude (Levizzani et al. 2011). Warmer temperatures imply a smaller portion of precipitation falling as snow. Decreased snow cover has been found to affect near surface air temperature (Vavrus 2007; Betts et al. 2014), which can affect large scale circulation patterns (Vavrus 2007; Xiao and Duan 2016) through radiative changes at the land-atmopshere interface (Cohen and Rind 1991). Additionally, warmer temperatures can change the seasonality of water supply by shifting the timing of snowmelt earlier in the spring or winter leading to reduced water availability during the summer. In fact, these changes have been documented in observational datasets (Stewart et al. 2005; Musselman et al. 2021). With one-sixth of the world's population, or over a billion people, depending on snowmelt to meet their freshwater needs (Barnett et al. 2005), it is important to accurately monitor snow across it's lifecycle.

In regions with seasonal snow, the snow's lifecycle consists of accumulation, evolution, and ablation (melt) phases before returning water to the environment. More often than not, the accumulation and melt phases do no exist distinctly while the evolution phase occurs at all times immediately after the snow reaches the surface. Different satellite sensors offer value during different times of the snow lifecycle. For example in Figure 1.1, the Global Precipitation Measurement

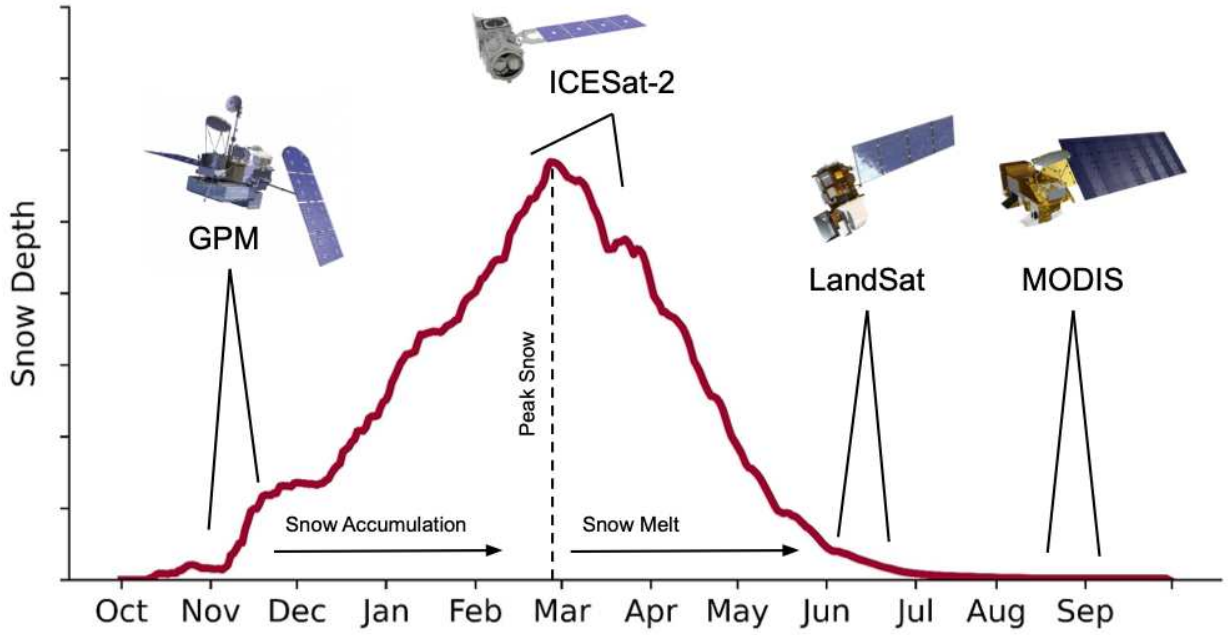

Fig. 1.1. The evolution of snow depth for a single water year (October - September), highlighting the value of different satellites throughout the snow life cycle.

(GPM, Hou et al. (2014)) mission is able to provide estimates of snowfall during the accumulation season using passive microwaves (PMW), ICESat-2 (Neumann et al. 2019) is able to measure the accumulated snow depth using laser altimetry, and LandSat/MODIS (Rittger et al. 2013) can detect the timing of snow disappearance using visible wavelengths. Together, these estimates could be used in conjunction with a snow evolution model to estimate the snow's lifecycle.

Mountain snowfall accounts for 4-5% of global snowfall (Daloz et al. 2020) with mountains able to store a larger amount of water in snow than non-mountainous regions (Wrzesien et al. 2018). Additionally, up to 70% of water in the western United States originates as snow (Li et al. 2017), highlighting the importance mountain snow has on total water storage. However, of the sensors shown in Figure 1.1, passive microwave sensors estimating snowfall currently have the largest limitation in mountain regions for multiple reasons. First, even the smallest PMW field of views at the highest frequencies, which are most sensitive to snow particles (You et al. 2017), are much too large to accurately represent snow processes in the mountains. Second, the reference snowfall dataset used in PMW retrieval algorithms, such as the Goddard Profiling Algorithm (GPROF,

Kummerow et al. (2015)), struggles to correctly capture the snowfall patterns in mountainous regions, which has a significant impact on the retrieval estimates. Third, as snow accumulates and evolves, the surface emissivity can significantly alter the scattering signature PMW are typically reliant on for snowfall detection and estimation. Finally, the interaction of radiation with snow crystals is more complex than liquid rain drops and a large range of crystal types can exist in nature that depend on local environmental conditions such as temperature, liquid water content, and ice nuclei (Nakaya 1951; Libbrecht 2007).

Therefore, this dissertation represents an attempt to improve GPROF-based snowfall retrievals using a data-driven approach. The effort focuses on improving the magnitude of snowfall estimates in mountainous terrain in the western United States. Understanding the ability of satellite snowfall estimates to reproduce snowpack characteristics is important when fitting snowfall into the larger picture of the snow's lifecycle.

1.2 Outline

This dissertation consists of three papers that are either submitted or in preparation. The focus of Chapter 2 is on the constant reflectivity-snowfall (Z-S) relationship used by the reference dataset for GPROF retrievals. In-situ data from the contiguous United States is used in combination with radar reflectivity to produce temperature dependent Z-S relationships that improve GPROF snowfall estimates. Chapter 3 focuses on GPROF snowfall estimates for complex terrain in the western United States. The reference data used in GPROF is adjusted in order to correctly capture the snow accumulation climatology across the study domain. Finally, in Chapter 4, a snow evolution model, SnowModel, is forced using satellite precipitation estimates. Two additional SnowModel simulations with different precipitation forcing data are used to assess the ability of satellite snowfall estimates to capture snow lifecycle characteristics.

CHAPTER 2

An Evaluation of GPROF-based Snowfall Retrievals and Their Training Data[1]

2.1 Introduction

Global observations of snowfall are important for a range of science and socioeconomic purposes. Falling snow accounts for 5% of Earth's annual precipitation (Levizzani et al. (2011)) and 31% of Earth's land surface is covered by seasonal snow (Vuyovich (2024)). Snowfall reaching the surface ultimately affects land-atmosphere feedbacks through energy exchanges at the land-atmosphere interface. Cohen and Rind (1991) discussed many of these exchange properties including high albedo, high thermal emissivity, low thermal conductivity, and a sink of latent heating during melting. Socioeconomically, the importance of snow resources has been estimated to be worth trillions of dollars with one-sixth of the world's population dependent upon seasonal snow (Sturm et al. (2017); Barnett et al. (2005)). Snowfall uncertainty is currently the largest source of error when it comes to modeling the life cycle of snow (Raleigh et al. (2016)). Even in ideal conditions, snowfall estimates are susceptible to a host of errors including instrument sensitivity, observational constraints, and the retrieval algorithm itself.

Accurate retrievals of snowfall from satellite passive microwave (PMW) sensors are hindered by variable land surface emissivity, complex scattering properties of ice crystals, and absorption by supercooled liquid cloud water. At the surface the snowpack exhibits large changes in emissivity as it accumulates, evolves, and melts (Mätzler (1993); Cordisco et al. (2006); Hirahara et al. (2020); Takbiri et al. (2021)), weakening the surface radiometric signal. Rahimi et al. (2022) created probability of vulnerability maps for snowfall retrievals derived from snowpack properties such as snow depth, snow grain size, and snow density. In the atmosphere, many ice crystal types exist in nature and depend on the local temperature and humidity (Nakaya (1951)). Scattering databases for such ice crystals have been produced to aid in radiative transfer computations (Liu (2008); Kuo et al. (2016)). Additionally, the ice scattering signature snowfall retrievals rely on can be completely masked by supercooled liquid water (Takbiri et al. (2019); Mroz et al. (2021)). Many

[1]Gonzalez, R., and C. Kummerow: An Evaluation of GPROF-based Snowfall Retrievals and Their Training Data. *Accepted in the Journal of Hydrometeorology*

snowfall retrieval algorithms have been created for different sensors using a range of techniques (Noh et al. (2006); Kummerow et al. (2015); Meng et al. (2017); Tang et al. (2018); Rysman et al. (2019); Pfreundschuh et al. (2022a)). Each algorithm is ultimately reliant on the assumptions made about snowfall microphysical properties; whether it is snow particle sizes, shapes, and distributions in physically based models or the reference snowfall rates derived from radar reflectivity-snowfall (Z-S) relationships (Heymsfield et al. 2017; Mai et al. 2023).

Noh et al. (2006) created a snowfall retrieval algorithm using high-frequency channels from the Advanced Microwave Sounding Unit-B (AMSU-B). The a priori database was comprised of brightness temperatures (TBs) computed from a radiative transfer model (RTM) with airborne and ground-based radar data as input. Two non-spherical (sector snowflakes and dendrite snowflakes) are assumed to be the primary snowflake types with an exponential size distribution. The conversion of radar reflectivity to snowfall rate for the a priori database was performed in two ways: using an empirically based Z-S from Aonashi et al. (2003) for ground-based radars and using the backscatter cross sections of the snowflakes along with an assumption of terminal velocity from Rutledge and Hobbs (1983) for the airborne radar; each Z-S was found to be within the range of Z-S in the literature. A Bayesian technique was used to retrieve snowfall rates with correlations between 0.6 and 0.96 compared to ground-based radars. The comparisons also indicate the retrieval works better for dry and heavy snow cases rather than wet and light snow cases.

Meng et al. (2017) developed a physically-based snowfall rate (SFR) algorithm that incorporates measurements from Advanced Microwave Sounding Unit-B (AMSU-A), Microwave Humidity Sounder (MHS), and the Advanced Technology Microwave Sounder (ATMS). The SFR algorithm is preceded by a snowfall detection algorithm using a regression-based statistical model from Kongoli et al. (2015) and numerical weather prediction (NWP) filters followed by a four step process to estimate the snowfall rate. The most important aspect of the SFR algorithm is the retrieval of cloud properties achieved using a one-dimensional (1DVAR) model followed by estimating ice particle velocities, adjusting ice water content, and finally determining the snowfall rate. The 1DVAR retrieval quantities include land surface emissivity, total precipitable water, ice water path, ice particle effective size, surface temperature, and cloud temperature, which are used

to compute particle velocity and ice water content. Errors in the 1DVAR retrieval can propagate into the determination of snowfall rate. Uncertainty analysis determined the largest sources of error to the algorithm originate from the assumption of the ice particle size distribution and ice particle shape. However, the algorithm is able to achieve correlations up to 0.6 compared to radar and gauge estimates of snowfall for several case studies throughout CONUS.

The Snow retrievaL ALgorithm fOr gMi (SLALOM) uses the Global Precipitation Measurement (GPM) Microwave Imager (GMI) to detect snowfall and estimate snow water path and surface snowfall rate (Rysman et al. (2018, 2019)). The a priori database consists of 820,000 coincident observations between CloudSat and GMI that are mainly near 60°N and 60°S. The advantage of this dataset is the ability to leverage CloudSat observations with the larger swath provided by GMI. Of seven different statistical approaches tested, gradient boosting resulted in this highest correlation coefficient and lowest root-mean-square error. SLALOM is able to achieve a correlation of 0.7, relative bias of -13%, and a root mean squared error (RMSE) of 0.08mm/hr compared to CloudSat. Notable sources of uncertainty exist in the CloudSat reference snowfall data including CloudSat not providing measurements close to the surface, underestimation (overestimation) of heavy (light) snowfall, limited database size constrained by coincident measurements, and the use of a single frequency W-band radar. The CloudSat reference snowfall is computed using profiles of retrieved snow size parameters with assumptions about the particle masses and fall speeds and, therefore, SLALOM is constrained by these assumptions.

The Goddard Profiling Algorithm (GPROF, Kummerow et al. (2015)) is a fully parametric Bayesian retrieval algorithm developed for the constellation of radiometers within the GPM mission. Continued improvements have been made to the a priori databases with the current version incorporating data from the Combined Radar/Radiometer Algorithm. Pfreundschuh et al. (2022a) introduced two neural network versions of GPROF that use the same a priori database and have identical outputs to the Bayesian version of GPROF. The authors demonstrate the ability of neural-networks to improve the accuracy of GPROF retrievals. The GPROF algorithm is used to retrieve precipitation from the suite of PMW sensors within the GPM constellation. It is an integral piece of the global satellite observation network that monitors Earth's hydrological cycle and thus is the

focus of this study. The reference snowfall rates for snow-covered land surfaces has evolved from using the dual precipitation radar (DPR) on GPM to using the Multi-Radar Multi-Sensor (MRMS) network (Zhang et al. 2016). This change removed the low bias introduced due to the 12dBZ sensitivity from the DPR. However, MRMS uses a constant Z-S relationship that is applicable for the mean state of atmospheric snow states, but is known to be unsuitable for all snow states (e.g. mountain snow or lake-effect snow).

Each of the above snowfall algorithms uses a different approach to retrieve snowfall that ultimately relies on the relationship to the ice particle size distribution. Because of the structure of the algorithm, Meng et al. (2017) was able to provide analysis on how the assumptions of different ice particle characteristics impacted the resulting retrieved snowfall. The other algorithms use a reference snowfall dataset with an assumed Z-S relationship or retrieved ice particle model without investigation into how either impacts the retrieved snowfall. Therefore the purpose of this paper is two-fold:

GPROF Snowfall Retrievals With the advent of two neural network versions of GPROF, this paper investigates the extent to which GPROF-based snowfall can leverage information provided by the a priori databases.

Z-S Temperature Dependence Using a snowfall gauge dataset, this paper derives a distribution of Z-S relationships from the MRMS data used in GPROF a priori databases

The neural network versions of GPROF offer increased complexity compared to the Bayesian version of GPROF through dynamic uncertainty estimates and the inclusion of spatial information and are expected to provide better quantitative and qualitative statistics. This analysis represents the first effort to compare snowfall estimates across the suite of GPROF algorithms. This work also investigates the response of the GPROF snowfall retrieval statistics to a change in the assumptions about the Z-S relationship used in the MRMS reference data. Coincidences between MRMS and the United States Climate Reference Network (USCRN) are used to create a distribution of snowfall rates for a given radar reflectivity, and temperature is used as a proxy for the local snow microphysics. Section 2.2 describes the creation of the GPROF retrieval databases, the GPROF

retrieval algorithms, and presents the analysis from the snowfall retrievals. An investigation into the MRMS snowfall data is presented in Section 2.3. A discussion of how a change in the reference snowfall modifies the GPROF retrievals is presented in Section 2.4 and a summary of the work is given in Section 2.5

2.2 GPROF Retrieval Benchmarks

2.2.1 GPROF Models

Three types of retrieval algorithms are used in this study: one Bayesian version and two neural network versions. The Bayesian version is the current NASA Precipitation Processing System (PPS) operational version of GPROF, or GPROF Version 7. The two neural network algorithms, recently developed by Pfreundschuh et al. (2022a) (SP22), were created with the intent to provide similar output as GPROF V7 by using the same training data, but with a more sophisticated inversion technique.

GPROF V7 uses a Bayesian method to identify the most likely precipitation rate and hydrometeor profile that is radiometrically consistent with the observations. It does so by searching an a priori database of hydrometeor profiles and computing a weighted average of the profiles based on the mean-square-difference between observed brightness temperatures (TBs) and simulated TBs in the database. The mathematics of the Bayesian method are well understood and the major sources of error arise from the creation of the a priori database and the error covariance matrices used in the Bayesian inversion. The error covariance matrix is a fixed, diagonal matrix that attempts to quantify sensor noise, calibration errors, or forward modeling errors within the database. The static covariance matrix is unable to account for all the errors in the retrieval that are likely dynamic and regime dependent. A complete description of the Bayesian GPROF model can be found in Kummerow et al. (2015).

The two neural network GPROF models are discussed in detail in SP22 and a brief description is given here. GPROF-NN 1D is a fully connected neural network that uses single pixels to predict the reference precipitation and hydrometeor profiles. GPROF-NN 3D builds upon GPROF-NN 1D by using a convolutional neural network (CNN) to include spatial information from the scene (221 pixels x 221 scans) to predict the corresponding precipitation and hydrometeor profiles. Both

models are trained for 80 epochs with initial learning rates of 0.001 and 0.0005 for GPROF-NN 1D and GPROF-NN 3D respectively. Importantly for this study, hydrometeor profiles are not retrieved and the only variable retrieved is surface snowfall.

2.2.2 Retrieval Databases

There are two retrieval databases used in this study depending on the the retrieval algorithm used. They are both comprised of similar data, but built in a slightly different manner.

Both retrieval databases used in this study are built by matching the input data with a reference precipitation rate for five water years (October 1, 2016 - September 30, 2021). In this case, the input data is made of coincident PMW observations from the Global Precipitation Measurement's (GPM) Microwave Image (GMI) and ancillary data from ERA5 reanalysis (Hersbach et al. (2020)) at the pixel spatial level. The reference precipitation data is provided by the Ground-Validated MRMS (GV-MRMS) (Kirstetter et al. 2018), which is spatially weighted to fit the 18.7 GHz channel field-of-view. All surface types are considered when creating the databases, but only pixels with snow at the surface, as defined by the NOAA Autosnow product, are evaluated.

The first retrieval database used by GPROF V7 must be condensed into self-similar clusters using ancillary data in order to make the retrieval more efficient. First, all profiles are binned by the 2m temperature, total column water vapor, air lifting index, and surface class. Next, the profiles in each bin are combined into clusters that correspond to self-similar TB vectors. After the binning and clustering procedures are performed, all clusters for each surface type are written to a single file to be used in the retrieval. The second retrieval database used by the GPROF-NN models can leverage the full observation space without the need for compression or clustering. The retrieval database for the GPROF-NN models is divided into training, validation, and test data. The test data is not used during model training and thus can be used for evaluation of GPROF V7, GPROF-NN 1D, and GPROF-NN 3D.

2.2.3 GPROF Model Comparisons

2.2.3.1 Snowfall Detection

Snowfall detection is an important first indicator of the ability of the snowfall retrieval algorithm. The following detection metrics widely used in precipitation retrieval studies are used: the

probability of detection (POD), false alarm rate (FAR), and the Heidke Skill Score (HSS). These detection metrics are defined as

$$POD = \frac{h}{h+m}, \tag{2.1}$$

$$FAR = \frac{f}{f+r}, \tag{2.2}$$

$$HSS = \frac{2(hr-fm)}{(h+m)(m+r)+(h+f)(f+r)} \tag{2.3}$$

where h, m, f, and r are the number of correct detections, missed detections, false alarms, and correct rejections. Table 2.1 shows the list of snowfall detection metrics of each GPROF algorithm for the different snow surface types in GPROF. The GPROF surface classes are intended to separate surfaces with similar surface emissivity in an attempt to provide information to the retrieval. In this way, a large range of HSS exists across snow surface types with the largest scores found for Minimum Snow surface types and GPROF-NN 3D. A monotonic increase in HSS is observed from Maximum Snow to Minimum Snow for GPROF V7 and GPROF-NN 1D. GPROF-NN 3D shows little dependence on surface type for all snow surface types. GPROF V7 has significant issues estimating snowfall for Maximum Snow and to an extent Moderate Snow likely due to the the surface snow impacting the TBs. GPROF-NN 3D has the highest HSS for all surface types and even has a higher HSS for Maximum Snow than GPROF V7 does for Minimum Snow indicating a significant increase in skill. In general there is improvement in the algorithms' ability to correctly detect snowfall as the surface snow decreases from Maximum Snow to Minimum Snow due to less scattering at the surface.

The POD for GPROF V7 and GPROF-NN 1D increases markedly between Moderate Snow and Low Snow whereas there is a smoother transition between the snow surface types in GPROF-NN 3D. GPROF-NN 3D also has noticeably smaller FAR across the snow surface types compared to the other two algorithms suggesting the spatial information included in GPROF-NN 3D plays an important factor in increasing detection capabilities and reducing false alarms. The highest POD

Table 2.1. Heidke Skill Score (HSS), probability of detection (POD), and false alarm rate (FAR) for each of the GPROF snow surface classes for GPROF V7, GPROF-NN 1D, and GPROF-NN 3D.

Surface Type	HSS (%)			POD (%)			FAR (%)		
	GPROF V7	GPROF-NN 1D	GPROF-NN 3D	GPROF V7	GPROF-NN 1D	GPROF-NN 3D	GPROF V7	GPROF-NN 1D	GPROF-NN 3D
Maximum Surface Snow	0.5	35.4	59.4	6.6	71.5	70.1	91.4	69.6	42.8
Moderate Surface Snow	7.3	37.0	56.2	4.8	48.5	54.7	66.8	65.1	37.8
Low Surface Snow	30.8	45.8	58.4	31.9	56.3	62.7	62.7	56.3	41.0
Minimum Surface Snow	34.4	51.4	60.2	34.9	63.5	67.7	58.7	52.2	41.8

occurs in the Maximum Snow surface type for the GPROF-NN models, which is opposite to what is expected given the Maximum Snow surface type is assumed to have the largest surface scattering signal that can decrease the atmospheric scattering signal.

The detection statistics shown in Table 2.1, along with comparisons to other studies, are sensitive to the threshold used to define snowing versus non-snowing. A liquid equivalent snowfall of 0.1 mm/hr is used here. Figure 2.1 shows the snowfall detection statistics as a function of the threshold used to define snowing versus non-snowing. The reference liquid equivalent snowfall of 0.1 mm/hr roughly corresponds to the maximum of HSS for GPROF V7 and GPROF-NN 3D; GPROF-NN 1D peak HSS occurs at a slightly higher snowfall rate. This could suggest GPROF-NN 1D is more sensitive to higher snowfall rates in the retrieval database compared to GPROF-NN 3D, likely due to the differences in model architecture. POD decreases monotonically with increasing snowfall rate for the GPROF-NN algorithms and a minimum in FAR is observed for a reference snowfall rate near 0.7 mm/hr. Above 0.7 mm/hr, the FAR increases sharply mainly due to a decrease in the number of correct detections. The retrieval algorithms have difficulty in retrieving high snowfall rates leading to an increase in FAR, due to these high snowfall rates being underrepresented in the training databases. An interesting artifact is observed for all snow detection statistics for GPROF V7 between 0.1-0.7 mm/hr. HSS, POD, and FAR are not as smooth as the GPROF-NN models between these snowfall rates. An investigation into the frequency of occurrence in the 2D-histogram between GPROF V7 and MRMS shows the algorithm has difficulty in estimating snowfall in this range of snowfall rates leading to similar snow detection statistics across these snowfall reference values. This artifact is likely due to the static error covariance matrices for GPROF V7 (Section 2) not optimized for these snowfall ranges.

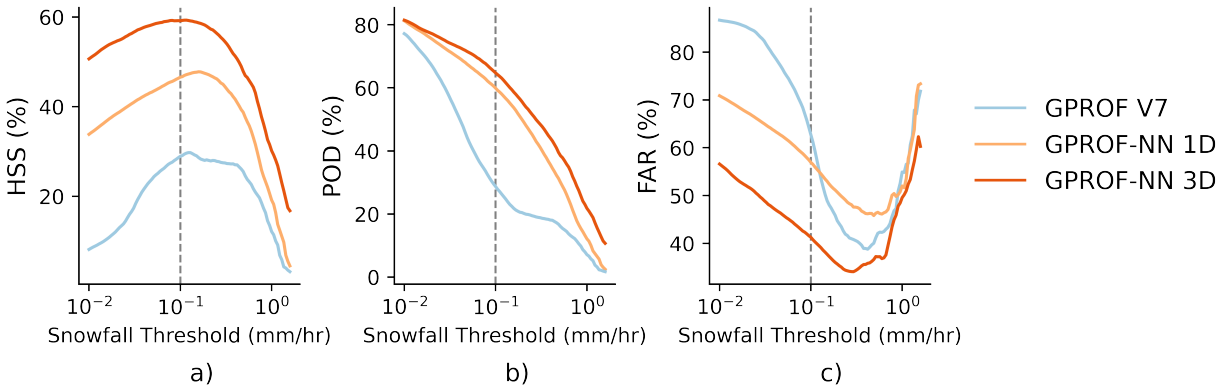

Fig. 2.1. Heidke Skill Score (HSS) (a), probability of detection (POD) (b), and false alarm rate (FAR) (c) as a function of the snowfall rate used to determine snowing versus non-snowing for GPROF V7, GPROF-NN 1D, and GPROF-NN 3D. The vertical dotted line is the threshold snowfall rate used throughout this study.

MRMS filters out radar echos lower than 5 dBz, which corresponds to an estimated liquid equivalent snowfall rate of 0.2 mm/hr. Smaller ($<$ 0.2 mm/hr) snowfall rates can exist in the retrieval database due the footprint averaging that is done to match the native MRMS data with a GMI footprint, and thus can be retrieved. The snowfall thresholds used in Figure 2.1 come from the retrieval output rather than the MRMS data.

2.2.3.2 Quantitative Snowfall Estimates

Given a common a priori database, the algorithm design can significantly contribute to differences in the retrieval outcome. Figure 2.2 shows the correlation and root mean squared error for the three different retrieval algorithms when all snowfall rates are considered. The neural network versions of GPROF have a larger correlation than GPROF V7 indicating a better comparison to MRMS across the range of snowfall rates. Additionally, the GPROF-NN algorithms perform better than GPROF V7 in terms of RMSE. GPROF-NN 3D performs the best among the the three retrieval algorithms for correlation and RMSE reinforcing that the spatial information included in the retrieval improves snowfall retrievals.

The snowfall retrieval algorithms presented here show a common pattern of overestimation of low snowfall rates and underestimation of higher snowfall rates, which creates interesting artifacts when evaluating model performance. Because the retrieval algorithms are unbiased against themselves, evaluation metrics can be skewed when using a threshold to define snowing versus

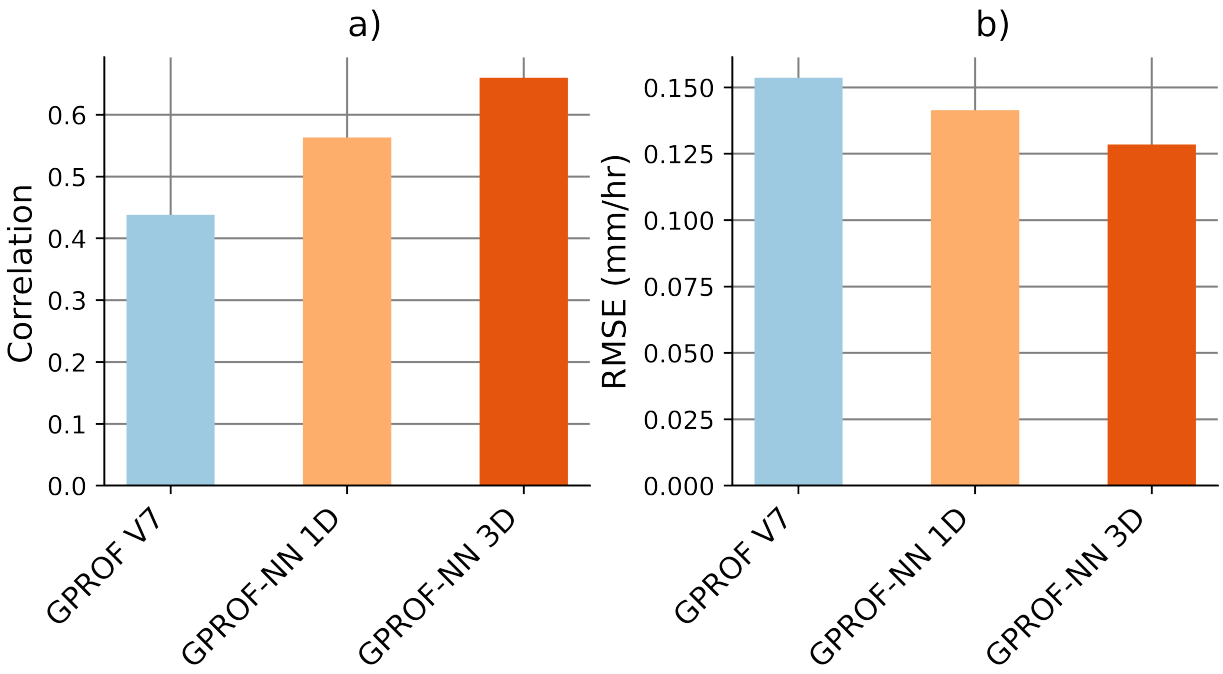

Fig. 2.2. Correlation (a) and root mean squared error (RMSE) (b) for GPROF V7, GPROF-NN 1D, and GPROF-NN 3D across all snow surface types.

non-snowing pixels. Using a threshold essentially discards a portion of the data, which introduces an artificial threshold bias that are not inherent to the retrieval algorithm. Figure 2.3 shows the thresholding bias, correlation, and RMSE as a function of the threshold used to define snowing versus non-snowing, similar to Figure 2.1. Across the tested snowfall rate thresholds, the GPROF-NN statistics are relatively smooth compared to GPROF V7. GPROF V7 exhibits a large range of thresholding bias that does not closely follow the GPROF-NN curves until snowfall rates above 0.3 mm/hr. The correlation and RMSE curves for GPROF V7 diverge from the GPROF-NN curves between 0.05 - 0.5 mm/hr which could indicate the channel error covariances used in the Bayesian inversion are not optimized for this range of snowfall rates, whereas the GPROF-NN algorithms are able to better estimate snowfall across the range of snowfall rates. The discussed range of snowfall rates does not contribute significantly to annual snowfall accumulations, but investigating the snow versus no-snow threshold is an important factor when evaluating and comparing retrieval model performance.

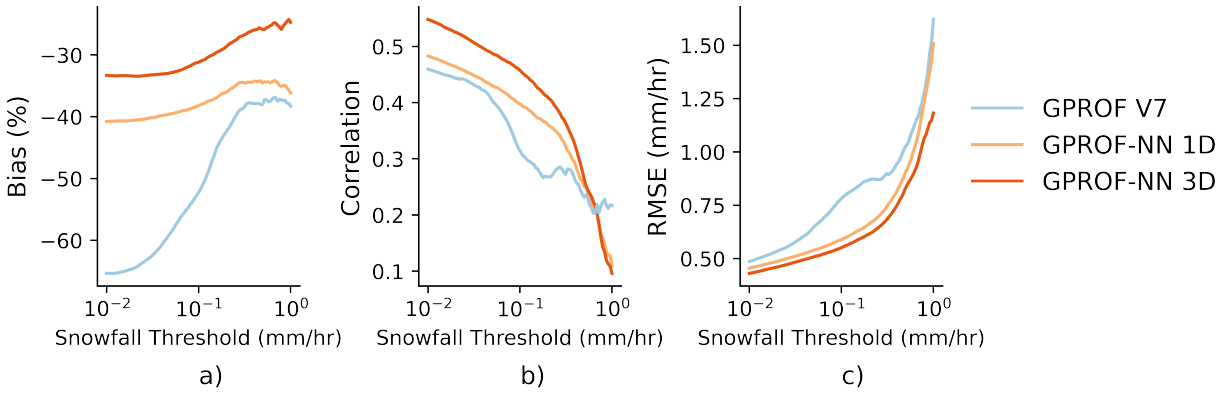

Fig. 2.3. Bias (a), correlation (b), and root mean squared error (RMSE) (c) as a function of the snowfall rate used to determine snowing versus non-snowing for GPROF V7, GPROF-NN 1D, and GPROF-NN 3D.

A similar threshold analysis of GPROF performance with respect to 2m temperature was performed for temperatures between 250-270 Kelvin. GPROF-NN 3D again outperformed the two other retrievals in terms of bias, correlation, and RMSE across all 2m temperatures and the retrieval bias and and correlation did not have a strong relationship with 2m temperature. Similar to the relationship RMSE has with snowfall rate threshold in Figure 2.3c, the RMSE increased with increasing temperature for all GPROF retrievals likely from increasing snow rates at warmer temperatures.

2.3 MRMS and In-Situ Z-S Distributions

MRMS currently utilizes a constant reflectivity (Z) to snowfall (S) relationship of $Z = 75S^2$. Measurement of the “true" snowfall is difficult, in part due to impacts from the local meteorology that can impact the vertical structure of the snowing cloud and the snowfall at the surface. This section uses in-situ snowfall measurements from the United States Climate Reference Network (USCRN) and MRMS to create distributions of Z-S relationships, similar to Kirstetter et al. (2015). Because MRMS uses a constant Z-S relationship, the USCRN data is used to provide reference snowfall estimates to the MRMS reflectivity data. This ultimately allows for a range of Z-S relationships to be computed. Additionally, an investigation into how temperature affects the computed Z-S relationships is conducted. Figure 2.4 shows the MRMS and USCRN site locations used for this analysis.

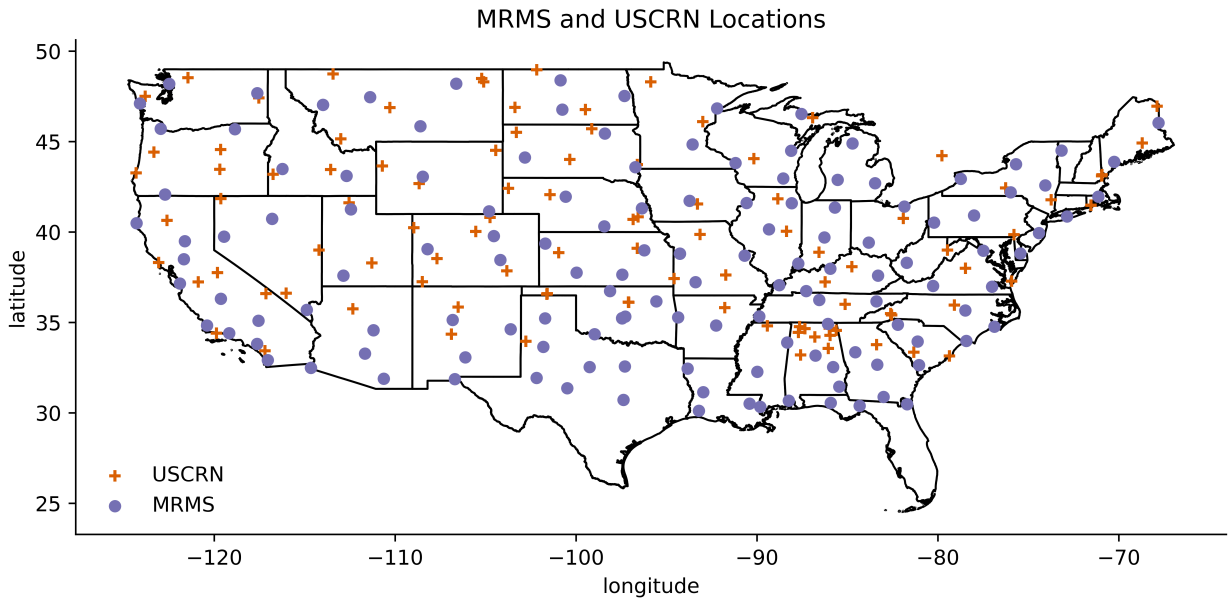

Fig. 2.4. Locations of MRMS radars (purple circles) and USCRN station (orange crosses) locations used to create the coincident database.

2.3.1 MRMS

The MRMS product integrates about 180 operational radars to produce precipitation estimates for the contiguous United States (CONUS) and portions of southern Canada. Importantly for this study, precipitation type is defined by the surface wet-bulb temperature from the High-Resolution Rapid Refresh (HRRR) model. A surface temperature of less than 2°C and a wet-bulb temperature of less than 0°C indicates snowfall is possible at the surface. The MRMS system estimates snow through a constant Z-S relationship that is known to not be generalizable to all snowing scenes. The MRMS radar quality index (RQI) attempts to account for sources of uncertainties in the radar estimates due to beam height, beam blockages, and beam position with respect to the freezing level.

2.3.2 USCRN

The United States Reference Network (USCRN) was designed in order to provide high-quality climate data (Diamond et al. (2013)). The network was conceived to reduce uncertainty in long-term in-situ data records that arise from station location, instrumentation, metadata, and operators, among others. A total of 114 USCRN stations exist throughout the contiguous United States and

careful consideration is taken to place stations in as open an area as possible to reduce data contamination. Air temperature, precipitation, and soil moisture are the primary environmental variables collected by each station. Three independent vibrating-wire weighing transducers are used to measure the weight of the precipitation that has fallen into an all-weather weighing precipitation gauge along with a tipping-bucket precipitation gauge for redundancy. Importantly for snowfall, an Alter shield and Double Fence Intercomparison Reference windscreen are used at each station to reduce the effect of wind on precipitation collection. While snowfall is not a primary product output by the USCRN, the configuration and instrumentation of each station is well-suited to measure liquid equivalent snowfall. Therefore, we consider the USCRN data to be suitable for this study.

2.3.3 MRMS and USCRN Matching

In order to create a database of collocations, the MRMS and USCRN data products are first matched in time and space. One year (2021) of data is matched at the hourly timescale set by the USCRN data. The hourly liquid equivalent precipitation is considered to be snowfall if the reported hourly average wet-bulb temperature is below -1°C. A wet-bulb temperature of -1°C corresponds to a 96% probability the precipitation is snowfall according to the lookup table of Sims and Liu (2015), which is used in GPROF for precipitation type determination. The reflectivity and snowfall rate of the nearest MRMS gridbox is extracted for each 2-minute timestep within the specified hour. The MRMS snowfall is then accumulated and a simple ratio is computed between the USCRN and MRMS hourly snowfall. This ratio is then retroactively applied to the 2-minute MRMS data that creates an MRMS snowfall rate that is scaled by the USCRN station. Ratios are only applied to the MRMS data if they fall between 0.01 - 100.0. This range aims to capture a large proportion of available data while removing large inconsistencies between the datasets. This methodology creates a distribution of snowfall rates for the measured reflectivity instead of a single snowfall rate for a given reflectivity (i.e. constant Z-S relationship). The MRMS RQI is included in the database to constrain the analysis to high-quality datapoints. For this study, RQI is used as a proxy for beam height that can, at high beam heights, cause the radar snowfall estimates to decouple with the surface snowfall. Thus, an RQI above 0.5 is used in this study to constrain the datapoints to relatively good quality pixels only.

The MRMS seamless hybrid scan reflectivity (SHSR) data product is used to provide the temporal variability while the USCRN hourly precipitation is used as the“truth" for the time period. Figure 2.5 shows the distribution of Z-S relationships produced from the database of collocations between USCRN and MRMS. As expected, the snowfall rate increases with increasing reflectivity. The corresponding percentiles show a similar behavior, but the interquartile range (IQR) increases with increasing reflectivity indicating larger uncertainty in the Z-S reflectivity at higher reflectivity. The MRMS Z-S shows conditional biases at all reflectivity levels, but more closely resembles the 50th percentile of the Z-S distribution for reflectivities below 12dBz and the 75th percentile of the Z-S distribution above 12dBz. This is a different result compared to Figure 3c from Kirstetter et al. (2015) likely associated with the USCRN stations being less influenced by topography. The MRMS Z-S falls within the observation space, but there is still considerable spread set by the USCRN snowfall that is consistent with other Z-S reported in literature (Kirstetter et al. 2015; Bukovčić et al. 2018).

2.3.4 Temperature Dependence of Z-S

Most Z-S relationships are not constrained by temperature even though it is understood that temperature plays a role in ice particle size distributions (Heymsfield et al. (2013)) and snow particle density (Judson and Doesken (2000); Sturm and Liston (2021)). It has been theorized there is a positive relationship between temperature and α parameter in the form $Z = \alpha S^2$ (Cooper (1986)). Using the reflectivity and snowfall rate from the database built in Section 2.3.3, the following equations are used to compute α and its standard deviation:

$$\alpha = \exp(\langle \ln Z_e \rangle - \langle \ln S^2 \rangle), \tag{2.4}$$

$$\sigma_\alpha = \sqrt{\alpha^2 \frac{\sum(\ln Z_e - \ln \alpha - \ln S^2)^2}{N^2}} \tag{2.5}$$

where Z_e is the equivalent radar reflectivity, S is the USCRN-scaled MRMS snowfall rate, and N is the total number of data points. This equation minimizes the deviation of lnZ_e from the line $lnZ_e = ln\alpha + lnS^2$ and was found to produce the the best fit to the data in Wolfe and Snider (2012)

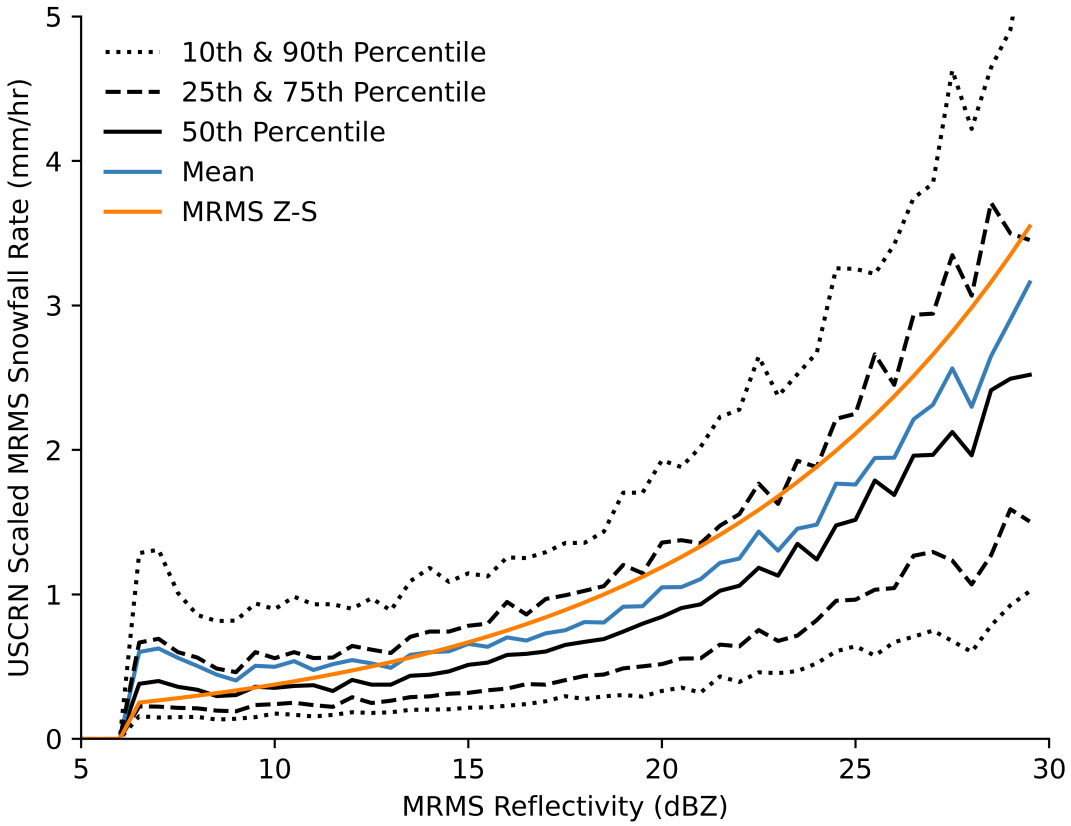

Fig. 2.5. USCRN Scaled MRMS snowfall rate distributions as a function of MRMS reflectivity (e.g. Z-S relationship). The dotted lines correspond to the 10th and 90th percentiles, the dashed lines correspond to the 25th and 75th percentiles, the solid black line corresponds to the 50th percentile, the solid blue line corresponds to the mean, and the solid orange line corresponds to the MRMS Z-S relationship $Z = 75S^2$

Figure 2.6a shows a map of the computed α parameter for each USCRN station while Figure 2.6b shows the corresponding Z-S relationship. USCRN stations with less than 30 datapoints are excluded from this analysis. As seen in Figure 2.6b, most of the Z-S relationships fall below the MRMS Z-S and there does not appear to be a dependence on station location to its corresponding Z-S. For example, no north-south gradient exists in the Z-S α as may be expected from the north-south temperature gradient during the winter. Additionally, there is no clear pattern for stations that exist within the Intermountain West region.

However, if we look at temperature ranges throughout the entirety of the database rather than individual stations, a temperature dependence on α appears. α is computed for three temperature

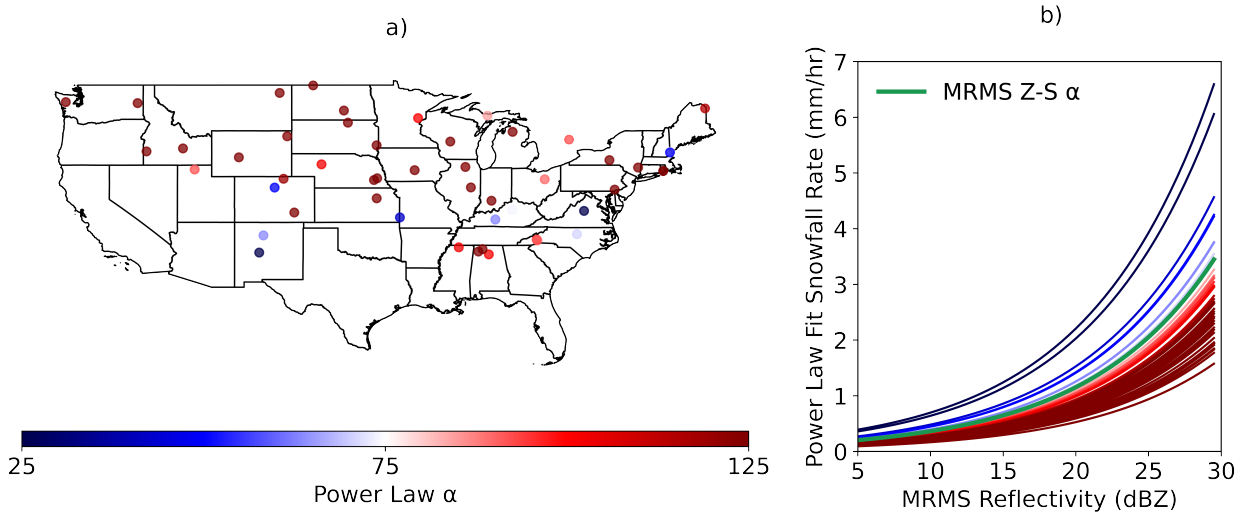

Fig. 2.6. Map of USCRN station Z-S α parameter (a) and the corresponding Z-S relationship (b). The MRMS Z-S of $Z = 75S^2$ is shown as a yellow line in (b). Only stations that have over 30 datapoints are used.

ranges: 0 to -7°C, -7 to -12°C, and -12 to -17°C. Figure 2.7 shows a clear decrease in α with a corresponding decrease in temperature that suggests a possible link between the observations and snow particle microphysics that manifests itself in the Z-S relationship. The database created here does not offer a detailed investigation into the microphysics, but sensitivity tests were performed using MRMS RQI and beam height to understand how the α-temperature relationship behaved at different thresholds. It was only for RQI thresholds above 0.5 that separation of α began to occur for the different temperature ranges. Below 0.5 RQI, the α-temperature relationship was flat likely due to noise introduced from the radar sampling characteristics when computing α. The MRMS beam height did not provide much information content to the α-temperature relationship. For all beam heights the relationship was flat, but there was a decrease in all αs with an increase in beam height likely due to the decorrelation between the reflectivity and surface snowfall.

These temperature dependent α parameters can be retroactively applied to the MRMS reflectivity to produce new snowfall rates. This process dynamically updates the Z-S used depending on the USCRN hourly temperature. Figure 2.8 shows the distribution of hourly scale factors between the MRMS snowfall and USCRN snowfall as a proxy for how the dynamic temperature Z-S relationship performs. Because the 50th percentile Z-S in Figure 2.5 is below the MRMS Z-S and a

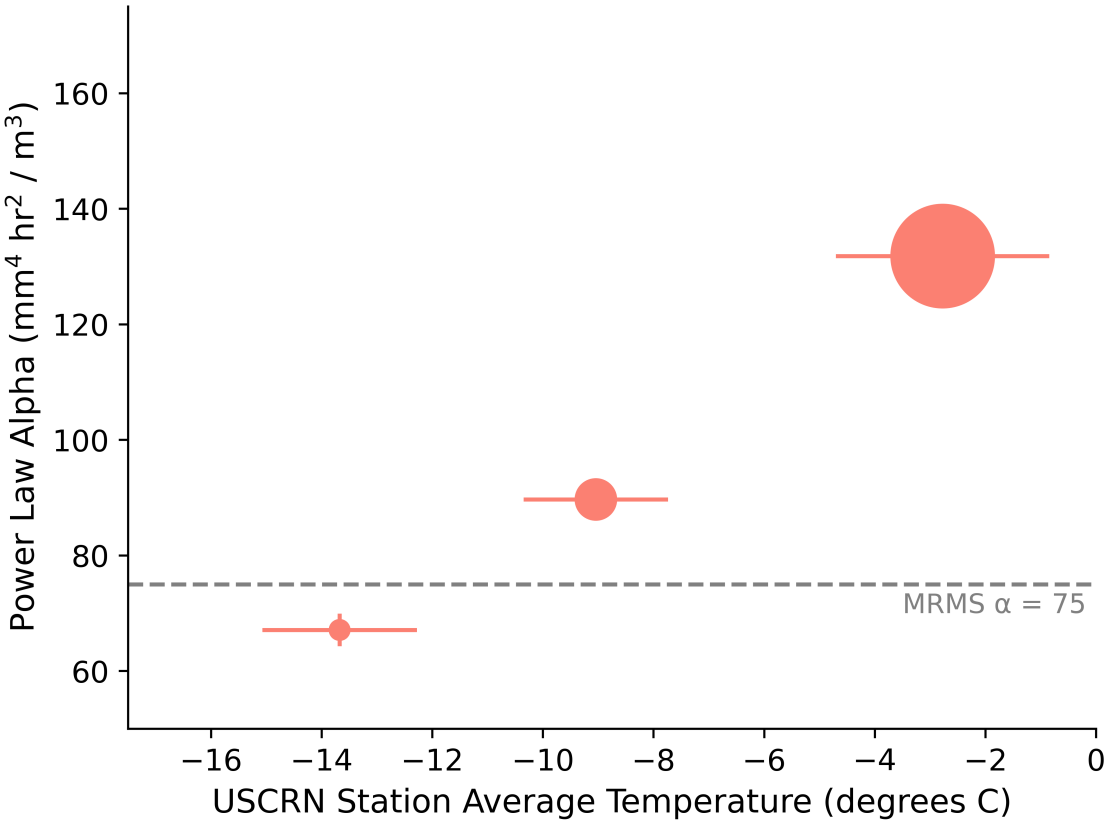

Fig. 2.7. The α from the power law $Z = \alpha S^2$ for three different temperature ranges: 0 to -7°C, -7 to -12°C, and -12 to -17°C. The size of the circles correspond to the relative amount of data within that temperature range. Horizontal (vertical) bars show the range of temperature (α uncertainty). The horizontal dotted line shows the α corresponding the operational MRMS Z-S.

large proportion of data points are above the MRMS Z-S in Figure 2.7, it is not surprising the scale factors needed to match the USCRN data are larger for the temperature dependent Z-S. In general, this property acts to decrease the snowfall rates compared to those using the nominal Z-S.

Figure 2.9 shows scatterplots of yearly snowfall accumulations for (a) the nominal MRMS Z-S relationship and (b) the temperature dependent Z-S. The colors in Figure 2.9b show the change in snowfall accumulation between the different Z-S relationships applied. There is a slight improvement in the bulk statistics in terms of MAE, SD, and CC when using the temperature dependent Z-S. The temperature dependent Z-S increased the snowfall accumulation for 51 stations and decreased the snowfall accumulation for 42 stations.

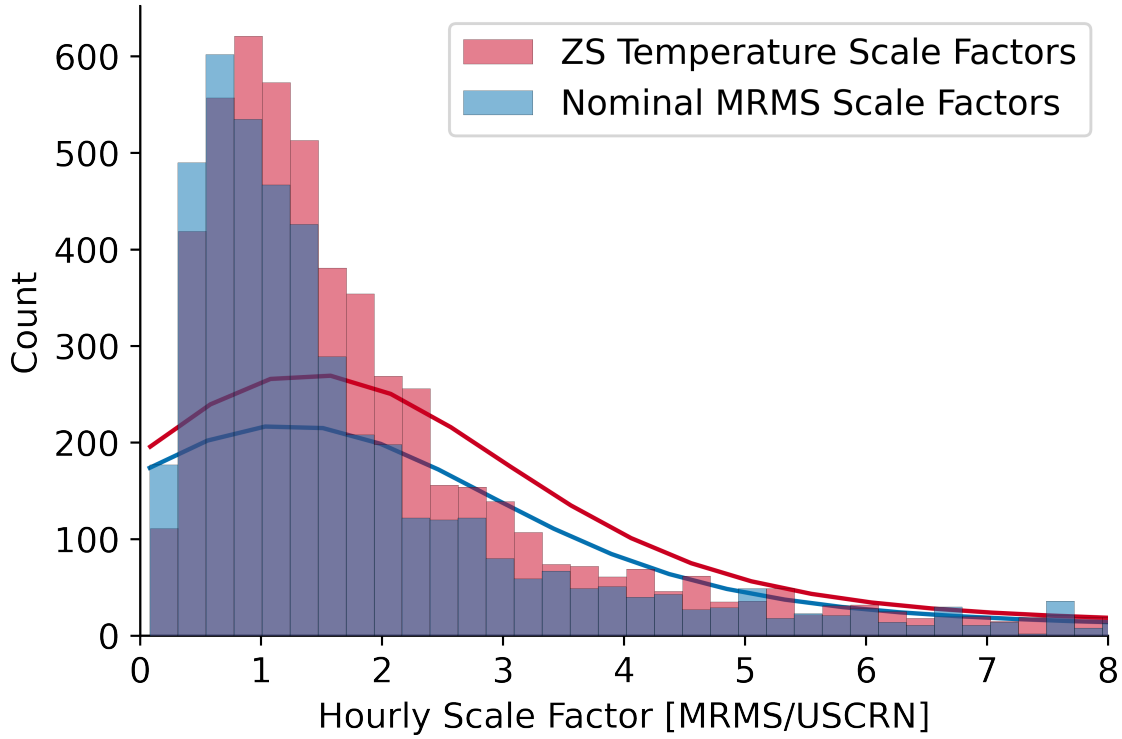

Fig. 2.8. Histograms of the hourly scale factors between hourly MRMS snowfall accumulation and hourly USCRN snowfall accumulation. The solid lines estimate the probability density function (PDF) provided by the histograms.

2.4 GPROF with Temperature Dependent Z-S

Section 2.3 showed how the USCRN and MRMS data can be used in conjunction to create a distribution of Z-S relationships and Figure 2.9 showed how a dynamic temperature dependent Z-S applied to the MRMS data can slightly improve snowfall accumulations compared to USCRN. Here we apply the temperature dependent Z-S to the MRMS data used in the construction of the GPROF retrieval databases described in Section 2.2.2. Table 2.2 shows the same snowfall detection metrics as Table 2.1. The value in parentheses is the change compared to the snowfall detection metrics shown in Table 2.1. Overall, the inclusion of a temperature dependent Z-S improved the snowfall detection metrics; improved HSS and FAR with decreased POD. The temperature dependent Z-S decreased the HSS of GPROF V7 for Moderate and Maximum Snow classes, but improved the HSS for Low and Minimum snow classes. The GPROF-NN models' HSS mostly

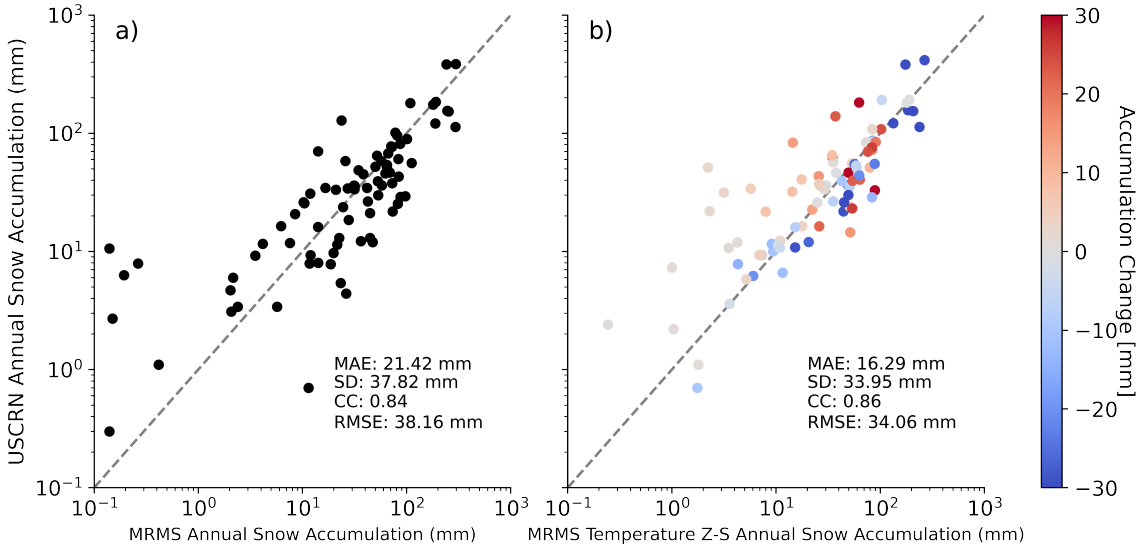

Fig. 2.9. Scatterplots of yearly snowfall accumulation between USCRN stations and MRMS. a) uses the nominal MRMS Z-S of $Z = 75S^2$ and b) uses the temperature dependent Z-S computed in Section 2.3.4. The colors in b) show the change in snowfall accumulation between the two MRMS Z-S relationships used.

Table 2.2. Heidke Skill Score (HSS), probability of detection (POD), and false alarm rate (FAR) for each of the GPROF snow surface classes for GPROF V7, GPROF-NN 1D, and GPROF-NN 3D for retrieval databases created using the temperature dependent Z-S computed in Section 2.3.4. The value in parentheses next to the snowfall detection metric shows the change of that metric compared to Table 2.1.

Surface Type	HSS (%)			POD (%)			FAR (%)		
	GPROF V7	GPROF-NN 1D	GPROF-NN 3D	GPROF V7	GPROF-NN 1D	GPROF-NN 3D	GPROF V7	GPROF-NN 1D	GPROF-NN 3D
Maximum Surface Snow	0.1 (-0.4)	36.3 (0.9)	62.2 (2.8)	9.6 (3.0)	69.5 (-2.0)	67.8 (-2.3)	91.9 (0.5)	68.5 (-1.1)	37.0 (-5.8)
Moderate Surface Snow	6.9 (-0.4)	37.0 (0.0)	56.2 (0.0)	4.3 (0.5)	45.0 (-3.5)	51.7 (-3.0)	59.7 (-7.1)	63.4 (-1.7)	33.9 (-3.9)
Low Surface Snow	34.0 (3.2)	46.6 (0.8)	59.3 (0.9)	32.7 (0.8)	54.9 (-1.4)	60.6 (-2.1)	56.8 (-5.9)	54.5 (-1.8)	37.7 (-3.3)
Minimum Surface Snow	35.4 (1.0)	52.3 (0.9)	60.3 (0.1)	33.8 (-1.1)	61.2 (-2.3)	64.9 (-2.8)	55.1 (-3.6)	49.9 (-2.3)	39.6 (-2.2)

improved for all but the Moderate Snow class where it was unchanged. The negative modification for most PODs is related to a change in both the number of correct detections and missed detections whereas the positive modification for FAR is driven mainly by a decrease in the number of false alarms.

Figure 2.10 shows the GPROF detection metrics (a-c) and quantitative metrics (d-f) as a function of snowfall similar to Figure 2.1 and Figure 2.3, but for the temperature dependent Z-S retrieval database. There is minimal change in the detection statistics with respect to snowfall rate, which is also seen in Table 2. The temperature dependent Z-S retrieval database significantly changes the

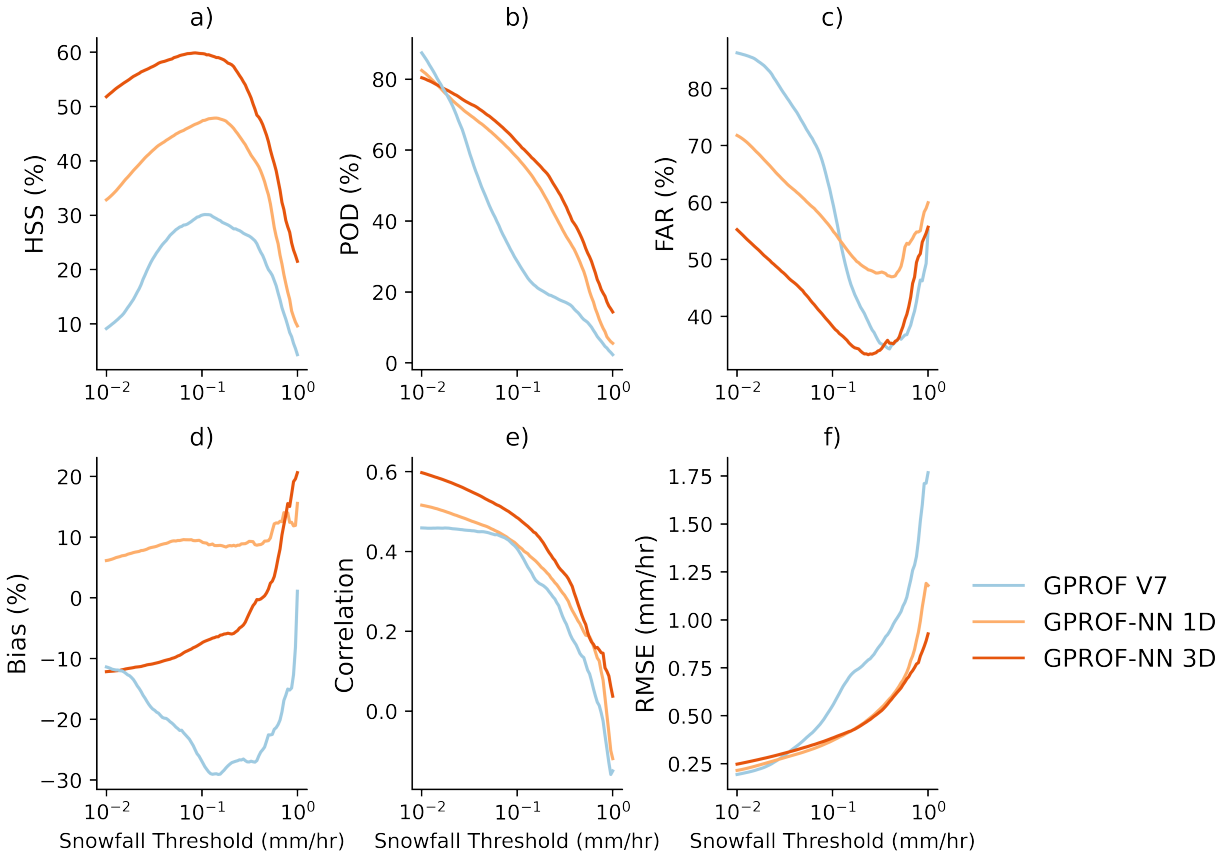

Fig. 2.10. (a) Heidke skill score, (b) probability of detection, (c) false alarm rate, (d) bias, (e) correlation, (f) RMSE for GPROF retrievals using the temperature dependent Z-S for MRMS.

biases in the GPROF algorithms with GPROF V7 showing negative biases across snowfall rates whereas the GPROF-NN 1D shows only positive biases for the snowfall rates and GPROF-NN 3D shows positive and negative biases. This ultimately changes the biases at smaller snowfall rates that are not shown here due to the minimal impact on snowfall accumulation. The largest change in quantitative metrics is the correlation of GPROF V7 across the snowfall rates which more closely follows the pattern in correlation seen in the GPROF-NN models. The temperature dependent Z-S improves the relationship between the reference snowfall and retrieved snowfall in GPROF V7, which is a notable difference compared to Figure 2.3b.

While the change in snowfall detection metrics is on the order of a few percentage points, the improvement seen is notable. The temperature dependent Z-S used in the retrieval database is able to improve detection scores likely due to the physical link between temperature and snow

microphysics noted in previous studies (Judson and Doesken (2000); Wolfe and Snider (2012); Heymsfield et al. (2013); Kirstetter et al. (2015); Sturm and Liston (2021)). Other physical links are likely present, but are difficult to uncover using the observationally generated database. An investigation into the vertical structure was performed by incorporating the ERA5 atmospheric profile into the USCRN/MRMS database, but did not provide a separation of the Z-S relationship similar to the USCRN temperature.

2.5 Conclusions

This study briefly investigated the limits of GPROF-based retrievals of snowfall over snow-covered surfaces. A common a priori database was built using 5 water year's of MRMS data matched to the footprint of GMI pixels. GPROF V7, GPROF-NN 1D, and GPROF-NN 3D were evaluated on a held out portion of the a priori database using snowfall detection and quantitative metrics. GPROF-NN 3D performed the best in terms of both metrics. It was found GPROF V7 did not contain much skill for the Maximum Snow and Moderate Snow surface types. The neural network versions of GPROF had a much better probability of detection for the same surface types. There was a clear dependence of surface class on the retrievals for GPROF V7, a slight dependence of surface class for GPROF-NN 1D, and little-to-no dependence on surface class for GPROF-NN 3D. The spatial information built into GPROF-NN 3D provides more information content to the retrieval compared to GPROF V7 and GPROF-NN 1D. Underestimation occurs across all GPROF models when using snowfall thresholds, which can be important when evaluating and comparing retrieval performance. The GPROF-NN models outperform GPROF V7 in terms of correlation and RMSE.

The algorithms' ability to estimate snowfall correctly relies on the reference snowfall used in the a priori databases. A detailed investigation into the Z-S relationship used by MRMS is performed to determine if a dynamic Z-S can be created. The United States Climate Reference Network is used to create a distribution of snowfall rates for a given MRMS reflectivity value. A ratio between the USCRN and MRMS hourly snowfall accumulation is computed and retroactively applied to the 2-minute MRMS reflectivity data. The USCRN provides the accumulation information while the MRMS reflectivity provides the spatial variability information throughout the hour. Up

to an order of magnitude of spread was found for snowfall rates for a single reflectivity value. All Z-S relationships increased with increasing reflectivity, but there did not appear to be any regional pattern in Z-S's computed for individual USCRN stations. An investigation into how temperature may affect the snowfall characteristics and resulting Z-S relationship was performed. A positive dependence between α, in $Z = \alpha S^2$, and temperature was found by separating all points within the database into temperature bins. This result follows similar findings to Wolfe and Snider (2012), but on a much larger database. The temperature dependent relationships did not produce significantly different Z-S relationships compared to MRMS that would significantly change the MRMS snowfall rates. However, when these temperature dependent Z-S relationships were applied to the MRMS data, there was slight improvement in bulk statistics compared to USCRN yearly snowfall accumulations.

The temperature dependent Z-S's were applied to the a priori database for the GPROF retrieval algorithms discussed in Section 2.2. Overall, this improved the snowfall detection Heidke Skill Scores, and false alarm rate for each of the retrieval algorithms. The probability of detection scores decreased for the GPROF-NN across all snow surface types, which is attributed to an increased number of correct detections and missed detections. Furthermore, the correlation of GPROF V7 is improved across all snowfall rates, which is a notable change compared to GPROF V7 using the nominal MRMS retrieval database in Section 2.2.32.

Despite the known drawbacks in using MRMS to build a priori databases, it remains a common dataset used to evaluate precipitation retrievals. This study shows the importance Z-S relationships have in providing retrieval databases with reference snowfall rates and how these snowfall rates can vary drastically for a given reflectivity. Uncertainties such as assumptions about the snow particles and their vertical distributions, trustworthiness of in-situ gauges, and radar beam characteristics all factor into the determination of the reference snowfall rate. Dual polarization capabilities offer some insight into snow particle behavior, but the challenge remains of how to apply these techniques across a range of meteorological conditions needed for satellite retrieval algorithms.

CHAPTER 3

Improving Satellite Snowfall Magnitudes in the Western United States Using Machine Learning and a SWE-Reanalysis[2]

3.1 Introduction

Mountains perform an essential role in Earth's water cycle by storing water in the snowpack during the cool season for use during the warm season. About 5% of Earth's annual precipitation falls in the form of snow (Levizzani et al. 2011) and 4-5% of the global snow falls over mountain regions (Daloz et al. 2020). However, mountains store a larger proportion of water as snow than non-mountainous regions (Mudryk et al. 2015; Wrzesien et al. 2018). The total amount of snowfall that reaches the surface ultimately affects the magnitude and timing of snowmelt, which has subsequent impacts for agriculture, power generation, and domestic needs. For example, it is estimated that between 60-70% of the water in the Western United States originates from snow (Li et al. 2017), highlighting the importance of monitoring snow resources. However, quantifying the amount of mountain snow across the snow life cycle remains a tremendous challenge for many communities (Bormann et al. 2018).

Snow cover extent (SCE) products derived from space-borne visible sensors are the most mature satellite snow products dating back to 1966 (Estilow 2013). In 1999, the interactive multisensor snow and ice mapping system (IMS) was developed to create daily SCE maps (Ramsay 1998). Since then, with the advancement and increase in number of space-borne sensors, many upgrades have been made to the IMS (Ramsay 2000; Helfrich et al. 2007), with the current product produced twice daily at 1km spatial resolution. The IMS product is used extensively for model snow initialization for the National Centers for Environmental Prediction (NCEP) (Meng et al. 2012; Dawson et al. 2016), the European Centre for Medium-Range Weather Forecasts (ECMWF) (Drusch et al. 2004), and the United Kingdom Met Office (Pullen et al. 2011). Proper initialization of snow states is crucial for subseasonal to seasonal (S2S) forecasts and has been proven to improve prediction skill (Lee et al. 2023). Orsolini et al. (2013) and Li et al. (2019) found improved surface

[2]Gonzalez, R., and C. Kummerow: Satellite Snowfall Magnitudes in the Intermountain West Using Machine Learning and a Western United States SWE-Reanalysis. *In review at Journal of Hydrometeorology*

temperature prediction over snow-covered surfaces within the first 10 days of the forecast when using IMS snow-cover products. Despite the importance of these products, SCE updates are only available when the measurement is cloud free. NOAA's AutoSnow product combines visible and microwave observations that allow for SCE updates when there are non-precipitating clouds in the satellite field-of-view (Lindeman 2022). AutoSnow has known difficulties with shallow snow and snow in forested areas when using the microwave observations. Comparisons between AutoSnow and IMS show good agreement in terms of binary snow cover, snow cover extent, and snow cover duration.

Snow depth (SD) and snow water equivalent (SWE), while hydrologically more important than SCE, are more challenging to monitor and estimate. Passive microwave (PMW), satellite SD and SWE algorithms have a considerable amount of uncertainty (Dawson et al. 2018; Gonzalez and Kummerow 2020) and are typically not used without some sort of augmentation from ground observations (Pulliainen et al. 2020), snow models (Josberger and Mognard 2002; Tedesco and Jeyaratnam 2016), or land surface models (De Lannoy et al. 2012). NASA's Airborne Snow Observatory (ASO) has shown promising results of estimating snow depth using airborne LIDAR (Painter et al. 2016) and recent work has highlighted the ability of space-borne laser altimetry sensors from ICESat-2 to retrieve snow depth (Lu et al. 2022; Deschamps-Berger et al. 2023; Besso et al. 2024). These altimetry snow depth estimates are typically within mountain watersheds where high resolution digital terrain models (DTMs) are available. ICESat-2 has the ability to produce global snow depth estimates, but without high resolution DTMs in snow covered regions, this technique remains limited.

Snowfall and the resulting snowpack (SCE, SD, and SWE) are inherently linked within the snow life cycle, but some observationally generated SWE datasets fail to address the physical mechanisms that constrain the SWE, namely snowfall accumulation. Liu et al. (2022), Cho et al. (2022a), and Fang et al. (2023) highlight the importance of snowfall on SWE by showing the variance of peak SWE during the accumulation season is mostly explained by snowfall accumulation. Similarly, Broxton et al. (2016a) constrain in-situ SWE estimates by snowfall accumulation before interpolating across the domain and found this linkage to significantly improve SWE estimates

compared to the Snow Data Assimilation System (SNODAS) product (National Operational Hydrologic Remote Sensing Center 2004). Liljedahl et al. (2017) shows that conventional snowfall gauges in tundra Alaska underestimate end-of-winter snowfall accumulation by 23-56% compared to collocated snow depth and density measurements. The authors highlight that errors in snowfall measurements from conventional snowfall gauges can be corrected by end-of-winter snow depth observations, which ultimately connects the instantaneous snowfall to snow accumulation. Failing to correct the snowfall errors led to a propagation of errors through the water budget and a false net-negative storage term.

PMW satellite algorithms have the ability to estimate snowfall and SWE, but the two have remained separate with no documented attempts to combine them into a single snow algorithm. Both algorithms consist of their own uncertainties and Gonzalez and Kummerow (2020) show there is a significant disagreement between the snowfall and SWE estimates from the Advanced Microwave Scanning Radiometer-EOS (AMSR-E). Even the highest standard product IMERG is unable to account for the snowfall accumulation in the Olympic Mountains (Cao et al. 2018). Currently, snow hydrologists do not consider using satellite snowfall as input forcing data to hydrological models due to their errors in snowfall accumulation. Doing so is hypothesized to provide increased information content to the retrieval algorithms. Therefore, the first step in increasing the usability of satellite snowfall estimates in the mountains is to improve the snowfall magnitudes from satellite snowfall algorithms.

Space-borne passive microwave (PMW) instruments offer the most encouraging results for long-term, global snowfall measurements (Levizzani and Cattani 2019). They are able to penetrate clouds, unlike infrared and visible sensors, and have adequate spatiotemporal sampling with their large swaths and relatively short-term repeat overpass times. PMW sensors have been widely used for snowfall detection and rate estimation (Noh et al. 2006; Kummerow et al. 2015; Meng et al. 2017) and machine-learning snowfall algorithms are becoming more widespread (Tang et al. 2018; Rysman et al. 2018; Pfreundschuh et al. 2022b). Each of these algorithms have been successful in retrieving snowfall with each limited by the specific approach used. For example, Meng et al. (2017) uses a one-dimensional variational (1DVAR) model to first estimate cloud properties to

derive the snowfall rate. The authors determined that the largest contributor to errors in snowfall rate estimation arises from assumptions about the ice particle shape and ice particle distribution. The Snow retrievaL Algorithm fOr gMi (SLALOM), developed by Rysman et al. (2018), uses a gradient boosting approach that is limited by assumptions in the CloudSat reference snowfall rate as well as biased sampling patterns driven by the polar orbit of CloudSat. Similar to SLALOM, the Goddard Profiling Algorithm (GPROF, Kummerow et al. (2015); Pfreundschuh et al. (2022b)) is dependent on the reference snowfall rate used and is the focus of this paper.

With the goal of addressing satellite snowfall quantification in the mountains, this study aims to answer the following question:

(1) Can a neural network retrieval algorithm, combined with the Western U.S. Snow Reanalysis, improve the accuracy, spatial distribution, and accumulation of retrieved snowfall?

This study uses three versions of the GPROF: GPROF V7, GPROF-NN 1D, and GPROF-NN 3D applied to observations from the Global Precipitation Measurement (GPM) Microwave Imager (GMI). The GPROF-NN algorithms are neural network versions of GPROF V7 and are outlined in Pfreundschuh et al. (2022b). Two retrieval databases are constructed: the first uses the nominal Multi-Radar Multi-Sensor (MRMS) snow water equivalent precipitation rate as the reference snowfall rate and the second incorporates snowfall accumulation climatology into the MRMS reference snowfall rate. The retrieval performance is assessed for both retrieval databases and all three retrieval algorithms against a held-out portion of the retrieval training database as well as for snowfall accumulation over the course of a water year (WY, October 1 - September 30). A case study of WY2016 snowfall accumulation is assessed along with an analysis of the temporal variability of the retrieval algorithms across five WYs. Finally, Snow Telemetry (SNOTEL) data is used as an independent verification tool to gauge the improvement of GPROF-based WY snowfall accumulations against high-quality, in-situ snow accumulation data.

3.2 Data and methods

This section introduces the MRMS and Western U.S. Snow Reanalysis (WUS-SR) datasets that are used to create the two databases used by the GPROF retrieval algorithms. All snowfall

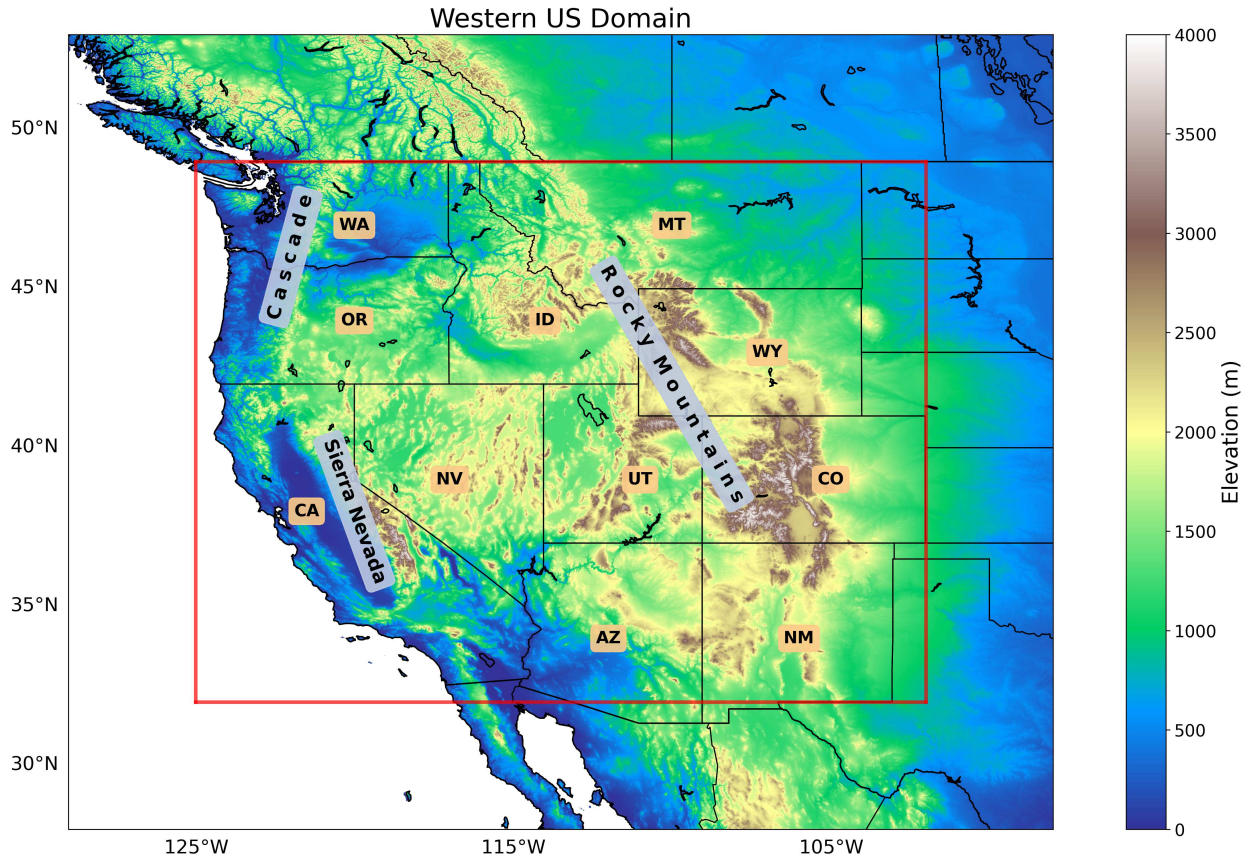

Fig. 3.1. Elevation of parts of North America. The Western United States domain used in this study is outlined in red.

discussed in this study is defined in units of liquid water equivalent. Six WY's (2017-2022) of data are processed for a region referred to as the Western U.S. (WUS; 32.0°N to 49.0°N, -102.0°W to -125.0°W) shown in Figure 3.1. This region encompasses a range of topography and snowfall regimes where MRMS is known to have difficulties in estimating snowfall.

3.2.1 MRMS

The MRMS system was created to produce precipitation products with a subfocus on hydrology (Zhang et al. 2016). It integrates about 180 operational weather radars at S-band to create gridded data for the contiguous United States (CONUS) and southern Canada. The MRMS grid has a horizontal resolution of 1 km (∼0.01°) and a time resolution of 2 minutes. Precipitation type is defined by the surface wet-bulb temperature ingested from the High-Resolution Rapid Refresh (HRRR) model with snowfall defined for pixels with a surface temperature of less than 2°C and

a wet-bulb temperature of less than 0°C. The snowfall rate is defined by a constant reflectivity-snowfall relationship of $Z = 75S^2$ where S is the snowfall rate in mm/hr. Bukovčić et al. (2018) and Kirstetter et al. (2015) discuss some of the uncertainties in this relationship and the standard product is not tuned for mountain snowfall estimates. The original MRMS snowfall rates are limited to reflectivity values above 5 dBz or a snowfall rate of 0.2 mm/hr.

The MRMS radar quality index (RQI) attempts to account for sources of uncertainties such as beam blockages and uncertainties in the empirical radar precipitation relationships. The RQI product provides and index between 0 and 1 with an RQI of 1 corresponding to the highest quality a gridpoint can have within the MRMS system. Other work utilizing the MRMS system for snowfall typically uses the RQI to constrain snowfall estimates (Wen et al. 2017; Mroz et al. 2021). Importantly for this work, all RQI values are considered in order to be able to use the MRMS snowfall estimates in the mountains within the WUS domain. An alternative quality control measure is discussed in Section 3.2.3.

3.2.2 Western U.S. Snow Reanalysis

The Western U.S. Snow Reanalysis (WUS-SR) uses a Bayesian framework to estimate snow depth, snow water equivalent (SWE), and fractional snow-covered area (fSCA) that are constrained by remotely-sensed fSCA measurements from MODIS and Landsat (Margulis et al. 2016; Fang et al. 2022, 2023). The snow reanalysis framework uses a land surface model (LSM) to generate an ensemble of prior snow states that is driven by topographic data, landcover data, and MERRA2 meteorological forcing. The precipitation, snow depletion curve, and snow albedo are perturbed to create distributions of snow states for the WY. Assimilation of Landsat and MODIS fSCA measurements provide posterior snow state estimates that update the prior ensemble weights, which indicate the likelihood the prior ensemble member resembles the fSCA measurements. The updated posterior weights combined with the prior ensemble members are used to derive posterior snow depth, SWE, and fSCA. The SWE-Reanlaysis product performs the best for regions with seasonal snow where there is a strong link between snow disappearance and fSCA. Regions with non-seasonal snow can contain errors because the additional constraint of fSCA and snow disappearance is unable to be applied.

Fang et al. (2022) performed in-situ verification of the WUS-SR SWE product with SNOTEL sites using peak SWE for each dataset. The correlation coefficient, mean difference, and root mean squared difference were found to be 0.77, -0.15m, and 0.28m, respectively. Comparisons of peak SWE against independent lidar-derived SWE via ASO showed similar results. The authors do acknowledge the lack of representative comparisons between point scale in-situ data and the grid-averaged WUS-SR data. These results provide confidence in the ability of the WUS-SR to provide proper climatologies of SWE to use for scaling the MRMS data. The WUS-SR is downsampled from its original spatial resolution of $\sim$500 m (0.004444°) to the MRMS spatial resolution using the mean of the WUS-SR gridpoints contained within the MRMS gridpoint.

3.2.3 MRMS Scaling

High radar beam heights or beam blockage can cause poor radar sampling in the mountains that leads to snowfall underestimation by MRMS. Depending on the topographic characteristics and radar location, this underestimation of instantaneous snowfall rates can lead to severe underestimation of snowfall climatologies. The WUS-SR SWE is used to correct the MRMS snowfall climatologies, which can then be used to retroactively scale the instantaneous snowfall rates estimated by MRMS to be used in the GPROF database. Specifically, snowfall from both datasets is accumulated (S_{acc}) over the course of a WY for each gridpoint and a simple ratio is computed according to Equation 3.1:

$$S_{acc\_ratio} = \frac{S_{acc\_wussr}}{S_{acc\_mrms}}, \tag{3.1}$$

where S_{acc_wussr} and S_{acc_mrms} are the snowfall accumulation for the WUS-SR and MRMS, respectively. The MRMS WY snowfall accumulation (S_{acc_mrms}) is calculated by averaging the snowfall rate (mm/hr) at each grid point and multiplying it by the total number of hours in the WY. All MRMS files corresponding to GPM overpasses within the MRMS domain are used to generate the MRMS WY snowfall accumulation, ensuring it is representative of the MRMS snowfall data used in constructing the GPROF retrieval databases. The MRMS precipitation type product is used to restrict accumulation to precipitation that is defined as snow. The WUS-SR WY snowfall accumulation (S_{acc_wussr}) is computed using inferred snowfall from the SWE data, which is defined

as a positive change in daily SWE. While there is the potential for snow ablation during each timestep that would affect the SWE value, especially outside of the mountains, this is assumed to be negligible.

Initially the unconstrained ratios spanned many orders of magnitude. A balance is needed between allowing large enough ratios to sample mountainous terrain while not producing unphysical snowfall rates. It was found that the distribution of ratios are more sensitive to WY MRMS snow accumulation than the RQI. This result is similar to Wen et al. (2017) where they found that the beam height did not constrain snowfall comparisons between MRMS and SNOTEL. Given these results, we constrained the ratios to gridpoints with a MRMS snow accumulation greater than 25 mm/WY. A threshold of 50 mm/WY was also tested, but did not significantly change the results. These constraints still allow for snowfall from different snowfall regimes to populate the retrieval databases.

Figure 3.2 shows the average WY snowfall accumulation for the WUS-SR and MRMS and the computed ratios between the two for the study period. The black circles indicate the MRMS radar locations. The WUS-SR WY average shows an expected spatial pattern of higher snowfall accumulation in the mountains and higher snowfall accumulation outside of the mountains with increasing latitude. The MRMS WY average shows some interesting patterns. In general, MRMS is able to accumulate more snow in the mountains, but there are mountain regions where the radar is unable to do so (e.g. Colorado, northwest Wyoming, and central Idaho). Additionally, MRMS hardly produces average snowfall accumulation magnitudes above $\sim$600 mm. This is likely due to terrain beam blockage and shallow, local scale orographic snowfall. Compared to the WUS-SR, it is clear MRMS is unable to reproduce WY snow magnitudes across the study domain.

Figure 3.2 also shows the computed ratio between WUS-SR and MRMS. The average ratio is 1.5 indicating an underestimation of the MRMS snowfall estimates across the domain. The radar locations and beam dynamics are evident in the ratio map. The computed ratios show the expected pattern of increasing with distance from the radar, or increasing beam height. An unexpected outcome from the ratio computation is the low ratios (WUS-SR < MRMS) closer to the radar sites. This is hypothesized to be a snow regime dependency that the single MRMS Z-S is unable to

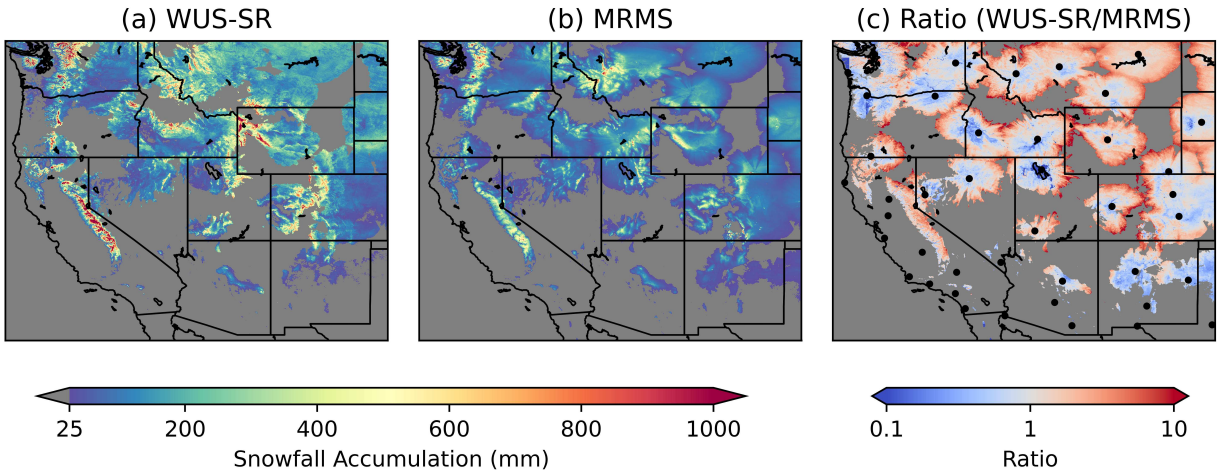

Fig. 3.2. Water year average snowfall accumulation for the Western US Snow Reanalysis and MRMS and the average accumulation ratio (WUS-SR/MRMS) for WY2017 - WY2022. Gray regions indicate where the ratios were masked out by MRMS WY snow accumulation less than 25 mm. Black circles in (c) indicate the location of the radars that are used in the MRMS product.

account for or could be related to the WUS-SR product that has better performance for seasonal snow due to the information added by observations of snow melt patterns. For example, the lower ratios are not as prevalent for radars in the northeast part of the study domain where seasonal snow is more common as compared to radars in the southeast and north west portion of the study domain.

The S_{acc_ratio} from Equation 3.1 allows for the scaling of instantaneous snowfall rates estimated by MRMS to better reproduce the snowfall climatologies across the WUS domain. Specifically, the S_{acc_ratio} is applied to each MRMS precipitation rate file within the WY for snowing gridpoints only, as defined by the MRMS precipitation type. The final WUS-SR-scaled MRMS product combines the strengths of the radar for snowfall detection with the added benefit of the snowfall climatology of the WUS-SR.

3.2.4 GPROF Algorithms

GPROF V7 is the current operational version of GPROF that uses a Bayesian framework to weight precipitation profiles in an a priori database that are radiometrically consistent with the GMI PMW observations (Team 2022; Kummerow et al. 2015). The retrieval searches self-similar retrieval databases given the attached ancillary data to the designated pixel. A fixed, diagonal error covariance matrix attempts to quantify sensor noise, calibration errors, or forward modeling errors

within the database. The static covariance matrix is known to be unable to account for all errors in the retrieval that are likely dynamic and regime dependent (Team 2022).

In addition to the Bayesian version of GPROF, two neural network versions of GPROF are used: GPROF-NN 1D and GPROF-NN 3D. A full description of the model architecture can be found in Pfreundschuh et al. (2022b). GPROF-NN 1D was designed to be similar to GPROF V7 in that single pixels are used to estimate target variable, or in this case, snowfall. GPROF-NN 3D improves on GPROF-NN 1D by incorporating spatial information into the retrieval through a convolutional neural network (CNN). In this study, unlike in Pfreundschuh et al. (2022b), we perform convolutions on the input ancillary data because spatial information in the mountains (i.e. elevation) is known to be important for snowfall accumulation. All three models have identical training, validation, and test datasets. Evaluation of the three GPROF models is performed on the test dataset that was held out during model training and is discussed in Section 3.3.1.

The retrieval database used by the GPROF algorithms is composed of PMW observations, ancillary data, and a corresponding reference snowfall data. The MRMS reference snowfall is coarsened from its original spatial resolution and spatially weighted to fit the GMI 18.7 GHz channel field-of-view. Table 3.1 provides a list of all the input variables and output retrieval variables. The 2 meter temperature, total column water vapor (TCWV), and air lifting index are either taken or derived from ERA5 (Hersbach et al. 2020), elevation data is provided by the United States Geological Survey (USGS) digital elevation model (DEM), and the surface type is defined according to Team (2022).

Table 3.1. Input and retrieval variables for GPROF models.

Input Variables	unit	Retrieval Variables	unit
13 GMI channels	K	MRMS Snowfall	mm/hr
2m temperature	K	WUS-SR Scaled MRMS Snowfall	mm/hr
TCWV	mm		
Surface type	-		
Air lifting index	-		
Elevation	m		

3.3 Results

The results of the GPROF retrievals are presented in this section. Due to the reference snowfall scaling described in Section 3.2.3, six GPROF algorithms are presented: GPROF V7, GPROF-NN 1D, and GPROF-NN 3D using the nominal MRMS retrieval database and the WUS-SR scaled MRMS database. First, an evaluation of the retrieval is performed on the held-out test dataset. Second, estimated snowfall from the retrieval algorithms is accumulated and evaluated for WY accumulations and spatiotemporal patterns. Finally, estimated snowfall is evaluated against independent Snow Telemetry (SNOTEL) sites.

3.3.1 Test dataset evaluation

The GPROF retrieval databases are comprised of pairs of input data and output data as described in Table 3.1. The input data consists of PMW observations with corresponding ancillary data and the output consists of either MRMS snowfall or WUS-SR scaled MRMS snowfall. The test dataset consists of observations from the first, second, and third day of each month from the retrieval database, but have not been used during the model training. The following discussion and comparisons evaluate the retrieved snowfall from the GPROF algorithms against the reference snowfall in the test dataset. Figure 3.3 shows 2D histograms of the retrieved snowfall versus the reference snowfall for retrievals using the nominal MRMS database (top row) and WUS-SR scaled MRMS database (bottom row). The histogram counts have been normalized by the reference snowfall in order to highlight the retrieval results across the range of snowfall rates. The following discussion, unless otherwise stated, can be applied to the GPROF retrieval algorithms for both retrieval databases.

Noticeable improvement in the accuracy of the retrieved snowfall is seen between GPROF V7 and the GPROF-NN models. GPROF V7 has little skill in estimating snowfall below $\sim$1 mm/hr as seen by the horizontal distribution of the histogram counts. Above 1 mm/hr, GPROF V7 continues to have difficulty in retrieving the reference snowfall with a majority of the pixels underestimating the reference snowfall. This result is not surprising given the complexities of retrieving mountain snowfall and the static error covariance matrix used in GPROF V7 discussed in Section 3.2.4. The GPROF-NN models are both able to estimate snowfall with more accuracy across the range

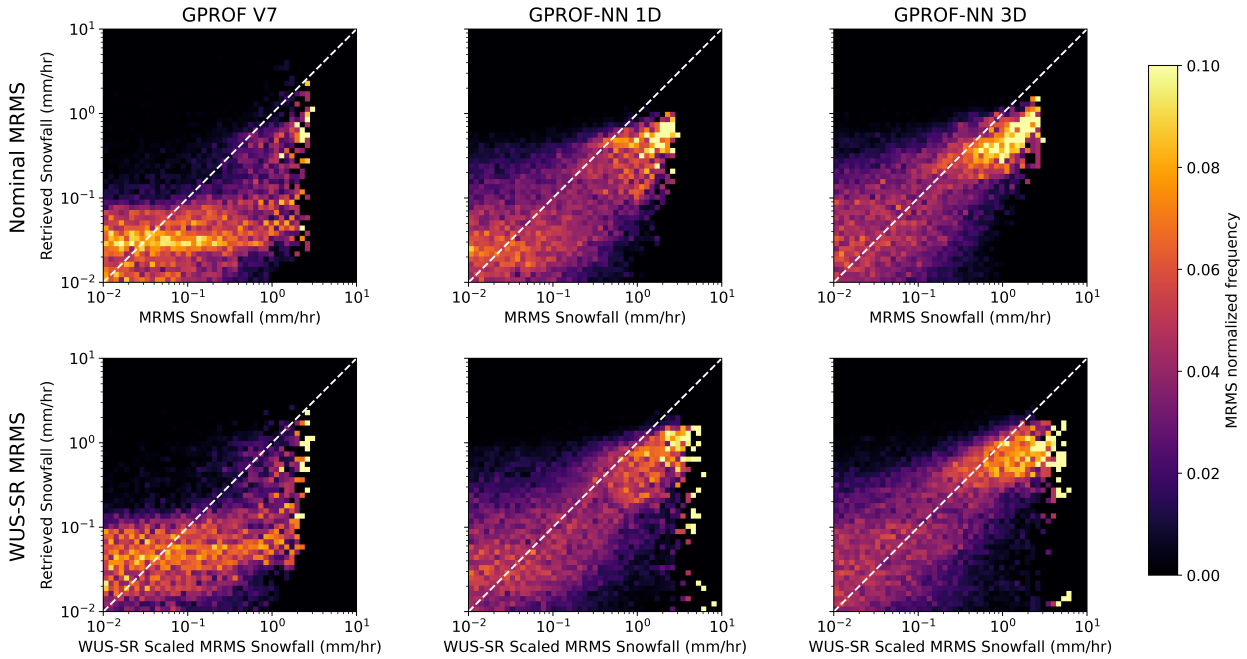

Fig. 3.3. 2D histograms of the GPROF retrieval variables: MRMS snowfall (top) and WUS-SR scaled MRMS snowfall (bottom). The columns correspond to the retrieval results from GPROF V7, GPROF-NN 1D, and GPROF-NN 3D. The histogram counts have been normalized by the reference snowfall.

of snowfall rates, and especially at the higher snowfall rates, compared to GPROF V7. This is a notable result as higher snowfall rates are more important for snowfall accumulation. GPROF-NN 3D also has more skill in the higher snowfall rates as compared to GPROF-NN 1D as seen by a higher frequency of occurrence above $\sim$0.5 mm/hr. All of the algorithms show a similar pattern of underestimation at the highest snowfall rates likely due to fewer instances of these snowfall rates within the retrieval database.

The results from the GPROF-NN models outperform those from GPROF V7, and GPROF-NN 3D appears to outperform GPROF-NN 1D. Inclusion of spatial information into the retrieval through a CNN for GPROF-NN 3D allows for more accurate retrievals within the mountains. These results provide internal consistency that the models, albeit less so for GPROF V7, can reproduce data from the same distribution as the training data. However, it is expected the retrievals using the WUS-SR scaled MRMS retrieval database will more accurately reproduce the expected spatial patterns of snowfall accumulation in the WUS.

The scaling of MRMS snowfall by the WUS-SR introduces noise into retrieval database that affects the GPROF-NN models more so than GPROF V7 as seen in the bottom row of Figure 3.3. Above 3 mm/hr, there is severe underestimation in snowfall rates by the GPROF-NN models. These pixels have TBs that are not consistent with the average TBs for higher snowfall rates, but are assumed to not significantly impact the retrieval results given they account for less than 1% of the test dataset. Despite this, the retrieval results from the WUS-SR scaled retrieval database produce the intended results of the scaling process discussed in Section 3.2.3; the GPROF models using the WUS-SR scaled MRMS retrieval database produce higher snowfall totals.

Quantitative statistics of the GPROF snowfall retrievals are provided in Table 3.2. The table gives statistics of correlation, absolute percent bias, and root-mean squared error (RMSE) for the three algorithms. The statistics computed using the nominal MRMS retrieval database is shown as the first value with the statistics computed using the WUS-SR scaled MRMS retrieval database is shown in parentheses. The GPROF-NN algorithms outperform GPROF V7 in all statistics for both retrieval databases and GPROF-NN 1D shows the best performance for absolute bias. The noise induced from the scaling of MRMS by WUS-SR is clearly seen in the RMSE, as previously discussed.

Table 3.2. Mean statistics for GPROF snowfall retrievals. The values shown are for the two retrieval databases with the WUS-SR scaled database shown in parentheses.

Metric	GPROF V7	GPROF-NN 1D	GPROF-NN 3D
Correlation	0.402 (0.408)	0.543 (0.543)	0.621 (0.599)
Percent Bias (%)	23.977 (25.287)	3.186 (12.913)	22.970 (17.001)
RMSE (mm/hr)	0.072 (0.111)	0.066 (0.108)	0.063 (0.103)

The test data are also evaluated in terms of each algorithms' detection capabilities in terms of widely used metrics such as the probability of detection (POD), false alarm rate (FAR), and Heidke Skill Score (HSS). The detection scores are shown in Table 3.3 and are computed based on a snowfall rate threshold of 0.1 mm/hr. The detection metrics are sensitive to this threshold and should be considered when comparing to other retrieval algorithms. The value in parentheses corresponds to detection metrics computed using the WUS-SR scaled retrieval database. The detection scores between the retrieval algorithms show larger differences than the detection scores

between the retrieval databases. Similar to the quantitative statistics, the GPROF-NN models show better detection capabilities than GPROF V7 with GPROF-NN 3D producing the highest POD and HSS. The FAR is comparable across the GPROF models for the MRMS retrieval database, but GPROF-NN 3D has the lowest FAR for the WUS-SR scaled retrieval database.

Table 3.3. Detection metrics for GPROF snowfall retrievals. The values shown are for the two retrieval databases with the WUS-SR scaled database shown in parentheses.

Retrieval Algorithm	POD (%)	FAR (%)	HSS (%)
GPROF V7	16.4 (15.8)	52.6 (69.2)	23.3 (19.5)
GPROF-NN 1D	46.8 (56.2)	57.9 (61.8)	42.8 (43.5)
GPROF-NN 3D	60.5 (63.9)	53.7 (54.2)	51.1 (51.8)

An important aspect of this study, as discussed in Section 3.2.3, is the scaling the MRMS snowfall accumulation to create an updated reference snowfall. The retrieval algorithms should emulate the implemented scaling during the retrieval process, but it is not guaranteed the GPROF models identify the correct scaling given the complex relationships between the TBs and associated scaled snowfall. Figure 3.4 shows the ratio between the scaled and unscaled reference snowfall divided into the GPROF surface classes. The ratios are computed for the reference snowfall within the retrieval databases as well as between the snowfall retrieved by the GPROF algorithms. Ideally, the ratios computed for the retrieved snowfall would be consistent with those from the retrieval database. GPROF V7 has difficulty in reproducing the scaling in all of the surface classes except fro the medium vegetation surface class. While there are some differences for the GPROF-NN models across all surface types, both do an acceptable job of reproducing the scaling. Importantly, both GPROF-NN models do well in reproducing the scaling in the snow surface classes and mountains classes that make up the largest portion (67%) of the surface classes within the study domain. It should be noted that GPROF can still retrieve snowfall in the Mountain Rain surface class which correspond to mountain gridpoints that do not have snow cover at the surface.

3.3.2 Case Study: Western U.S. Water Year 2016

The results above assess the retrieval statistics from the instantaneous snowfall for the held-out test dataset. Figure 3.5 shows the snowfall accumulation for the GPROF models as well as the WUS- SR for WY2016, which was not used for model training. All GMI orbits within the

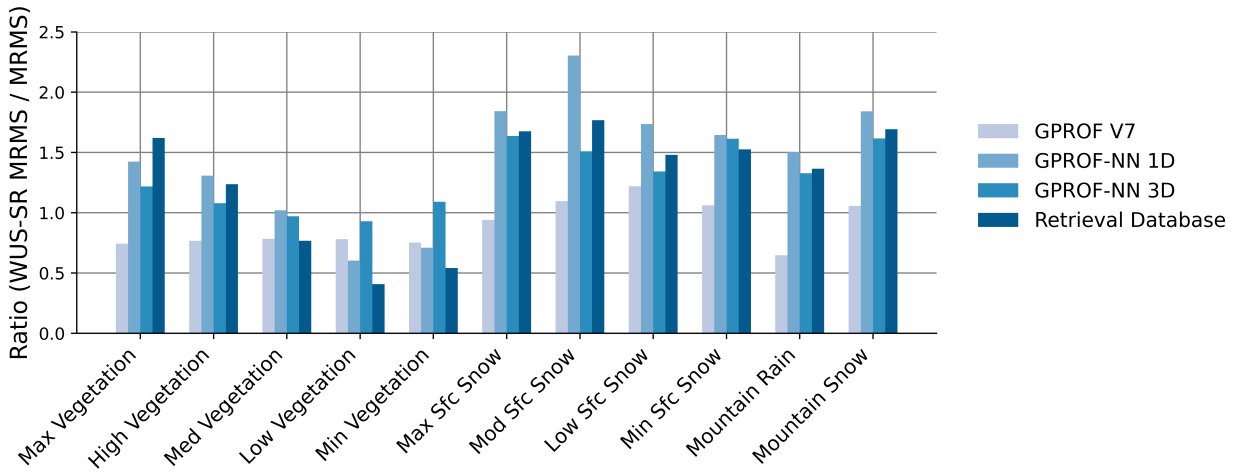

Fig. 3.4. Ratio of the unscaled and scaled snowfall for the training data and target variables in the three GPROF models.

WY are used to produce the WY snowfall accumulation. The retrieved snowfall accumulation shows relatively similar patterns for all models with higher snowfall accumulation in the mountains. However, the snowfall accumulation magnitudes exhibit large differences between models, even for the same retrieval database. GPROF V7 is unable to correctly estimate WY snowfall magnitudes compared to WUS-SR for both retrieval databases. While the GPROF-NN models are a significant improvement over GPROF V7, they are also unable to produce the WY snowfall magnitudes as seen in the WUS-SR when using the nominal MRMS retrieval database. The GPROF-NN models using the WUS-SR scaled MRMS retrieval database produce more similar WY snowfall accumulations to the WUS-SR, although not a completely independent comparison. All of the GPROF models are unable to capture the fine spatial details observed in the WUS-SR due to the larger spatial resolution of the GMI observations, and the inherent averaging that occurs when sensors have limited information content.

Figure 3.7 shows the snowfall accumulation for WY2016 from Figure 3.5 in boxplot format in order to see the accumulation distribution. The horizontal, dashed gray line shows the 25th, 50th, and 75th percentiles from the WUS-SR distribution to aid in comparison with the other model's distributions. All of the GPROF models underestimate the 25th percentile compared to WUS-SR while the median, 75th percentile, and upper quartile vary between the models and retrieval databases. The GPROF model interquartile range (25th to 75th percentile) increases noticeably

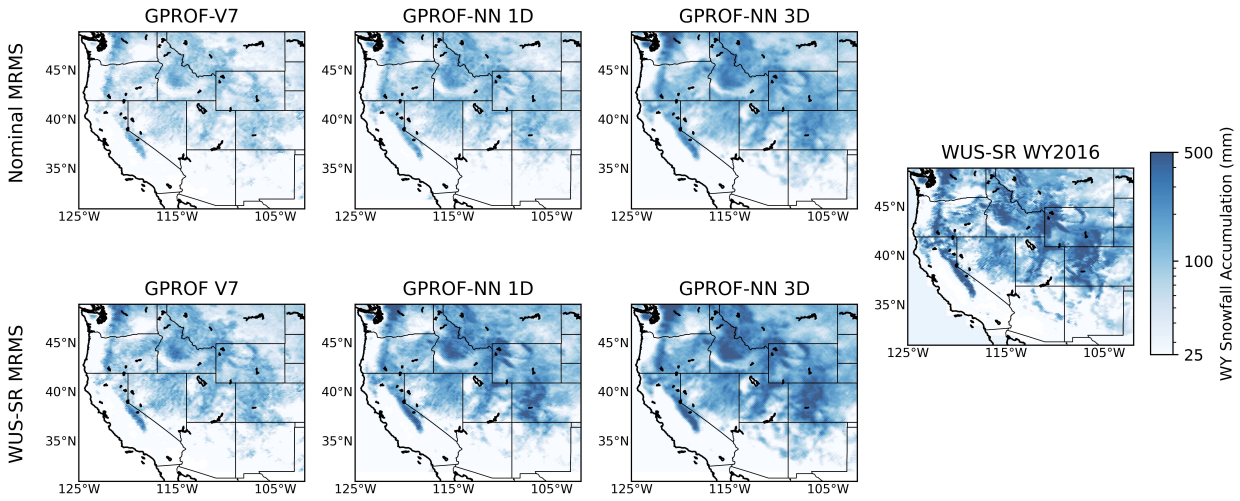

Fig. 3.5. WY2016 snowfall accumulation for the GPROF retrieval variables: MRMS snowfall (top) and WUS-SR scaled MRMS snowfall (bottom). The WUS-SR WY2016 snowfall accumulation is included on the far right.

from GPROF V7 to GPROF-NN 3D and from the MRMS to WUS-SR scaled MRMS retrieval database. GPROF-NN 3D overestimates the median, 75th percentile, and upper quartile compared to the WUS-SR when using the WUS-SR scaled retrieval database. This overestimation can be explained in part by the larger ratios compared to the retrieval database seen in the Medium, Low, and Minimum Vegetation surface classes in Figure 3.4. Investigation into the dependency of snowfall accumulation on elevation shown in Figure 3.6 shows the GPROF models tend to overestimate (underestimate) snowfall accumulation at lower (higher) elevation compared to WUS-SR. GPROF-NN 3D overestimates the distribution of snowfall accumulation compared to WUS-SR at elevations below $\sim$2500 m, which propagates into the domain-wide distribution in Figure 3.7. All GPROF models underestimate snowfall accumulation compared to the WUS-SR and do so at different elevations. The GPROF models using the WUS-SR scaled retrieval database are able to capture snowfall accumulations at higher elevations compared to those using the nominal MRMS retrieval database.

3.3.3 SWE Anomaly Correlations

While the results in Section 3.3.2 are notable in terms of the improvements seen in GPROF snowfall accumulations, it is also important that the models are able to capture the interannual variability across the domain. For this, five WYs are considered: WY2016-WY2020. Snowfall

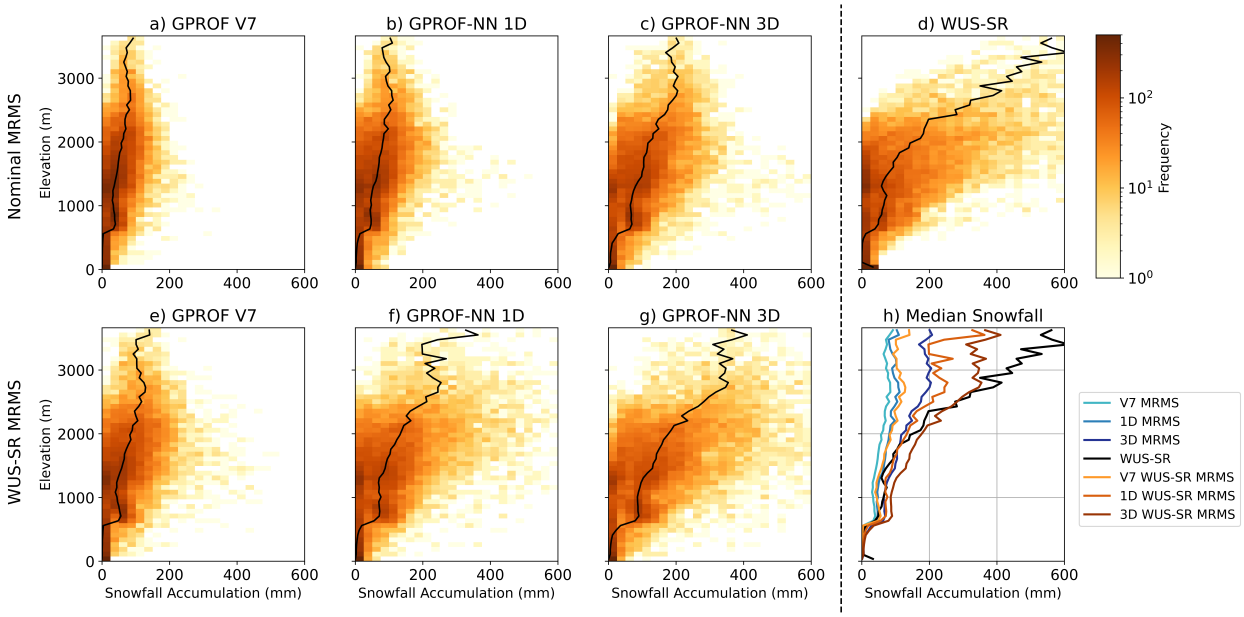

Fig. 3.6. 2D-histograms of GPROF snow accumulation for the nominal MRMS retrieval database (a,b,c) and WUS-SR scaled retrieval database (e,f,g). The solid black lines indicate the mean snowfall accumulation as a function of elevation. d) shows the WUS-SR data for reference and h) shows the individual mean lines.

is first accumulated for the GPROF models and WUS-SR for each WY and constrained to a WY accumulation of 50 mm before being standardized using the WY mean and standard deviation for each gridbox. This process produces a time series of standardized anomalies, which can be used to compute the correlation between the GPROF models and WUS-SR to assess the ability of the models to capture year-to-year snowfall accumulation variability. Figure 3.8 shows these time series correlations with the domain average correlation shown in parentheses.

The correlations between each model and the WUS-SR show some mixed results. GPROFV7 has difficulty in reproducing the interannual variability as shown by a relatively flat line of standardized anomalies shown in Figure 3.8d. The GPROF-NN models have higher domain average correlations than GPROF V7 with GPROF-NN 1D having the highest correlation of 0.42. Higher correlations are seen in the Sierra Nevada and Cascade mountain ranges and degrade in the Rocky Mountains. The Sierra Nevada and Cascade Mountains receive most of their snowfall from atmospheric river (AR) events whereas the Rocky Mountains are more dependent upon storm track and elevation. The study domain encompasses a range of topography and snowfall regimes, which

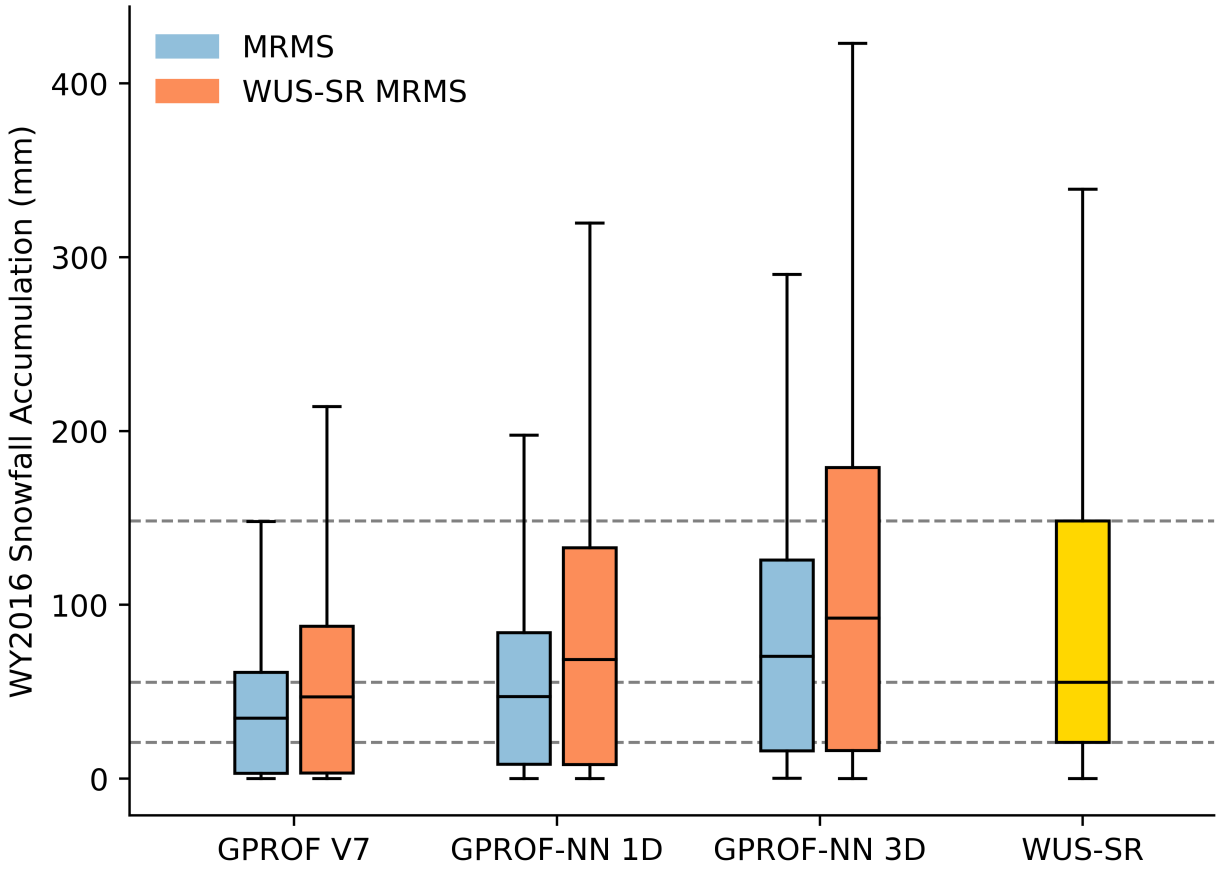

Fig. 3.7. Box and whisker plots for the WY2016 snowfall accumulation for each retrieval database. The box represents the interquartile range (IQR), with the line in the middle corresponding to the median. The whiskers extend to the maximum and minimum values excluding outliers. The horizontal dashed gray lines represent the IQR and median for the WUS-SR distribution.

appear to drive the differences in the models' ability to reproduce the interannual variability. For example, the temperature differences between the Sierra Nevada and Rocky Mountains alter the snowfall characteristics in terms of crystal size, shape, and density. These characteristics are variable, but the more synoptically driven AR events in the Sierra Nevada are likely easier for the GPROF-NN models to reproduce.

All of the GPROF models have difficulty reproducing the interannual variability along the borders of Montana, Idaho, and Wyoming as well as north, central Wyoming. Both of these regions are not well sampled due to the high radar beam that leads to low MRMS snowfall accumulation.

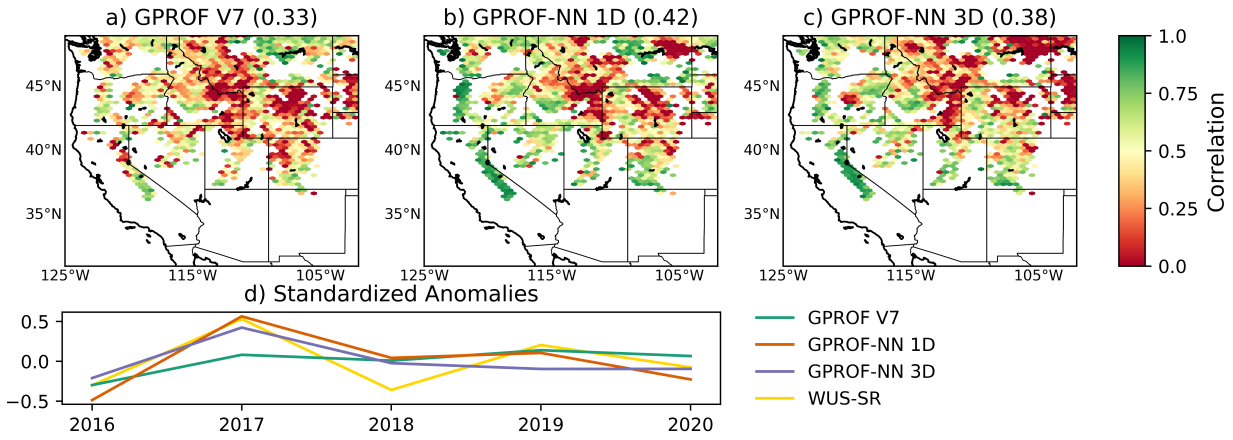

Fig. 3.8. Correlations between the standardized WY snowfall accumulations from the GPROF models and the WUS-SR (a-c). The domain average standardized WY snowfall accumulation time series (d).

These regions are filtered out of the retrieval databases as discussed in Section 3.2.3. The GPROF-NN models, and especially GPROF-NN 3D, also have difficulty in northeast Montana where the retrieval database sampling is reasonable. This region receives between 200 and 400mm of WY SWE, but is subject to multiple snow accumulation and ablation cycles throughout the WY, which changes the surface characteristics the retrieval must contend with.

3.3.4 Water Year SNOTEL Comparisons

In this section, grid averaged liquid equivalent snowfall accumulation from the GPROF algorithms is compared to point-scale SNOTEL SWE accumulation. These comparisons have unavoidable representativeness complications with the GPROF footprint size and SNOTEL point-scale area. The GPROF footprint field-of-view is unable to account for the sub-grid variability in snow accumulation driven by elevation, terrain, and forest cover. Additionally, SNOTEL stations are meant to provide an upper limit on the SWE for hydrological purposes. Given these differences, it is not expected the two snow accumulations match each other, but rather provide an independent verification of the improvement seen from different GPROF algorithms and the associated retrieval databases. Additionally, SNOTEL are used extensively as comparisons for snow and mountain hydrology studies and less so for satellite retrieval comparison, therefore, the comparison provided here tries to bridge this gap.

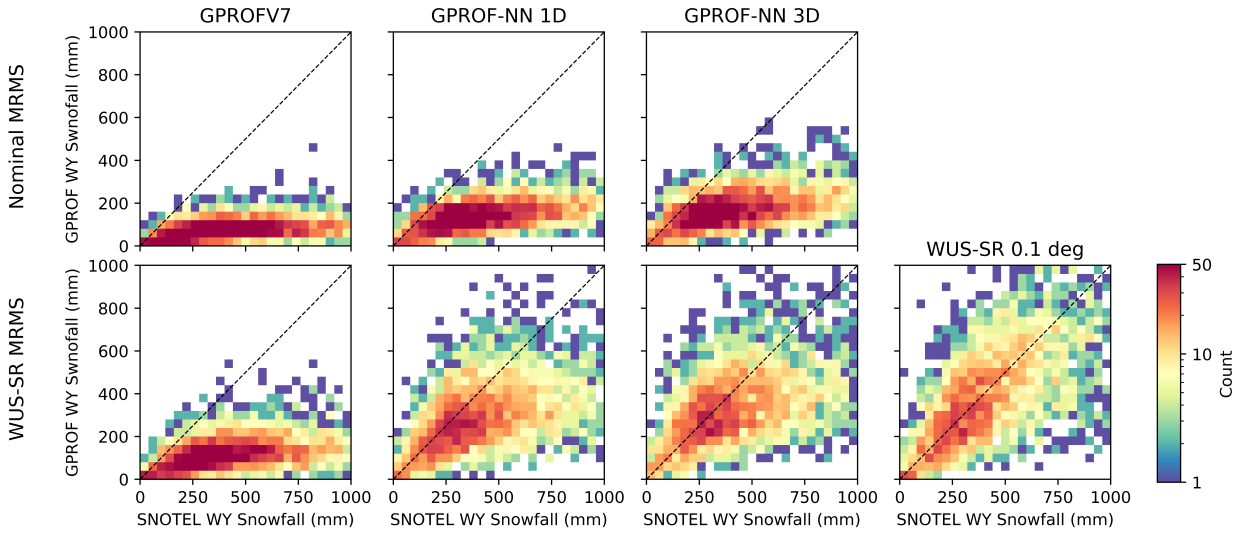

Fig. 3.9. 2D histograms between WY snowfall accumulations and SNOTEL SWE for the three GPROF algorithms and WUS-SR (columns) and the two retrieval databases (rows).

Figure 3.9 shows 2D-histograms of the GPROF liquid equivalent snowfall accumulation and SNOTEL SWE accumulation for WY2016 - WY2020. A 2D-histogram of the WUS-SR SWE accumulation is also included as a reference. The GPROF retrievals using the nominal MRMS retrieval database have significant underestimation in WY snowfall compared to the SNOTEL sites and retrievals using the WUS- SR scaled MRMS retrieval database show an improvement from this underestimation as seen in the error metrics shown in Table 3.4. Despite the seemingly reasonable correlations from the GPROF models using the nominal MRMS retrieval database, the RMSE shows the large deviation from the SNOTEL values. All of the GPROF models, irrespective of the retrieval database used, appear to have a threshold for which the retrieval cannot exceed. This is clearly evident in both GPROF V7 models as well GPROF-NN 1D and GPROF-NN 3D for the nominal MRMS retrieval database, but less so in GPROF-NN 1D and GPROF-NN 3D for the WUS-SR scaled retrieval database. This artifact is likely explained by a lack of information content and underestimation of higher snowfall rates as shown in Figure 3.3.

3.4 Discussion and Conclusions

This study evaluated the improvement of satellite passive microwave snowfall retrievals in the Western United States by incorporating snowfall accumulation information into the retrieval

Table 3.4. Mean statistics for WY snowfall accumulation between between the GPROF and SNOTEL. The values shown are for the two retrieval databases with the WUS-SR scaled database shown in parentheses.

Metric	GPROF V7	GPROF-NN 1D	GPROF-NN 3D	WUS-SR
Correlation	0.40 (0.40)	0.58 (0.49)	0.58 (0.55)	0.64
Bias (mm/WY)	423.91 (375.31)	354.07 (186.55)	319.44 (135.47)	68.96
RMSE (mm/WY)	543.44 (500.66)	476.71 (364.28)	445.57 (325.03)	282.66

databases. This study also made use of two neural-network versions of the GPROF algorithm, GPROF-NN 1D and GPROF-NN 3D. In total, three GPROF retrieval models were run on two retrieval databases and assessed on instantaneous snowfall statistics, WY snowfall accumulation, and interannual snowfall accumulation variability.

The answer to our research question,"Can a neural network retrieval algorithm, combined with the Western U.S. Snow Reanalysis improve the accuracy, spatial distribution, and accumulation of retrieved snowfall?", is a yes, but with some caveats. Throughout this study it was shown the GPROF-NN models perform better compared to GPROF V7 irrespective of which retrieval database was used. A possible explanation for this result could be from how the different models account for the observation error covariances. These errors are prescribed in GPROF V7 and likely do no account for regime dependent errors that are present across the study domain. The GPROF-NN models are able to dynamically infer these errors as long as the retrieval database represents them accordingly. Additionally, the GPROF-NN models are able to operate on the full retrieval database as opposed to GPROF V7 requiring a compressed database for efficiency purposes.

Models using the WUS-SR scaled retrieval database showed increases in WY snowfall accumulation as seen in Figure 3.5 and Figure 3.7. As discussed, the GPROF-NN models are better able to use the scaled retrieval database than GPROF V7 and are more comparable to the WUS-SR for WY2016, which was excluded from the retrieval databases. The overestimation observed in the distribution of GPROF-NN 3D WY snowfall accumulations compared to WUS-SR in Figure 3.7 is explained by overestimation for lower elevations (< 2500 m) and for specific GPROF surface types. These features are likely exacerbated in the GPROF-NN 3D models because of the CNN that accounts for the spatial structure of the scene.

The three GPROF models reproduced the temporal patterns in WY snowfall compared to the WUS-SR with mixed results. The models better reproduced the interannual variability of WY snowfall accumulation in the Sierra Nevada and Cascade Mountains compared to the Rocky Mountains. A likely explanation for this is the differences in snowfall regimes and topography characteristics for the different regions. It appears atmospheric river events that are synoptically driven are better captured year-to-year compared to snow events in the Rocky Mountains that are more dependent on storm track and the corresponding elevation. Surface characteristics can also impact the accuracy of the retrievals where multiple cycles of snow accumulation and melt occur throughout the WY significantly changing the surface emissivity the retrievals must contend with.

Independent validation of the GPROF models against SNOTEL showed some promising results for satellite snowfall retrievals in the mountains. GPROF-NN 1D and GPROF-NN 3D using the WUS-SR scaled retrieval database had the best comparisons to SNOTEL. Using the WUS-SR scaled significantly reduces the bias and RMSE for the GPROF-NN models. Despite the uncertainties associated with comparing satellite and point-scale measurements of snowfall, this validation provides a positive step towards showing satellite snowfall retrievals in the mountains can be useful and that incorporating snow climatology information into the retrieval database is important.

This study focused on the Western United States due to the availability of the WUS-SR, MRMS, and SNOTEL data. The results are underpinned by the snow information provided by the SWE-Reanalysis. The SWE-Reanalysis product has been produced for the Western U.S., Andes Mountains, and High Mountain Asia (HMA) with extensive validation for the Western U.S. and Andes. The HMA region has virtually no validation data available so positive results are inferred, with caveats, given the successful validation in the other regions. This is somewhat encouraging from a satellite point-of-view that the methodology outlined in this work to be extended outside of the Western U.S..

To test the generalizability of the GPROF models trained on data in the WUS to mountains outside of the WUS domain, Figure 3.10 shows the WY2015 snowfall accumulation for the Andes from the GPROF-NN 3D model that was trained using the WUS-SR scaled MRMS retrieval database and the snowfall accumulation for the Andes derived from the Andes SWE-Reanalysis

(Andes-SR). The Andes are on the order of $\sim$2000 m higher than the WUS mountains so an elevation adjustment was made to the retrieval database elevation. Specifically, all elevation in the retrieval database was increased by $\sim$2000 m. It should be noted this adjustment is not intended to be perfect, but rather allow for comparisons between GPROF-NN 3D trained in the WUS and the Andes. Figure 3.10 shows GPROF-NN 3D is able to, with some limitations, reproduce the magnitudes and south-to-north gradient of snowfall accumulation in the Andes. There is a tendency for GPROF-NN 3D to overestimate snowfall accumulation in the valleys as seen in the Andes-SR. An investigation into the input ancillary data for GPROF-NN 3D in the Andes compared to the WUS shows higher elevation, despite the adjustment, and warmer temperatures that could both lead to higher snowfall rates and explain the overestimation in GPROF-NN 3D. While not expected to a perfect comparison, these results show promise that SWE-Reanalysis datasets could be produced for targeted mountain ranges that provide a robust enough description of global mountains to be used for the creation of a global mountain PMW retrieval database for snowfall retrievals.

Earth's snow resources have been and will continue to be affected in a changing climate and, therefore, is important to accurately monitor snow across its lifecycle. This study focuses on mountain snowfall that drives snow accumulation and the hydrologically more important snow water equivalent. Due to their limitations, satellite snowfall estimates have historically not been used in hydrologic applications. Ultimately, this study sheds a positive light on the ability of PMW snowfall retrievals to provide more useful snowfall estimates that could be used as forcing for a snow evolution model.

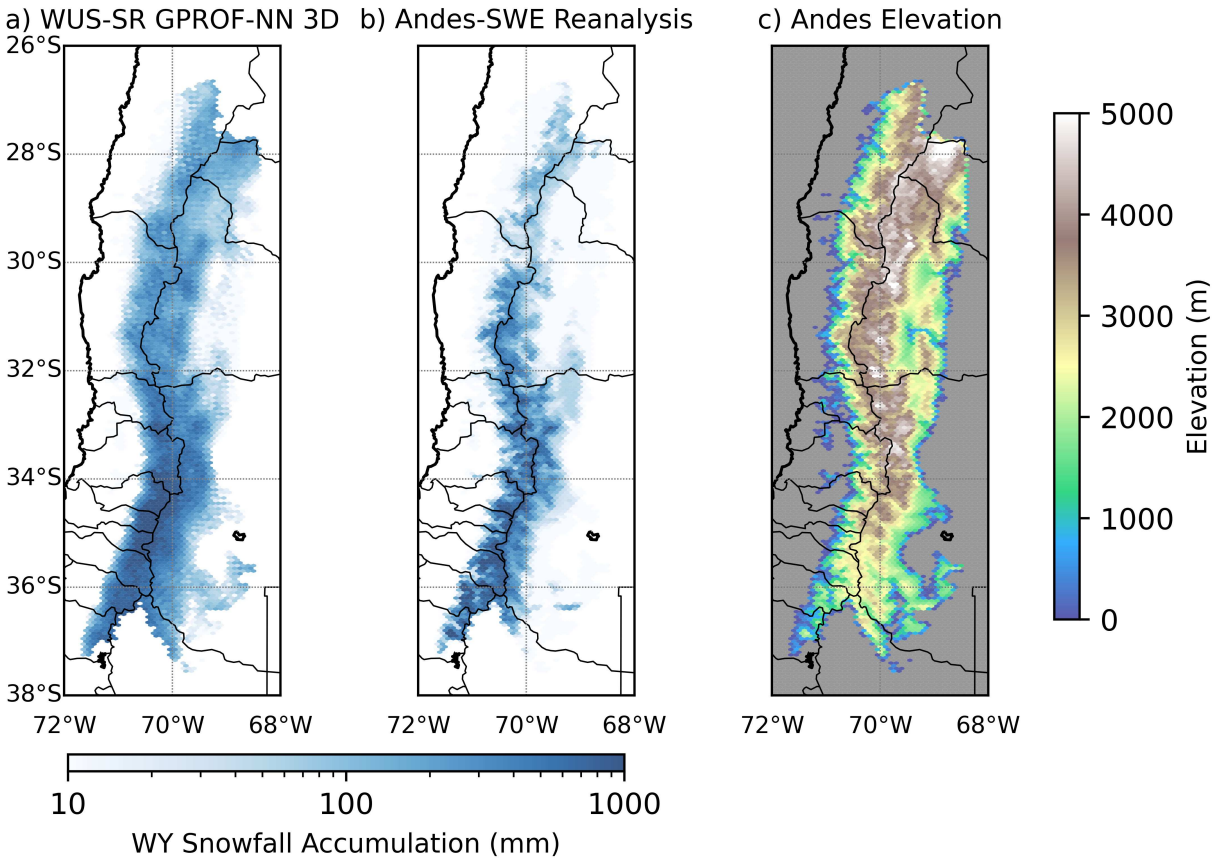

Fig. 3.10. WY2015 snowfall accumulation from a) GPROF-NN 3D trained on the WUS-SR scaled MRMS retrieval database and b) Andes SWE-Reanalysis. A mask has been applied to only allow pixels with valid snow accumulation as given by the Andes-SR.

CHAPTER 4

Integrating Satellite Precipitation Retrievals with a Snow Evolution Model in the Western United States[3]

4.1 Introduction

Proper representation of mountain snowfall is critical in understanding the availability of snow water that is fundamental for runoff predictions, hydroelectric power generation, and recreation among others (Pagano et al. 2004; Sturm et al. 2017). In the western United States, where an estimated 53% of runoff originates as snow (Li et al. 2017), there have been declining trends in snow resources recorded over the past 7 decades (Mote et al. 2005, 2018). A warming climate has been linked to decreases in snowpack magnitudes (Mote et al. 2005; Liston and Hiemstra 2011), snow extent (Kunkel et al. 2016), earlier snowmelt onset (Pederson et al. 2011; Kapnick and Hall 2012), and decreased warm-season streamflows (Pederson et al. 2011) with similar trends projected for the upcoming century (Intergovernmental Panel On Climate Change (Ipcc) 2022; Shulgina et al. 2023). These changes increase the stress on outdated reservoir and dam infrastructure that must compete with earlier snowmelt and larger hydrologic variability (Pagano and Garen 2005; Abatzoglou et al. 2014). The region's reliance on a limited and decreasing water supply, coupled with increasing demand, creates significant challenges for all stakeholders that depend on accurate representations of snow throughout its lifecycle.

Models have difficulty accurately representing snow water equivalent (SWE) in mountainous terrain (Broxton et al. 2016b), with precipitation widely considered the primary source of error (Raleigh et al. 2016; Günther et al. 2019). Global climate models, global reanalyses, and regional mesoscale models, frequently used to estimate SWE, generate precipitation at spatial resolutions too coarse to capture the complex orographic processes that drive small scale precipitation patterns (Cho et al. 2022b). High resolution models (< 10km) produce significantly more SWE than their global counterparts, but these models are not readily available (Wrzesien et al. 2019; Hammond et al. 2023). Precipitation downscaling methods can be employed in order to use global model

[3]To be submitted to *Journal of Hydrometeorology* in 2025 after further revisions. Gonzalez, R., and C. Kummerow: Integrating Satellite Precipitation Retrievals with a Snow Evolution Model in the Western United States.

and reanalysis data for SWE estimation (Gutmann et al. 2012; Tabari et al. 2021; Rugg et al. 2023). These approaches either use a dynamical downscaling method that incorporates a regional climate model with higher resolution physics (Komurcu et al. 2018; Yoshikane and Yoshimura 2023; Hughes et al. 2024) or a statistical downscaling method that establishes relationships to the surrounding landscape (Daly et al. 1993; Liston and Elder 2006a). However, these methods assume the larger scale precipitation data is error free and has been shown otherwise (Lundquist et al. 2015). More recently, repeatable snow depth patterns have been used adjust precipitation inputs (Vögeli et al. 2016; Brauchli et al. 2017; Pflug et al. 2021).

Rather than downscale input precipitation data, satellite remote sensing observations have been used extensively to adjust the state variables in hydrologic models to better match observations. Binary snow covered area observations obtained from visible sensors offer limited information on the amount of snow present for a given observation, but have often been coupled with a snow depletion curve (SDC) to infer SWE (Essery and Pomeroy 2004; Andreadis and Lettenmaier 2006; Durand et al. 2008; Margulis et al. 2019). The choice of the SDC can lead to errors in the associated hydrologic processes involving snowmelt (Arsenault and Houser 2018; Schneider et al. 2021). SWE observations from passive microwave (PMW) sensors have been assimilated with varying success, mostly dependent on snow depth (Andreadis and Lettenmaier 2006; De Lannoy et al. 2012; Fletcher et al. 2012; Gonzalez and Kummerow 2020). Techniques for retroactive SWE reconstruction have been proposed and recently implemented in operational products that connect input model forcing data to the snow state constrained by fractional SCA observations (Margulis et al. 2016; Fang et al. 2022). Most snow modeling studies using remote sensing data focus on the assimilation of snow states such as SCA and SWE, as these are directly related to the temporal change of snowpack over time. To the authors knowledge, there is limited literature that specifically focuses on the incorporation of satellite remote sensing of precipitation into snow models.

One of the challenges of using satellite precipitation estimates for snow modeling is the relatively infrequent observation times compared to ground based measurements. Early techniques to combine PMW and infrared (IR) observations used the temporal strength of IR based precipitation estimates with the instantaneous strength of PMW estimates (Adler et al. 1992; Huffman

et al. 1997; Kummerow and Giglio 1995). Kidd et al. (2003) improved upon these early combined techniques to shorter temporal estimates at half-hourly resolution. Today, the Integrated Multi-Satellite Retrievals for GPM (IMERG) product provides a solution that compensates for the limited sampling of a single satellite by using all available satellites supplemented by geostationary IR observations (Huffman et al. 2023). Recent studies have shown promise using IMERG estimates for hydrologic studies (Eckert et al. 2022; Seo et al. 2023), but do not investigate snow as only the most recent version of IMERG, IMERG V07, includes PMW precipitation estimates over snow and ice surfaces (Huffman et al. 2024). Spaceborne radars, such as the Dual Precipitation Radar (DPR) onboard the Global Precipitation Measurement (GPM) Core Observatory, provide a more direct estimate of precipitation compared to passive microwave or IR estimates (Battaglia et al. 2020), but are limited to revisit times (2-3 days) that are unable to properly capture snowfall accumulation. CloudSat's Cloud Profiling Radar (CPR) is better suited for estimating snowfall compared to the DPR (Casella et al. 2017), but has a significantly increased revisit time given its narrow (1.7 km) swath (Rysman et al. 2018).

Another challenge of using satellite PMW precipitation estimates for snow modeling is the complex interactions of radiation with snow particles (Ulaby and Stiles 1980). The scattering signal of snow and ice particles in precipitating clouds has been the basis of snowfall retrieval algorithms (Kulie et al. 2010), with the degree of scattering related to the shape and density of the snow particles (Liu 2008). These snow characteristics are difficult to account for due to their strong sensitivity to local conditions (Pettersen et al. 2020b,a). Cold and mountainous regions that exhibit seasonal snow complicate the retrieval of snowfall by decreasing the upwelling radiation through volume scattering in the snowpack (Mätzler 1993) that changes as the snowpack evolves (Cordisco et al. 2006) and with vegetation (Kelly 2009). A combination of low and high frequencies sensitive to the surface and atmosphere, respectively, can offer increased snowfall detection skill (Takbiri et al. 2019). Snowfall retrieval algorithms have been created using a range of approaches from physical algorithms (Meng et al. 2017), empirical algorithms (Noh et al. 2006), Bayesian algorithms (Kummerow et al. 2015), and machine learning algorithms (Rysman et al.

2018; Pfreundschuh et al. 2022a), but none have been constructed specifically for mountain snowfall.

This work looks to bridge this gap and apply precipitation estimates within the mountainous western United State to a snow model. Specifically, this study aims to answer the following research questions:

(1) Are satellite precipitation estimates suitable for use as input to a snow model in the western United States?

(2) What are the associated errors in modeled snowpack across different watersheds when using satellite precipitation estimates as inputs?

To address these questions, we employ a machine learning precipitation retrieval algorithm that incorporates the currently available PMW and IR observations with a ground-based radar snowfall product that has itself been adjusted to match climatological observations. The algorithm is designed to leverage the strengths of both microwave and IR observations to create a continuous precipitation product that can be used as forcing data to a snow model. This approach represents a novel framework that provides a direct way of incorporating satellite snowfall estimates into the snow life cycle, offering a new approach to snow modeling that can be complimentary or even replace the PMW SWE product that is notoriously difficult to measure (Gonzalez and Kummerow 2020). Section 4.2 describes the machine learning model, the snow evolution model, and the experimental setup used to assess multiple aspects of our research question. Sections 4.3 and 4.4 present the results and conclusions highlighting the performance of the retrieval algorithm and using its precipitation estimates to force a snow model.

4.2 Data and methods

Snow evolution models depend primarily on the input precipitation to drive snow accumulation. The spatial variability, timing, and intensity of snowfall are key contributors to the accuracy of modeled snow properties. Because of its importance, this section provides an overview of the satellite precipitation retrieval algorithm as well as the snow evolution model, SnowModel, used

in this study. This is followed by a discussion of the experimental setup of coupling the satellite precipitation estimates with SnowModel.

4.2.1 Chalmers/CSU Integrated Multi-Satellite Retrieval Platform

The Chalmers/Colorado State University Integrated Multi-Satellite Retrieval Platform (CHIMP, Pfreundschuh (2024)) is used to retrieve precipitation over the study domain. The version of CHIMP used here combines the strengths of three model architectures and techniques: a UNet-based encoder-decoder (Ronneberger et al. 2015), EfficientNet-V2 Sequential (Tan and Le 2021), and a quantile regression loss function (Pfreundschuh et al. 2018). The UNet-based architecture is able to capture multi-scale features and has been applied extensively to precipitation (Sadeghi et al. 2020; Hilburn et al. 2021; Rahimi et al. 2024). EfficientNet-V2 Sequential, along with its reduced computational complexity, is used to create higher efficiency and quicker processing times compared to traditional convolutional neural networks (CNNs). The model's training is guided by a quantile regression loss function that is able to estimate the a posteriori distribution, or quantiles, of the retrieved precipitation. This is similar in output to Bayesian formulations of precipitation retrievals, but with the added benefit of providing consistent uncertainty estimates for non-Gaussian retrieval errors and increased computational efficiency.

The first step in using CHIMP is the extraction of data used for training and testing. An area projection is required to extract and grid the input satellite observations and reference precipitation. Brightness temperature (TB) observations for input sensors are regridded from the orbit overpasses to the specified domain using nearest neighbor sampling. Sensor data is extracted at a time interval of 30 minutes; the most recent overpass is used for sensors with multiple domain overpasses within the 30 minute interval. Data from five passive microwave sensors and one infrared sensor are used to build the training database and include the Advanced Microwave Scanning Radiometer 2 (AMSR2), Advanced Technology Microwave Sounder (ATMS), Global Precipitation Measurement (GPM) Microwave Imager (GMI), Microwave Humidity Sounder (MHS), Special Sensor Microwave Imager/Sounder (SSMIS), and NOAA's Climate Prediction Center (CPC) Merged Infrared (CPCIR). The Multi-Radar Multi-Sensor (MRMS) dataset provides the input sensors with a reference precipitation target used during the training and validation processes. Data is extracted

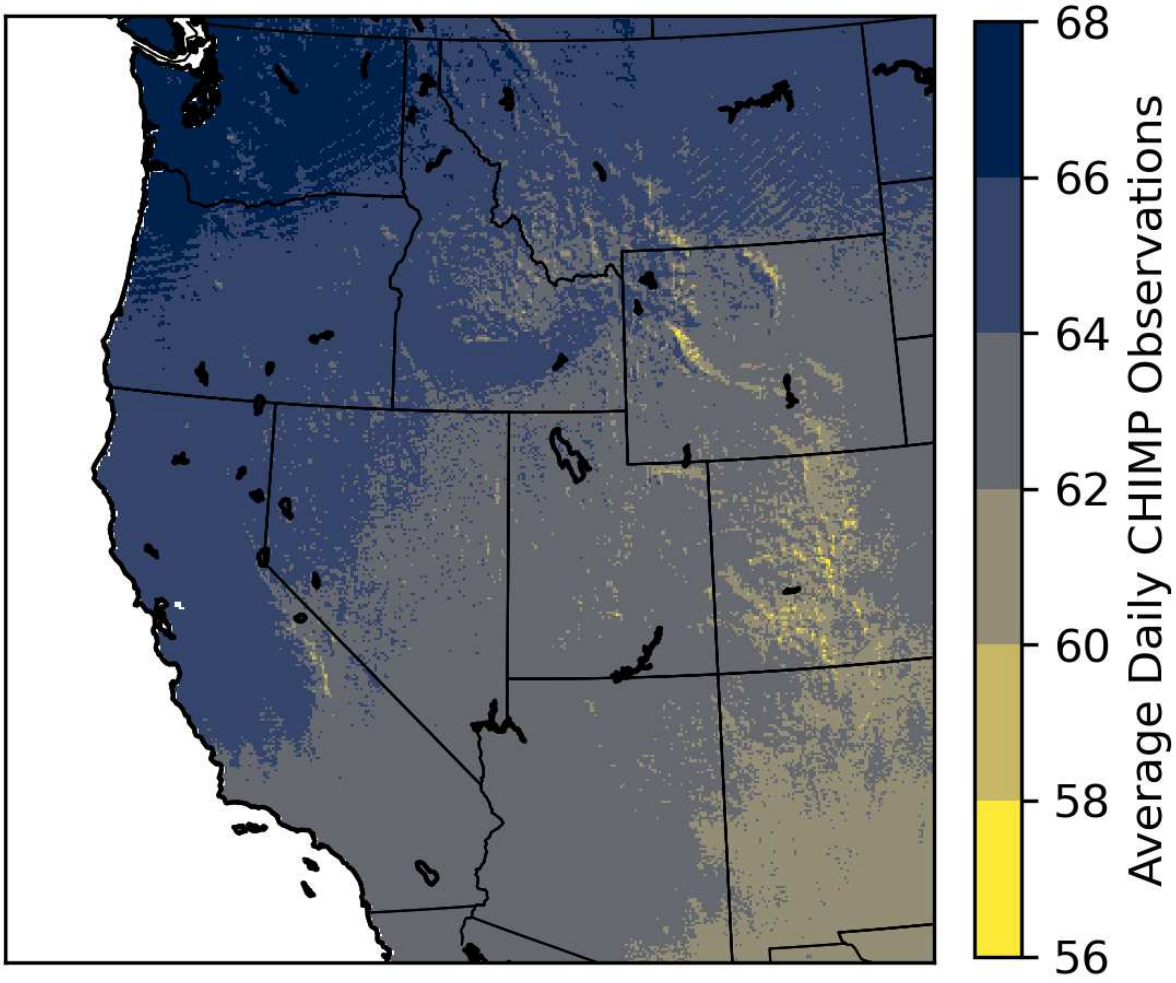

Fig. 4.1. Daily average number of CHIMP input observations between AMSR2, ATMS, GMI, MHS, SSMIS, and CPCIR for WY2021.

for 4 water years (WY, October 1 - September 30) with WY2018-WY2020 used for training while WY2021 is used for testing with SnowModel. The first 3 days of each month in the training data are used as validation to monitor the training performance and data from July of 2018 and January 2019 are excluded from training to independently evaluate CHIMP. Figure 4.1 shows the average number of daily input observations for WY2021 that highlights the ability for CHIMP to use all available sensors to provide precipitation estimates at relatively high resolution. The majority of the total daily inputs correspond to the geostationary CPCIR data available at 30 minute resolution with the PMW sensors adding input information as their orbit allows.

An important aspect of this study is the scaling of the MRMS reference precipitation according to Gonzalez and Kummerow (2024) (in review). The MRMS snowfall rates and corresponding snow accumulations are known to be unrepresentative of the snow climatology in complex terrain. Thus, the scaling procedure adjusts the MRMS liquid equivalent snowfall rates to match the climatology of snow water equivalent (SWE) from the Western United States UCLA Daily Snow Reanalysis (UCLA-WUS-SR) (Fang et al. 2022). The UCLA-WUS-SR has been extensively validated in the Western US (Fang et al. 2023) and is considered to be representative of the study domain's snow climatology.

The CHIMP model training uses a staged approach, alternating between warmup phases and primary training stages. Each warmup phase runs for 10 epochs on 10% of the training data using a warmup scheduler to stabilize initial training dynamics. Following each warmup phase, a corresponding primary training stage provides a more comprehensive training with progressively increasing epochs. A total of four primary trainings stages are used in this study.

4.2.2 SnowModel

SnowModel is a spatially distributed snow-evolution modeling system that can simulate snow evolution processes (Liston and Elder 2006b). Four submodels are used to simulate different physical processes known to drive snow evolution: MicroMet, EnBal, SnowPack, and SnowTran-3D (Liston and Hall 1995; Liston and Elder 2006a; Liston et al. 2007). At each timestep within a SnowModel simulation, the meteorological forcing data are distributed over the model domain using MicroMet and then the following processes are simulated for each gridcell: 1) perform near-surface energy balance calculations using EnBal, 2) evolve the snowpack properties using SnowPack, and 3) redistribute the snowpack by wind-driven properties using SnowTran-3D. A fifth submodel named SnowAssim can constrain the modeled SWE using observations (Liston and Hiemstra 2008). Each submodel is described below except for SnowTran-3D as this submodel is excluded for this study due to the gridsize of the SnowModel domain.

SnowModel has been used for many applications across a range of snow regimes with different domain sizes, but uncertainties remain depending on the SnowModel simulation setup. An assumption about homogeneous grid cell processes creates large uncertainties, especially for larger

SnowModel grids. For example, snow transport by wind-driven processes occur at spatial scales of less than a few hundred meters. Large SnowModel grids are also susceptible to uncertainties driven by the assumed temperature and precipitation relationship with elevation that are important to defining snow accumulation and ablation across elevation ranges. Finally, cloud cover is determined by converting the surface air temperature and relative humidity to 700mb and computing the cloud fraction. This cloud cover computation is a first-order estimate and is not valid under all meteorological conditions such as radiative cooling under clear skies that acts to increase the surface relative humidity.

4.2.2.1 MicroMet

MicroMet is a quasi-physically based, high-resolution meteorological distribution model that is designed to produce the required meteorological forcing input to SnowModel. It has the ability to distribute in-situ observations, remote sensing data, or gridded data from reanalyses or models over the specified SnowModel domain with individual submodels for each variable designed to improve the data based on the surrounding landscape, primarily topography. MicroMet outputs air temperature, relative humidity, wind speed, wind direction, incoming solar radiation, incoming longwave radiation, surface pressure, and precipitation at the SnowModel grid.

4.2.2.2 EnBal

EnBal computes surface energy balances in response to the near-surface meteorological variables provided by MicroMet. The energy balance is determined by the incoming solar and longwave radiation, emitted longwave radiation, turbulent latent heat exchange, and conductive energy transport being equal to the melt energy. Each term is formulated with surface temperature as the only unknown variable. When snow is present and the surface temperature is above 0°C, surface temperature is set to 0°C and the residual energy is the energy available for melting the snowpack.

4.2.2.3 SnowPack

SnowPack calculates the changes in the snowpack in response to the input meteorological data from MicroMet and the melt energy from EnBal. It is designed to handle single or multi-layered snowpack with multiple pathways for evolving the snowpack density. Snow density can change

with time due to the snowpack temperature and weight of the overlying snow, through redistribution of melt water, and sublimation of snow from EnBal. SnowModel has the ability to assimilate snow density observations to help reproduce the SWE, snow depth, and density observations.

4.2.2.4 SnowAssim

SnowAssim uses in-situ, remotely sensed, or gridded SWE data to constrain the modeled outputs of snowfall. Deficiencies in SnowModel can arise due to the limitation in the modeled snow physics, uncertainties in the meteorological forcing data, or the topography and vegetation. The differences between modeled and observed SWE are used to adjust the precipitation needed to better reproduce the observed SWE. For example, if the observed April 1 SWE is higher than the modeled April 1 SWE then SnowAssim will scale the precipitation for all timesteps before April 1 in order to replicate the observed SWE. In addition to precipitation scaling, the snow melt can also be scaled to match observed SWE during the melt season or the snow free date.

4.2.3 Experimental Design

This section is used the present the experimental design of this study. First, a description of the study domain used by CHIMP and SnowModel is presented followed by details of how the CHIMP retrieval output is processed and ingested into SnowModel. Finally, a description of the different SnowModel configurations is given.

There are many advantages to using CHIMP as an input source of precipitation for SnowModel. One of the advantages is the ability to define the model domain and spatial resolution for both CHIMP and SnowModel. The study domain is referred to as the Western United States (WUS) and is defined using an Albers Equal Area (AEA) projection within a bounding box of 31°N to 49°N and -102°W to -125°W with a spatial resolution of $\sim$5000 m. Figure 4.2 provides an example of WUS domain with static elevation and vegetation data used in SnowModel as well as the Hydrologic Unit Code 2 (HUC-2) watersheds and major mountain ranges referenced throughout this work. The elevation and vegetation data is provided by the United States Geological Survey's National Elevation Dataset and the North American Land Change Monitoring System. Table 4.1 provides details of the domain projection.

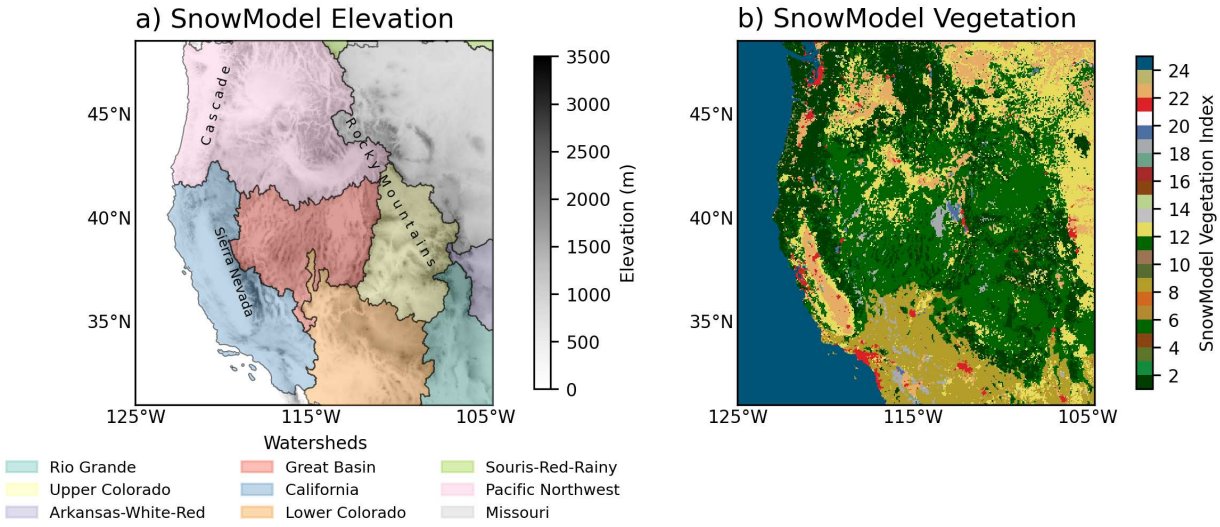

Fig. 4.2. SnowModel (a) elevation and (b) vegetation maps of the western United States (WUS) domain. (a) includes shaded regions delineating the HUC-2 watersheds with labels of the major mountain ranges for reference.

Table 4.1. Western United States domain projection parameters.

Parameter	Value
Projection	Albers Equal Area (AEA)
Datum	WGS84
Units	meters
Gridsize	5000 m
Central Latitude	40 N
Central Longitude	113.5 W
Standard Parallel 1	34 N
Standard Parallel 2	46 N
Lower Left Corner X	-1101962.29 m
Lower Left Corner Y	-929160.32 m

Additionally, CHIMP is able to operate on any number of input sensors, which allows for continuous precipitation estimates rather than irregularly spaced observations from a single sensor. Information on precipitation type (e.g. rain or snow) is not included in the CHIMP estimates as this is handled directly by SnowModel's MicroMet submodel. As discussed in Section 4.2.1, CHIMP retrievals are output every half hour for the entire WY2021 study period. To use these precipitation estimates in SnowModel, they must first be converted to the 3 hourly model timestep.

Table 4.2. SnowModel configurations for the three SnowModel runs.

SnowModel Run	Meteorology Forcing	Precipitation Forcing	Assimilation Data
CHIMP SnowModel	MERRA-2	CHIMP (Section 4.2.1)	-
MERRA-2 SnowModel	MERRA-2	MERRA-2	-
WUS-SR SnowModel	MERRA-2	MERRA-2	UCLA-WUS-SR (Section 4.2.1)

This is performed by accumulating the precipitation retrieval estimates for each 3 hourly timestep and directly replacing the model precipitation forcing with the CHIMP precipitation forcing.

SnowModel is used to simulate the snowpack characteristics at 3 hourly timescales and 5km spatial resolution. The simulations assume a single snow layer without wind-driven redistribution; the snowpack is affected solely by the input precipitation and snowmelt and sublimation derived from input meteorological forcing data. The CHIMP SnowModel simulation (hereafter, CHIMP-SM) uses NASA's Modern-Era Retrospective analysis for Research and Applications (MERRA-2) for all meteorological forcing data except for precipitation, which is supplied by the CHIMP retrieval. Terrain-based MicroMet lapse rates were used to redistribute the input CHIMP precipitation, but this interpolation scheme had minor to no effect given the similar model grids used by SnowModel and CHIMP. In addition to the CHIMP-SM simulation, two additional SnowModel simulations were performed: one using the precipitation input from MERRA-2 (hereafter, MERRA-2) and a second assimilating SWE data from the UCLA-WUS-SR described in Section 4.2.1. The WUS-SR SnowModel (hereafter, WUS-SR) assimilates UCLA-WUS-SR on the first day of December, January, February, March, and April in order to replicate UCLA-WUS-SR SWE throughout the snow accumulation season. The MERRA-2 and WUS-SR SnowModel simulations are used as a reference for the CHIMP-SM. Table 4.2 shows a list of the main SnowModel components that change across the three SnowModel simulations.

CHIMP-SM is assessed against MERRA-2 and WUS-SR and includes comparisons of the spatial distributions of peak SWE, the timing of peak SWE, the snowpack water storage for HUC-2 watersheds, and the relationship between SWE and elevation. The study period of WY2021 is summarized for context. Throughout the winter season, many areas of the western United

Table 4.3. Mean evaluation metrics for surface precipitation estimated by CHIMP.

Metric	January 2019
Bias (CHIMP - MRMS) (mm/hr)	-0.0117
Bias (%)	-19.204
MAE (mm/hr)	0.0731
MSE (mm/hr)	0.1366
Correlation	0.4152

States recorded below-average snowfall. Above average snowfall improved snowpack in the central Rocky Mountains during March, but by April 1st, observed SWE was above average mainly for the Pacific Northwest. Additionally, persistent Spring high-pressure systems led to accelerated and record breaking snow melt across the entire WUS domain (McEvoy and Hatchett 2023)

4.3 Results

4.3.1 CHIMP Evaluation

Evaluation of the CHIMP model is performed for the month of January 2019 that was not used during training. Table 4.3 provides quantitative measures of the retrieval accuracy for the test data. The percent bias (CHIMP - MRMS) shows a systematic underestimation in the CHIMP precipitation estimates compared to the reference precipitation of MRMS.

It is also important to assess the ability of CHIMP to estimate surface precipitation spatial patterns. Figure 4.3 shows the MRMS reference precipitation (Figs. 4.3a,b,c) compared to the CHIMP estimated precipitation (Figs. 4.3d,e,f) for January 17-18, 2019. Two times (January 18 0030 and January 17 1730 UTC) are shown along with the precipitation accumulation for the whole period. This event is notable as a large area of low pressure moved inland over California producing heavy snowfall in the Sierra Nevada mountains before continuing west and producing snowfall across the Rocky Mountains and Great Plains. Overall, CHIMP is able to reproduce the location of precipitation well for this example. This is especially noticeable in Figures 4.3c,f showing the precipitation accumulation. CHIMP captures the precipitation in California and Utah well, but slightly struggles in Colorado and Montana. There is a clear underestimation of CHIMP precipitation compared to MRMS, which was shown in Table 4.3.

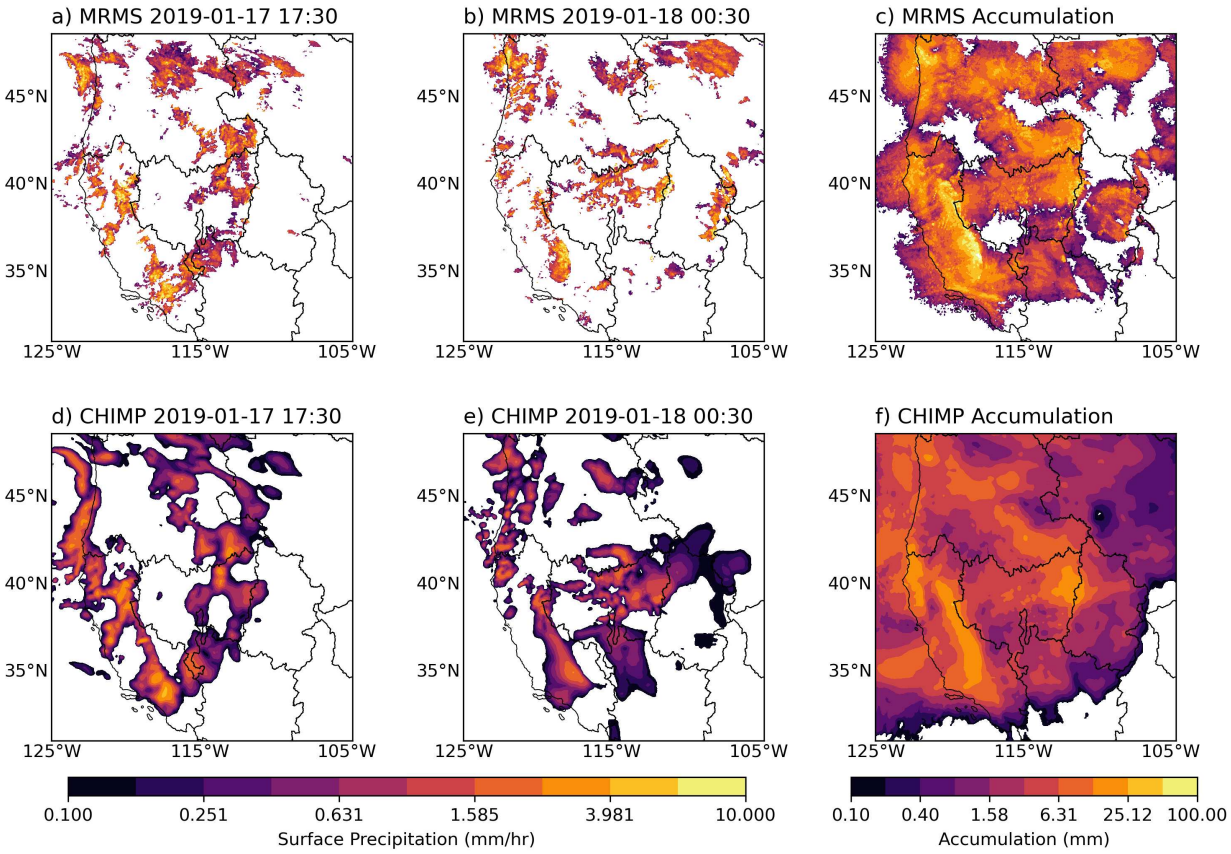

Fig. 4.3. WUS January 17-18, 2019 case for (a), (b), (c) MRMS extracted reference precipitation according to Section 4.2.1 and (d), (e), (f) CHIMP retrieved precipitation. (a) and (d) correspond to a single time step on January 17 1730 UTC, (b) and (e) correspond to a single time step on January 18 0030 UTC, and (c) and (f) correspond to accumulation for the entire case.

Evaluation using a threshold precipitation rate of 10^{-3} mm/hr (the minimum threshold used during CHIMP training) degraded the statistics indicating CHIMP overestimates precipitation where MRMS indicates little or no precipitation. This results in an underestimation of CHIMP at higher precipitation rates that are more important for snow accumulation and is thought to be an artifact of the model setup. CHIMP uses the MRMS radar quality index (RQI) in order to mask pixels based on the specified RQI. For this study, CHIMP is permitted to use MRMS precipitation data for RQI values above 0.1 to allow the reference precipitation and UCLA-WUS-SR scaling to influence model training; if RQI is set too high, reference data in complex terrain would be masked and not used during training. An investigation into where CHIMP estimates non-zero precipitation and MRMS estimates zero precipitation shows that $\sim$60% of these points have an RQI value less

than 0.3. Given this, it is thought these gridpoints are largely composed of where CHIMP inputs have signal indicative of precipitation despite MRMS indicating no precipitation. Additionally, it is thought the quantile fractions could introduce a systematic retrieval bias. Currently, CHIMP uses 32 evenly spaced quantiles between 0 and 1, with the uppermost quantile being 0.96. Because the UCLA-WUS-SR scaling of the MRMS precipitation emphasizes higher precipitation values, limiting the quantile estimation to 0.96 may reduce the expected precipitation rate estimate by the CHIMP quantiles. This hypothesis is currently being implemented and tested.

Additionally, the quantile fractions, or the spacing of the quantiles, is thought to cause a systematic bias. Currently, CHIMP implements quantiles that are linearly spaced between 0 and 1 using 32 quantiles resulting in a highest quantile of 0.96. The UCLA-WUS-SR scaling of the MRMS database increases the higher end of the precipitation distribution such that truncating the estimated quantiles at a value of 0.96 results in a smaller expected precipitation rate retrieved by CHIMP. This hypothesis is currently being implemented and tested.

In order to account for the systematic bias in the CHIMP model discussed above, a quantile mapping (QM) approach is taken to adjust the distribution of the CHIMP precipitation to better resemble the reference precipitation. QM is a widely used bias correction technique that has been extensively applied to precipitation for hydrologic studies (Teutschbein and Seibert 2012; Chen et al. 2013; Abbas and Xuan 2019). The QM procedure uses a blending approach that adjusts the model outputs to the MRMS reference precipitation while also preserving aspects of the CHIMP model outputs using Equation 4.1,

$$\mathrm{CHIMP}_{\mathrm{QM}}(t) = \alpha \cdot \mathrm{MRMS}_{\mathrm{obs}}(t) + (1-\alpha) \cdot \mathrm{CHIMP}_{\mathrm{nominal}}(t), \tag{4.1}$$

where α is the blending parameter. A blending parameter between 0.15 and 0.30 was found to produce realistic snowpack characteristics without over reliance on the MRMS reference data. This study uses a blending parameter value of 0.23. Snow accumulation and evolution are sensitive to input snowfall; an underestimation can lead to earlier snowmelt and surface warming that has a positive feedback that promotes accelerated melting of any additional snowfall. Figure 4.4 shows the distributions of the CHIMP and MRMS precipitation for a 6 month period of October 2020

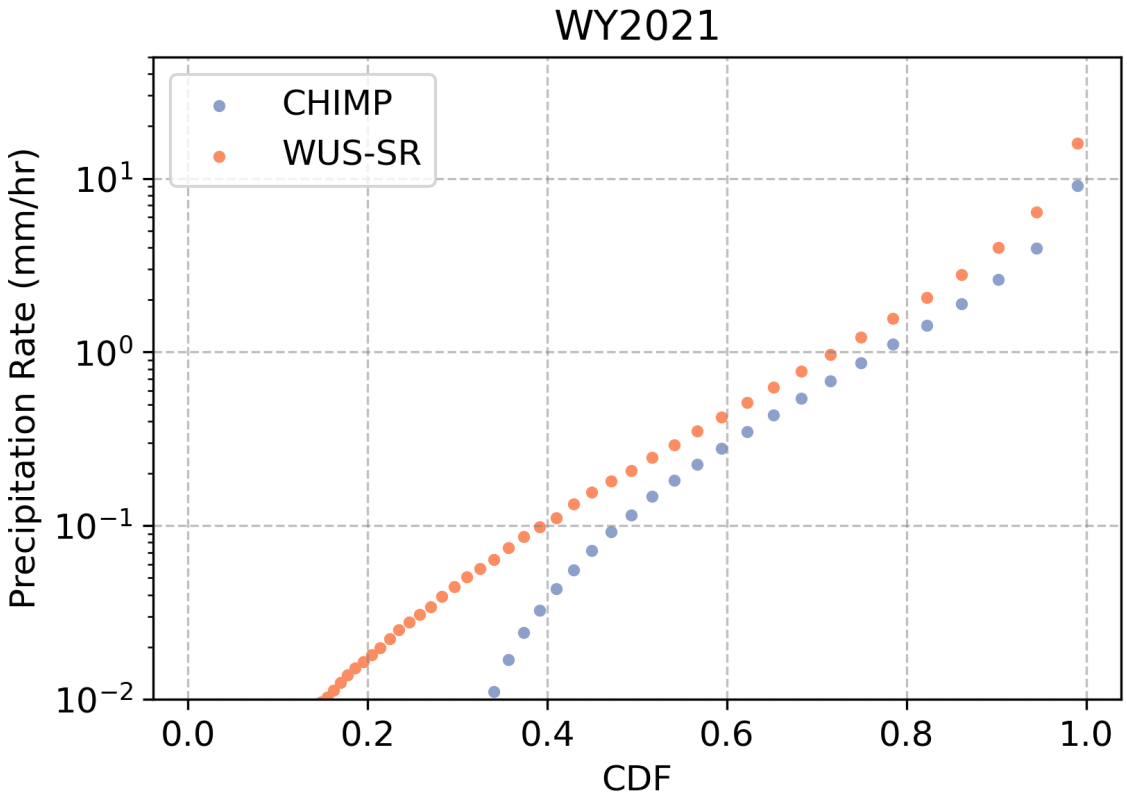

Fig. 4.4. Empirical CDFs of the reference MRMS precipitation and CHIMP retrieved precipitation for WY2021.

- March 2021. This time period was chosen to broadly cover the snow accumulation season. A nonlinear bias is observed for different quantiles of the distribution with the highest bias observed around the 10th percentile of the distribution. The QM procedure used here adjusts these non-uniform biases ensuring the adjusted CHIMP precipitation estimates better match the statistical distribution of the reference MRMS precipitation data. Additionally, the QM procedure adjusts the magnitude of CHIMP's precipitation estimates without altering the spatiotemporal patterns, a critical aspect of precipitation retrievals. This adjustment ultimately provides more realistic CHIMP precipitation inputs to SnowModel that, without the adjustment, would significantly underestimate snowfall accumulation. It should be noted the adjustment is possible due to the availability of reference MRMS data for WY2021. The biases observed in Figure 4.4 are likely not stationary across WYs due to the interannual variability of snowfall and cannot be applied generally.

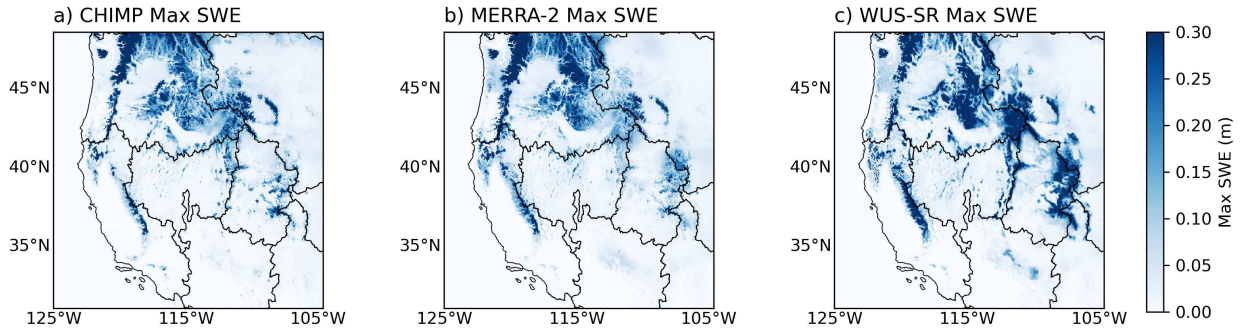

Fig. 4.5. SnowModel gridpoint peak SWE during WY2021 for (a) CHIMP-SM, (b) MERRA-2, and (c) WUS-SR.

4.3.2 SnowModel Simulations

The following section presents the results of the three SnowModel simulations described in Section 4.2.3. The spatial distribution of SWE is important when assessing snow water storage and for most of the WUS domain snow accumulates and completely melts during a single water year. Given this seasonal signature, peak SWE serves as an indicator of the maximum amount of water contained within the snowpack. Figure 4.5 shows the peak SWE for each SnowModel run with the HUC-2 watersheds within the WUS domain overlayed. The three models exhibit the expected spatial patterns of SWE with the largest SWE observed in the mountains. The WUS-SR estimates the largest max SWE values for all of the watersheds especially for the Upper Colorado, the Rocky Mountains within the Pacific Northwest, and California watershed. MERRA-2 estimates the largest max SWE in the Cascade mountains in the Pacific Northwest. CHIMP-SM max SWE estimates more closely resemble the MERRA-2 estimates, but does have some similarities to the WUS-SR estimates in the Cascades and Missouri watershed.

Peak SWE timing is driven by precipitation inputs as well as topographic features and shows significant variability across the WUS domain. Figure 4.6 shows the day-of-water-year (DOWY) that corresponds to the estimated peak SWE. For reference, the DOWY 100, 130, and 180 correspond to January 8th, February 7th, and March 29, respectively. Despite the differences seen in the magnitude of peak SWE between the SnowModel runs, the timing of peak SWE is quite similar indicating CHIMP-SM is able to accurately reproduce the timing of precipitation events with sufficient magnitude as to not produce unrealistic timing of peak SWE. For example, if CHIMP-SM

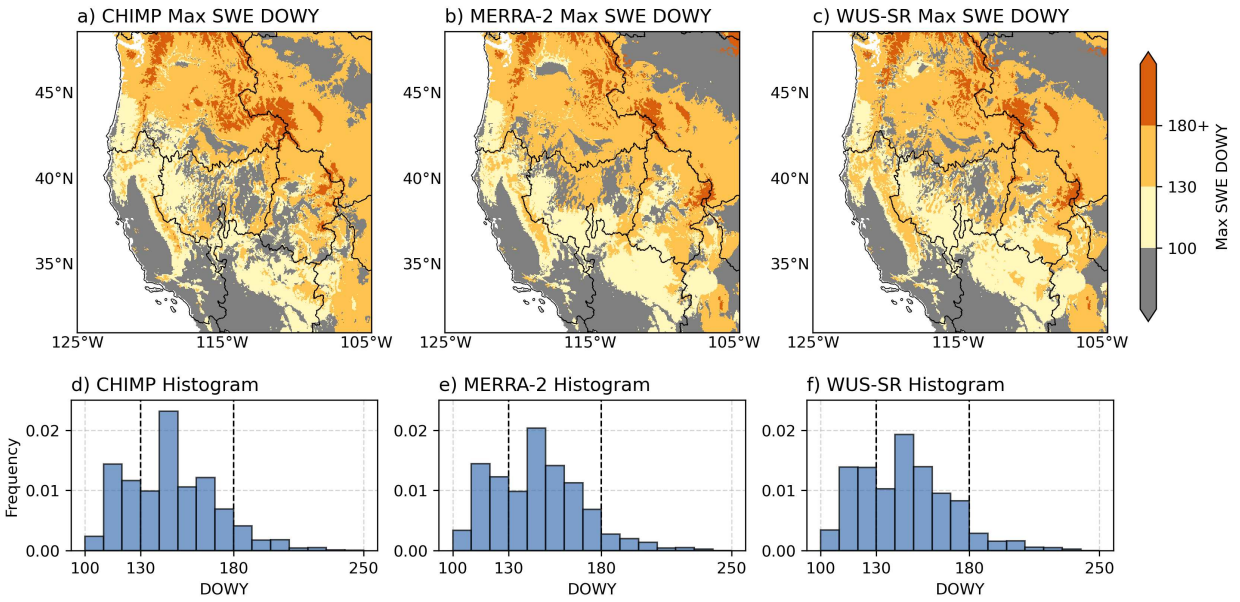

Fig. 4.6. SnowModel peak SWE day-of-water-year (DOWY) and corresponding histogram for (a),(d) CHIMP-SM , (b),(e) MERRA-2, and (c),(f) WUS-SR. The DOWY 100, 130, and 180 correspond to January 8th, February 7th, and March 29, respectively. Land gridpoints with a peak SWE DOWY less than 100 are masked out (gray).

significantly underestimated precipitation compared to MERRA-2 and WUS-SR then the snowpack would be thinner leading to earlier snow melt and earlier peak SWE timing. Some notable differences in CHIMP-SM peak SWE timing are observed in the Great Basin, Upper Colorado, and Lower Colorado watersheds where CHIMP-SM appears to have difficulties. Gridpoints with a peak SWE DOWY greater than 180 are indicative of deep snowpacks and are observed for gridpoints in the Rocky Mountains as well as the norther Cascade mountains. The CHIMP-SM peak SWE DOWY has 24% (6%) less gridpoints for DOWY greater than 180 in the Upper Colorado watershed and 14% (11%) less gridpoints for DOWY greater than 180 in the Missouri watershed compared to WUS-SR (MERRA-2). This result is an example of the difficulty CHIMP has in producing deep snowpacks in these two watersheds. On the other hand, CHIMP overestimates the number of gridpoints with a peak SWE DOWY greater than 180 in the eastern part of the Pacific Northwest watershed by $\sim$25% compared to both MERRA-2 and WUS-SR.

The integrated SWE volume, or snow water storage (SWS), is defined as the summation of SWE over all pixels multiplied by the spatial area of those pixels and is typically presented in units of cubic kilometers. It is a popular metric used to describe the total amount of water held

in the snowpack for watersheds or mountain ranges. Comparisons of the daily SWS for each SnowModel run are shown in Figure 4.7a with the SWS separated into HUC-2 watersheds in Figure 4.7b,c,d. Table 4.4 shows additional peak SWS statistics for the domain and watersheds as well as comparisons to the WUS-SR Snowmodel run. For WY2021, the Pacific Northwest watershed is the largest contributor to total SWS for all of the SnowModel runs. CHIMP-SM is able to reproduce the temporal variation of SWS for the domain and individual watersheds albeit with up to 59% underestimation observed in the Upper Colorado watershed. This underestimation is a common theme for CHIMP-SM across all watersheds ranging between 1% and 59%. CHIMP-SM has the best performance relative to WUS-SR in the Missouri watershed where less complex terrain is thought to allow the CHIMP-SM retrieval to have better performance. Besides the Upper Colorado watershed, CHIMP-SM is able to reproduce the spatial distribution of SWS relatively well across the domain as seen in column 5 of Table 4.4.

Snowfall and snow accumulation are strongly affected by elevation, with higher elevation more likely to experience colder temperatures, leading to more snow and deeper snowpacks. The impact of elevation is assessed as a cause of the underestimation of SWE and SWS observed in CHIMP-SM compared to WUS-SR (Figure 4.8). For each watershed, the elevation profile is divided into elevation bins that represent every 5th-percentile of the distribution in order to illustrate the differences in SWE across elevation bands similar to Smith and Bookhagen (2018); Liu et al. (2021). Across all watersheds and model runs, there is a non-linear relationship between elevation and SWE that is dependent upon the watershed's elevation profile and meteorological influences. The local maximum in SWE at elevations around 1000 m in the California and Pacific Northwest watersheds is indicative of orographic enhancement of precipitation that is not observed in the other watersheds. The lack of orographic signal in the other watersheds is likely an artifact of the 5 km gridsize of SnowModel rather than no orographic enhancement at all.

The overestimation of MERRA-2 compared to WUS-SR in the Pacific Northwest shown in Table 4.4 is dominated by elevations below $\sim$1500 m. The elevation-SWE relationships between CHIMP-SM and MERRA-2 are relatively similar for all watersheds except the California, Pacific Northwest, and Lower Colorado watersheds. The underestimation of CHIMP-SM compared to

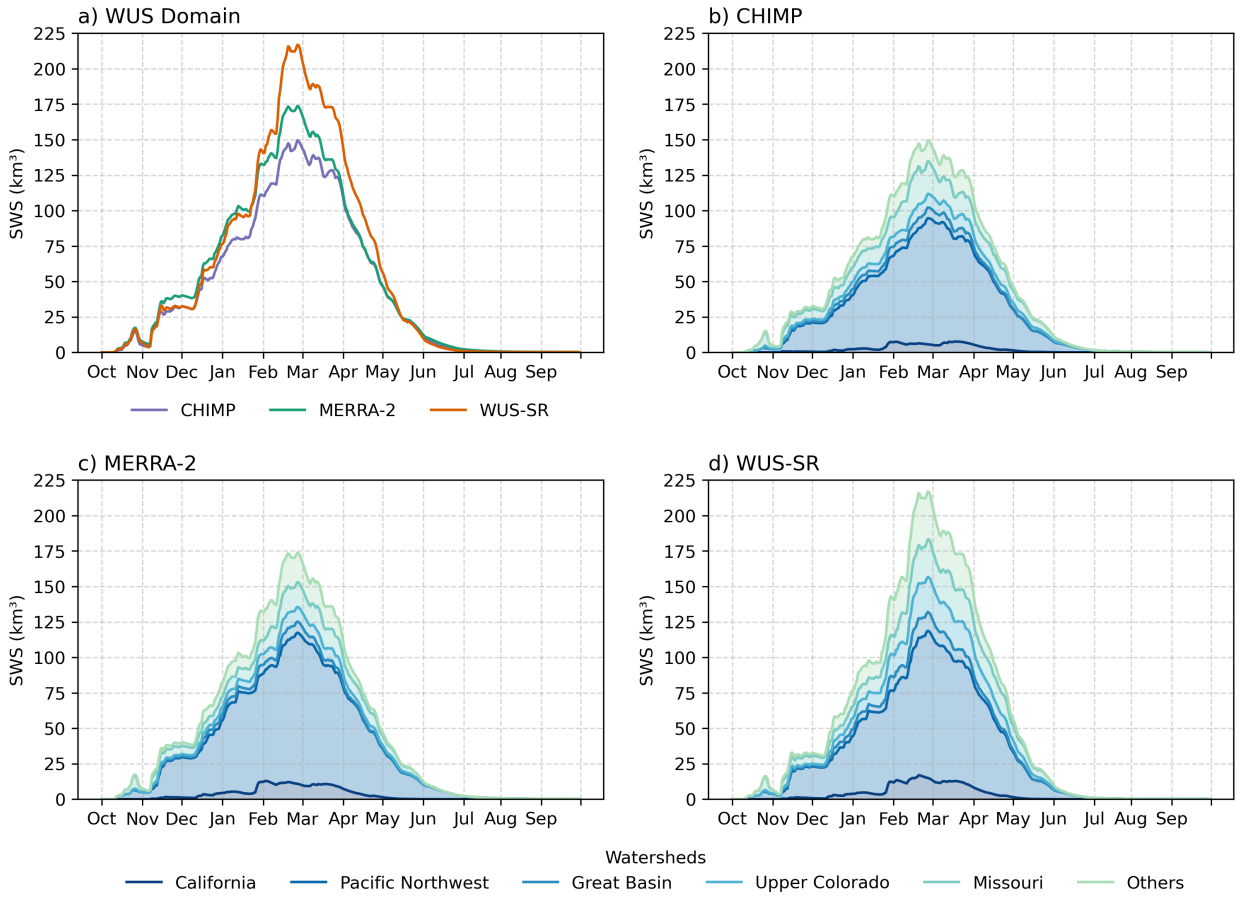

Fig. 4.7. WY2021 snow water storage (SWS) for the (a) WUS domain and cumulative SWS of individual watersheds for (b) CHIMP-SM, (c) MERRA-2, and (d) WUS-SR.

MERRA-2 is mostly seen in the fourth quartile of elevation. An interesting result from this analysis is the significant increase in SWE for elevations above the 80th percentile for most watersheds with the increase occurring at the 40th percentile of elevation for the Upper Colorado watershed. This phenomenon could be the result of the nominal resolution of the UCLA-WUS-SR product that has a resolution of $\sim$500 m that can resolve higher elevations. The underestimation of CHIMP-SM and MERRA-2 SWE compared to WUS-SR predominantly comes from the highest quartile of elevation with underestimation decreasing as elevation decreases.

Figure 4.9a shows the SWE time series of each SnowModel run for gridpoints within the Upper Colorado watershed where max SWE is greater than 2.55 cm (0.0255 m). The average SWE time series (bold line) and day of initial melt (vertical dashed line) are also shown. The individual

Table 4.4. Summary statistics of snow water storage for the WUS domain and HUC-2 watersheds.

Watershed	SnowModel Run	Peak SWS (km^3)	Error (%)	Watershed Percentage (%)
Total	MERRA-2	180.32	-19.04	100
	WUS-SR	222.71	-	100
	CHIMP-SM	157.19	-29.42	100
California	MERRA-2	12.94	-23.29	7.18
	WUS-SR	16.87	-	7.58
	CHIMP-SM	7.73	-54.18	4.92
Pacific Northwest	MERRA-2	106.45	2.58	59.03
	WUS-SR	103.77	-	46.60
	CHIMP-SM	88.92	-14.31	56.57
Great Basin	MERRA-2	8.35	-39.11	4.63
	WUS-SR	13.72	-	6.16
	CHIMP-SM	7.55	-44.98	4.80
Upper Colorado	MERRA-2	10.81	-56.33	6.00
	WUS-SR	24.76	-	11.12
	CHIMP-SM	10.01	-59.59	6.37
Missouri	MERRA-2	18.75	-30.40	10.40
	WUS-SR	26.95	-	12.10
	CHIMP-SM	26.54	-1.53	16.88
Others	MERRA-2	23.01	-37.21	12.76
	WUS-SR	36.64	-	16.45
	CHIMP-SM	16.45	-55.11	10.46

time series are a mixture of gridpoints with distinct accumulation and ablation seasons, intermittent ablation after a distinct accumulation season, and isolated accumulation and ablation that is synonymous with single storms for lower elevation gridpoints. Figure 4.9b shows an example of the time series of an individual gridpoint that shows a distinct accumulation and ablation. For this example, and on average, the underestimation of SWE results in an earlier initial melt out date, defined as the first date SWE disappears. Some gridpoints for CHIMP-SM, because of the underestimation, completely melt out during the accumulation season and then have difficulty accumulating additional SWE given the positive feedback a snow free surface has on the surface temperature. On average, MERRA-2 and CHIMP-SM begin to underestimate SWE starting in mid-December and continue to underestimate SWE throughout the remainder of the WY, which leads to an earlier

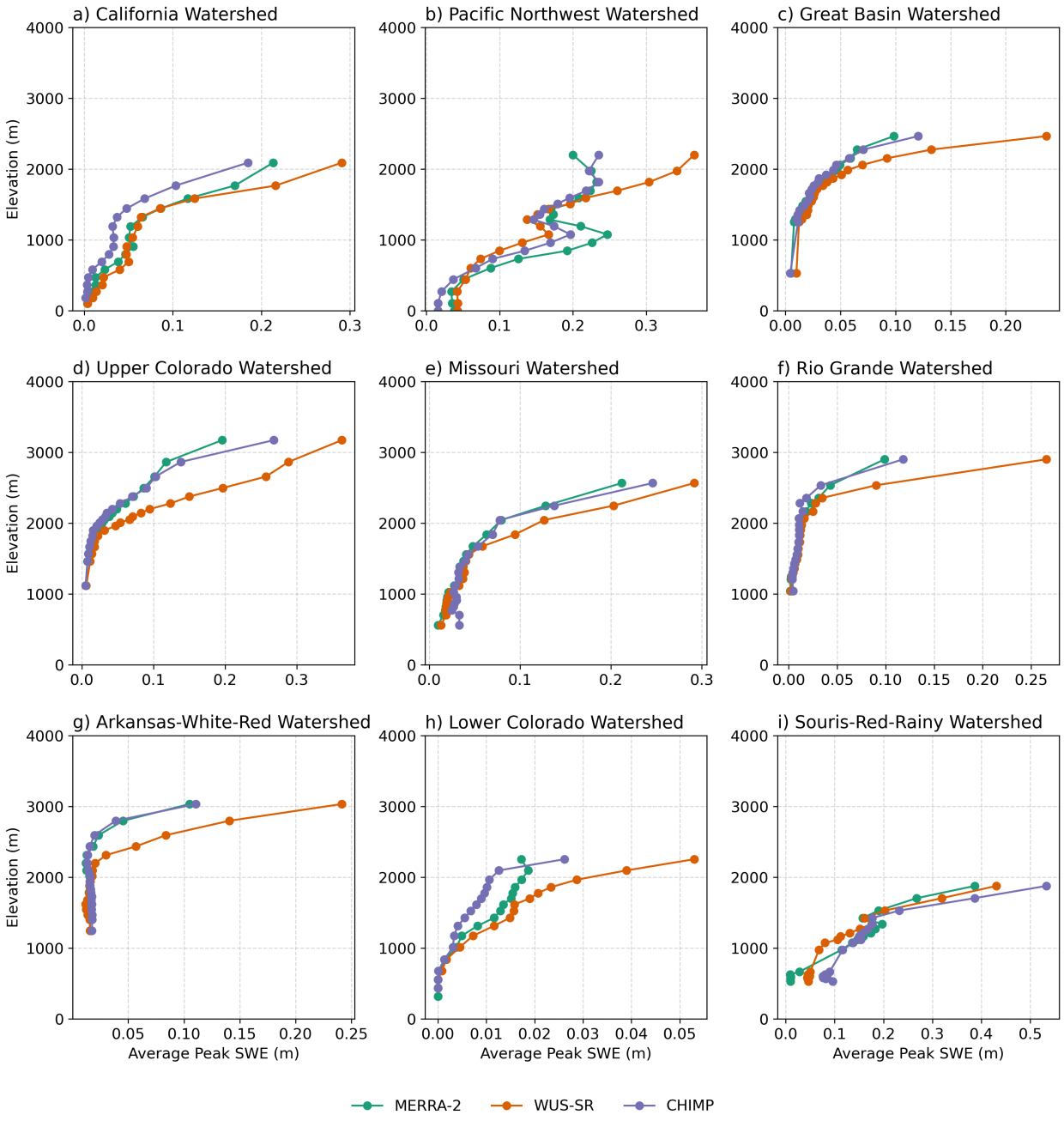

Fig. 4.8. Average peak SWE in elevation bands for the three SnowModel runs for the HUC-2 watersheds. The watershed elevation profile is separated into 20 elevation bins based on 5-percentile increments of elevation.

initial melt out date of 11 (1) days for MERRA-2 (CHIMP-SM) compared to WUS-SR. The mean initial melt out date for MERRA-2, WUS-SR, and CHIMP-SM are April 4th, April 15th, and April 1st. respectively. The earlier melt out is ultimately caused by errors in the input precipitation that SnowModel propagates into the snowpack evolution resulting in a compounding of errors.

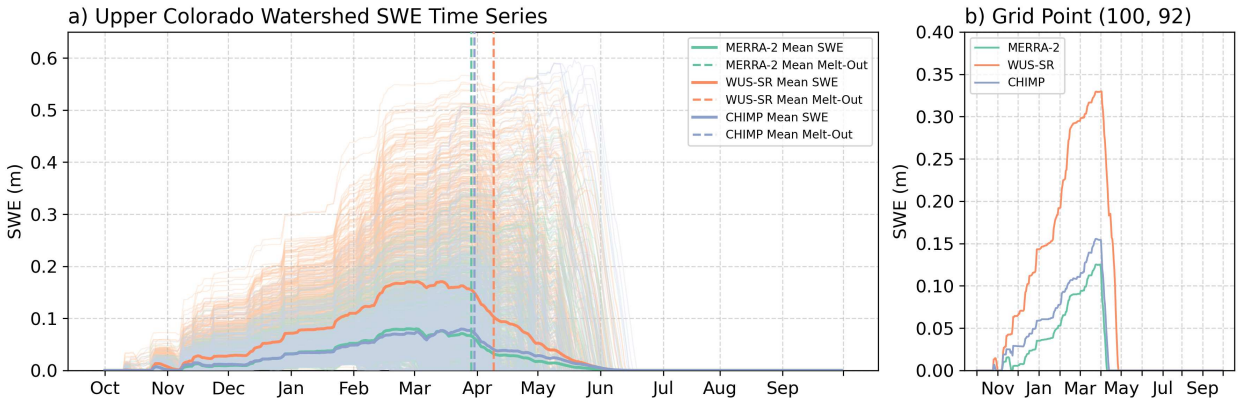

Fig. 4.9. (a) Individual SnowModel gridpoint SWE time series for the Upper Colorado watershed. The darker, thick lines show the average SWE for each SnowModel run with the vertical dashed lines indicating the initial SWE free day. (b) Single gridpoint SWE time series to highlight earlier SWE melt out due to SWE underestimation.

4.4 Discussion and Conclusion

The first research question, “Are satellite precipitation estimates suitable for use as input to a snow model in the western United States?", was answered by employing a machine learning precipitation retrieval algorithm, CHIMP, that is trained on PMW inputs from AMSR2, ATMS, GMI, MHS, and SSMIS and IR input from CPCIR. This model is able to estimate precipitation at a time resolution of 30 minutes, which is more than sufficient to be used as forcing data to large scale snow models (e.g. SnowModel) that are typically tun with a time step of 3 hours or even up to a daily time step. The ability to create a user-specific domain is especially intriguing from a snow modeling point of view as research is typically focused on different watershed scales or for specific locations. Therefore, from a technical point of view, the CHIMP model setup is considered suitable for use with SnowModel.

From an applied perspective, the nominal CHIMP precipitation estimates are currently not considered suitable for use with SnowModel. For this study, the reference precipitation used by CHIMP comes from MRMS with a climatological scaling applied to snowing gridpoints as defined by the MRMS precipitation type. Despite the adjustments made to the reference dataset, there are errors clearly persist in the retrieval model that result in underestimation of precipitation which propagate into later stages of the snow lifecycle. These biases can only be removed

because of the availability of reference precipitation data during the study period, but cannot be applied generally for other water years. Even when reference precipitation is available, the CHIMP precipitation estimates likely suffer from poor effective spatial resolution that is not able to resolve the finer scale precipitation distributions. The concept of effective resolution describes the smallest spatial scale at which a retrieval algorithm can accurately replicate the spatial variability of precipitation.Guilloteau et al. (2017) showed GPROF retrievals can accurately represent spatial scales coarser than 80km for ocean and land gridpoints, except for mountainous regions. Even then, GPROF is only able to represent precipitation spatial variability on the order of 40-80 km over land. In order for CHIMP precipitation estimates to be suitable for from both a technical and applied point of view, the CHIMP precipitation estimates require higher spatial resolution information in order to better guide the retrieval. Furthermore, CHIMP models could also be trained for targeted watersheds or mountain ranges in order to focus on regional scale patterns in the input data.

To account for the systematic bias between the nominal CHIMP estimates and the reference MRMS precipitation, a quantile mapping procedure was implemented that bias corrects quantiles of the precipitation distribution. This adjustment was performed in such a way that blends the distributions in order to retain aspects of the CHIMP precipitation estimates. This bias corrected precipitation estimates ultimately produced realistic snowpacks when used as input precipitation forcing data to SnowModel.

The second research question, “What are the associated errors in modeled snowpack across different watersheds when using satellite precipitation estimates as inputs?”, was answered by processing and ingesting the quantile mapped CHIMP precipitation estimates into SnowModel. The results were evaluated for snow statistics across the domain and HUC-2 watersheds against two additional SnowModel runs forced with MERRA-2 and an assimilated WUS-SR SnowModel run. An ongoing theme throughout the evaluation of the CHIMP forced SnowModel is underestimation of SWE compared to WUS-SR and to a lesser extent MERRA-2. CHIMP-SM is able to reproduce the expected spatial distribution of max SWE across the domain with notable underestimation in across the Rocky Mountains and California watershed. Despite the noted underestimation, CHIMP

produces a similar distribution of the timing of peak SWE. For gridpoints with deeper snowpacks, as defined by peak SWE DOWY greater than 180, CHIMP-SM contains 15% and 11% less gridpoints than WUS-SR and MERRA-2, respectively, excluding the Pacific Northwest, indicating CHIMP-SM has difficulty estimating deeper snowpacks. CHIMP underestimates domain-wide SWS by 30% with underestimation in individual watersheds between 1 - 55%. MERRA-2 shows a similar pattern of underestimation for all watersheds except for the Pacific Northwest. The distribution of SWE across elevation zones was assessed for each watershed. The upper quartile of elevation bins was found to be the biggest contributor to the underestimation of SWE and associated SWS with the upper half of elevation bins causing the underestimation in the Upper Colorado and California watersheds. The elevation-SWE dependence could be solved with changes to the CHIMP retrieval model discussed above that would likely lead to improved comparisons with MERRA-2 and WUS-SR.

An agreement in snow statistics such as peak SWE or SWS, for example, does not indicate the estimates are correct. In fact, there are multiple studies showing similar variability and biases of SWE and SWS for a variety of regional and global datasets (Dutra et al. 2011; Broxton et al. 2016b; Dawson et al. 2016; Wrzesien et al. 2017, 2018). The framework presented here allows for the addition of snow models forced with satellite precipitation estimates to be compared in parallel to global and regional models given the variability between widely used model products. The framework can be easily adapted to different study regions with appropriate care taken for the reference precipitation used. The merging of remote sensing and model data at higher resolutions could be a key area for future research to improve the initial precipitation retrieval estimates that would lead to improved snowpack representation.

CHAPTER 5

Conclusions

As discussed in Chapter 1, snowfall serves as the primary source of input for snowpack accumulation processes, but is the most difficult to estimate using passive microwave retrieval algorithms. Mountainous regions, characterized by deep snowpacks, present significant challenges for accurate snowfall estimation, especially in terms of snowfall accumulation. The goal of this dissertation was to use a data-driven approach to improve GPROF-based retrievals of snowfall with a focus on mountain snowfall. Evaluating how well satellite-based snowfall estimates can capture snowpack characteristics is crucial for integrating snowfall into the broader context of the snow lifecycle.

Chapter 2 evaluated three GPROF-based retrievals of snowfall over snow-covered surfaces. The neural network versions, GPROF-NN 1D and GPROF-NN 3D, outperformed the Bayesian version, GPROF V7, in terms of snowfall detection and quantitative metrics. GPROF V7 exhibited a clear dependence on surface snow class, which decreased for GPROF-NN 1D and GPROF-NN 3D. The three models were again evaluated using a different reference snowfall dataset that used a temperature-dependent Z-S relationship to estimate snowfall. The Z-S was derived using observations of in-situ snowfall and temperature from USCRN stations combined with radar reflectivity from MRMS. Up to an order of magnitude of spread in snowfall was found for a given reflectivity, highlighting the variability of snowfall characteristics. Surface temperature data from USCRN stations was used to demonstrate the temperature dependence on the functional form of the Z-S equation. Retroactively incorporating the temperature-dependent Z-S into the MRMS data resulted in improved quantitative snowfall accumulation statistics between USCRN and MRMS as well as improved detection statistics in the GPROF algorithms.

The focus of Chapter 3 shifted towards evaluation of mountain snowfall from the three GPROF retrieval algorithms. Each retrieval model was run on two different retrieval databases and assessed on their ability to reproduce water year snowfall magnitudes in the western United States. It was

observed that the nominal MRMS data does not reliably reproduce the spatial patterns or magnitudes in water year snowfall accumulation, resulting in similar issues in the retrieval algorithm estimates. Across the entire study domain, MRMS required an average correction factor of 1.5, though this value varied considerably between gridpoints due to radar location and proximity to mountainous terrain. By incorporating better representations of the water year snowfall into the reference snowfall data using a SWE-Reanalysis dataset, significant improvements are observed in the GPROF algorithms. Compared to GPROF V7, the GPROF-NN models showed improved performance in representing interannual snowfall accumulation variability, albeit with regional variations. Differences are observed in the models' ability to reproduce interannual variability between the coastal mountain ranges and intermountain Rocky Mountains hypothesized to be driven by differences in synoptic driven snowfall events. The evaluation of the GPROF-NN 3D model, trained in the western U.S. and tested in the Andes mountains, revealed the need for more robust snow climatologies for global mountain ranges.

Chapter 4 aimed to assess the ability of satellite-based snowfall estimates in depicting snow water equivalent (SWE). The objective of this study is to present an approach that can be complimentary to or even replace the direct estimation of SWE from passive microwave sensors, specifically AMSR-E. In order to achieve this, a UNet-based precipitation retrieval algorithm, CHIMP, was coupled with a snow evolution model, SnowModel. This method provided a straightforward way to integrate satellite precipitation estimates into SnowModel, given the flexibility to specify the study domain and projection for both CHIMP and SnowModel. This approach proved to be useful, but was limited in terms of application that stemmed from a systematic underestimation in precipitation from the retrieval model that was likely introduced from the MRMS quality index or the quantile loss function used in the CHIMP model. Even after the retrieval was bias corrected, the CHIMP forced SnowModel was unable to reproduce SWE magnitudes in the western United States. The underestimation in SWE resulted from the inability of CHIMP to adequately capture snowfall at higher elevations that leads to lower snowfall accumulations and faster snow melt out. While this dissertation falls short of being able to use satellite snowfall estimates to estimate

SWE for mountainous terrain, it provides approaches that can be used to address the remaining challenges.

Several potential directions for future work emerge from the findings of these studies. It would be interesting to extend the Z-S derivation between USCRN and MRMS to additional study years or even different in-situ stations. The difficulty of this is, however, the lack of proper station configuration for accurate observations of snowfall due to wind-driven undercatch issues. One such candidate dataset could be the World Meteorological Organization Solid Precipitation Intercomparison Experiment (WMO-SPICE) that has fewer stations compared to USCRN, but covers a wider range of climate zones. Critical to each Chapter of the dissertation was the inclusion of additional information to the retrieval algorithms through the reference precipitation. Chapter 2 used a temperature-dependent Z-S and Chapter 3 and Chapter 4 included adjustments to snowfall accumulations. In Chapter 4 especially, the adjustment to the reference precipitation did not provide enough information content to the retrieval. Higher resolution data such as repeatable snow patterns may need to be incorporated as well. And finally, as discussed at the end of Chapter 3, extension of the retrieval algorithms presented here to global mountains would be an important step for global snow quantification from satellite sensors. The SWE-Reanalysis currently is only available for the western U.S., Andes, and High Mountain Asia, and it is not straightforward how to incorporate these data into the reference precipitation without accompanying ground based radars currently used in GPROF retrieval databases.

REFERENCES

Abatzoglou, J. T., R. Barbero, J. W. Wolf, and Z. A. Holden, 2014: Tracking Interannual Streamflow Variability with Drought Indices in the U.S. Pacific Northwest. *Journal of Hydrometeorology*, doi: 10.1175/JHM-D-13-0167.1.

Abbas, S. A., and Y. Xuan, 2019: Development of a New Quantile-Based Method for the Assessment of Regional Water Resources in a Highly-Regulated River Basin. *Water Resources Management*, **33 (9)**, 3187–3210, doi: 10.1007/s11269-019-02290-z.

Adler, R. F., A. Negri, P. Keehn, and I. Hakkarinen, 1992: Estimation of Monthly Rainfall over Japan and Surrounding Waters from a Combination of Low-Orbit Microwave and Geosynchronous IR Data. *JOURNAL OF APPLIED METEOROLOGY*, **32**, 335–356.

Andreadis, K. M., and D. P. Lettenmaier, 2006: Assimilating remotely sensed snow observations into a macroscale hydrology model. *Advances in Water Resources*, **29 (6)**, 872–886, doi: 10.1016/j.advwatres.2005.08.004.

Aonashi, K., Y. Shoji, H. Fujii, T. Koike, S. Shimizu, K. Nakamura, Y.-J. Noh, and G. Liu, 2003: The structural and microphysical characteristicsof solid precipitation derived from Wakasa Bay field experiment. *Fall Meeting, Meteorol. Soc. of Jpn.*, Sapporo, Japan.

Arsenault, K. R., and P. R. Houser, 2018: Generating Observation-Based Snow Depletion Curves for Use in Snow Cover Data Assimilation. *Geosciences*, **8 (12)**, 484, doi: 10.3390/geosciences8120484.

Barnett, T. P., J. C. Adam, and D. P. Lettenmaier, 2005: Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, **438 (7066)**, 303–309, doi: 10.1038/nature04141.

Battaglia, A., and Coauthors, 2020: Spaceborne Cloud and Precipitation Radars: Status, Challenges, and Ways Forward. *Reviews of Geophysics*, **58 (3)**, e2019RG000 686, doi: 10.1029/2019RG000686.

Besso, H., D. Shean, and J. D. Lundquist, 2024: Mountain snow depth retrievals from customized processing of ICESat-2 satellite laser altimetry. *Remote Sensing of Environment*, **300**, 113 843, doi: 10.1016/j.rse.2023.113843.

Betts, A. K., R. Desjardins, D. Worth, S. Wang, and J. Li, 2014: Coupling of winter climate transitions to snow and clouds over the Prairies. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013JD021168, doi: 10.1002/2013JD021168.

Bormann, K. J., R. D. Brown, C. Derksen, and T. H. Painter, 2018: Estimating snow-cover trends from space. *Nature Climate Change*, **8 (11)**, 924–928, doi: 10.1038/s41558-018-0318-3.

Brauchli, T., E. Trujillo, H. Huwald, and M. Lehning, 2017: Influence of Slope-Scale Snowmelt on Catchment Response Simulated With the Alpine3D Model. *Water Resources Research*, **53 (12)**, 10 723–10 739, doi: 10.1002/2017WR021278.

Broxton, P. D., N. Dawson, and X. Zeng, 2016a: Linking snowfall and snow accumulation to generate spatial maps of SWE and snow depth: USING SNOWFALL TO INTERPOLATE SWE. *Earth and Space Science*, **3 (6)**, 246–256, doi: 10.1002/2016EA000174.

Broxton, P. D., X. Zeng, and N. Dawson, 2016b: Why Do Global Reanalyses and Land Data Assimilation Products Underestimate Snow Water Equivalent? *Journal of Hydrometeorology*, **17 (11)**, 2743–2761, doi: 10.1175/JHM-D-16-0056.1.

Bukovčić, P., A. Ryzhkov, D. Zrnić, and G. Zhang, 2018: Polarimetric Radar Relations for Quantification of Snow Based on Disdrometer Data. *Journal of Applied Meteorology and Climatology*, **57 (1)**, 103–120, doi: 10.1175/JAMC-D-17-0090.1.

Cao, Q., T. H. Painter, W. R. Currier, J. D. Lundquist, and D. P. Lettenmaier, 2018: Estimation of Precipitation over the OLYMPEX Domain during Winter 2015/16. *Journal of Hydrometeorology*, doi: 10.1175/JHM-D-17-0076.1.

Casella, D., G. Panegrossi, P. Sanò, A. C. Marra, S. Dietrich, B. T. Johnson, and M. S. Kulie, 2017: Evaluation of the GPM-DPR snowfall detection capability: Comparison with CloudSat-CPR. *Atmospheric Research*, **197**, 64–75, doi: 10.1016/j.atmosres.2017.06.018.

Chen, J., F. P. Brissette, D. Chaumont, and M. Braun, 2013: Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America. *Water Resources Research*, **49 (7)**, 4187–4205, doi: 10.1002/wrcr.20331.

Cho, E., C. M. Vuyovich, S. V. Kumar, M. L. Wrzesien, R. S. Kim, and J. M. Jacobs, 2022a: Precipitation biases and snow physics limitations drive the uncertainties in macroscale modeled snow water equivalent. *Hydrology and Earth System Sciences*, **26 (22)**, 5721–5735, doi: 10.5194/hess-26-5721-2022.

Cho, E., C. M. Vuyovich, S. V. Kumar, M. L. Wrzesien, R. S. Kim, and J. M. Jacobs, 2022b: Precipitation biases and snow physics limitations drive the uncertainties in macroscale modeled snow water equivalent. *Hydrology and Earth System Sciences*, **26 (22)**, 5721–5735, doi: 10.5194/hess-26-5721-2022.

Cohen, J., and D. Rind, 1991: The Effect of Snow Cover on the Climate. *Journal of Climate*, **4 (7)**, 689–706, 26196419.

Cooper, W. A., 1986: Ice Initiation in Natural Clouds. *Meteorological Monographs*.

Cordisco, E., C. Prigent, and F. Aires, 2006: Snow characterization at a global scale with passive microwave satellite observations. *Journal of Geophysical Research: Atmospheres*, **111 (D19)**, doi: 10.1029/2005JD006773.

Daloz, A. S., and Coauthors, 2020: How much snow falls in the world's mountains? A first look at mountain snowfall estimates in A-train observations and reanalyses. *The Cryosphere*, **14 (9)**,

3195–3207, doi: 10.5194/tc-14-3195-2020.

Daly, C., R. Neilson, and D. Phillips, 1993: A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain. *Journal of Applied Meteorology*, **33**, 140–158.

Dawson, N., P. Broxton, and X. Zeng, 2018: Evaluation of Remotely Sensed Snow Water Equivalent and Snow Cover Extent over the Contiguous United States. *Journal of Hydrometeorology*, **19 (11)**, 1777–1791, doi: 10.1175/JHM-D-18-0007.1.

Dawson, N., P. Broxton, X. Zeng, M. Leuthold, M. Barlage, and P. Holbrook, 2016: An Evaluation of Snow Initializations in NCEP Global and Regional Forecasting Models. *Journal of Hydrometeorology*, doi: 10.1175/JHM-D-15-0227.1.

De Lannoy, G. J. M., R. H. Reichle, K. R. Arsenault, P. R. Houser, S. Kumar, N. E. C. Verhoest, and V. R. N. Pauwels, 2012: Multiscale assimilation of Advanced Microwave Scanning Radiometer–EOS snow water equivalent and Moderate Resolution Imaging Spectroradiometer snow cover fraction observations in northern Colorado. *Water Resources Research*, **48 (1)**, doi: 10.1029/2011WR010588.

Deschamps-Berger, C., S. Gascoin, D. Shean, H. Besso, A. Guiot, and J. I. López-Moreno, 2023: Evaluation of snow depth retrievals from ICESat-2 using airborne laser-scanning data. *The Cryosphere*, **17 (7)**, 2779–2792, doi: 10.5194/tc-17-2779-2023.

Diamond, H. J., and Coauthors, 2013: U.S. CLIMATE REFERENCE NETWORK AFTER ONE DECADE OF OPERATIONS. *Bulletin of the American Meteorological Society*, doi: DOI:10.1175 / BAMS -D -12- 0 0170.1.

Drusch, M., D. Vasiljevic, and P. Viterbo, 2004: ECMWF's Global Snow Analysis: Assessment and Revision Based on Satellite Observations. *Journal of Applied Meteorology and Climatology*.

Durand, M., N. P. Molotch, and S. A. Margulis, 2008: A Bayesian approach to snow water equivalent reconstruction. *Journal of Geophysical Research: Atmospheres*, **113 (D20)**, doi: 10.1029/2008JD009894.

Dutra, E., S. Kotlarski, P. Viterbo, G. Balsamo, P. M. A. Miranda, C. Schär, P. Bissolli, and T. Jonas, 2011: Snow cover sensitivity to horizontal resolution, parameterizations, and atmospheric forcing in a land surface model. *Journal of Geophysical Research: Atmospheres*, **116 (D21)**, doi: 10.1029/2011JD016061.

Eckert, E., and Coauthors, 2022: Validation of the Final Monthly Integrated Multisatellite Retrievals for GPM (IMERG) Version 05 and Version 06 with Ground-Based Precipitation Gauge Measurements across the Canadian Arctic. *Journal of Hydrometeorology*, doi: 10.1175/JHM-D-21-0040.1.

Essery, R., and J. Pomeroy, 2004: Implications of spatial distributions of snow mass and melt rate for snow-cover depletion: Theoretical considerations. *Annals of Glaciology*, **38**, 261–265, doi: 10.3189/172756404781815275.

Estilow, T., 2013: Climate Algorithm Theoretical Basis Document (C-ATBD) Northern Hemisphere Snow Cover Extent.

Fang, Y., Y. Liu, D. Li, H. Sun, and S. A. Margulis, 2023: Spatiotemporal snow water storage uncertainty in the midlatitude American Cordillera. *The Cryosphere*, **17 (12)**, 5175–5195, doi: 10.5194/tc-17-5175-2023.

Fang, Y., Y. Liu, and S. A. Margulis, 2022: A western United States snow reanalysis dataset over the Landsat era from water years 1985 to 2021. *Scientific Data*, **9 (1)**, 677, doi: 10.1038/s41597-022-01768-7.

Fletcher, S. J., G. E. Liston, C. A. Hiemstra, and S. D. Miller, 2012: Assimilating MODIS and AMSR-E Snow Observations in a Snow Evolution Model. *Journal of Hydrometeorology*, doi: 10.1175/JHM-D-11-082.1.

Gonzalez, R., and C. Kummerow, 2024: Improving Satellite Snowfall Magnitudes in the Intermountain West Using Machine Learning and a Western United States SWE-Reanalysis. *Journal of Hydrometeorology*.

Gonzalez, R., and C. D. Kummerow, 2020: AMSR-E Snow: Can Snowfall Help Improve SWE Estimates? *Journal of Hydrometeorology*, **21 (11)**, 2551–2564, doi: 10.1175/JHM-D-20-0066.1.

Guilloteau, C., E. Foufoula-Georgiou, and C. D. Kummerow, 2017: Global Multiscale Evaluation of Satellite Passive Microwave Retrieval of Precipitation during the TRMM and GPM Eras: Effective Resolution and Regional Diagnostics for Future Algorithm Development. *Journal of Hydrometeorology*, **18 (11)**, 3051–3070, doi: 10.1175/JHM-D-17-0087.1.

Günther, D., T. Marke, R. Essery, and U. Strasser, 2019: Uncertainties in Snowpack Simulations—Assessing the Impact of Model Structure, Parameter Choice, and Forcing Data Error on Point-Scale Energy Balance Snow Model Performance. *Water Resources Research*, **55 (4)**, 2779–2800, doi: 10.1029/2018WR023403.

Gutmann, E. D., R. M. Rasmussen, C. Liu, K. Ikeda, D. J. Gochis, M. P. Clark, J. Dudhia, and G. Thompson, 2012: A Comparison of Statistical and Dynamical Downscaling of Winter Precipitation over Complex Terrain. *Journal of Climate*, **25 (1)**, 262–281, doi: 10.1175/2011JCLI4109.1.

Hammond, J. C., and Coauthors, 2023: High Resolution SnowModel Simulations Reveal Future Elevation-Dependent Snow Loss and Earlier, Flashier Surface Water Input for the Upper Colorado River Basin. *Earth's Future*, **11 (2)**, e2022EF003 092, doi: 10.1029/2022EF003092.

Helfrich, S. R., D. McNamara, B. H. Ramsay, T. Baldwin, and T. Kasheta, 2007: Enhancements to, and forthcoming developments in the Interactive Multisensor Snow and Ice Mapping System (IMS). *Hydrological Processes*, **21 (12)**, 1576–1586, doi: 10.1002/hyp.6720.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146 (730)**, 1999–2049, doi: 10.1002/qj.3803.

Heymsfield, A., M. Krämer, N. B. Wood, A. Gettelman, P. R. Field, and G. Liu, 2017: Dependence of the Ice Water Content and Snowfall Rate on Temperature, Globally: Comparison of in Situ Observations, Satellite Active Remote Sensing Retrievals, and Global Climate Model Simulations. *Journal of Applied Meteorology and Climatology*, **56 (1)**, 189–215, doi: 10.1175/JAMC-D-16-0230.1.

Heymsfield, A. J., C. Schmitt, and A. Bansemer, 2013: Ice Cloud Particle Size Distributions and Pressure-Dependent Terminal Velocities from In Situ Observations at Temperatures from 0° to -86°C. *Journal of the Atmospheric Sciences*, **70 (12)**, 4123–4154, doi: 10.1175/JAS-D-12-0124.1.

Hilburn, K. A., I. Ebert-Uphoff, and S. D. Miller, 2021: Development and Interpretation of a Neural-Network-Based Synthetic Radar Reflectivity Estimator Using GOES-R Satellite Observations. *Journal of Applied Meteorology and Climatology*, **60 (1)**, 3–21, doi: 10.1175/JAMC-D-20-0084.1.

Hirahara, Y., P. de Rosnay, and G. Arduini, 2020: Evaluation of a Microwave Emissivity Module for Snow Covered Area with CMEM in the ECMWF Integrated Forecasting System. *Remote Sensing*, **12 (18)**, 2946, doi: 10.3390/rs12182946.

Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement Mission. *Bulletin of the American Meteorological Society*, **95 (5)**, 701–722, doi: 10.1175/BAMS-D-13-00164.1.

Huffman, G., D. Bolvin, R. Joyce, O. Kelley, E. Nelkin, J. Tan, D. Watters, and J. West, 2023: Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation. NASA Global Precipitation Measurement.

Huffman, G. J., and Coauthors, 1997: The Global Precipitation Climatology Project (GPCP) Combined Precipitation Dataset. *Bulletin of the American Meteorological Society*, **78 (1)**, 5–20, doi: 10.1175/1520-0477(1997)078<0005:TGPCPG>2.0.CO;2.

Huffman, G. J., and Coauthors, 2024: IMERG V07 Release Notes. NASA Global Precipitation Measurement.

Hughes, M., J. D. Lundquist, and B. Henn, 2024: Dynamical downscaling improves upon gridded precipitation products in the Sierra Nevada, California. *Climate Dynamics*.

Intergovernmental Panel On Climate Change (Ipcc), 2022: *The Ocean and Cryosphere in a Changing Climate: Special Report of the Intergovernmental Panel on Climate Change*. 1st ed., Cambridge University Press, doi: 10.1017/9781009157964.

Josberger, E. G., and N. M. Mognard, 2002: A passive microwave snow depth algorithm with a proxy for snow metamorphism. *Hydrological Processes*, **16 (8)**, 1557–1568, doi: 10.1002/hyp.1020.

Judson, A., and N. Doesken, 2000: Density of Freshly Fallen Snow in the Central Rocky Mountains. *Bulletin of the American Meteorological Society*, **81 (7)**, 1577–1587, doi: 10.1175/1520-0477(2000)081<1577:DOFFSI>2.3.CO;2.

Kapnick, S., and A. Hall, 2012: Causes of recent changes in western North American snowpack. *Climate Dynamics*, **38 (9)**, 1885–1899, doi: 10.1007/s00382-011-1089-y.

Kelly, R., 2009: The AMSR-E Snow Depth Algorithm: Description and Initial Results. *Journal of The Remote Sensing Society of Japan*, **29 (1)**, 307–317.

Kidd, C., D. R. Kniveton, M. C. Todd, and T. J. Bellerby, 2003: Satellite Rainfall Estimation Using Combined Passive Microwave and Infrared Algorithms. *Journal of Hydrometeorology*, **4 (6)**, 1088–1104, doi: 10.1175/1525-7541(2003)004<1088:SREUCP>2.0.CO;2.

Kirstetter, P., W. A. Petersen, and J. J. Gourley, 2018: GPM Ground Validation Multi-Radar/Multi-Sensor (MRMS) Precipitation Reanalysis for Satellite Validation Product V1. https://cmr.earthdata.nasa.gov/search/concepts/C1980955793-GHRC_DAAC.html.

Kirstetter, P.-E., J. J. Gourley, Y. Hong, J. Zhang, S. Moazamigoodarzi, C. Langston, and A. Arthur, 2015: Probabilistic precipitation rate estimates with ground-based radar networks. *Water Resources Research*, **51 (3)**, 1422–1442, doi: 10.1002/2014WR015672.

Komurcu, M., K. A. Emanuel, M. Huber, and R. P. Acosta, 2018: High-Resolution Climate Projections for the Northeastern United States Using Dynamical Downscaling at Convection-Permitting Scales. *Earth and Space Science*, **5 (11)**, 801–826, doi: 10.1029/2018EA000426.

Kongoli, C., H. Meng, J. Dong, and R. Ferraro, 2015: A snowfall detection algorithm over land utilizing high-frequency passive microwave measurements—Application to ATMS. *Journal of Geophysical Research: Atmospheres*, **120 (5)**, 1918–1932, doi: 10.1002/2014JD022427.

Kulie, M. S., R. Bennartz, T. J. Greenwald, Y. Chen, and F. Weng, 2010: Uncertainties in Microwave Properties of Frozen Precipitation: Implications for Remote Sensing and Data Assimilation. *Journal of the Atmospheric Sciences*, **67 (11)**, 3471–3487, doi: 10.1175/2010JAS3520.1.

Kummerow, C., and L. Giglio, 1995: A Method for Combining Passive Microwave and Infrared Rainfall Observations. *Journal of Atmospheric and Oceanic Technology*.

Kummerow, C. D., D. L. Randel, M. Kulie, N.-Y. Wang, R. Ferraro, S. Joseph Munchak, and V. Petkovic, 2015: The Evolution of the Goddard Profiling Algorithm to a Fully Parametric Scheme. *Journal of Atmospheric and Oceanic Technology*, **32 (12)**, 2265–2280, doi: 10.1175/JTECH-D-15-0039.1.

Kunkel, K. E., D. A. Robinson, S. Champion, X. Yin, T. Estilow, and R. M. Frankson, 2016: Trends and Extremes in Northern Hemisphere Snow Characteristics. *Current Climate Change Reports*, **2 (2)**, 65–73, doi: 10.1007/s40641-016-0036-8.

Kuo, K.-S., and Coauthors, 2016: The Microwave Radiative Properties of Falling Snow Derived from Nonspherical Ice Particle Models. Part I: An Extensive Database of Simulated Pristine Crystals and Aggregate Particles, and Their Scattering Properties. *Journal of Applied Meteorology and Climatology*, **55 (3)**, 691–708, doi: 10.1175/JAMC-D-15-0130.1.

Lee, J., M.-I. Lee, S. Tak, E. Seo, and Y.-K. Lee, 2023: Assimilation of snow water equivalent from AMSR2 and IMS satellite data utilizing the local ensemble transform Kalman filter. doi:

10.5194/gmd-2023-221.

Levizzani, V., and E. Cattani, 2019: Satellite Remote Sensing of Precipitation and the Terrestrial Water Cycle in a Changing Climate. *Remote Sensing*, **11 (19)**, 2301, doi: 10.3390/rs11192301.

Levizzani, V., S. Laviola, and E. Cattani, 2011: Detection and Measurement of Snowfall from Space. *Remote Sensing*, **3**, 145–166.

Li, D., M. L. Wrzesien, M. Durand, J. Adam, and D. P. Lettenmaier, 2017: How much runoff originates as snow in the western United States, and how will that change in the future? *Geophysical Research Letters*, **44 (12)**, 6163–6172, doi: 10.1002/2017GL073551.

Li, F., Y. J. Orsolini, N. Keenlyside, M.-L. Shen, F. Counillon, and Y. G. Wang, 2019: Impact of Snow Initialization in Subseasonal-to-Seasonal Winter Forecasts With the Norwegian Climate Prediction Model. *Journal of Geophysical Research: Atmospheres*, **124 (17-18)**, 10 033–10 048, doi: 10.1029/2019JD030903.

Libbrecht, K. G., 2007: The Formation of Snow Crystals. *American Scientist*, 9.

Liljedahl, A. K., L. D. Hinzman, D. L. Kane, W. C. Oechel, C. E. Tweedie, and D. Zona, 2017: Tundra water budget and implications of precipitation underestimation. *Water Resources Research*, **53 (8)**, 6472–6486, doi: 10.1002/2016WR020001.

Lindeman, J., 2022: Global Multisensor Automated Snow and Ice Mapping System (GMASI) AutoSnow External Users Manual. *NOAA*.

Liston, G. E., and K. Elder, 2006a: A Distributed Snow-Evolution Modeling System (SnowModel). *Journal of Hydrometeorology*, **7 (6)**, 1259–1276, doi: 10.1175/JHM548.1.

Liston, G. E., and K. Elder, 2006b: A Meteorological Distribution System for High-Resolution Terrestrial Modeling (MicroMet). *Journal of Hydrometeorology*, **7 (2)**, 217–234, doi: 10.1175/JHM486.1.

Liston, G. E., R. B. Haehnel, M. Sturm, C. A. Hiemstra, S. Berezovskaya, and R. D. Tabler, 2007: Simulating complex snow distributions in windy environments using SnowTran-3D. *Journal of Glaciology*, **53 (181)**, 241–256, doi: 10.3189/172756507782202865.

Liston, G. E., and D. K. Hall, 1995: An energy-balance model of lake-ice evolution. *Journal of Glaciology*, **41**, 373–382.

Liston, G. E., and C. A. Hiemstra, 2008: A Simple Data Assimilation System for Complex Snow Distributions (SnowAssim). *Journal of Hydrometeorology*, **9 (5)**, 989–1004, doi: 10.1175/2008JHM871.1.

Liston, G. E., and C. A. Hiemstra, 2011: The Changing Cryosphere: Pan-Arctic Snow Trends (1979–2009). *Journal of Climate*, **24 (21)**, 5691–5712, doi: 10.1175/JCLI-D-11-00081.1.

Liu, G., 2008: A Database of Microwave Single-Scattering Properties for Nonspherical Ice Particles. *Bulletin of the American Meteorological Society*, **89 (10)**, 1563–1570, doi:

10.1175/2008BAMS2486.1.

Liu, Y., Y. Fang, D. Li, and S. A. Margulis, 2022: How Well do Global Snow Products Characterize Snow Storage in High Mountain Asia? *Geophysical Research Letters*, **49 (16)**, e2022GL100 082, doi: 10.1029/2022GL100082.

Liu, Y., Y. Fang, and S. A. Margulis, 2021: Spatiotemporal distribution of seasonal snow water equivalent in High Mountain Asia from an 18-year Landsat–MODIS era snow reanalysis dataset. *The Cryosphere*, **15 (11)**, 5261–5280, doi: 10.5194/tc-15-5261-2021.

Lu, X., and Coauthors, 2022: Deriving Snow Depth From ICESat-2 Lidar Multiple Scattering Measurements: Uncertainty Analyses. *Frontiers in Remote Sensing*, **3**, doi: 10.3389/frsen.2022.891481.

Lundquist, J. D., M. Hughes, B. Henn, E. D. Gutmann, B. Livneh, J. Dozier, and P. Neiman, 2015: High-Elevation Precipitation Patterns: Using Snow Measurements to Assess Daily Gridded Datasets across the Sierra Nevada, California. *Journal of Hydrometeorology*, **16 (4)**, 1773–1792, doi: 10.1175/JHM-D-15-0019.1.

Mai, L., S. Yang, Y. Wang, and R. Li, 2023: Impacts of Shape Assumptions on Z–R Relationship and Satellite Remote Sensing Clouds Based on Model Simulations and GPM Observations. *Remote Sensing*, **15 (6)**, 1556, doi: 10.3390/rs15061556.

Margulis, S. A., G. Cortés, M. Girotto, and M. Durand, 2016: A Landsat-Era Sierra Nevada Snow Reanalysis (1985–2015). *Journal of Hydrometeorology*, **17 (4)**, 1203–1221, doi: 10.1175/JHM-D-15-0177.1.

Margulis, S. A., Y. Liu, and E. Baldo, 2019: A Joint Landsat- and MODIS-Based Reanalysis Approach for Midlatitude Montane Seasonal Snow Characterization. *Frontiers in Earth Science*, **7**, 272, doi: 10.3389/feart.2019.00272.

Mätzler, 1993: Passive microwave signatures of landscapes in winter. *Meteorology and Atmospheric Physics*.

McEvoy, D. J., and B. J. Hatchett, 2023: Spring heat waves drive record western United States snow melt in 2021. *Environmental Research Letters*, **18 (1)**, 014 007, doi: 10.1088/1748-9326/aca8bd.

Meng, H., J. Dong, R. Ferraro, B. Yan, L. Zhao, C. Kongoli, N.-Y. Wang, and B. Zavodsky, 2017: A 1DVAR-based snowfall rate retrieval algorithm for passive microwave radiometers. *Journal of Hydrometeorology*, **122 (12)**, 6520–6540, doi: 10.1002/2016JD026325.

Meng, J., R. Yang, H. Wei, M. Ek, G. Gayno, P. Xie, and K. Mitchell, 2012: The Land Surface Analysis in the NCEP Climate Forecast System Reanalysis. *Journal of Geophysical Research: Atmospheres*, doi: 10.1175/JHM-D-11-090.1.

Mote, P. W., A. F. Hamlet, M. P. Clark, and D. P. Lettenmaier, 2005: Declining Mountain Snowpack in Western North America. *Bulletin of the American Meteorological Society*, **86 (1)**, 39–50, doi: 10.1175/BAMS-86-1-39.

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

Mroz, K., M. Montopoli, A. Battaglia, G. Panegrossi, P. Kirstetter, and L. Baldini, 2021: Cross Validation of Active and Passive Microwave Snowfall Products over the Continental United States. *Journal of Hydrometeorology*, **22 (5)**, 1297–1315, doi: 10.1175/JHM-D-20-0222.1.

Mudryk, L. R., C. Derksen, P. J. Kushner, and R. Brown, 2015: Characterization of Northern Hemisphere Snow Water Equivalent Datasets, 1981–2010. *Journal of Climate*, **28 (20)**, 8037–8051, doi: 10.1175/JCLI-D-15-0229.1.

Musselman, K. N., N. Addor, J. A. Vano, and N. P. Molotch, 2021: Winter melt trends portend widespread declines in snow water resources. *Nature climate change*, **2021**, 10.1038/s41 558–021–01 014–9, doi: 10.1038/s41558-021-01014-9.

Nakaya, U., 1951: The Formation of Ice Crystals. *Compendium of Meteorology: Prepared under the Direction of the Committee on the Compendium of Meteorology*, T. F. Malone, Ed., American Meteorological Society, Boston, MA, 207–220, doi: 10.1007/978-1-940033-70-9_18.

National Operational Hydrologic Remote Sensing Center, 2004: Snow Data Assimilation System (SNODAS) Data Products at NSIDC, Version 1. NSIDC, doi: 10.7265/N5TB14TC.

Neumann, T. A., and Coauthors, 2019: The Ice, Cloud, and Land Elevation Satellite – 2 mission: A global geolocated photon product derived from the Advanced Topographic Laser Altimeter System. *Remote Sensing of Environment*, **233**, 111 325, doi: 10.1016/j.rse.2019.111325.

Noh, Y.-J., G. Liu, E.-K. Seo, J. R. Wang, and K. Aonashi, 2006: Development of a snowfall retrieval algorithm at high microwave frequencies. *Journal of Geophysical Research: Atmospheres*, **111 (D22)**, doi: 10.1029/2005JD006826.

Orsolini, Y. J., R. Senan, G. Balsamo, F. J. Doblas-Reyes, F. Vitart, A. Weisheimer, A. Carrasco, and R. E. Benestad, 2013: Impact of snow initialization on sub-seasonal forecasts. *Climate Dynamics*, **41 (7)**, 1969–1982, doi: 10.1007/s00382-013-1782-0.

Pagano, T., and D. Garen, 2005: A Recent Increase in Western U.S. Streamflow Variability and Persistence. *Journal of Hydrometeorology*, doi: 10.1175/JHM410.1.

Pagano, T., D. Garen, and S. Sorooshian, 2004: Evaluation of Official Western U.S. Seasonal Water Supply Outlooks, 1922–2002. *Journal of Hydrometeorology*.

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

Pederson, G. T., S. T. Gray, T. Ault, W. Marsh, D. B. Fagre, A. G. Bunn, C. A. Woodhouse, and L. J. Graumlich, 2011: Climatic Controls on the Snowmelt Hydrology of the Northern Rocky Mountains. *Journal of Climate*, **24 (6)**, 1666–1687, doi: 10.1175/2010JCLI3729.1.

Pettersen, C., M. S. Kulie, L. F. Bliven, A. J. Merrelli, W. A. Petersen, T. J. Wagner, D. B. Wolff, and N. B. Wood, 2020a: A Composite Analysis of Snowfall Modes from Four Winter Seasons in Marquette, Michigan. *Journal of Applied Meteorology and Climatology*, **59 (1)**, 103–124, doi: 10.1175/JAMC-D-19-0099.1.

Pettersen, C., and Coauthors, 2020b: The Precipitation Imaging Package: Assessment of Microphysical and Bulk Characteristics of Snow. *Atmosphere*, **11 (8)**, 785, doi: 10.3390/atmos11080785.

Pflug, J. M., M. Hughes, and J. D. Lundquist, 2021: Downscaling Snow Deposition Using Historic Snow Depth Patterns: Diagnosing Limitations From Snowfall Biases, Winter Snow Losses, and Interannual Snow Pattern Repeatability. *Water Resources Research*, **57 (8)**, e2021WR029 999, doi: 10.1029/2021WR029999.

Pfreundschuh, S., 2024: Simonpf/chimp: The Chalmers Multi-Satellite Precipitation Retrieval. https://github.com/simonpf/chimp.

Pfreundschuh, S., P. J. Brown, C. D. Kummerow, P. Eriksson, and T. Norrestad, 2022a: GPROF-NN: A neural-network-based implementation of the Goddard Profiling Algorithm. *Atmospheric Measurement Techniques*, **15 (17)**, 5033–5060, doi: 10.5194/amt-15-5033-2022.

Pfreundschuh, S., P. J. Brown, C. D. Kummerow, P. Eriksson, and T. Norrestad, 2022b: GPROF-NN: A neural-network-based implementation of the Goddard Profiling Algorithm. *Atmospheric Measurement Techniques*, **15 (17)**, 5033–5060, doi: 10.5194/amt-15-5033-2022.

Pfreundschuh, S., P. Eriksson, D. Duncan, B. Rydberg, N. Håkansson, and A. Thoss, 2018: A neural network approach to estimating a posteriori distributions of Bayesian retrieval problems. *Atmospheric Measurement Techniques*, **11 (8)**, 4627–4643, doi: 10.5194/amt-11-4627-2018.

Pullen, S., C. Jones, and G. Rooney, 2011: Using Satellite-Derived Snow Cover Data to Implement a Snow Analysis in the Met Office Global NWP Model. *Journal of Applied Meteorology and Climatology*, doi: 10.1175/2010JAMC2527.1.

Pulliainen, J., and Coauthors, 2020: Patterns and trends of Northern Hemisphere snow mass from 1980 to 2018. *Nature*, **581 (7808)**, 294–298, doi: 10.1038/s41586-020-2258-0.

Rahimi, R., A. Ebtehaj, G. Panegrossi, L. Milani, S. E. Ringerud, and F. J. Turk, 2022: Vulnerability of Passive Microwave Snowfall Retrievals to Physical Properties of Snowpack: A Perspective From Dense Media Radiative Transfer Theory. *IEEE Transactions on Geoscience and Remote Sensing*, **60**, 1–13, doi: 10.1109/TGRS.2022.3184530.

Rahimi, R., P. Ravirathinam, A. Ebtehaj, A. Behrangi, J. Tan, and V. Kumar, 2024: arXiv:2307.10843. Global Precipitation Nowcasting of Integrated Multi-satellitE Retrievals for GPM: A U-Net Convolutional LSTM Architecture. arXiv, doi: 10.48550/arXiv.2307.10843, 2307.10843.

Raleigh, M. S., B. Livneh, K. Lapo, and J. D. Lundquist, 2016: How Does Availability of Meteorological Forcing Data Impact Physically Based Snowpack Simulations? *Journal of Hydrometeorology*, **17 (1)**, 99–120, doi: 10.1175/JHM-D-14-0235.1.

Ramsay, B. H., 1998: The interactive multisensor snow and ice mapping system. *Hydrological Processes*, **12**, 1537–1546.

Ramsay, B. H., 2000: Prospects for the Interactive Multisensor Snow and Ice Mapping System (IMS).

Rantanen, M., A. Y. Karpechko, A. Lipponen, K. Nordling, O. Hyvärinen, K. Ruosteenoja, T. Vihma, and A. Laaksonen, 2022: The Arctic has warmed nearly four times faster than the globe since 1979. *Communications Earth & Environment*, **3 (1)**, 168, doi: 10.1038/s43247-022-00498-3.

Rittger, K., T. H. Painter, and J. Dozier, 2013: Assessment of methods for mapping snow cover from MODIS. *Advances in Water Resources*, **51**, 367–380, doi: 10.1016/j.advwatres.2012.03.002.

Ronneberger, O., P. Fischer, and T. Brox, 2015: arXiv:1505.04597. U-Net: Convolutional Networks for Biomedical Image Segmentation. arXiv, doi: 10.48550/arXiv.1505.04597, 1505.04597.

Rugg, A., E. D. Gutmann, R. R. McCrary, F. Lehner, A. J. Newman, J. H. Richter, M. R. Tye, and A. W. Wood, 2023: Mass-Conserving Downscaling of Climate Model Precipitation Over Mountainous Terrain for Water Resource Applications. *Geophysical Research Letters*, **50 (20)**, e2023GL105 326, doi: 10.1029/2023GL105326.

Ruiz-Barradas, A., and S. Nigam, 2006: IPCC's Twentieth-Century Climate Simulations: Varied Representations of North American Hydroclimate Variability. *Journal of Climate*, doi: 10.1175/JCLI3809.1.

Rutledge, S. A., and P. Hobbs, 1983: The Mesoscale and Microscale Structure and Organization of Clouds and Precipitation in Midlatitude Cyclones. VIII: A Model for the "Seeder-Feeder" Process in Warm-Frontal Rainbands. *Journal of the Atmospheric Sciences*.

Rysman, J.-F., G. Panegrossi, P. Sanò, A. C. Marra, S. Dietrich, L. Milani, and M. S. Kulie, 2018: SLALOM: An All-Surface Snow Water Path Retrieval Algorithm for the GPM Microwave Imager. *Remote Sensing*, **10 (8)**, 1278, doi: 10.3390/rs10081278.

Rysman, J.-F., and Coauthors, 2019: Retrieving Surface Snowfall With the GPM Microwave Imager: A New Module for the SLALOM Algorithm. *Geophysical Research Letters*, **46 (22)**, 13 593–13 601, doi: 10.1029/2019GL084576.

Sadeghi, M., P. Nguyen, K. Hsu, and S. Sorooshian, 2020: Improving near real-time precipitation estimation using a U-Net convolutional neural network and geographical information. *Environmental Modelling & Software*, **134**, 104 856, doi: 10.1016/j.envsoft.2020.104856.

Schneider, D., N. P. Molotch, J. S. Deems, and T. H. Painter, 2021: Analysis of topographic controls on depletion curves derived from airborne lidar snow depth data. *Hydrology Research*, **52 (1)**, 253–265, doi: 10.2166/nh.2020.267.

Seo, B.-C., F. Quintero, and W. F. Krajewski, 2023: Hydrologic Assessment of IMERG Products Across Spatial Scales Over Iowa. *Journal of Hydrometeorology*, **24 (6)**, 997–1015, doi: 10.1175/JHM-D-22-0129.1.

Shulgina, T., and Coauthors, 2023: Observed and projected changes in snow accumulation and snowline in California's snowy mountains. *Climate Dynamics*, **61 (9)**, 4809–4824, doi: 10.1007/s00382-023-06776-w.

Sims, E. M., and G. Liu, 2015: A Parameterization of the Probability of Snow–Rain Transition. *Journal of Hydrometeorology*, doi: 10.1175/JHM-D-14-0211.1.

Smith, T., and B. Bookhagen, 2018: Changes in seasonal snow water equivalent distribution in High Mountain Asia (1987 to 2009). *Science Advances*, **4 (1)**, e1701 550, doi: 10.1126/sciadv.1701550.

Stewart, I. T., D. R. Cayan, and M. D. Dettinger, 2005: Changes toward Earlier Streamflow Timing across Western North America. *Journal of Climate*, **18 (8)**, 1136–1155, doi: 10.1175/JCLI3321.1.

Sturm, M., M. A. Goldstein, and C. Parr, 2017: Water and life from snow: A trillion dollar science question: SNOW AND LIFE. *Water Resources Research*, **53 (5)**, 3534–3544, doi: 10.1002/2017WR020840.

Sturm, M., and G. E. Liston, 2021: Revisiting the Global Seasonal Snow Classification: An Updated Dataset for Earth System Applications. *Journal of Hydrometeorology*, **22 (11)**, 2917–2938, doi: 10.1175/JHM-D-21-0070.1.

Tabari, H., S. M. Paz, D. Buekenhout, and P. Willems, 2021: Comparison of statistical downscaling methods for climate change impact analysis on precipitation-driven drought. *Hydrology and Earth System Sciences*, **25 (6)**, 3493–3517, doi: 10.5194/hess-25-3493-2021.

Takbiri, Z., A. Ebtehaj, E. Foufoula-Georgiou, P.-E. Kirstetter, and F. J. Turk, 2019: A Prognostic Nested k -Nearest Approach for Microwave Precipitation Phase Detection over Snow Cover. *Journal of Hydrometeorology*, **20 (2)**, 251–274, doi: 10.1175/JHM-D-18-0021.1.

Takbiri, Z., L. Milani, C. Guilloteau, and E. Foufoula-Georgiou, 2021: Quantitative Investigation of Radiometric Interactions between Snowfall, Snow Cover, and Cloud Liquid Water over Land. *Remote Sensing*, **13 (13)**, 2641, doi: 10.3390/rs13132641.

Tan, M., and Q. Le, 2021: EfficientNetV2: Smaller Models and Faster Training. *Proceedings of the 38th International Conference on Machine Learning*, PMLR, 10 096–10 106.

Tang, G., D. Long, A. Behrangi, C. Wang, and Y. Hong, 2018: Exploring Deep Neural Networks to Retrieve Rain and Snow in High Latitudes Using Multisensor and Reanalysis Data. *Water Resources Research*, **54 (10)**, 8253–8278, doi: 10.1029/2018WR023830.

Team, P. M. A., 2022: GPM ATBD. NASA.

Tedesco, M., and J. Jeyaratnam, 2016: A New Operational Snow Retrieval Algorithm Applied to Historical AMSR-E Brightness Temperatures. *Remote Sensing*, **8 (12)**, 1037, doi: 10.3390/rs8121037.

Teutschbein, C., and J. Seibert, 2012: Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *Journal of Hydrology*, **456–457**, 12–29, doi: 10.1016/j.jhydrol.2012.05.052.

Ulaby, F. T., and W. H. Stiles, 1980: The active and passive microwave response to snow parameters: 2. Water equivalent of dry snow. *Journal of Geophysical Research*, **85 (C2)**, 1045, doi: 10.1029/JC085iC02p01045.

Vavrus, S., 2007: The role of terrestrial snow cover in the climate system. *Climate Dynamics*, **29 (1)**, 73–88, doi: 10.1007/s00382-007-0226-0.

Vögeli, C., M. Lehning, N. Wever, and M. Bavay, 2016: Scaling Precipitation Input to Spatially Distributed Hydrological Models by Measured Snow Distribution. *Frontiers in Earth Science*, **4**, doi: 10.3389/feart.2016.00108.

Vuyovich, C., 2024: SnowEx | Snow. https://snow.nasa.gov/campaigns/snowex.

Wen, Y., P. Kirstetter, J. J. Gourley, Y. Hong, A. Behrangi, and Z. Flamig, 2017: Evaluation of MRMS Snowfall Products over the Western United States. *Journal of Hydrometeorology*, **18 (6)**, 1707–1713, doi: 10.1175/JHM-D-16-0266.1.

Wolfe, J. P., and J. R. Snider, 2012: A Relationship between Reflectivity and Snow Rate for a High-Altitude S-Band Radar. *Journal of Applied Meteorology and Climatology*, **51 (6)**, 1111–1128, doi: 10.1175/JAMC-D-11-0112.1.

Wrzesien, M. L., M. T. Durand, T. M. Pavelsky, I. M. Howat, S. A. Margulis, and L. S. Huning, 2017: Comparison of Methods to Estimate Snow Water Equivalent at the Mountain Range Scale: A Case Study of the California Sierra Nevada. *Journal of Hydrometeorology*, **18 (4)**, 1101–1119, doi: 10.1175/JHM-D-16-0246.1.

Wrzesien, M. L., M. T. Durand, T. M. Pavelsky, S. B. Kapnick, Y. Zhang, J. Guo, and C. K. Shum, 2018: A New Estimate of North American Mountain Snow Accumulation From Regional Climate Model Simulations. *Geophysical Research Letters*, **45 (3)**, 1423–1432, doi: 10.1002/2017GL076664.

Wrzesien, M. L., T. M. Pavelsky, M. T. Durand, J. Dozier, and J. D. Lundquist, 2019: Characterizing Biases in Mountain Snow Accumulation From Global Data Sets. *Water Resources Research*, **55 (11)**, 9873–9891, doi: 10.1029/2019WR025350.

Wu, Y., C. Miao, X. Fan, J. Gou, Q. Zhang, and H. Zheng, 2022: Quantifying the Uncertainty Sources of Future Climate Projections and Narrowing Uncertainties With Bias Correction Techniques. *Earth's Future*, **10 (11)**, e2022EF002 963, doi: 10.1029/2022EF002963.

Xiao, Z., and A. Duan, 2016: Impacts of Tibetan Plateau Snow Cover on the Interannual Variability of the East Asian Summer Monsoon. *Journal of Climate*, **29 (23)**, 8495–8514, doi: 10.1175/JCLI-D-16-0029.1.

Yoshikane, T., and K. Yoshimura, 2023: A downscaling and bias correction method for climate model ensemble simulations of local-scale hourly precipitation. *Scientific Reports*, **13 (1)**, 9412, doi: 10.1038/s41598-023-36489-3.

You, Y., N.-Y. Wang, R. Ferraro, and S. Rudlosky, 2017: Quantifying the Snowfall Detection Performance of the GPM Microwave Imager Channels over Land. *Journal of Hydrometeorology*, **18 (3)**, 729–751, doi: 10.1175/JHM-D-16-0190.1.

Zhang, J., and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) Quantitative Precipitation Estimation: Initial Operating Capabilities. *Bulletin of the American Meteorological Society*, **97 (4)**, 621–638, doi: 10.1175/BAMS-D-14-00174.1.