Title

Towards large-scale implementation of a high resolution snow reanalysis over midlatitude montane ranges

Author

Baldo, Elisabeth

Publication Date

2017

Peer reviewed|Thesis/dissertation

University of California
Los Angeles

Towards large-scale implementation of a high resolution snow reanalysis over midlatitude montane ranges

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Civil Engineering

by

Elisabeth Baldo

2017

Abstract of the Dissertation

Towards large-scale implementation of a high resolution snow reanalysis over midlatitude montane ranges

by

Elisabeth Baldo

Doctor of Philosophy in Civil Engineering

University of California, Los Angeles, 2017

Professor Steven A. Margulis, Chair

Accurately representing the spatial variability of montane snowpack is challenging due to the high degree of complexity in the terrain's topography, and the lack of good quality in-situ data. To explicitly resolve snow processes in montane environment while taking into account the uncertainties in the system's high spatial and temporal resolution snow reanalyses are required. Ensemble-based approaches assimilating Landsat VIS-NIR remote sensing data at spatial resolutions of 100 m or less are, however, prohibitively expensive to run at large scales, and sub-optimal given that only the most complex parts of a montane range require such fine resolution. In addition, the assimilation of remote sensing data from a single source can also be unsatisfactory due to a lack of global coverage, hardware malfunction etc. In order to optimize computational needs while preserving the accuracy of ~ 100 m reanalyses, a raster-based multi-resolution approach was first developed and successfully implemented for a headwater catchment in the Colorado River Basin over the full length of Landsat record (30+ years). The potential use of MODIS-derived snow cover information in addition to Landsat in snow reanalyses was then investigated over three different regions in the Western U.S. and High Mountain Asia in order to make up for Landsat's shortcomings over midlatitude snowpacks. The key findings of this dissertation can be summarized as follows: 1) The physiographic complexity of a terrain can be characterized by its standard deviations of elevation, northness index and forested fraction. Using such a complexity metric to discretize

the terrain into different spatial resolutions via a multi-resolution approach can significantly reduce computational needs, while mitigating errors in snow processes representation. 2) The multi-resolution approach did not significantly impact the remote sensing observations assimilated, and the posterior snow water equivalent (SWE) ensemble median and standard deviation matched the 90 m reanalysis, thus leading to a robust implementation in the context of a data assimilation framework. 3) A MODIS-derived snow cover product was found to be a useful complementary source of remote sensing data to be simultaneously assimilated with Landsat. Off-nadir-looking observations first had to be screened out of the reanalysis due to the distorting and snow obscuring effects of the sensor viewing geometry at high zenith angles. Ultimately, the methods developed in this work can be applied to all midlatitude montane ranges over the full lengths of Landsat and MODIS records to generate a fine spatial resolution SWE reanalysis dataset that will be useful to snow hydrologists to solve many unanswered science questions over such challenging regions.

The dissertation of Elisabeth Baldo is approved.

Michael Durand

Yongkang Xue

Dennis P. Lettenmaier

William Yeh

Steven A. Margulis, Committee Chair

University of California, Los Angeles

2017

To my wonderful parents Ilona and Pascal ...
for their love and endless support since the beginning,
including my terrible teenage years.

Table of Contents

List of Figures

List of Tables

Acknowledgments

First of all, I would like to express my thanks and gratitude to my advisor, Professor Steven A. Margulis for his precious and unlimited patience, guidance and support. Many thanks to all my current and previous coworkers at UCLA for their friendship and encouragement, especially Manuela for inspiring me to do this Ph.D. in the first place, Rike for being both an amazing roommate and lab mate, Akash for introducing me to so many great movies and books, Mahdi for always sharing your drawer of snacks with us, and Gonzalo for our countless conversations / coffee breaks, as well as the fun we had working on our drone Albert. I would like to extend a special acknowledgement to my committee members, Professor Yeh, Professor Xue, Professor Lettenmaier and Professor Durand for their insightful comments. A special thanks to Professor Xue for providing us with the land surface model used in this dissertation. The sponsorship of this research by the NASA NEWS project (grant NNX15AD16G) and the National Science Foundation (grant number EAR-1246473) is also greatly appreciated.

This work would have never been possible without the support of my invaluable group of friends with whom I have shared so many good memories over the years. I want to thank in particular Camille, Caroline, Chloé, Angelique, Bertrand, Sam, Edouard, Lucie and Erika for warmly welcoming me, advising me and encouraging me since I arrived in LA, and Vladimir, Nanar, Justine and Claudia for all our much needed fun conversations, movie nights and dinners. Special thanks to Hortense, Aurélien and Franky for always being there for me, and Josépha and Valentin for feeding me delicious food and keeping me company while I wrote this dissertation.

Last but not least, I would like to express my thanks and gratitude to my parents, Pascal and Ilona, my brother Alexandre, and my sister Clémentine for sending me so much love and support from afar. I owe a very special thanks to my boyfriend Henri-Louis for being my rock, and occasionally my rubber duck when debugging codes.

Vita

2011	B.S. (Civil Engineering), École Spéciale des Travaux Publics, Paris, France.
2013	M.S. (Civil and Environmental Engineering), UCLA, Los Angeles, California.
2013	M.S. (Civil Engineering), École Spéciale des Travaux Publics, Paris, France.
2012–2015	Teaching Assistant, French and Francophone Studies Department, UCLA. Instructed beginner (FRNCH 1 and FRNCH3) and intermediate (FRNCH4 and FRNCH5) French courses.
2013–2016	Teaching Assistant, Civil and Environmental Engineering Department, UCLA. Instructed Applied numerical analysis and modeling (CEE103) and Introduction to Hydrology (CEE150) courses.
2013–present	Graduate student researcher, Civil and Environmental Engineering Department, UCLA. Focus on Snow processes, remote sensing and data assimilation.

Publications

Baldo, E., and S. A. Margulis (2017), Implementation of a physiographic complexity-based multiresolution snow modeling scheme, Water Resour. Res., 53, doi:10.1002/2016WR020021.

Baldo, E., and S. A. Margulis (2017, in review), Assessment of a multi-resolution snow reanalysis framework: a multi-decadal reanalysis case over the Upper Yampa River basin, Colorado, Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-664.

Baldo, E., and S. A. Margulis (under preparation), Joint assimilation of Landsat and MODIS-based fractional snow cover area products in a probabilistic SWE reanalysis framework

CHAPTER 1

Introduction

1.1 Motivation

Seasonal montane snowpacks (highlighted in red in Figure 1.1) play a significant role in providing fresh water to approximately 2 billion people worldwide (Barnett et al. [2005], Mankin et al. [2015], especially during the dry season when runoff from snow melt constitute a vital water resource.

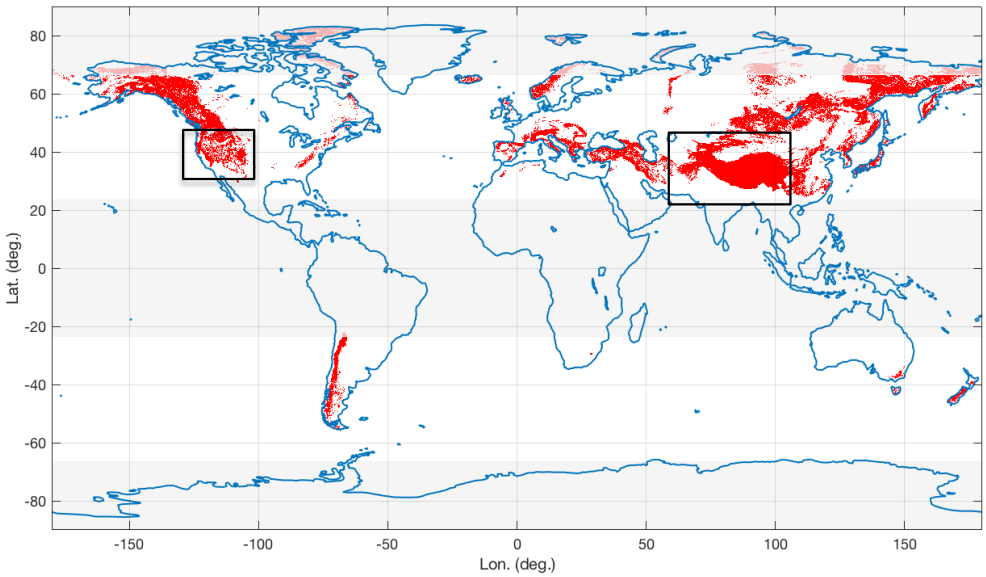

Figure 1.1: Map highlighting montane areas subject to seasonal snow processes (red regions). The specific domains to be studied in this work are the Western U.S. and High Mountain Asia (HMA) ranges shown in black boxes

However, many hydrologic science questions remain unanswered over theses regions, given their topographic complexity and the lack of in-situ data. Estimating the spatial distribution of snow water equivalent (SWE) has therefore been deemed the most important unsolved problem in snow hydrology (Dozier et al. [2016]). Addressing the issue of complex terrain representation is challenging, because land surface models have to be run at very fine spatial resolutions on the order of 100 m or less (Clark et al. [2011]), which is computationally expensive. In addition, deterministic forward modeling itself, even at high-resolution, is often insufficient to accurately represent snow processes due to errors in model inputs (most notably precipitation, Luo et al. [2003]) that are poorly characterized in montane regions. Consequently, ensemble-based data assimilation (DA) methods are the preferred approach to estimate snow states (Clark et al. [2006], Andreadis and Lettenmaier [2006], Su et al. [2008], De Lannoy et al. [2010], Liu et al. [2013], Arsenault et al. [2013], Girotto et al. [2014b], Margulis et al. [2015], Kumar et al. [2015]). Over in-situ data scarce regions such as the High Mountain Asia (HMA, framed in black in Figure 1.1), assimilating remotely sensed observations offers spatially and temporally continuous estimates, with the additional advantage of providing a measure of their uncertainty. However, the cost of running DA frameworks (also referred to as reanalysis in this work) at very fine spatial resolutions can be prohibitive due to their ensemble nature.

The current global SWE reanalysis product, GLDAS (Rodell et al. [2004]) therefore simulated snow processes at a resolution of 0.25°, and the contiguous United States product, SNODAS (Barrett [2003]) is available at 1 km. Simulating at these scales can solve the computational issue, but inherently sacrifices valuable information related to sub-grid heterogeneities in montane regions, and GLDAS has been shown to consistently underestimate SWE (Broxton et al. [2016]). Using spatial resolutions of 1 km or more is sub-optimal since relevant remote sensing data streams that can act as model constraints (e.g. Lidar, Landsat, MODIS, etc.) are available at higher resolution (from meter - to hundreds of meter scale). Instead of using a "one size fits all" approach that is computationally expensive at fine resolutions, this work aims to develop a multi-resolution (MR) approach that focuses the use

of high spatial resolutions only over areas that are complex enough to require it. The goal is to generate a SWE reanalysis dataset with the accuracy of 100 m SWE estimates, while significantly reducing the computational needs.

In addition to the issue of spatial resolution, the quality and availability on the assimilated remote sensing images is also crucial to generate a SWE reanalysis dataset that accurately represents the spatial variability of the seasonal montane snowpack. Margulis et al. [2016] and Cortés and Margulis [2017] successfully assimilated Landsat-derived fractional snow cover area (*fsca*) observations (Cortés et al. [2014]), to generate a 30 + year SWE reanalysis dataset at a resolution of 90 m over the Sierra Nevada and 180 m over the Andes respectively. The main advantage of using Landsat imagery is the length of its record period (1985 - present), and its fine spatial resolution of 30 m with a nadir looking geometry. However, the presence of clouds, combined with its 16 days temporal resolution can significantly reduce the number of usable observations. In addition, gaps in the Landsat-era record exist in some regions of the globe (Kovalskyy and Roy [2013], Wulder et al. [2016]) during some years, which is especially problematic over the High Mountain Asia (HMA, Figure 1.1) region. In order to overcome Landsat limitations in the context of a global SWE reanalysis dataset, the goal of this work is to explore the simultaneous assimilation of both Landsat-derived *fsca* and MODIS-derived MODSCAG *fsca* product by Painter et al. [2009]. The MODSCAG product is indeed complementary to the Landsat-derived *fsca*, with a good coverage over the seasonal montane snowpack regions shown in Figure 1.1 since 2001, a daily temporal resolution, and a slightly coarser spatial resolution of ~500 m. One potential restriction when using MODIS-based observations is the effect of the sensor viewing geometry on SWE estimates. Given the scanning nature of the MODIS sensor, acquisitions can be made at zenith viewing angle varying from 0° to over 60°, which could be problematic over tall vegetated areas.

1.2 Objective of the dissertation

The objective of this work is to facilitate the implementation of a high spatial resolution SWE reanalysis dataset over all montane areas subject to seasonal snow (Figure 1.1), while taking into consideration the key challenges mentioned above:

1. What is the impact of the land surface model spatial resolution on snow estimates? Can a complexity metric be derived from the physiographic characteristics of a terrain and serve as a proxy for the modeled SWE spatial variability? How can a multi-resolution (MR) approach optimize snow modeling to reduce computational constraints while preserving the accuracy of high resolution estimates?

2. How are remotely sensed *fsca* observations, and the data assimilation framework impacted by the MR approach? How well does the MR approach perform in terms of estimating the central tendency (i.e. ensemble median) and uncertainty of the posterior snow state distribution in space and time?

3. What is the impact of MODIS viewing geometry on its derived MODSCAG *fsca* and fractional forested cover (*fveg*) products? How does MODSCAG observations differ from Landsat over montane seasonal snowpack? Can both sensors be simultaneously assimilated in the SWE reanalysis framework?

While the focus of this work was on snow processes and the assimilation of snow cover area products, the techniques presented could be generalized to include other state variables such as soil moisture over non snow dominated areas. The present dissertation is therefore expected to provide a set of useful tools for hydrologist to implement complex and computationally costly frameworks over large-scale regions.

1.3 Organization of the dissertation

This dissertation is organized as follows: Chapter 2 describes in details the development of a multi-resolution algorithm that adapts the spatial resolution of a domain to its physiographic complexity in the context of snow processes modeling. Chapter 3 investigates the effect of the multi-resolution approach on the assimilated *fsca* observations and the resulting SWE reanalysis dataset for a large headwater catchment over a 31-yr period. Chapter 4 then addresses the issue of the scarcity of Landsat coverage by exploring the joint assimilation of the complementary MODIS-based *fsca* product. Finally, a summary of this work, concluding remarks and ideas for future work are given in Chapter 5.

CHAPTER 2

Implementation of a physiographic complexity-based multi-resolution snow modeling scheme

Using a uniform model resolution over a domain is not necessarily the optimal approach for simulating hydrologic processes when considering both model error and computational cost. Fine -resolution simulations at 100 m or less can provide fine-scale process representation, but can be costly to apply over large domains. On the other hand, coarser spatial resolutions are more computationally inexpensive, but at the expense of fine-scale model accuracy. Defining a multi-resolution (MR) grid spanning from fine-resolutions over complex mountainous areas to coarser resolutions over less complex regions can conceivably reduce computational costs, while preserving the accuracy of fine-resolution simulations on a uniform grid. A MR scheme was developed using a physiographic complexity metric (CM) that combines surface heterogeneity in forested fraction, elevation, slope and aspect. A data reduction term was defined as a metric (relative to a uniform fine-resolution grid) related to the available computational resources for a simulation. The focus of the effort was on the melt season where physiographic complexity is known to have a significant signature. MR simulations were run for different data reduction factors to generate melt rate estimates for three representative water years over a test headwater catchment in the Colorado River Basin. The MR approach with data reductions up to 47% led to negligible cumulative snowmelt differences compared to the fine-resolution baseline case, while tests with data reductions up to 60% showed differences lower than 2%. Large snow-dominated domains could therefore benefit from a MR approach to be more efficiently simulated while mitigating error.

2.1 Background and motivation

Since most processes represented in distributed land surface models are nonlinear, a change in spatial resolution can introduce bias in the basin-average variables Blöschl [1999]. In the context of snow process modeling in complex terrain, the effects of increasing spatial resolution were studied by Cline et al. [1998] and Winstral et al. [2014] over small alpine catchments (of a few km^2 or less). Cline et al. [1998] found that a coarsening of the model resolution resulted in a loss of information regarding the distribution of snow water equivalent (SWE) within the basin due to smoothing of the basin terrain. Winstral et al. [2014] tested meteorological forcing resolutions ranging from 10 m to 1500 m and concluded that simulations using forcing inputs processed at 100 m or less led to biases in surface water input (which includes snowmelt, rain falling on bare soil, and rain passing through an isothermal snow cover) that were less than 4%. However, further degradation of the forcing inputs as a result of coarsened resolution led to increasing errors. The extensive review of previous studies using field data to evaluate the spatial variability of SWE done by Clark et al. [2011] concluded that spatial resolutions of ~100 m or less are required to capture heterogeneity of snow in montane regions where physiography is complex.

In lieu of the coarse resolutions of ~10 to 100 km traditionally employed over continental domains (Reichle et al. [2011], Mesinger et al. [2006]), much finer resolutions on the order of a 100 m or less may therefore be desirable in order to accurately assess and monitor the terrestrial water and energy cycle. In this vein, fine-resolution land surface modeling has been defined as the next grand challenge for the hydrologic community (Wood et al. [2011], Bierkens et al. [2015], Beven et al. [2015]). However, implementing fine-resolution deterministic modeling over large domains is computationally expensive, with increasing expense when ensemble-based data assimilation methods are employed (Clark et al. [2006], De Lannoy et al. [2010], Girotto et al. [2014b], Margulis et al. [2015]). Furthermore, such fine resolutions may not be necessary everywhere, making a one size fits all approach sub-optimal because significant computational effort will be consumed unnecessarily. Snow processes provide a useful illustrative example of this issue, where very high resolution is likely warranted in

regions of high physiographic complexity (i.e. high surface heterogeneity in topography and vegetation cover), but may be less so in other areas.

The main hypothesis tested in this chapter is that even for snow-dominated watersheds with complex topography, only subsets of the domain may require a very fine model scale. Sub-regions with less physiographic complexity will generally show less spatial variability (i.e. in snowmelt), and can therefore be treated at coarser resolutions without introducing significant errors. Previous works used clustering algorithms to divide basins into hydrologic response units (HRUs) of topographic/hydrologic similarities (Beven and Kirby [1979], U. S. Geological Survey et al. [1983], Sivapalan et al. [1987] and Chaney et al. [2016]). By lumping together proxies of the drivers of spatial heterogeneity, HRUs improved the representation of sub-grid variability in land surface models (LSMs), especially when modeling snow states over complex terrain (Pomeroy et al. [2007], Fiddes and Gruber [2012] and Newman et al. [2014]).

Another approach that allows for variable spatial resolutions uses triangulated irregular networks (TINs; Tucker et al. [2001]). For example, the TIN based Real-time Integrated Basin Simulator, or tRIBS, LSM developed by Ivanov et al. [2004] attempts to capture key hydrologic features using only 5-10% of the original grid elements. The discretization is based on either a slope preserving criterion (rugged terrain is highly resolved) or a hydrologic similarity criterion (areas frequently saturated due to a rising water table are highly resolved), and the specification of a data reduction factor (Vivoni et al. [2004]). As shown by Mascaro et al. [2015], tRIBS can be used over large basins (~3800 km^2) to perform fine-resolution simulations at a very fine nominal resolution of (~78 m). The Penn State hydrologic model (PIHM, Qu and Duffy [2007]) is another example of a distributed hydrologic model using unstructured meshes instead of a regular grid. Marsh et al. [2012] showed that TINs were superior to structured grids when representing mountain shading, which means that uniform grids have to use a higher resolution DEM than TINs in order to avoid spatial artifacts of slope. Many studies therefore linked PIHM with a regular gridded snow model in order to model melt flux (Kumar et al. [2013], Wang et al. [2013], Chen et al. [2016] and Jepsen et al.

[2016]).

The main objective of this work is to develop a relatively parsimonious multi-resolution (MR) approach, that is an alternative to TINs and HRUs, using finer resolutions over areas of complex physiography (which includes both topography and vegetation) and coarser resolutions over less complex areas. The MR technique is applied herein in the context of raster-based models given their more trivial implementation using two-dimensional arrays. Our goal is to illustrate how snowmelt errors driven by model resolution can be mitigated while optimizing the computational resources available. This paper focuses on snowmelt during the ablation season because it is the main driver of runoff and is the snow process expected to be most sensitive to model resolution (Cline et al. [1998], Winstral et al. [2014]). The MR method described herein was developed and tested over a representative basin in the Western U.S. to address the following questions: 1) How does the average ablation season melt rate error vary with increasingly coarser spatial resolutions? 2) Can a simple physiographic complexity metric serve as a criterion to discretize a domain in order to mitigate the modeled snowmelt error caused by coarsening the spatial resolution? 3) Can the MR scheme be successfully applied over a large snow-dominated mountainous region to maintain accurate snowmelt estimates while decreasing computational expense? While the focus is on snow-dominated processes, the approach could be used to provide a mechanism for fine-resolution modeling in other contexts as well.

The rest of the chapter is organized as follows: Section 2.2 presents the study area and illustrates the MR approach based on the physiographic complexity metric. Section 2.3 describes the methodology employed to apply the MR approach to generate snowmelt estimates at varying spatial resolutions. Results and discussion related to the science questions listed above are presented in Section 2.4. Finally, Section 2.5 summarizes the key points of this work.

2.2 Development and implementation of the multi-resolution scheme over a snow-dominated watershed

2.2.1 Physiographic characteristics of the study area

This study used the snow-dominated Upper Yampa River headwater catchment (6770 km^2; Figure 2.1) in the Colorado River basin (CRB; Figure 2.1) as a representative test-bed. This watershed is located in the northeastern part of the CRB where snowmelt is highly correlated to streamflow, and its size is large enough to span a wide array of topographic and vegetation characteristics to implement and test the proposed MR scheme.

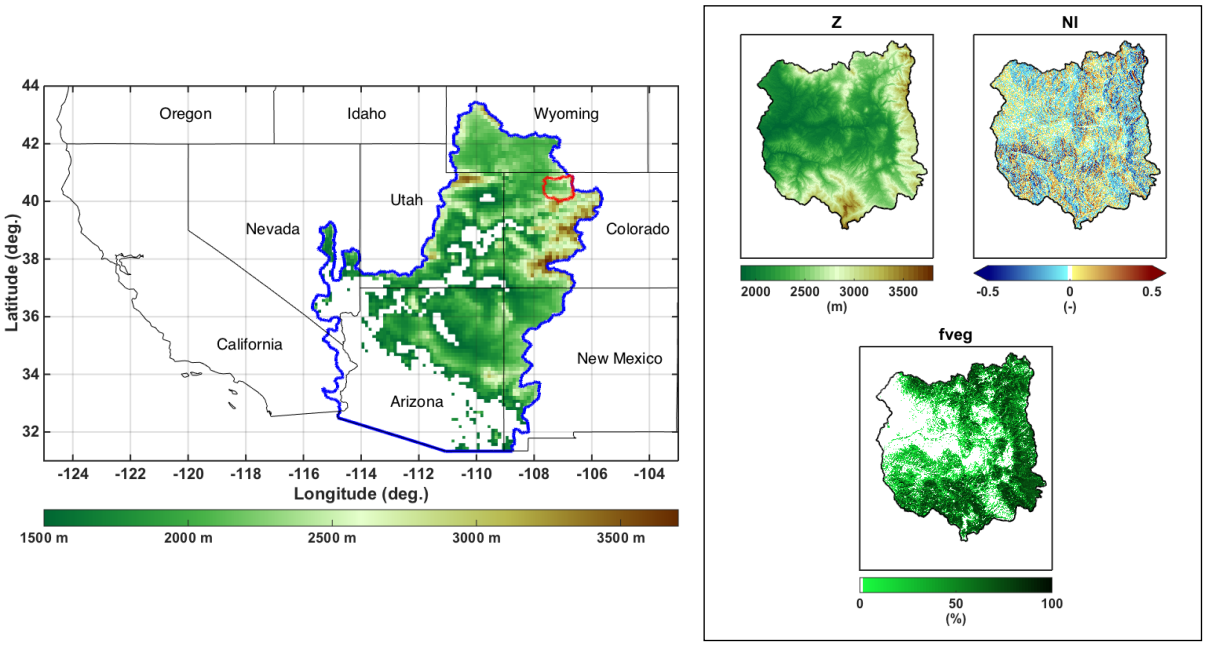

Figure 2.1: Map of the Southwestern United States showing the Colorado River Basin (CRB) elevation (left panel) with the CRB outlined in blue and the Upper Yampa River Basin (UYRB) outlined in red. Enclosed in the right panel are the basin maps of elevation (Z), northness index (NI) and forested fraction ($fveg$) for the Upper Yampa River Basin.

The advanced spaceborne thermal emission and reflection (ASTER) DEM (JPL [2009]),

which is available at a spatial resolution of 30 m was used to extract elevation, slope, aspect, sky-view-factor and shading by neighboring terrain for each pixel. Land cover and forest cover fraction data were obtained from the National Land Cover Dataset (NLCD, Homer et al. [2007]) at a spatial resolution of 30 m. As seen in Figure 2.1, the Upper Yampa River Basin (UYRB) is at a relatively high elevation, with the entire basin above 1500 m, peaks above 3750 m and a mean elevation of 2446 m. High elevation areas are densely covered with forest while most of the downstream flat and low elevation areas are sparsely covered with shrubs and grassland. As a metric to identify the combined impact of slope and aspect on radiative fluxes, the following northness index (NI) from Molotch et al. [2004] was used:

$$NI = \cos aspect \times \sin slope \tag{2.1}$$

NI is by construction constrained between -1 (vertical south facing) and 1 (vertical north facing).

2.2.2 Construction of a physiographic complexity metric

A metric representative of the underlying heterogeneity of the basin needs to be constructed in order to differentiate complex parts of the regions requiring fine model spatial resolution from less complex regions that could be more coarsely simulated. Zhou et al. [2008] defined a complexity index constructed from normalized total curvature, rugosity, local relief and local standard deviation based on a DEM raster grid, and was found to adequately represent terrain complexity. Revuelto et al. [2014] performed statistical analyses to identify which terrain characteristics were primarily responsible for controlling snow depth distribution. Snow depth distribution was linked to the topographic position index (difference between the pixel elevation and the mean elevation of its neighbors), maximum upwind slope, northing and elevation. The study area was however not vegetated, explaining the absence of vegetation control variables. The topographic position index was found to most accurately represent

the terrain complexity, and could be used as predictor of snow depth distribution.

This study similarly used the topographic complexity metrics of standard deviation of elevation (Z) and northness index (NI) in order to characterize the complexity of the underlying topographical characteristics of the domain that have the most impact on snowpack states. Since vegetation can often have a non-negligible impact in snow-dominated region (Trujillo et al. [2007]), the standard deviation of forested fraction ($fveg$) was also included. In this application, the metric was built at a fine-resolution of 90 m. Specifically, standard deviations of Z, NI, and $fveg$ were calculated for each 90 m pixel i ($\sigma_{Z_i}, \sigma_{NI_i}$, and σ_{fveg_i}) using the nine sub-grid 30 m cells from the maps described in Section 2.2.1 (Z_k, NI_k, and $fveg_k$,as well as their mean values (μ_{Z_i}, μ_{NI_i}, and μ_{fveg_i}):

$$\begin{aligned}
\sigma_{Z_i} &= \sqrt{\frac{1}{9}\sum_{k=1}^{9}(Z_k - \mu_{Z_i})^2} \\
\sigma_{NI_i} &= \sqrt{\frac{1}{9}\sum_{k=1}^{9}(NI_k - \mu_{NI_i})^2} \\
\sigma_{fveg_i} &= \sqrt{\frac{1}{9}\sum_{k=1}^{9}(fveg_k - \mu_{fveg_i})^2}
\end{aligned} \tag{2.2}$$

The standard deviations were then normalized using their respective minimum and maximum values to be rescaled to a [0 -1] range as follows:

$$\begin{aligned}
\hat{\sigma_{Z_i}} &= \frac{\sigma_{Z_i} - min(\sigma_Z)}{max(\sigma_Z) - min(\sigma_Z)} \\
\hat{\sigma_{NI_i}} &= \frac{\sigma_{NI_i} - min(\sigma_{NI})}{max(\sigma_{NI}) - min(\sigma_{NI})} \\
\hat{\sigma_{fveg_i}} &= \frac{\sigma_{fveg_i} - min(\sigma_{fveg})}{max(\sigma_{fveg}) - min(\sigma_{fveg})}
\end{aligned} \tag{2.3}$$

where $\hat{\sigma_{Z_i}}$, $\hat{\sigma_{NI_i}}$, and, $\hat{\sigma_{fveg_i}}$ are the normalized standard deviations for each 90 m pixel i and $min(\sigma_Z)$, $max(\sigma_Z)$, $min(\sigma_{NI})$, $max(\sigma_{NI})$, $min(\sigma_{fveg})$, and $max(\sigma_{fveg})$ are the respective

minimum and maximum standard deviations over the study area. The three normalized variables were then summed to form the physiographic complexity metric CM:

$$CM_i = \hat{\sigma_{Z_i}} + \hat{\sigma_{NI_i}} + \hat{\sigma_{fveg_i}} \tag{2.4}$$

where CM_i is the physiographic complexity metric for a 90 m pixel i. The normalized standard deviation maps and the resulting complexity metric over the UYRB are illustrated in Figure 2.2.

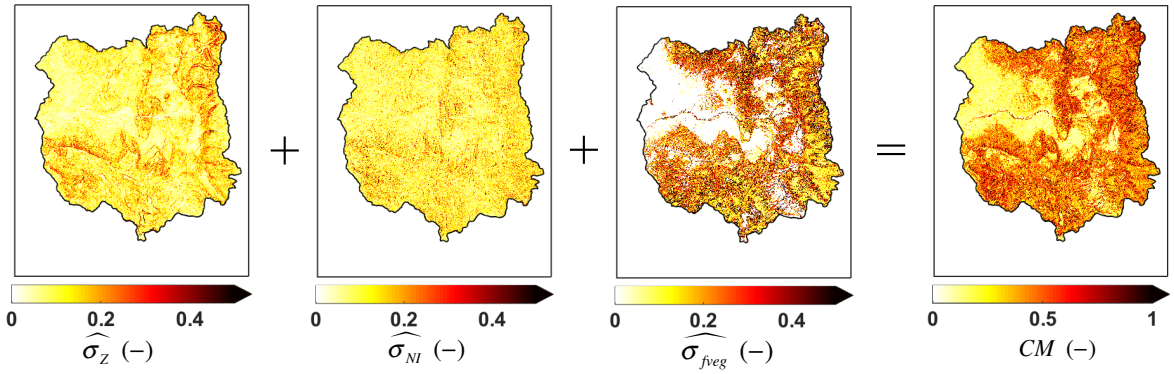

Figure 2.2: Maps of $\hat{\sigma_{Z_i}}$, $\hat{\sigma_{NI_i}}$, and, $\hat{\sigma_{fveg_i}}$ at 90 m over the Upper Yampa River Basin. The resulting physiographic complexity map is illustrated in the far right panel.

In addition to elevation and northness index, the CM metric also takes into account the heterogeneity in forested fraction. Higher CM values indicate higher degrees of physiographic complexity, and the metric may therefore serve as a good predictor of spatial patterns of melt rate errors because it is physically consistent with the possible sources of snowmelt errors in a snow model. Specifically, high σ_Z, σ_{NI}, and σ_{fveg} values are indeed indicators of drifting and avalanching, variability in melt energy, snow interception, sublimation and non-uniform unloading by the canopy, etc. (Clark et al. [2011]). It is thus reasonable to expect that areas of high CM would likely need to be resolved at higher resolutions than pixels with low CM

if the goal is to mitigate errors.

2.2.3 Multi-resolution domain discretization

The MR scheme developed herein is defined by the specification of a maximum CM threshold (CM_{max}) that is chosen based on the computational resources available to model pixels at a fine-resolution (i.e. 90 m in this study). For instance, choosing a CM_{max} value corresponding to the 90^{th} percentile aims to resolve the 10% most complex pixels at 90 m, while the remaining pixels are modeled at equal or coarser resolutions. In doing so, the method allows for tying the choice of the complexity threshold to the computational resources available. Choosing the resolution of a uniform grid is most commonly done based on such practical constraints, however the MR scheme instead targets the resources to modeling pixels at higher resolutions where doing so will likely mitigate errors. Here, for simplicity in the aggregation process, we chose to use multiples of the baseline fine-resolution (i.e. 90 m, 180 m, 360 m or 720 m), but alternative or larger resolutions could be used instead. As shown in Figure 2.3, the basic idea behind the MR approach is to span from coarsest to finest resolutions using a parsimonious scheme to identify which pixels should be modeled at which resolution. In this application, if all sub-grid 90 m pixels within a 720 m pixel have a CM below the threshold CM_{max}, all 64 sub-grid pixels are aggregated to one 720 m pixel. If not, we look at all 90 m pixels within each sub-grid 360 m pixel and proceed the same way until all of the original 90 m pixels have been assigned a resolution and aggregated where possible. This MR scheme is therefore a pre-processing step that identifies the spatial distribution of pixel resolution and their associated static model characteristics. Once determined, the offline modeling can proceed pixel-wise as would be done for a uniform grid.

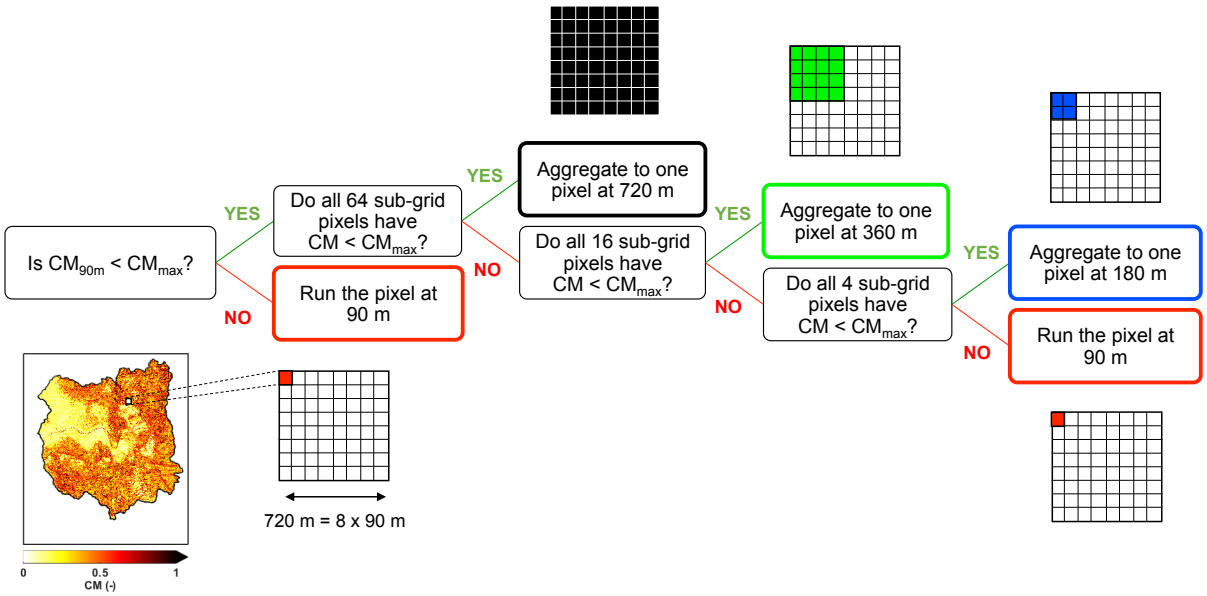

Figure 2.3: Illustration of the multi-resolution pre-processing scheme using (as an example) a fine-resolution of 90 m and integer aggregation to coarser resolutions depending on the physiographic complexity.

As done in Vivoni et al. [2004], the performance of the MR scheme in terms of computational resources can be quantified by a data reduction metric d:

$$d = (1 - \frac{n_{coarse}}{n_{baseline}}) \times 100 \tag{2.5}$$

where $n_{baseline}$ is the number of baseline pixels at the uniform fine-resolution (90 m) grid and n_{coarse} is the number of pixels from either coarser uniform or MR grids. For instance, modeling the basin with a uniform grid of 180 m would divide the number of original baseline pixels by a factor of four, leading to a data reduction of 75%. A MR case using a CM_{max} value corresponding to the 90th percentile will have a data reduction factor that depends on the spatial distribution of the basin physiographic complexity. The best case scenario in terms of computational resources savings for the UYRB (10% of the pixels resolved at 90 m and the remaining 90% modeled at 720 m) would lead to a data reduction factor of

89.4%, while the worst case scenario (10% of the pixels resolved at 90 m and the remaining 90% modeled at a combination of 90 m and 180 m) would lead to a data reduction factor of 45.5%.

2.3 Application of the MR approach to snowmelt modeling

The MR scheme presented in Section 2.2.3 was tested using modeled average ablation season melt rates derived from SWE estimates over the UYRB at different spatial resolutions. Three water years (WYs) representative of a normal, early and late start of the ablation season (WY 1988, WY 2004, and WY 1993 respectively) were modeled. WYs start on day of water year (DOWY) 1, or October 1st, and end on DOWY 365 (366 for 1988 which was a leap year), or September 30$^{\text{th}}$. Different sets of model estimates were generated using uniform grids at 180 m, 360 m and 720 m and using the MR scheme for CM_{max} corresponding to the 95$^{\text{th}}$, 90$^{\text{th}}$, 85$^{\text{th}}$, 80$^{\text{th}}$ and 75$^{\text{th}}$ CM percentiles (MR95, MR90, MR85, MR80, and MR75 respectively). The accuracy of each coarser resolution scheme was then assessed by comparison with a 90 m baseline case.

2.3.1 Model framework and forcing: generating a baseline of SWE estimates at 90 m

2.3.1.1 Model setup

This study is based on deterministic simulations that used the same modeling setup as Margulis et al. [2015] and Margulis et al. [2016], involving the Simplified simple biosphere (SSiB) model developed by Xue et al. [1991], coupled with a three-layer snow and atmosphere soil transfer (SAST) model (Sun and Xue [2001], Xue et al. [2003]). In addition to the static inputs described in Section 2.2.1, the dynamic meteorological forcings needed by the LSM were obtained from the Phase 2 North American Land Data Assimilation System (NLDAS-2, Cosgrove et al. [2003]; Xia et al. [2012]) hourly forcing dataset, which is available since

1979. NLDAS-2 variables include precipitation, incident shortwave radiation, near-surface air temperature, humidity, wind speed and pressure at a coarse spatial resolution of 1/8°. The NLDAS-2 forcings were downscaled to the model resolution using topographic correction methods that have been previously applied over the Sierra Nevada and the Andes (Girotto et al. [2014b], Margulis et al. [2016] and Cortés et al. [2016]). Elevation lapse rates of 6.5°K/km and 4.1°K/km were used for air temperature and dewpoint temperature respectively. Downscaling approaches for atmospheric pressure, specific humidity, and the incoming longwave and shortwave radiation fluxes are explained in detail in Appendix A of Girotto et al. [2014b]. The NLDAS-2 precipitation data has been shown to be biased in mountainous terrain (Pan et al. [2003], De Lannoy et al. [2010], Girotto et al. [2014b] and Margulis et al. [2015]). Therefore, in order to provide more realistic snow water equivalent (SWE) realizations at high spatial resolutions for the baseline case, the downscaling of NLDAS-2 precipitation forcings were first improved using the data assimilation framework developed by Margulis et al. [2015] (Supporting Information). The data assimilation was only used to generate realistic precipitation scaling coefficients that were used to downscale the NLDAS-2 precipitation forcing to the baseline model resolution. All results shown below are from deterministic forward simulations (no data assimilation) with the improved precipitation inputs.

2.3.1.2 SWE outputs

Deterministic simulations were run at a uniform spatial resolution of 90 m at an hourly time step to generate SWE estimates for WYs 1988, 1993 and 2004 that were chosen as test years representative of a normal, late and early start of the ablation season respectively. Figure 2.4 illustrates the temporal dynamics of the basin-averaged SWE estimates. WY 2004 shows a shorter accumulation season with a peak SWE of 0.18 m on DOWY 161, which is significantly smaller than for the other two years. WYs 1988 and 1993 experienced similar peak SWE values (0.32 m and 0.36 m respectively), with WY 1988 ablation starting earlier on DOWY 177, leading to a complete melt out by the end of the summer. WY 1993 shows

a delayed start of the ablation season on DOWY 202, which led to snow remaining on the ground for some high-elevation pixels at the end of the water year.

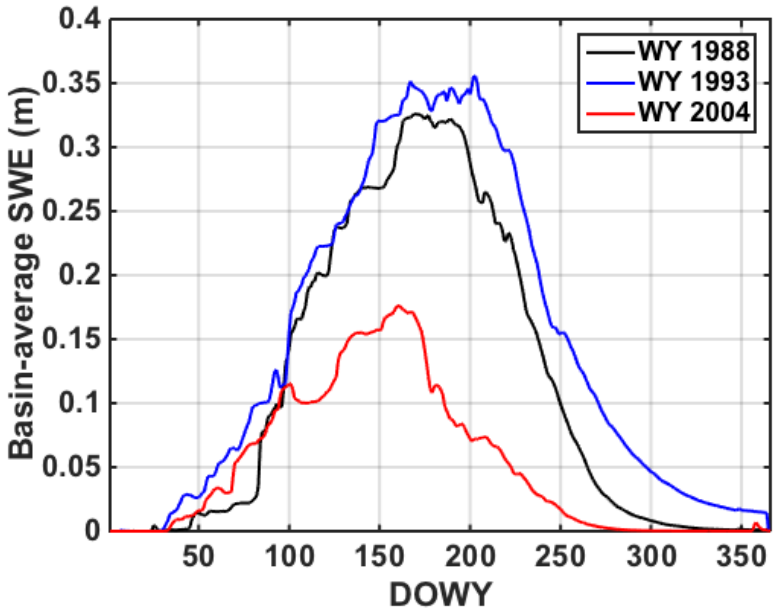

Figure 2.4: Baseline basin-averaged SWE time series.

2.3.2 Baseline ablation melt rate estimates

SWE outputs from the baseline scenario were used to evaluate the average amount of snowmelt during the ablation season and how it varies spatially. This snowmelt metric is referred to as the average ablation melt rate (AMR) and was calculated from peak SWE and the duration of melt season:

$$AMR_i^{Baseline} = \frac{SWE_{peak_i}^{90m}}{t_{m_i} - t_{peak}} \tag{2.6}$$

where $SWE_{peak_i}^{90m}$ is SWE modeled at 90 m on the day of peak basin-average SWE (t_{peak}; Figure 2.4) for pixel i, and t_{m_i} is the first day of melt-out following t_{peak} on pixel i. The spatial AMR fields generated at the uniform 90 m resolution were used as the reference (or

baseline) solution, throughout the study in order to ultimately characterize errors introduced in melt as a result of aggregation.

In WY 1988 the snowpack started melting on March 25$^{\text{th}}$ and the entire basin had melted out in 143 days (Table 2.1), while in WY 1993 the snowpack started melting 25 days later and snow partially remained on September 30$^{\text{th}}$. WY 1993 experienced the highest basin-averaged AMR of 0.63 cm/day, followed by WY 1988 and 2004 with values of 0.48 cm/day and 0.38 cm/day respectively. Since WYs 1988 and 1993 had similar SWE_{peak}, the difference in AMR observed was due to the duration of melt and not the magnitude of SWE. Even though some snow remained at the end of the year, the melt rate was higher for WY 1993, given the later start of the ablation season. In contrast, WY 2004 had a SWE_{peak} value 50% lower than for the other two years. Even though the snowpack started melting on March 9$^{\text{th}}$, 16 days earlier than for WY 1988, and completely melted out in 119 days, the shorter ablation season was not sufficient to counter the effect of the lower SWE_{peak}.

Water Year	**1988**	**1993**	**2004**
$\boldsymbol{SWE_{peak}}$ **(m)**	0.32	0.36	0.18
$\boldsymbol{t_{peak}}$	25 March 1988 DOWY 177	19 April 1993 DOWY 202	9 March 2004 DOWY 161
$\boldsymbol{AMR}$ **(cm/day)**	0.48	0.63	0.38
Day of disappearance	15 August 1988 DOWY 320	No complete melt over some pixels	6 June 2004 DOWY 280

Table 2.1: Simulated basin-averaged characteristics of the ablation season for the three representative water years (see Figure 2.4).

Figure 2.5 illustrates the baseline SWE_{peak}, melt duration and AMR maps generated for all three water years. For WY 1988, high SWE accumulation and melt duration are observed in the eastern and southwestern parts of the basin that correspond to high Z and

fveg areas (Figure 12.1). The eastern part of the basin also shows high *NI*. These areas showed large AMR as well, which means that the high melt rates observed were driven by larger SWE_{peak} and not a shorter ablation season. The same spatial pattern is observed for WY 1993, although more contrast is observed due to the more heterogeneous spatial SWE_{peak} distribution and later start of the ablation season. WY 2004, which experienced an early start of the ablation season, shows the opposite pattern. SWE_{peak} is lower than for the other WYs and more homogeneously distributed over the basin. AMR is at its highest for bare, flat pixels at low elevations located in the center of the watershed, which is driven by the shorter melt duration compared to the rest of the basin.

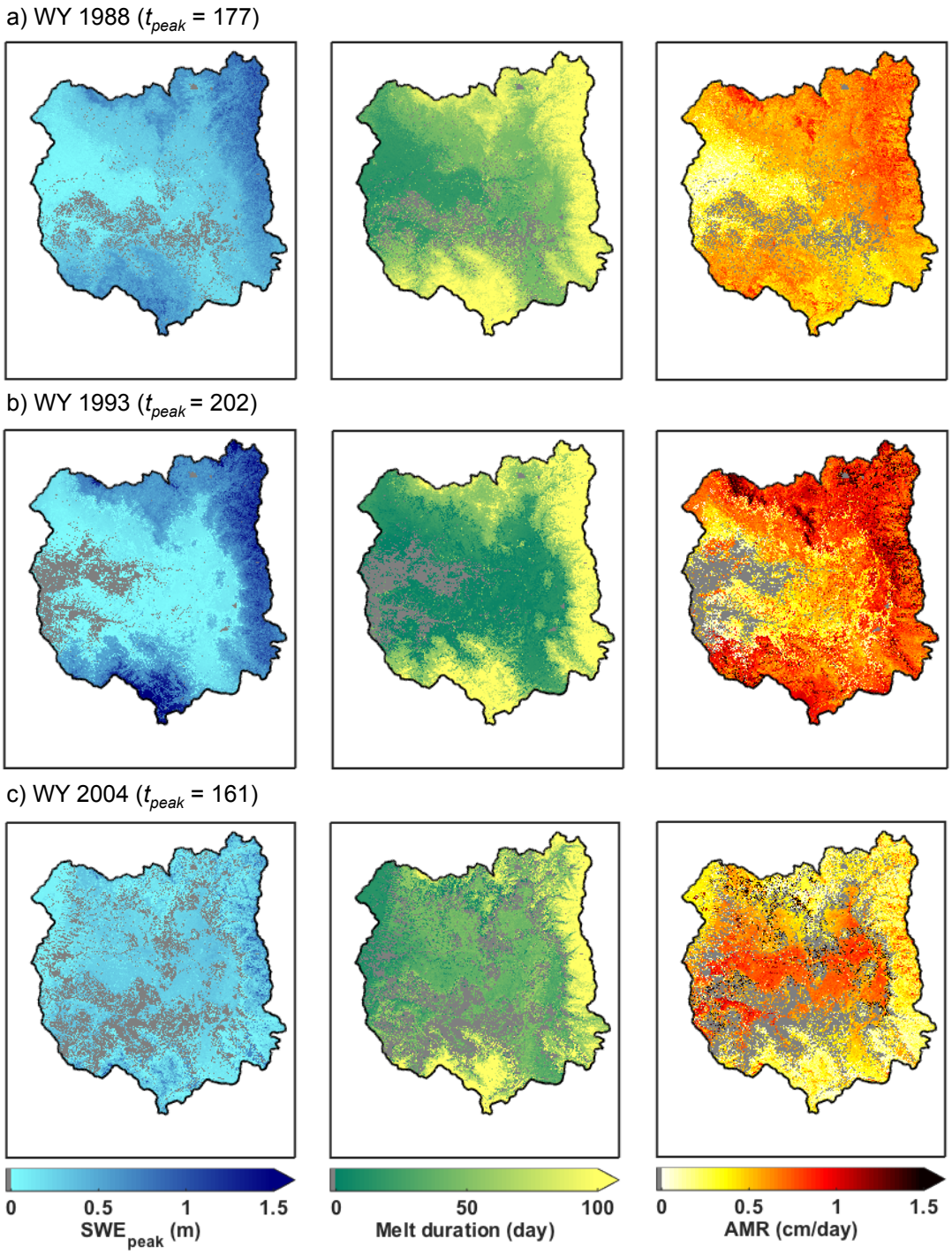

Figure 2.5: Baseline maps of SWE_{peak}k (left column) melt duration (middle column) and AMR (right column) for the Upper Yampa River Basin for WYs (a) 1988, (b) 1993, and (c) 2004. The dates of peak SWE are specified in DOWY for each year. Areas with zero-valued SWE are shaded in grey.

2.3.3 Characterizing the performance of the MR scheme compared to the uniform grid-based method

2.3.3.1 Coarser simulations setup

The baseline precipitation forcing at 90 m was aggregated to the following increasingly coarser resolutions: 180 m, 360 m, and 720 m. Elevation and forested fraction were first aggregated to coarser resolutions by averaging over the 4, 16, or 64 neighboring 90 m subgrid pixels in order to be upscaled to 180 m, 360 m, or 720 m respectively. Landcover was aggregated similarly, but by taking the mode instead of the mean since it is a classification-based variable. Coarser slope, aspect, sky view factor and shade were calculated from the aggregated elevation. Regarding the LSM parameters, conductivity and other soil parameters are based on lookup tables based on landcover classification and will therefore vary accordingly with the aggregated landcover classification. The dynamic meteorological forcings were downscaled to the pixel resolution (either 90 m, 180 m, 360 m, or 720 m) from their native 1/8° spatial resolution using the topographic correction methods described in Section 2.3.1.1. Regarding the downscaling of the precipitation forcing, the 90 m precipitation scaling coefficients (Supporting Information) were first averaged over the 4, 16, or 64 neighboring sub-grid pixels in order to be upscaled to 180 m, 360 m, or 720 m respectively. The coarser precipitation scaling coefficients were then used to downscale the raw NLDAS precipitation forcing as described in the Supporting Information. The coarser deterministic simulations were then performed for each water year using the aggregated inputs described above.

2.3.3.2 Comparison of the coarser outputs with the baseline

In order to quantify the difference between the baseline AMR and the coarser estimates, the coarser SWE_{peak}, melt duration and AMR outputs were first resampled to 90 m before being compared to the 90 m reference solution. The differences were calculated as follows:

$$
\begin{aligned}
(SWE_{peak}\ difference)_i^{90m} &= (SWE_{peak})_i^{Coarser} - (SWE_{peak})_i^{Baseline} \\
(Melt\ duration\ difference)_i^{90m} &= (t_{m_i} - t_{peak})^{Coarser} - (t_{m_i} - t_{peak})^{Baseline} \\
(AMR\ difference)_i^{90m} &= (AMR)_i^{Coarser} - (AMR)_i^{Baseline}
\end{aligned}
\tag{2.7}
$$

where $(SWE_{peak}\ difference)_i^{90m}$, $(\ Melt\ duration\ difference)_i^{90m}$ and $(\ AMR\ difference)_i^{90m}$ are respectively the differences in SWE_{peak}, melt duration and AMR between the coarser simulation and the 90 m baseline for pixel i. The Baseline superscript refers to the baseline simulation at 90 m and the Coarser superscript refers to the coarser simulation outputs. They both apply to the same pixel i since all simulations are resampled to 90 m. Hence a positive difference means that the baseline is overestimated and a negative difference means that it is underestimated. By construct, the MR cases have differences that are equal to zero for all pixels modeled at 90 m.

2.4 Results and discussion

2.4.1 Implementation of the MR scheme

All of the cases tested, as well as their respective number of pixels covering the UYRB and the resulting data reductions are presented in Table 2.2. For instance, the 20% most complex pixels were modeled at 90 m for the MR80 case, while the remaining 80% were resolved at 720 m, 360 m, 180 m or 90 m depending on the complexity of their neighboring sub-grid pixels, which resulted in a total data reduction of 53% compared to the baseline. The MR75 through MR90 cases had less data reduction compared to the uniform 180 m case, while the MR95 case showed a slightly superior data reduction. It is worth noting that the data reduction term d is directly correlated to memory and computation time because all pixels were run independently and the process was fully parallelized. Therefore, a data reduction of 50% led to memory and computation time being cut in half. As stated in Section 2.2.3, the MR scheme is a pre-processing step that has to be performed only once, and the increase

in time to implement the MR setup was negligible when compared to the resulting decrease in computational time required to run simulations.

Case	Uniform Grid				Multi-Resolution (MR)				
	90m	180m	360m	720m	MR95	MR90	MR85	MR80	MR75
Number of pixels	839970	209240	52304	13065	163080	258990	333800	395070	445770
Data reduction d	0%	75%	93.8%	98.4%	80.6%	69.2%	60.3%	53%	46.9%

Table 2.2: Simulated basin-averaged characteristics of the ablation season for the three representative water years (see Figure 2.4).

Figures 2.6a-e show maps of the different multi-resolution configurations that were generated using different CM_{max} thresholds. Figure 2.6a shows that a chosen CM_{max} value corresponding to the 95^{th} percentile led to only 12% of the total number of pixels resolved at 90 m, with the flat northwestern part of the basin modeled almost entirely at 720 m. As the specified CM_{max} threshold decreases, Figures 2.6b to 2.6e show that much larger areas are modeled at 90 m, while the portion of areas modeled at 360 m and 720 m decreased. This shows the flexibility of the MR scheme to be adapted to the computational constraints available for the modeling problem.

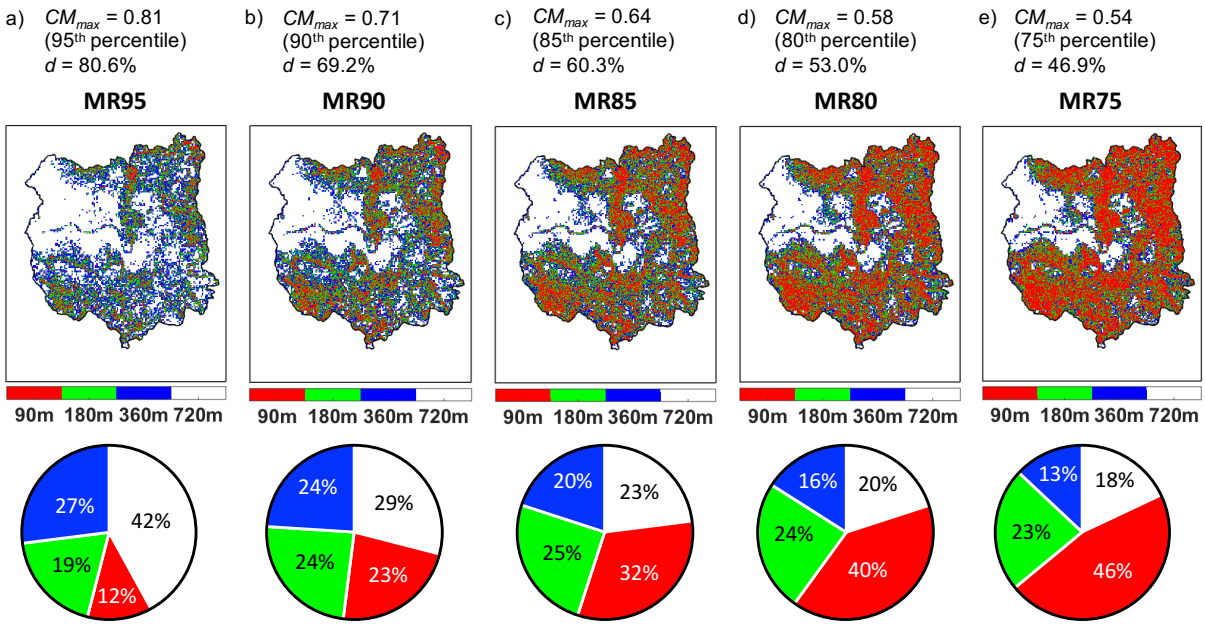

Figure 2.6: (a)–(e) Derived multi-resolution grids across the Upper Yampa River Basin using different CM_{max} values ranging from the 95th to the 75th percentile. The breakdowns of the spatial resolutions used in each scheme (as a fraction of the total number of pixels) are presented in pie charts below the maps.

2.4.2 Basin-averaged differences in SWE_{peak}, melt duration, and AMR

Figure 2.7 shows basin-averaged differences in SWE_{peak}, melt duration and AMR for each WY and case tested as a function of the data reduction metric. The absolute differences for all three metrics increase monotonically with aggregation. SWE_{peak} is on average underestimated by the coarser simulations with differences spanning from 0 cm for the MR75 case for WY 1993 to - 3 cm for the 720 m case for WYs 1988 and 2004. In contrast, the average melt duration is being overestimated with differences ranging from 0 day for the MR75 case for WY 1988 to 5.5 days for the 720 m case for WY 2004. The ratio of negative SWE_{peak} differences to positive melt duration differences hence leads to negative AMR differences on average, which vary from 0.005 cm/day for the MR75 case for WY 1988 to 0.12 cm/day. In

contrast to WYs 1988 and 1993, whose differences in AMR are influenced by the combination of both SWE_{peak} and melt duration, the mean AMR differences for WY 2004 follow the exact same pattern as the differences in SWE_{peak}, which means that the underestimation of SWE_{peak} with aggregation is the main driver of AMR differences for that particular year.

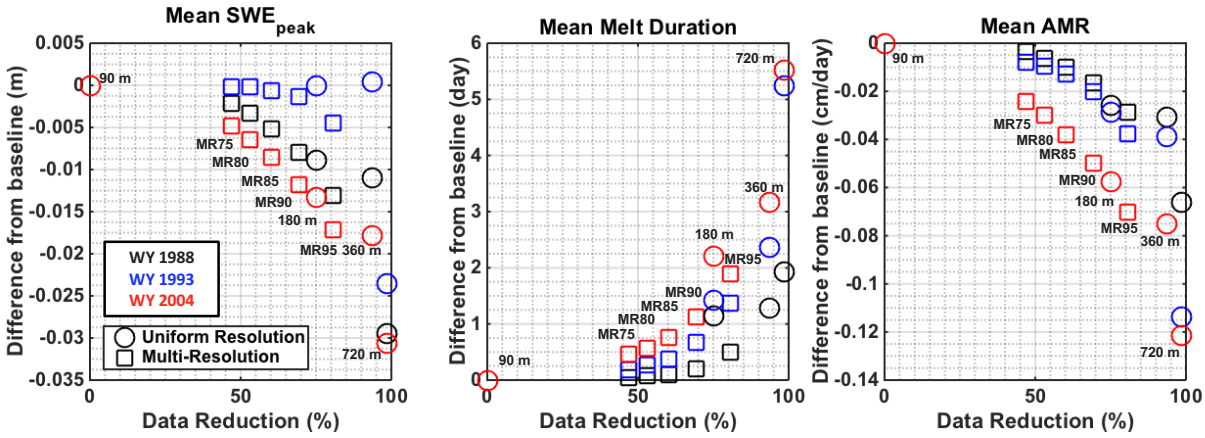

Figure 2.7: Basin mean difference in SWE_{peak} (left panel) melt duration (middle panel) and AMR (right panel) between coarser simulations and the baseline for each water year and resolution scheme tested.

Since the amount of precipitation was identical for each simulation (Supporting Information), the difference in SWE_{peak} can be explained by the changes in elevation and forested fraction seen in Figure 2.8. First, the number of high elevation pixels with Z larger than 3500 m decreases across the coarser uniform simulations, with the 180 m, 360 m and 720 m cases having respectively 10%, 15% and 30% less high elevation pixels than the baseline. Aggregating the model resolution therefore decreases the basin elevation, which means that more precipitation will be classified as rain instead of snow, leading to the underestimation of SWE_{peak} observed in Figure 2.7. This decrease of high elevation pixels is not observed in the MR cases, which means that the MR scheme improved the representation of the basins elevation. However, the MR cases also showed the same underestimation of SWE_{peak}, which means that the change in forested fraction also has a significant impact on SWE accumula-

tion. As seen in Figure 2.8, the number of densely forested pixels with $fveg > 75\%$ increases with aggregation for all MR and uniform cases, with values ranging from 5% to 20% more densely forested pixels for the MR75 and the 720 m cases respectively. Canopy interception has therefore a greater impact in coarser simulations compared to the baseline, which explains the underestimation of SWE_{peak} observed in Figure 2.7.

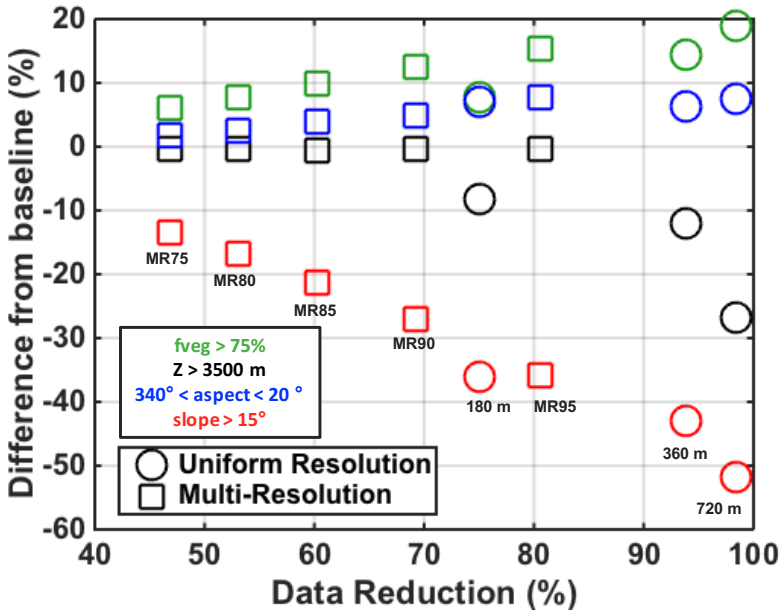

Figure 2.8: Basin mean relative difference in number of densely forested pixels (green), high elevation pixels (black), north facing pixels (blue) and steep pixels (red) between coarser simulations and the baseline.

The averaged overestimation of melt duration can be traced back to the changes in slope and aspect seen in Figure 2.8. As the model resolution coarsens, the terrain is smoothed out with more gentle slopes as a much faster pace than the change in elevation and forested fraction observed earlier. The number of steep pixels with slopes larger than 15° decreases from 15% for the MR75 case to over 50% for the 720 m case. Since the zenith angle at noon for the UYRB varies between 40° and 16° during the ablation season, more gentle slopes

will be therefore less perpendicular to the direct incoming solar radiation beam. In addition to the decrease in slope, Figure 2.8 shows that the number of north facing pixels increases with aggregation. This implies more shading effects, and therefore an average decrease in direct incoming shortwaves. The combination of more gentle and north facing slopes explains the average overestimation of melt duration observed in Figure 2.7 for the UYRB. As the basin becomes more forested with coarser simulations, the canopy shading of the incoming solar radiation can also have an impact and increase the average melt duration. However, the increase in canopy cover will also counterbalance the shading effect on incoming solar radiation by increasing the incoming longwave radiation received by the snowpack, which means that the changes in slope and aspect are the most important driver of melt duration overestimation with aggregation.

Focusing on average differences for uniform gridded simulations, the rate of increase with data reduction varied across the three years tested. WY 2004 experienced the steepest increase in terms of AMR, with differences twice as large as the differences in WY 1988 for the 720 m, 360 m and 180 m cases (-0.12 cm/day vs. -0.065 cm/day, 0.075 cm/day vs. -0.03 cm/day and -0.06 cm/day vs. 0.025 cm/day respectively). In addition to the inter-annual variation, there is a clear gap in performance between the 720 m case and the 180 m or 360 m cases. Differences in the three metrics increased slowly from 180 m to 360 m, with the exception of WY 1993 where both 180 m and 360 m SWE_{peak} were almost identical to the baseline. As seen in Figures 2.2 and 2.5, SWE_{peak} was concentrated in areas that have a homogeneous vegetation fraction, which will not change the canopy cover drastically with aggregation. Since aggregating to 180 m or 360 m does not impact high elevation pixels (Figure 2.8), there is no degradation of average SWE_{peak} representation with data reduction in that particular case.

Focusing on the MR cases, the MR75, MR80, MR85 and MR90 cases bridged the gap in accuracy between the baseline and 180 m for all years and all three metrics, while showing less inter-annual variability. MR simulations significantly improved melt duration estimates compared to uniform simulations, leading to the MR95 case being superior to the 180 m case

in terms of both data reduction and accuracy. The MR75 case showed negligible differences with respect to the baseline, while having a data reduction of 46.9%, and was therefore computationally superior to the baseline without inducing significant errors at the basin scale due to aggregation. Implementing the MR90 scheme over the UYRB would allow a saving in terms of computational resources comparable to running the entire basin at 180 m (69.3% vs. 75.0%) while being closer to the baseline (in fact twice as close in terms of melt duration) for all WYs tested.

2.4.3 Variation in *AMR* difference as a function of physiographic complexity and spatial aggregation

Maps of *AMR* difference for the three WYs are shown for selected coarser simulations in Figure 2.9. Areas of the basin with south facing slopes (negative *NI* in blue in Figure 2.1) tend to underestimate *AMR* (negative *AMR* difference from baseline shown in blue in Figure 2.9) because the decrease in steep slopes with aggregation reduces the amount of direct incoming solar radiation, which slows down snowmelt. In contrast, areas with north facing slopes (positive *NI* in yellow/red in Figure 2.1) tend to overestimate *AMR* (positive *AMR* difference from baseline shown in red in Figure 2.9) because the decrease in steep slopes with aggregation reduces shading effects, effectively increasing the amount of direct incoming solar radiation, and accelerating snowmelt. Significant amount of SWE_{peak} covered large most of the basin for WYs 1988 and 1993 (Figure 2.5), leading to both over and underestimation of *AMR*, which cancel out when averaging differences over the basin. In contrast, WY 2004 was very dry (Figure 2.4) and accumulated a significant amount of SWE_{peak} over only very complex regions of the basin. This leads to larger underestimation of SWE_{peak} with aggregation due to the loss of high elevation and the gain of densely forested pixels. *AMR* differences, driven by the underestimation of SWE_{peak} (Figure 2.7) are therefore mostly negative for WY 2004, which explains why the basin averages are the lowest over the three years.

The underperformance of the 720 m resolution can again be observed in Figure 2.9

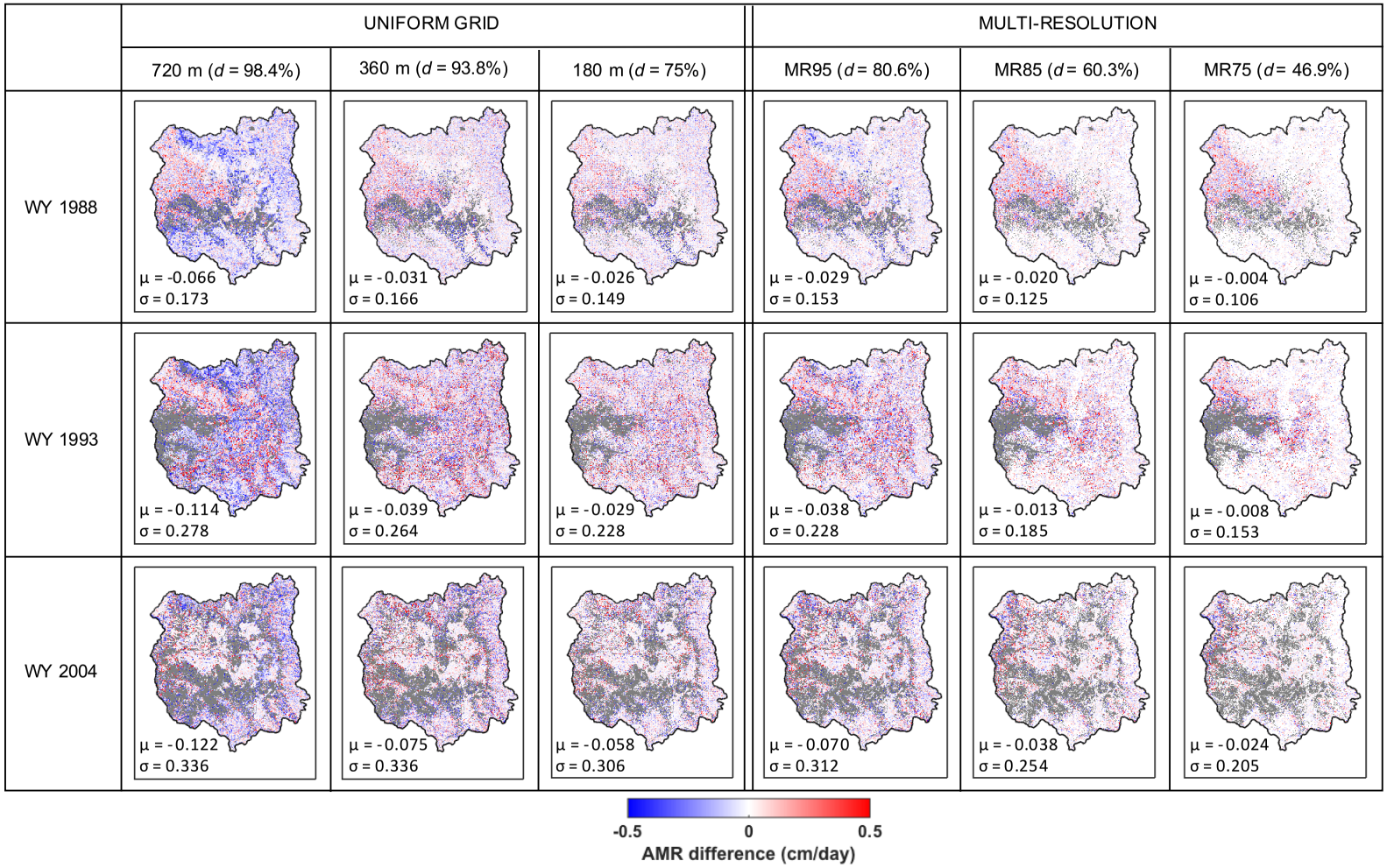

Figure 2.9: Maps of *AMR* difference from baseline (in cm/day) for WYs 1988, 1993 and 2004 at 720 m, 360 m, 180 m, and the MR95, MR85 and MR75 cases. The basin mean difference and standard deviation in cm/day are displayed in the lower left corner of each map. Areas with zero-valued SWE are shaded in grey.

for all three WYs, while the 360 m, 180 m and MR95 cases display similar differences, in terms of both spatial variability and basin averaged values. *AMR* difference maps for the MR85 and MR75 cases highlight the increase in accuracy compared to the coarser cases. Areas with low *CM* that were resolved at coarser resolutions constitute the main source of differences from the baseline, with values as high as 0.25 cm/day. However, these areas accumulated very little SWE (Figure 2.5), and therefore do not significantly contribute to the basin total snowmelt. The goal of the MR approach is to minimize *AMR* differences in areas with high SWE accumulation, which generate critical snowmelt runoff. Figure 2.9 shows that this is clearly the case for the MR85 and MR75 simulations, where the parts of the basin accumulating the most SWE (Figure 2.5 showed almost no difference from the baseline for all three WYs.

The relationship between physiographic complexity and snowmelt errors is illustrated in Figure 2.10, where *AMR* absolute difference is shown to generally increase with *CM* in all uniform simulations and for all three water years. For WY 1988, the median *AMR* absolute difference stayed constant around 0.05 cm/day before increasing with *CM* values above the 70^{th} percentile, reaching a maximum around 0.07 cm/day. WY 2004 displayed a similar behavior while WY 1993 experienced a constant and steep increase from around 0.05 cm/day to around 0.10 cm/day. Due to the nature of the MR approach, pixels with higher percentiles of *CM* are more likely to correspond to the 90 m baseline and hence the difference decreases. This explains why all MR cases for all three years showed a decrease of median *AMR* absolute difference with CM from around 0.05 cm/day to 0 cm/day.

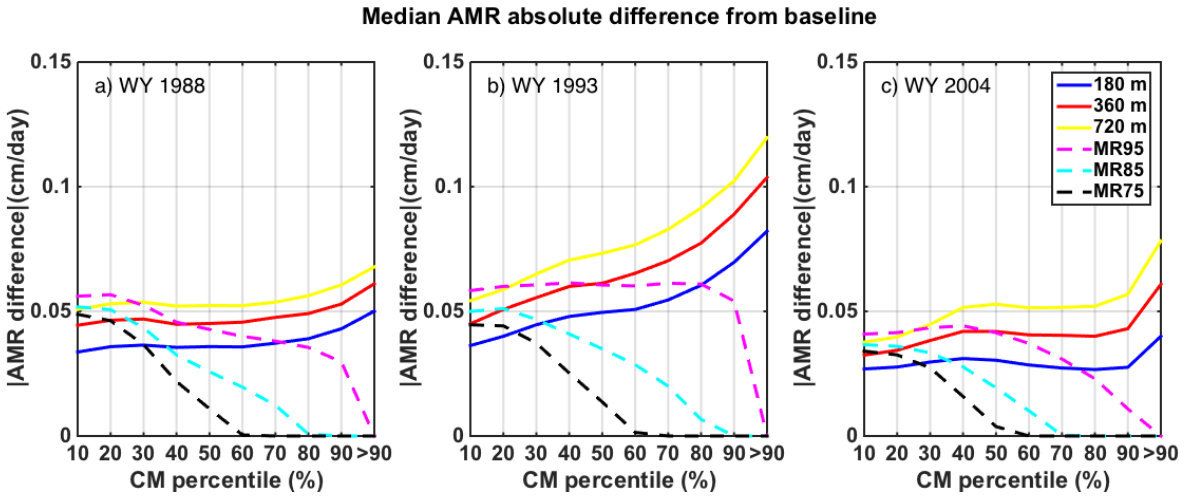

Figure 2.10: Median AMR absolute difference distribution as a function of CM for WYs (a) 1988, (b) 1993, and (c) 2004. The uniform 720 m, 360 m, and 180 m resolutions are displayed as solid lines, and the MR95, MR85 and MR75 cases are displayed as dashed lines.

2.4.4 Cumulative impact of aggregation on snowmelt

Snowmelt was calculated pixel-wise for each case and the end-of-WY cumulative value was extracted in order to quantify the total amount of snowmelt runoff available over the basin for WYs 1988, 1993 and 2004. The baseline simulation over the Upper Yampa River Basin yielded the largest cumulative snow melt of 3.4 km^3 for WY1993, WY 1988 showed a slightly lower cumulative snowmelt of 2.7 km^3 and WY 2004 led to the lowest cumulative snow melt of 1.7 km^3.

Figure 2.11 highlights the bulk effect of aggregation on snowmelt by showing the difference in total cumulative melt between the baseline and the coarser resolution cases. For all three years and coarser resolutions tested, the total amount of melt was underestimated. Aggregating the model resolution changed the physiographic distribution of the terrain, which resulted on average in less SWE, longer melt duration and therefore less snowmelt over the UYRB as explained in Section 2.4.2. The coarser simulations can be organized into three categories. At the end of each year, the 720 m uniform case always shows the highest

underestimation, from 7% for WY 2004 to almost 12% for WY 1993. The MR95, 360 m and 180 m cases then follow with underestimation errors on the order of 2-3% for WYs 1988 and 2004, and 4% for WY1993, with the exception of the MR95 case that shows a larger error of 7%. Finally, the MR85 and MR75 cases underestimate the baseline by only 1% for WYs 1988 and 2004, and 2% for WY 1993. Hence, melt estimates resolved using the MR95 scheme were comparable to the uniform 180 m and 360 m cases, while the MR85 and MR75 cases led to melt estimates very similar to the baseline for all three years.

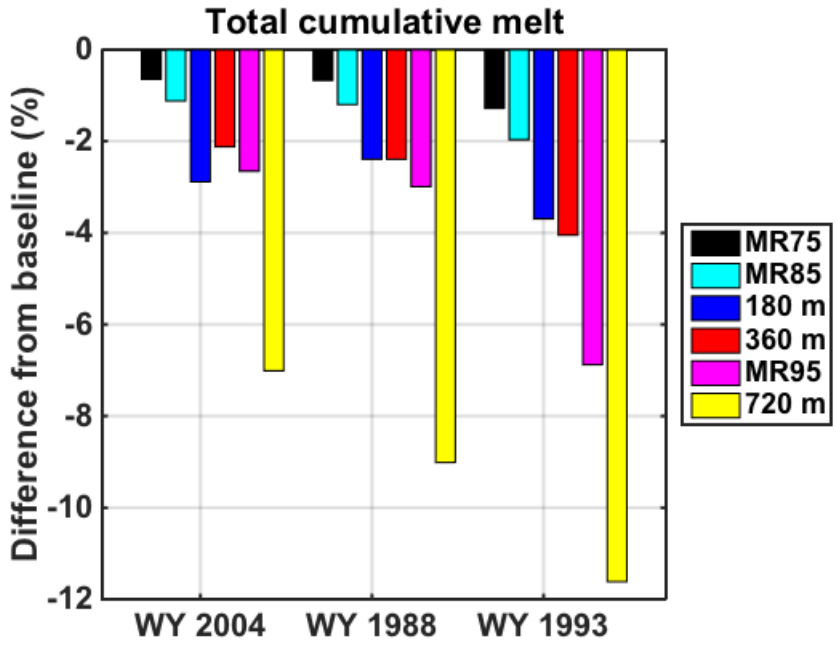

Figure 2.11: Difference in total cumulative snowmelt between the baseline and all coarser cases tested.

2.5 Summary and conclusions

This study presented a new MR grid-based terrain discretization method in order to facilitate fine-resolution modeling of snow-dominated mountainous domains. The physiographic *CM*

was constructed as a function of the standard deviations of elevation, northness index and forested fraction in order to be used as an indicator of snowmelt rate errors. The MR scheme is constrained by a user-specified input related to computational expense, and designed to mitigate the error of spatially distributed fine-resolution snow model applications. The MR technique was successfully applied over a large test watershed in the Colorado River Basin and proved to be superior in terms of modeled snowmelt accuracy to the uniform grid-based terrain discretizations (at 180 m, 360 m, and 720 m), while increasing computational efficiency.

This flexible MR approach could be applied at large scales for any offline modeling applications (including ensemble-based and data assimilation frameworks) where it would be expected to reduce snowmelt errors while simultaneously reducing computational expense. A direct future application of this work is for instance to generate fine-resolution SWE reanalysis estimates over the Colorado River Basin to extend previous applications over the Sierra Nevada (Margulis et al. [2016]) and the Andes (Cortés et al. [2016]) that were applied using a traditional uniform grid approach. Additionally, the MR method could be generalized to include complexity in other variables (e.g. soil properties) that might be important in watersheds where snow is not a dominant process (e.g. where soil moisture is the primary state variable of interest).

CHAPTER 3

Assessment of a multi-resolution snow reanalysis framework: a multi-decadal reanalysis case over the Upper Yampa River basin, Colorado

A multi-resolution (MR) approach was successfully implemented in the context of a data assimilation (DA) framework to efficiently estimate snow water equivalent (SWE) over a large head water catchment in the Colorado River Basin (CRB), while decreasing computational constraints by 60%. Thirty-one years of fractional snow cover area (*fsca*) images derived from Landsat TM, ETM+ and OLI sensors measurements were assimilated to generate two SWE reanalysis datasets, a baseline case at a uniform 90 m spatial resolution and another using the MR approach. A comparison of the two showed negligible differences in terms of snow accumulation, melt and timing for the posterior estimates (in terms of both ensemble median and standard deviation). The MR approach underestimated the baseline peak SWE by less than 2%, and day of peak and duration of the accumulation season by a day on average. The largest differences were, by construct, limited primarily to areas of low complexity, where shallow snowpacks tend to exist. The MR approach should allow for more computationally efficient implementations of snow data assimilation applications over large-scale mountain ranges with accuracies similar to those that would be obtained using ~100 m simulations. Such uniform resolution applications are generally infeasible due to the computationally expensive nature of ensemble-based DA frameworks.

3.1 Background and motivation

Spatial resolutions of 100 m or less are more commonly being recommended when using land surface models (Wood et al. [2011], Bierkens et al. [2015], Beven et al. [2015]), especially when trying to capture the heterogeneity of snowpack states in montane regions (Clark et al. [2011], Winstral et al. [2014]). Previous work using hydrologic response units (HRUs; Beven and Kirby [1979], U. S. Geological Survey et al. [1983], Sivapalan et al. [1987], Chaney et al. [2016]), or triangulated irregular networks (TINs; Tucker et al. [2001], Vivoni et al. [2004], Mascaro et al. [2015]), showed that simulating in a one size fits all (uniform grid) approach is not only computationally expensive, but also sub-optimal since only small subsets of watersheds actually require being resolved at fine spatial resolutions. Along these lines, Baldo and Margulis [2017] developed a multi-resolution (MR) scheme for raster-based models and tested it in the context of deterministic snow modeling. By adapting the grid size to the physiographic complexity of the terrain, runtime and storage needs were cut in half while preserving the accuracy of a 90 m baseline simulation.

Deterministic forward modeling itself, even at high-resolution, is often insufficient due to errors in model inputs (most notably precipitation) that are poorly characterized in montane regions. In lieu of deterministic modeling techniques, ensemble-based data assimilation (DA) methods are now frequently used to estimate snow states (Clark et al. [2006], Andreadis and Lettenmaier [2006], Su et al. [2008], De Lannoy et al. [2010], Liu et al. [2013], Arsenault et al. [2013], Girotto et al. [2014b], Margulis et al. [2015], Kumar et al. [2015]). The advantage of such approaches is to offer spatially and temporally continuous estimates, while also providing a measure of their uncertainty. However, due to their ensemble nature, such methods can be extremely expensive to run at high spatial resolutions, which at least partly explains why many of the large-scale studies cited above simulate snow processes at resolutions on the order of 1 km or greater. Simulating at these scales can solve the computational issue, but inherently sacrifices valuable information related to sub-grid heterogeneities in montane regions. This is undesirable since relevant remote sensing data streams that can act as model constraints (e.g. Lidar, Landsat, MODIS, etc.) are available at higher resolution

(from meter- to hundreds of meter scale).

The recently developed 30+ year Sierra Nevada and Andes snow reanalysis datasets by Margulis et al. [2016] and Cortés and Margulis [2017] successfully leveraged high-resolution Landsat data using a data assimilation framework applied at uniform resolutions of 90 and 180 m respectively. For these regional-scale domains, this resulted in 6 million and 5.5 million simulation pixels respectively, which were run in the context of a 100-member ensemble. For reference, given that Northern Hemisphere snow covered area is on the order of 8 million km^2 (Derksen and Brown [2012]), using a 100 m resolution would require the simulation of 8 billion pixels, a nearly four order of magnitude increase relative to the combined effort for the Sierra Nevada and Andes. Hence, extending these ensemble-based reanalysis methods to much larger scales using a uniform resolution on the order of 100 m is computationally prohibitive. Taking advantage of a MR approach to significantly reduce computational constraints might therefore greatly benefit ensemble-based DA frameworks and allow for applications at much larger scales. This paper aims to test the performance of the MR approach from Baldo and Margulis [2017] in the context of a probabilistic DA framework (Margulis et al. [2015]).

The MR approach as applied by Baldo and Margulis [2017] only impacted prior (model-based) snow estimates as a result of aggregation of model inputs. In the context of the DA framework used by Margulis et al. [2016] and Cortés et al. [2016], the MR approach will also coarsen the *fsca* observations derived from raw Landsat images (Cortés et al. [2014]), which can potentially additionally impact the accuracy of the posterior snow state estimates. We hypothesize that this additional source of aggregation error will have minimal impact on the posterior estimates because it is expected a priori that the heterogeneity of *fsca* in areas of low complexity will be minimal. Areas of high physiographic complexity typically correspond to areas of spatially heterogeneous snow accumulation and melt patterns, which drive *fsca* evolution. Applying the MR approach to *fsca* observations will therefore coarsen regions of the image where *fsca* is most likely homogeneous and refine regions where *fsca* is most likely heterogeneous, and should therefore mitigate the impact of reducing the number

of pixels on the reanalysis accuracy.

In this chapter, a high-resolution (90 m) uniform grid baseline SWE reanalysis dataset was compared to one derived using the MR scheme to address the following questions: 1) How does the MR approach impact the assimilated *fsca* observations? 2) How well does the MR approach perform in estimating the central tendency (i.e. ensemble median) of the posterior snow state distribution in space and time? 3) How well does the MR approach perform in estimating the uncertainty of the posterior snow state distribution in space and time?

The rest of the chapter is organized as follows: Section 3.2 illustrates the study area and the methodology used in this work. Section 3.3 compares the MR approach to the 90 m baseline case in order to answer the questions listed above. Finally, Section 3.4 summarizes the key points of this work.

3.2 Methodology

3.2.1 Study area

In order to maintain consistency with the work of Baldo and Margulis [2017], this study also used the Upper Yampa River Basin (UYRB, outlined in black in Figure 3.1) as a representative test domain of the Colorado River Basin (CRB). The CRB is large (6770 km^2) and snow-dominated, which makes it a critical source of fresh water for the 20 million people living downstream (Christensen et al. [2004]).

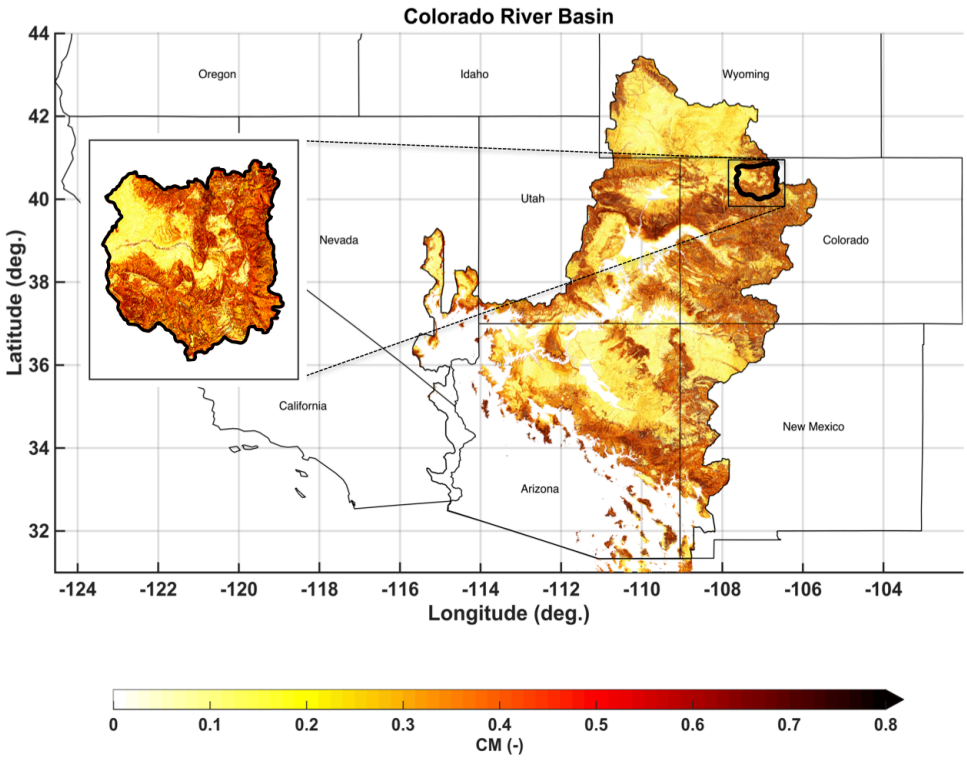

Figure 3.1: Complexity metric (CM) map of the Colorado River Basin (CRB) with the Upper Yampa River Basin (UYRB) outlined in black and displayed in more detail in the sub-panel.

In this study, the physiographic complexity metric (CM) was calculated for each 90 m pixel i across the CRB (Figure 3.1) following the approach described in Baldo and Margulis [2017]:

$$CM_i = \hat{\sigma_{Z_i}} + \hat{\sigma_{NI_i}} + \hat{\sigma_{fveg_i}} \tag{3.1}$$

where the normalized standard deviations of elevation ($\hat{\sigma_{Z_i}}$), and northness index ($\hat{\sigma_{NI_i}}$, Molotch et al. [2004]) were derived from the advanced spaceborne thermal emission and reflection (ASTER) global digital elevation model (DEM, JPL [2009]), and the normalized standard deviation of forested fraction ($\hat{\sigma_{fveg_i}}$) was derived from the National Land Cover

Dataset (NLCD, Homer et al. [2007]). Across the CRB, *CM* varies from 0 (bare and flat areas) to over 0.8 (steep and forested areas), with the UYRB sampling a similar range of complexity (Figure 3.1).

3.2.2 Multi-resolution approach

The MR algorithm begins with a pre-defined set of resolutions across which a raster-based model implementation will be applied. The finest baseline resolution is chosen to correspond to that deemed important for representing processes in high-complexity areas of a basin. The specific set of resolutions to be applied are chosen by the user; herein we use factor 2 multiples of a 90 m baseline resolution up to 720 m. The final spatial distribution of resolutions depends on the choice of a maximum *CM* threshold (CM_{max}), above which pixels are simulated at the finest resolution and below which pixels are simulated at a mix of coarser resolutions. The threshold is chosen based on available computational resources for an application. In this study we chose to use a CM_{max} of 0.65, which corresponds to the 90$^{\text{th}}$ percentile of the CRB *CM* values (Figure 3.2). Based on the benchmarking tests performed by Baldo and Margulis [2017], such a threshold leads to a decrease in total pixel numbers on the order of 60 to 70%, which corresponds to reasonable computational costs for a full CRB snow reanalysis.

By construct, all of the UYRB pixels with a *CM* value larger than 0.65 were resolved at the baseline spatial resolution of 90 m, while the less complex ones were assigned either 720 m, 360 m, 180 m or 90 m by the MR algorithm developed by Baldo and Margulis [2017]. The majority of the 720 m pixels are located in the northwestern part of the basin (Figure 3.2) corresponding to flat and grassy areas. Modeling almost a quarter of the pixels at this coarse resolution represents the main source of computational savings, while minimizing the impact on snow accumulation and melt patterns given the homogeneous physiography of the terrain. The remaining low *CM* pixels were assigned either 360 m, or 180 m depending on the complexity of their neighbors. In terms of the most complex pixels, 31% of the pixels are resolved at 90 m in order to preserve the accuracy of SWE estimates. In UYRB, these pixels

tend to be located at higher elevations, where the terrain is rugged and densely forested as described in Baldo and Margulis [2017] (Figure 3.2).

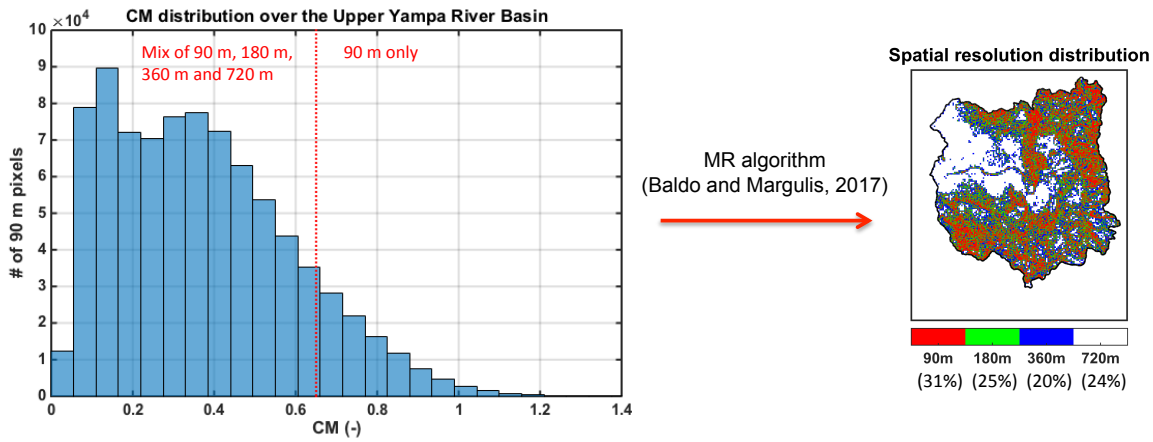

Figure 3.2: (left panel) Complexity metric distribution for the Upper Yampa River Basin. The choice of the maximum threshold CM_{max} of 0.65 represented as the red vertical line leads to (right panel) the spatial resolution distribution map.

3.2.3 SWE reanalysis framework

3.2.3.1 Model framework and forcings

The modeling setup used in this study is the same as described in Margulis et al. [2016]. The Simplified Simple Biosphere (SSiB) model developed by Xue et al. [1991], coupled with a three layer snow and atmosphere soil transfer (SAST) model (Sun and Xue [2001], Xue et al. [2003]) was used as the land surface model (LSM) to represent the interactions between the atmosphere, vegetation, and snow. A snow depletion curve (SDC) (Liston [2004]) was used to represent the sub-grid heterogeneity in SWE and the resulting *fsca*. The coupled LSM-SDC generates time series of SWE and *fsca* as a function of the sub-grid coefficient of variation (CV) and pixel-averaged cumulative snowfall and snowmelt.

The static inputs required by the LSM are latitude, longitude, elevation, slope, and

aspect, which were derived from the advanced spaceborne thermal emission and reflection (ASTER) DEM (JPL [2009]), as well as landcover derived from the National Land Cover Database (NLCD, Homer et al. [2007]). The static inputs were aggregated from their original 30 m resolution to the model resolution (either 90 m for the baseline or a mix of 90 m, 180 m, 360 m, and 720 m for the MR case). The dynamic meteorological forcings were obtained from the Phase 2 North American Land Data Assimilation System (NLDAS-2, Cosgrove et al. [2003], Xia et al. [2012]) hourly forcing dataset. NLDAS-2 variables include precipitation, incident shortwave radiation, near-surface air temperature, humidity, wind speed and pressure at a coarse spatial resolution of 1/8 °. The NLDAS-2 forcings were downscaled to the model resolution using topographic correction methods that have been previously applied over the Sierra Nevada and the Andes (Girotto et al. [2014b], Girotto et al. [2014a], Margulis et al. [2016] and Cortés et al. [2016]) as well as Upper Yampa in Baldo and Margulis [2017]. Lapse rates of 6.5°K/km and 4.1°K/km were used for air temperature and dewpoint temperature respectively. Downscaling approaches for atmospheric pressure, specific humidity, and the incoming longwave and shortwave radiation fluxes are explained in detail in Girotto et al. [2014b] (Appendix A). The downscaling is not deterministic, but also incorporates a priori uncertainty in the forcings (Girotto et al. [2014b]; Appendix A). It is important to note that the precipitation is not downscaled a priori, but treated as an uncertain random variable following a lognormal distribution with a mean of 2.25 and a standard deviation of 0.5 that is then implicitly downscaled and updated as part of the data assimilation framework.

3.2.3.2 Assimilation of Landsat-based fractional snow cover area using a particle batch smoother

The probabilistic DA framework used in this study is referred to as the Particle Batch Smoother, or PBS, and was developed by Margulis et al. [2015] in order to improve the probabilistic reanalysis framework used previously for SWE reanalysis in Durand et al. [2008], Girotto et al. [2014b] and Girotto et al. [2014a]. The coupled LSM-SDC provides a prior

ensemble estimate for all snow states and fluxes based on the specified input uncertainty and its propagation through the model. The prior ensemble treats each replicate as an equally likely (equal weight) realization based on the postulated input uncertainty. The goal of the PBS approach is to optimally weight the different uncertainty sources coming from the meteorological forcing and *fsca* retrievals in order to generate posterior snow estimates. Specifically, the reanalysis step is applied to a batch of the full set of *fsca* measurements (retrospectively) over the water year. A likelihood function updates the prior weights whereby the posterior weights can be used to determine the pdf or moments (i.e. mean, median, variance, inter-quartile range, etc.) of any of the snow states/fluxes. The mathematical framework is presented in detail in Margulis et al. [2015].

Landsat-5 thematic mapper (TM), Landsat-7 enhanced thematic mapper (ETM+), and Landsat-8 operational land imager (OLI) images from water year 1985 to 2015 were used to calculate *fsca* and fractional vegetation cover over each pixel. For a given sensor, measurements are available every 16 days at a spatial resolution of 30 m, and only clear-sky images were processed to obtain *fsca*. The raw data consist of multispectral top of atmosphere radiance measurements that are transformed into top of atmosphere reflectance before being atmospherically corrected. The spectral unmixing algorithm validated by Cortés et al. [2014] and based on Painter et al. [2009] then retrieves the fraction and type of constituent (snow, vegetation or bare rock/soil) within each pixel through a least-square-error optimization. The linear unmixing model estimates reflectances from each constituent and selects the combination of constituents leading to the lowest root mean square error (RMSE) between the modeled reflectance and a library of snow reflectances that have previously been calculated for different combinations of constituents within each pixel. The validation of the algorithm by Cortés et al. [2014] showed an *fsca* retrieval error of approximately 15%. The vegetation cover fraction (*fveg*) was also retrieved from the spectral unmixing algorithm and annually averaged and used within the LSM-SDC. The *fveg* derived from Landsat observations was chosen over the static NLCD for use in the LSM-SDC model to allow for inter-annual variability and because it is also, by construct, more consistent with the *fsca*

observations used in the assimilation step. Similar to the static input data, the *fsca* and *fveg* images at 30 m were then aggregated to either 90 m for the baseline case, or a mix of 90 m, 180 m, and 720 m for the MR case.

3.2.3.3 Verification of posterior SWE estimates

A posterior set of SWE reanalysis estimates was first generated for 31 years (WY 1985 WY 2015) at the baseline resolution of 90 m, and compared to in-situ measurements to assess its accuracy. A total of 203 peak SWE measurements from six SNOTEL stations and 1421 monthly manually sampled SWE from seven snow courses were used. Not all locations have full records for the full period, with two snow pillows / courses starting in 1986 and one in 1998. All snow pillows are collocated with snow courses and station 5 is a snow course only (Figure 3.3). All in-situ observations are taken at high elevations, between 2500 and 3200 m, in densely forested clearings; some representativeness errors are therefore expected when compared to grid-averaged SWE estimates.

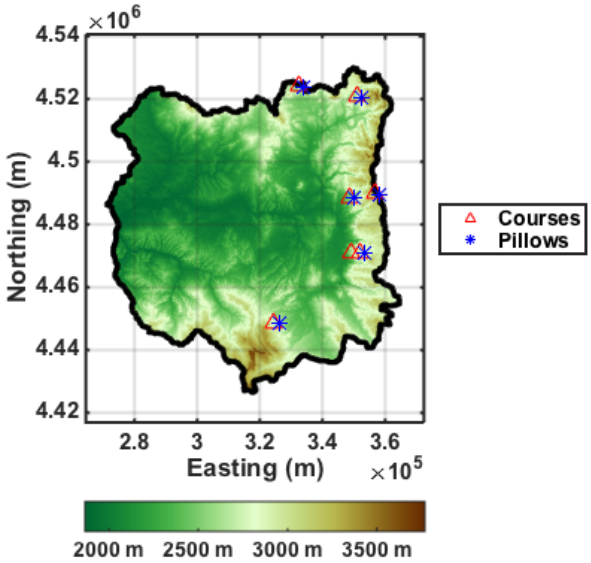

Figure 3.3: Elevation map of the Upper Yampa River Basin with the location of the seven snow courses shown in red and the location of the six snow pillows shown in blue.

The prior SWE estimates are highly uncertain by construct, and overestimated insitu observations from both snow courses and pillow (Figure 3.4). Prior estimates had a mean difference (MD) of 30 cm, and a root mean square difference (RMSD) of 41 cm for snow courses, and a MD of 43 cm, with a RMSD of 51 cm for snow pillows. Both showed a similar correlation coefficient (R^2) of 0.86. Note that, based on previous work (Luo et al. [2003], Girotto et al. [2014b]), the NLDAS-2 precipitation was assumed biased and therefore bias-corrected using the prior distribution (using a mean of 2.25 as indicated above). The fact that the prior SWE overestimates in situ data is an indication that there is likely an over-correction in the prior precipitation (at least at these sites). In contrast, the reanalysis generated posterior SWE estimates that are much more consistent with the in-situ data, are extremely well correlated to in-situ measurement and show limited mean differences. The MD is less than 2 cm for snow courses and less than 5 cm for snow pillows, with RMSD of 10 cm and R^2 higher than 0.95 for both. The small differences observed may be partly explained by undercatch problems with SNOTEL pillows measurements, and also by the fact that in-situ SWE measurements are usually made in easily accessible areas such as clearings and therefore not fully representative of the collocated 90 m pixel-average values. The difference in errors between the prior and posterior is primarily indicative of the data assimilation method properly selecting ensemble members with precipitation forcing that is consistent with the *fsca* observations. Based on the comparison with in situ data, the posterior SWE estimates generated at 90 m can be considered to be an accurate representation of the true underlying SWE for the UYRB and are thus used as a baseline throughout.

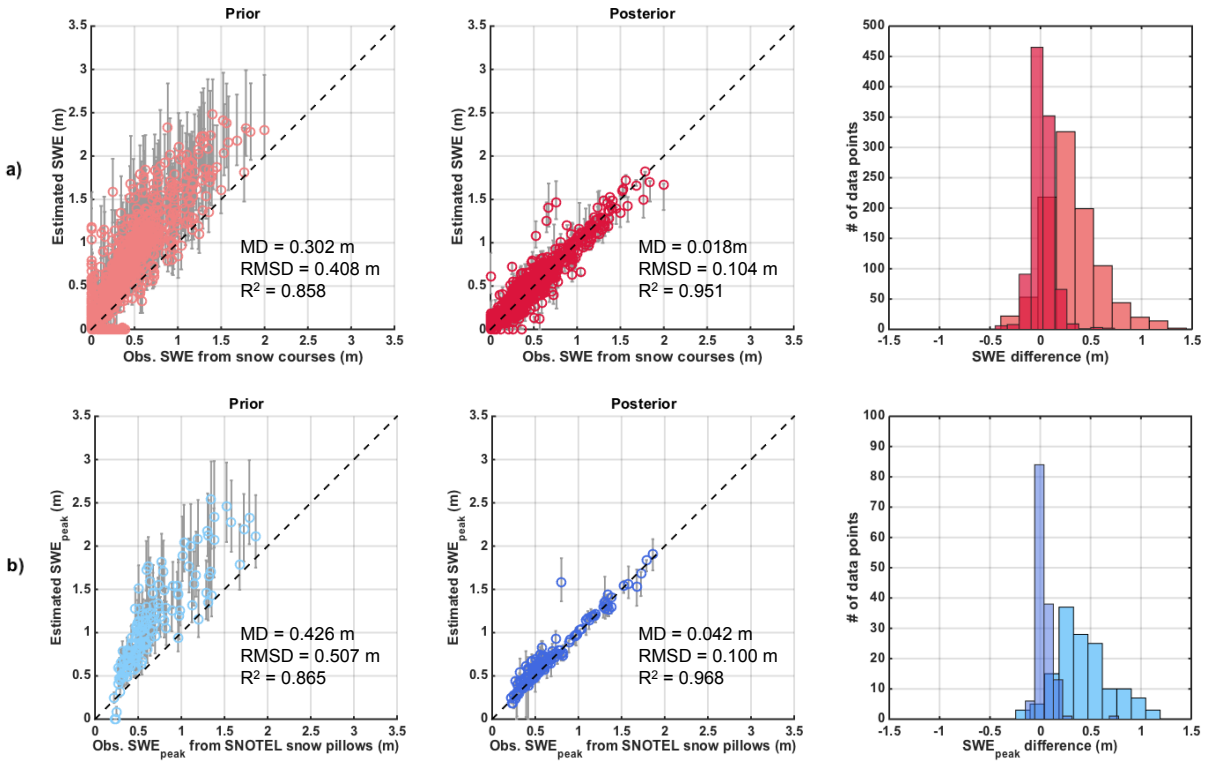

Figure 3.4: (left panel) Scatter plots of prior estimated snow water equivalent (SWE) vs. in situ, (middle panel) posterior estimated SWE vs. in situ and (right panel) histogram of the difference for (a) all snow courses and (b) snow pillows. The markers represent ensemble medians while the intervals represent the interquartile range (IQR). The mean difference (MD), root mean square difference (RMSD) and correlation coefficient (R^2) are displayed.

3.3 Performance of the MR SWE reanalysis compared to the 90 m baseline

As shown previously in section 3.2.3.3, performing a SWE reanalysis at 90 m yields an accurate reference solution for our test basin. However, such a simulation is very expensive in terms of computational resources. Modeling the basin uniformly at 90 m meant running almost 840,000 pixels with an ensemble size of 100 replicates, which took over a month on

the UCLA computer cluster and required 850G of space to store the resulting outputs. On the other hand, the MR approach decreased the number of pixels and storage need by 59%. Since pixels are simulated independently from each other, they are run in parallel, which is why runtime also decreased by 59% and took less than 2 weeks. Knowing that the MR SWE reanalysis can decrease computational constraints by a factor of two or more, the following section aims to assess its performance in terms of accuracy.

3.3.1 Impact of the MR approach on the assimilated *fsca* observations

The MR modeling approach as applied previously in Baldo and Margulis [2017] impacts the prior snow simulations, but in the context of a DA (reanalysis) framework as done herein, it also coarsens the *fsca* observations that provide the key constraint that generates the posterior estimates. Assessing the difference between the baseline and MR *fsca* is therefore crucial to understand the full effect of the MR approach on the data assimilation step. To this end, processed *fsca* images at the 90 m baseline and at the MR were first compared during the accumulation season, around day of peak (DOP) and during the ablation season. As seen in Figure 3.5, the MR approach does not significantly alter the *fsca* observations for the three sample dates chosen, and the mean absolute difference (MAD) is on the order of 5% (the MD is 0% for all three measurements) over the UYRB. The largest differences are concentrated over areas with partial snow cover (notably the Southeastern corner in Figure 3.5a and the more central parts of the basin in Figure 3.5b-c), which most likely correspond to snow ablation.

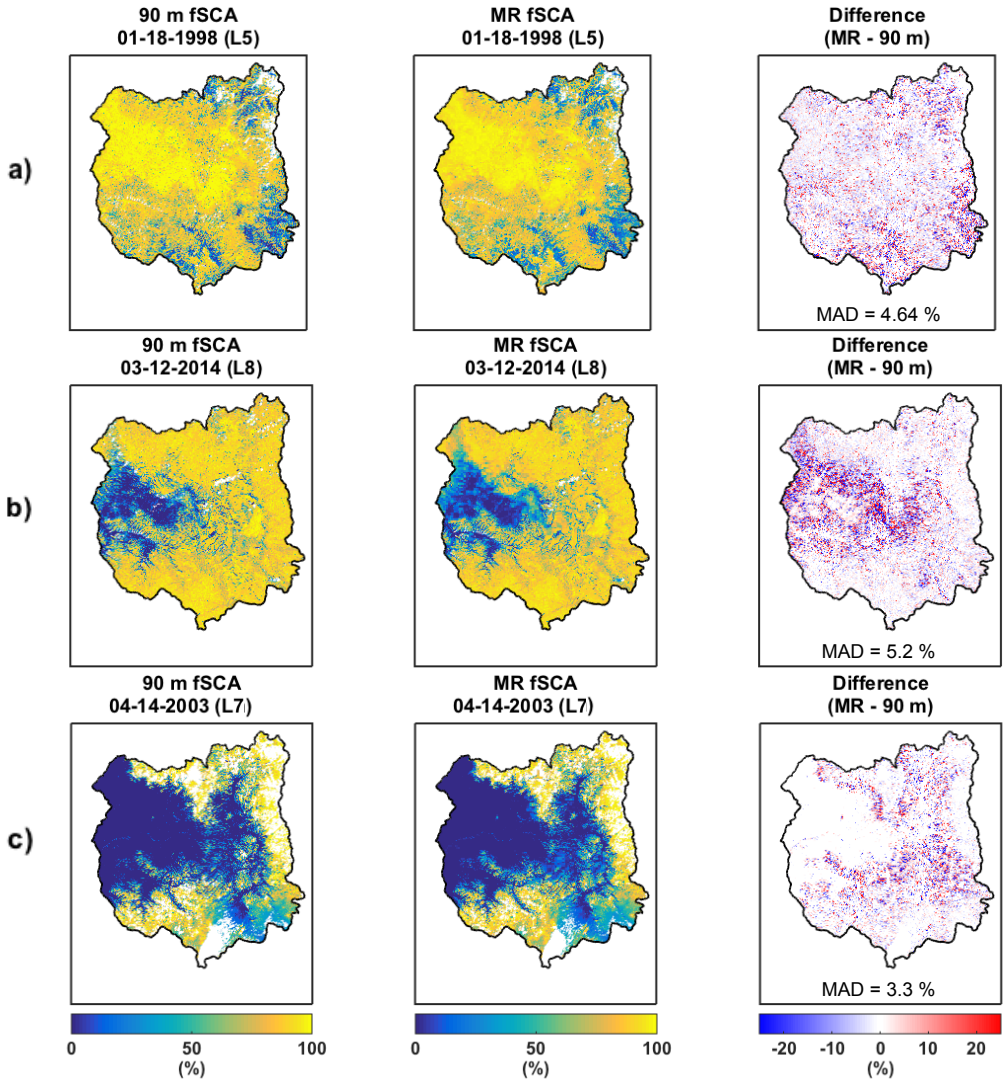

Figure 3.5: Maps of fractional snow cover area (*fsca*) during (a) the accumulation season, (b) at a time near day of peak (*DOP*), and (c) during the ablation season over the Upper Yampa River Basin for the 90 m baseline, the multi-resolution (MR) case and the difference between the two approaches. The exact acquisition day and sensor type (L5 for Landsat-5 TM, L7 for Landat-7 ETM+, and L8 for Landat-8 OLI) are displayed for all three samples. White areas inside the watershed bounds (in the left and middle panels) were covered by clouds. The basin-averaged absolute differences (MAD) are displayed at the bottom of each difference map.

In order to better understand the seasonality of the *fsca* differences, all observations were binned by month and averaged over the 31 years of record (Figure 3.6a). The differences are negligible between the 90 m baseline and the MR case during the accumulation season (October to January), while the MR method slightly overestimates the baseline *fsca* by 4% or less during the ablation season (February to August). The annual average difference is 0.87%. The expected impact of assimilating larger *fsca* values during the ablation season is an overestimation of the length of the snowmelt period, which, for the same amount of melt season energy inputs, would translate into larger posterior SWE estimates. As seen in Figure 3.6b-c, *fsca* from both the baseline and the MR case share a similar distribution with respect to *CM* and Peak SWE (SWE_{peak}). As expected, areas of high *fsca* correspond to areas of high SWE accumulation at the higher elevation of the basin, which also tend to be the most complex. By design, the MR approach does not coarsen areas of high physiographic complexity that can experience sharp differences in accumulation/ablation from one pixel to another. Hence, by construct, the MR *fsca* is identical to the baseline for CM larger than 0.65, and slightly differs from the baseline in low complexity areas as seen in Figure 3.6b. In addition, Figure 3.6c shows that the difference in fSCA over regions of high SWE accumulation is negligible as well (1.3% or less). Given the small differences observed, the effect of the MR approach on the assimilated *fsca* observations is minimal and therefore is not expected to significantly alter the performance of the data assimilation scheme (discussed in more detail below).

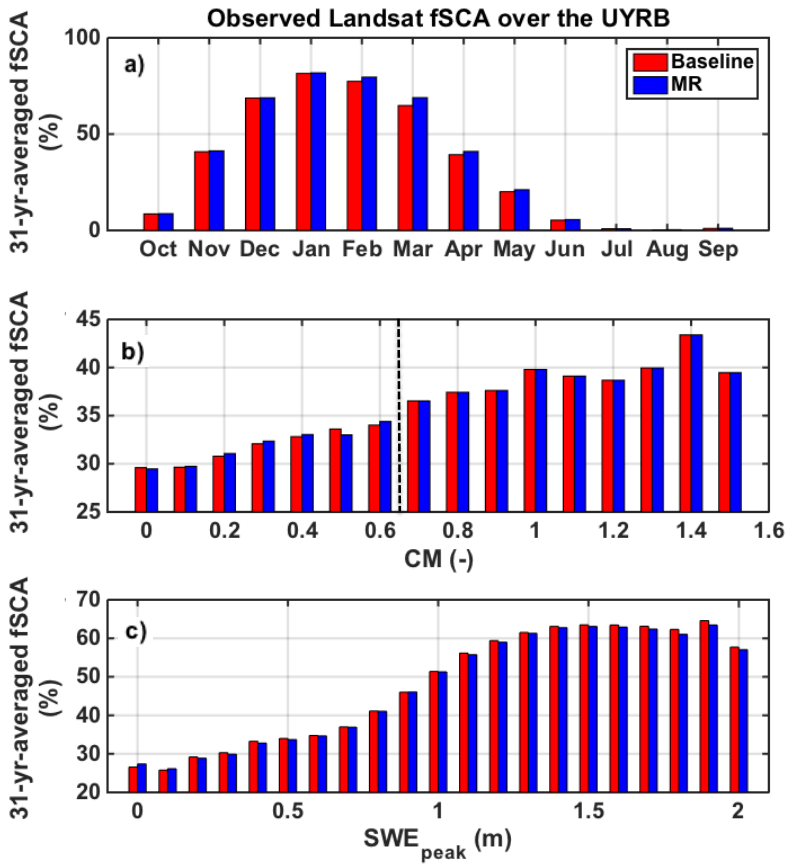

Figure 3.6: *fsca* climatology derived from the 31-yr record of Landsat observations over the Upper Yampa River Basin: bin-averaging of all observations across the range of (a) months of the water year, (b) CM values, and (c) peak SWE (SWE_{peak}) values for the 90 m baseline and MR case. The CM maximum threshold CM_{max} of 0.65 is represented by the vertical dashed line (b).

3.3.2 Impact of the MR approach on snow climatology metrics

The following analysis focuses on the comparison of the posterior ensemble median SWE estimates for the baseline and MR cases. Peak SWE (SWE_{peak}), day of peak (DOP), and duration of melt (DOM) were chosen for analysis. SWE_{peak} is defined as the maximum daily SWE in a given WY. DOP is defined for each WY as the day when SWE is equal to

SWE_{peak}. DOM is the difference between the melt-out day, defined as the day when only 1% of the original SWE_{peak} remains, and DOP, which effectively quantifies the duration of the ablation season. These metrics can be defined either pixel-wise or for basin-averaged values.

3.3.2.1 Mean spatial distribution

Figures 3.7a, 3.8a, and 3.9a show maps of the 31-yr average pixel-wise SWE_{peak}, DOP and DOM, while figures 3.7b, 3.8b, and 3.9b show the distribution of the respective 31-yr average relative differences binned by CM, elevation (Z), slope, $fveg$, and SWE_{peak}. In these figures, the baseline estimates were always subtracted from the MR estimates, which means that a positive difference represents an overestimation of the baseline by the MR case and vice versa.

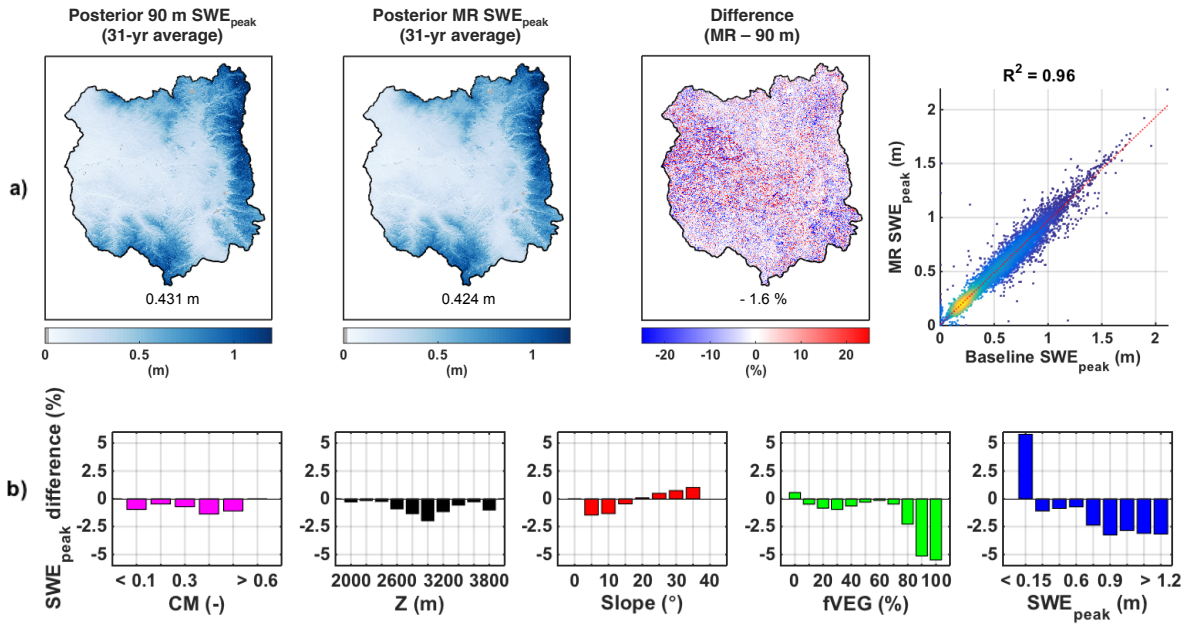

Figure 3.7: (a) Maps of pixel-wise 31-year average posterior peak SWE (SWE_{peak}) over the Upper Yampa River Basin for the 90 m baseline, the MR case, the percent difference between the two approaches (MR baseline), and the corresponding scatter plot. Basin averages are displayed at the bottom of each map. (b) Distribution of SWE_{peak} relative difference with complexity metric (CM), elevation (Z), slope, forested fraction ($fveg$), and SWE_{peak}. Pixels with a 31-yr average SWE_{peak} lower than 5 cm were discarded from the analysis.

As expected, the climatological SWE_{peak} shows significant spatial variability for both the MR and 90 m baseline with values ranging from zero to well over 1 m of SWE (Figure 3.7a). The middle and western parts of the basin that are not physiographically complex (see Figure 3.1) receive 25 cm or less on average. Given their location and relatively low elevation (less than 2000 m) the SWE accumulation is not orographically driven, but more heavily influenced by the few winter snowstorms occurring over the basin. The more complex areas in the eastern and southern edges of the basin accumulate a much larger amount of SWE (on the order of 1 m or more). On average, the MR approach underestimated pixel-wise SWE_{peak} by 7.2 mm or 1.6%, with the most complex areas showing no difference since they

were modeled at 90 m by design, and the less complex but high elevation areas showing larger differences on the order of 10 cm, or roughly 10% of SWE_{peak}. As seen in the density scatter plot, the majority of pixels have a SWE_{peak} around 20 cm, and the correlation between the baseline and the MR case is very strong with a correlation coefficient of 0.96. Figure 3.7b shows that the bin-averaged relative differences between the pixel-wise MR and baseline SWE_{peak} are constrained between -5 and 5%. By construct, the CM bands larger than 0.65 show no difference because all the MR pixels were simulated at the baseline resolution. All elevations bands show an underestimation of SWE_{peak}, with the largest differences observed at middle elevations between 2600 m and 3200 m. Since the UYRB is densely forested at these elevations, this is consistent with the largest underestimation occurring for the highest *fveg* bands. Regarding the distribution of the differences with slope, the lower slope bands (0 15) underestimate SWE_{peak} while the higher slope bands (20 35) show overestimation. As discussed in Baldo and Margulis [2017], the coarsening of pixel properties by the MR method leads to a slight increase in *fveg* for densely vegetated pixels, as well as an increase of more gentle sloped and north facing pixels. In the context of the SWE reanalysis, the magnitude of melt energy flux largely dictates the peak SWE that is consistent with a given *fsca* depletion time series. The increase in *fveg* as a result of the MR approach leads to an underestimation of the melt (energy) flux at the snow surface (as a result of attenuation of solar radiation), which decreases the posterior MR SWE_{peak} for these pixels. Since the minimum solar zenith angle during the ablation season over the UYRB is 16°, reducing gentle slopes (0° 15°) leads to an underestimation of the melt flux (as a result of becoming less perpendicular to the incoming direct beam solar radiation), which decreases the posterior MR SWE_{peak} for these pixels. Reducing steeper slopes (20° 35°) has the opposite effect and overestimates the melt flux, increasing the posterior MR SWE_{peak} for these pixels.

The posterior SWE_{peak} estimates are therefore impacted by the MR approach in two ways: i) an overestimation of the assimilated *fsca* during the ablation season and ii) a general underestimation of the melt flux due to the coarsening of the basin physiography, with the exception of steep pixels where the melt flux is overestimated. The basin-averaged

underestimation of SWE_{peak} observed in Figure 3.7 suggests that the effect of coarsening the static inputs and meteorological forcing on SWE_{peak} is more important than the effect from the coarsened assimilated *fsca* images. More importantly, the differences are the largest for the lowest SWE_{peak} band (less than 15 cm). The MR approach therefore concentrated the largest SWE_{peak} differences to areas of low CM that tend to accumulate less SWE.

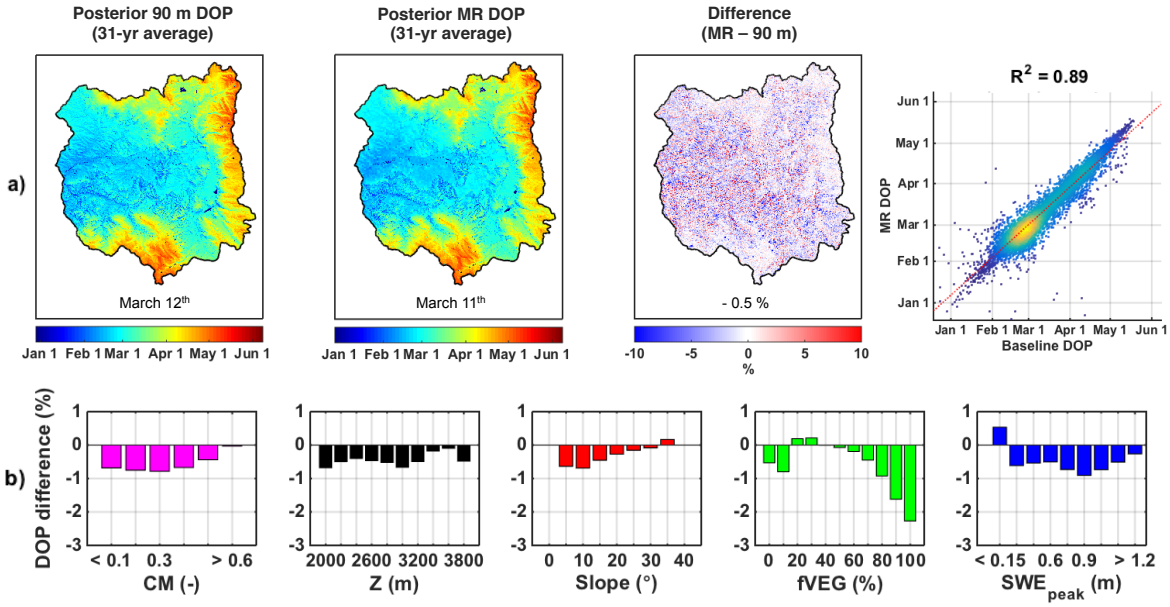

Figure 3.8: (a) Maps of pixel-wise 31-year average day of peak (DOP) over the Upper Yampa River Basin for the 90 m baseline, the MR case, the percent difference between the two approaches (MR baseline), and the corresponding scatter plot. Basin averages are displayed at the bottom of each map. (b) Distribution of DOP relative difference with complexity metric (CM), elevation (Z), slope, forested fraction ($fveg$), and SWE_{peak}. Pixels with a 31-yr average SWE_{peak} lower than 5 cm were discarded from the analysis.

Regarding DOP, Figure 3.8a shows that SWE in the middle and western regions of the basin that are not physiographically complex peaks early during the winter between January and March. In contrast, the more complex regions in the eastern and southern parts of the UYRB accumulated SWE until much later during the spring (April to June).

These complex regions show very good agreement between the baseline and MR case in term of timing, with larger differences over the rest of the basin. The average underestimation of 0.8 day or - 0.5% is negligible. As seen in the density scatter plot, the majority of pixels have peak values around March 1st, with a strong correlation coefficient of 0.89. Figure 3.8b shows DOP difference distributions with CM, elevation, slope, $fveg$ and SWE_{peak} similar to SWE_{peak} (Figure 3.7b), while the magnitude of the DOP differences is much smaller and ranges between 0.5% and -2%. The MR approach therefore preserves the accuracy of areas accumulating large amounts of SWE, that peak later in the spring.

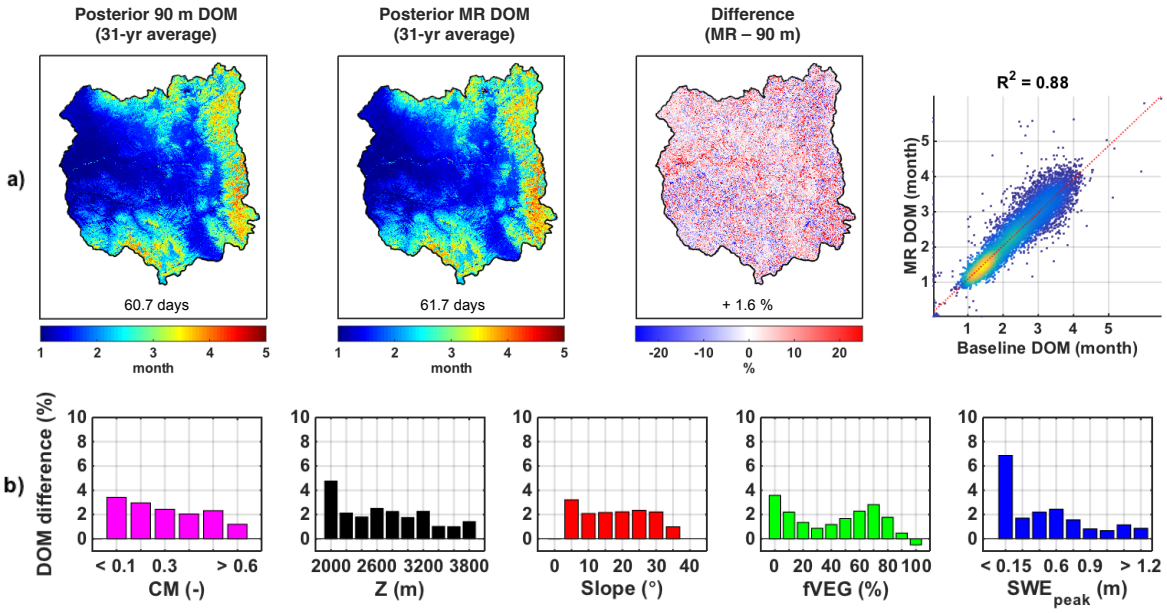

Figure 3.9: (a) Maps of pixel-wise 31-year average duration of melt (DOM) over the Upper Yampa River Basin for the 90 m baseline, the MR case, the percent difference between the two approaches (MR baseline), and the corresponding scatter plot. Basin averages are displayed at the bottom of each map. (b) Distribution of DOM relative difference with complexity metric (CM), elevation (Z), slope, forested fraction ($fveg$), and SWE_{peak}. Pixels with a 31-yr average SWE_{peak} lower than 5 cm were discarded from the analysis.

Regarding the duration of the ablation season, DOM can vary from less than a month

over the areas that accumulated little SWE and started melting as soon as the snowstorm events ended, to almost five months over the southwestern edge of the basin (Figure 3.9a). The average *DOM* is 61.7 days, or 2 months for the MR case, which overestimates the 90 m baseline by 1 day or 1.6%. The density scatter plot shows that the majority of pixels have a *DOM* between one and two months, with a strong correlation coefficient of 0.88. The slight overestimation of *DOM* by the MR case was expected, given the underestimation of melt fluxes from the increase of gentle north facing and densely forested pixels (Baldo and Margulis [2017]) and the higher assimilated *fsca* observations in the MR case. Figure 3.9b shows that the largest *DOM* overestimation occurs at the lowest band for all five variables. When looking at the distribution with SWE_{peak} specifically, pixels accumulating 15 cm of SWE of less show a *DOM* difference of 7%, while pixels accumulating 1 m or more only show a *DOM* difference of 1% or less. Pixels accumulating low amounts of SWE can be very intermittent in nature, without a clear SWE_{peak} or *DOP*, which can explain the higher difference seen in Figure 3.9b.

Based on these results, when applying the MR approach to the SWE reanalysis framework, we therefore expect the largest differences to occur over areas of low physiographic complexity. These types of areas tend to peak early during the winter, accumulate less SWE, and melt within a month and display lower levels of spatial variability that are easier to model at coarser resolutions.

3.3.2.2 Basin-average mean seasonal cycle

The mean seasonal cycle of MR SWE underestimates the baseline case by less than 1 cm as shown by the 31-yr average difference displayed in black in Figure 3.10b. Figure 3.10a-b show that the seasonal cycles for both the MR and baseline case closely match during the accumulation season (November to March) with differences in the range of +1 / -1 cm and a negative mean around 5 mm as shown by the grey shaded area and the black line respectively in Figure 3.10b. The basin-averaged mean SWE_{peak} is 0.374 m for the MR case, and 0.381 m for the 90 m baseline (Figure 3.10a), which leads to a mean difference of -7.1

mm (or -1.95%, Figure 3.10b). In terms of timing, *DOP* based on the mean seasonal cycle fits almost perfectly within a day, with the MR case peaking on March 15$^{\text{th}}$ and the baseline case on March 16$^{\text{th}}$ on average (Figure 3.10a). The underestimation is more pronounced during the early ablation season (March to June), where the difference in assimilated *fsca* observations is the largest (Figure 3.6a) with the entire range of WYs showing negative differences, and a maximum of -2.1 cm (or -5.4%) observed for the wettest year, WY 1996. Even though the MR case is assimilating slightly larger *fsca* observations during the ablation season (Figure 3.6a), the coarsening of the static inputs shown in Baldo and Margulis [2017] decreases the energy inputs, which ultimately lowers the posterior MR SWE estimates, and therefore explains the slight underestimation observed during the ablation season in Figure 3.10.

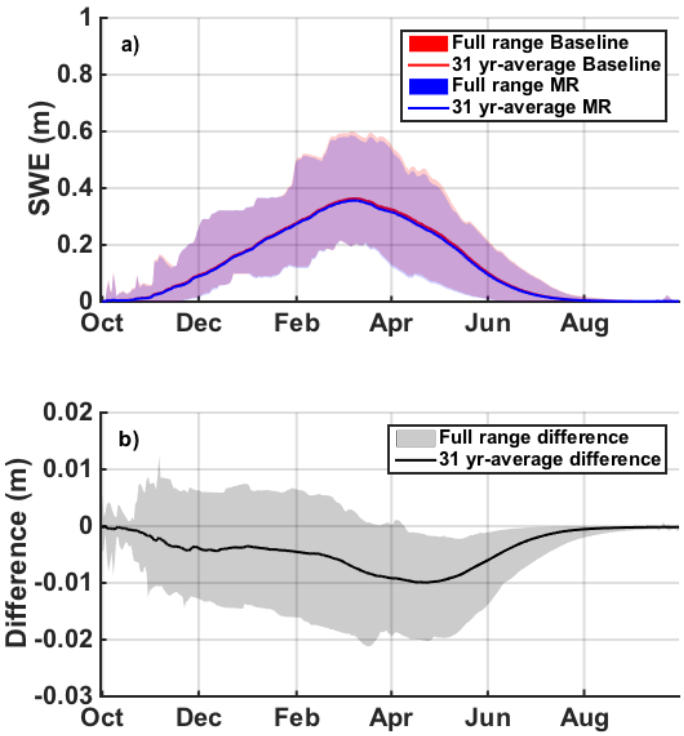

Figure 3.10: (a) Daily timeseries of basin-averaged posterior SWE from WY 1985 to WY 2015. The 31-yr averages are displayed in solid lines, while the shaded regions represent the full range across WYs. (b) The 31-yr averaged difference between the MR case and the baseline is displayed in black, with the full range of differences shaded in grey.

3.3.2.3 Inter-annual variability

The baseline and MR annual timeseries of SWE_{peak} show close agreement in inter-annual variations (Figure 3.11a). The scatter plot illustrates the positive performance of the MR case, including at both ends of the spectrum, which confirms that the MR case is estimating dry and wet years accurately. Figure 3.11b-c also illustrates the similarities in DOP and DOM inter-annual variability. WY 1985 shows the largest differences because there were two similar values of maximum SWE within 1 cm that occurred 15 days apart. The MR case identified the first peak as SWE_{peak}, while the baseline did the opposite, which does not impact the SWE_{peak}, estimate, but does impact both DOP and DOM. Beyond this single-

year, the MR case closely represents the inter-annual variability in the timing and length of accumulation and ablation seasons over the reanalysis period.

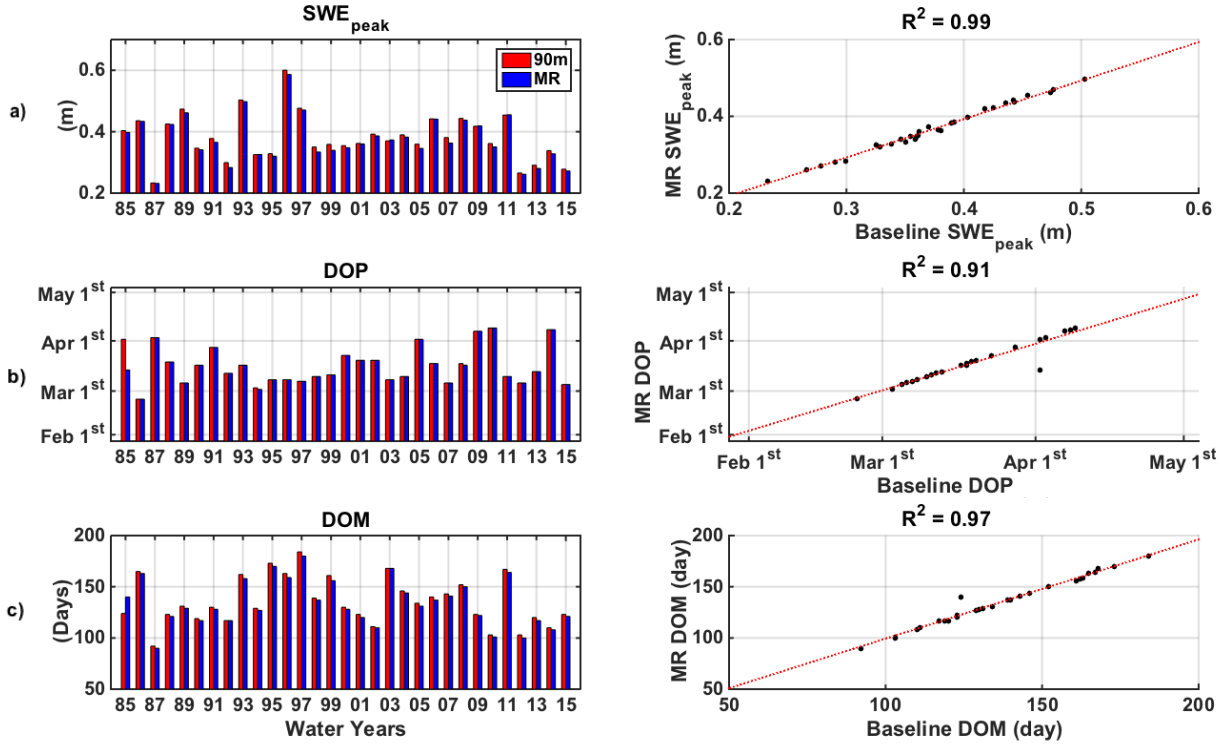

Figure 3.11: (left panel) Annual timeseries and (right panel) scatter plot with linear regressions of basin-averaged (a) peak SWE (SWE_{peak}), (b) day of peak (DOP), and (c) duration of melt (DOM) or the 90 m baseline and the MR case.

The fact that the MR case is capturing inter-annual variability correctly is further confirmed by the similar shapes of the baseline and MR empirical distribution functions (ECDFs) for all three metrics (Figure 3.12). The 10^{th} percentiles, medians and 90^{th} percentiles for both the MR and baseline cases are presented in Table 3.1. The statistics between the baseline and MR cases only differ by 1 cm for SWE_{peak}, and 2 days for DOM. Regarding DOP, the statistics are identical (Table 3.1), and the difference between the two ECDFs in Figure 3.12b is due to the special conditions in WY 1985 as explained above. Half of the WYs had

SWE_{peak} less than 0.37 m, while 10% had less than 0.29 m and 90% had less than 0.47 m for the baseline, or 0.36 m, 0.28 m, and 0.46 m for the MR case. The *DOP* distribution was identical between the baseline and MR cases, with the median, 10th and 90th percentiles on March 12, March 4 and April 2 respectively. Finally, *DOM* was shorter than 303 days for the baseline or 301 days for the MR case for half of the WYs simulated, with the 10th and 90th percentiles being 278 days / 276 days and 327 days / 325 days respectively.

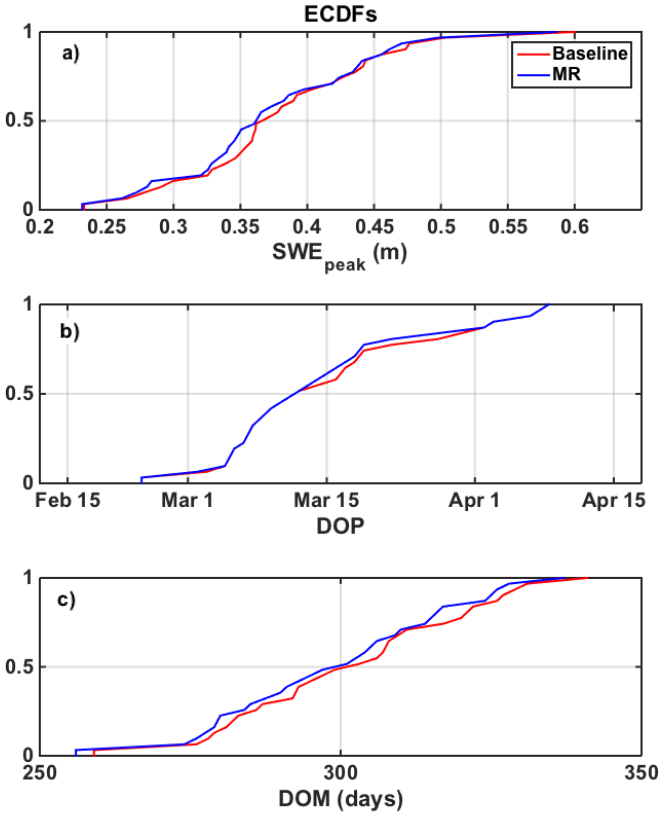

Figure 3.12: Empirical cumulative distribution functions (ECDFs) of (a) peak SWE (SWE_{peak}), (b) day of peak (*DOP*), and (c) duration of melt (*DOM*) for the 90 m baseline and the MR case.

Percentile	Baseline			Multi-Resolution (MR)		
	SWE_{peak} (m)	DOP	DOP (day)	SWE_{peak} (m)	DOP	DOP (day)
10^{th}	0.29	March 4	278	0.28	March 4	276
50^{th}	0.37	March 12	303	0.36	March 12	301
90^{th}	0.47	April 2	327	0.46	April 2	325

Table 3.1: Return period values for peak SWE (SWE_{peak}), day of peak (DOP), and duration of melt (DOM) for both the baseline and MR cases.

3.3.3 Impact of the MR approach on spatial variations of SWE uncertainty

The previous analysis focused on the impact of the MR approach on the posterior ensemble SWE median (i.e. a metric of central tendency). However, another strength of the reanalysis framework is to also provide a measure of uncertainty via the posterior ensemble. In this section the impact of the MR approach on the posterior ensemble SWE_{peak} standard deviation ($< \sigma >$) and coefficient of variation ($< CV >$) is examined, where the angle brackets ($<>$) are used to emphasize the ensemble operator.

In order to focus on the spatial distribution of the ensemble posterior SWE_{peak} uncertainty, the 31-yr average maps of $< \sigma >$ and $< CV >$ (Figure 3.13) were created by pooling $< \sigma >$ and $< CV >$ for each pixel over all 31 WYs as follows (Bingham and Fry [2010]):

$$\begin{aligned} \overline{< \sigma >_i} &= \sqrt{\frac{\sum_{y=1}^{31}(< \sigma >_i^y)^2 + \sum_{y=1}^{31}(< \mu >_i^y - \overline{< \mu >_i})^2}{31}} \\ \overline{< CV >_i} &= \frac{\overline{< \sigma >_i}}{\overline{< \mu >_i}} \end{aligned} \tag{3.2}$$

where the overbar notation denotes the 31-year average. $\overline{< \sigma >_i}$ is the 31-yr average ensemble SWE_{peak} standard deviation for pixel i, and $\overline{< \mu >_i}$ is the 31-year average ensemble SWE_{peak} mean for the same pixel i. $< \sigma >_i^y$ and $< \mu >_i^y$ are respectively the ensemble SWE_{peak} standard deviation and mean for each individual WY y. The 31-yr average SWE_{peak} coefficient of variation ($\overline{< CV >_i}$) for each pixel i was calculated as the ratio between the pixel 31-yr

average ensemble SWE_{peak} standard deviation and mean.

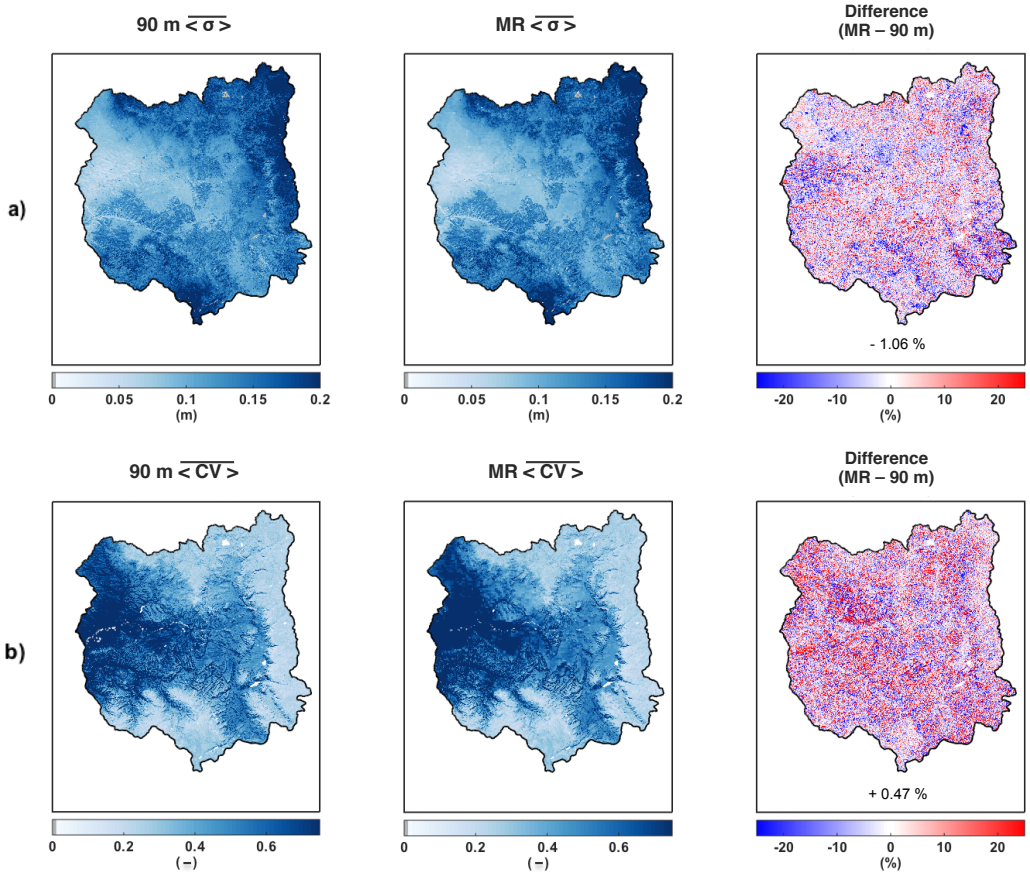

Figure 3.13: Maps of the 31-yr average ensemble SWE_{peak} (a) standard deviation ($\overline{<\sigma>_i}$), and (b) coefficient of variation ($\overline{<CV>_i}$) for (left panel) the 90 m baseline, and (middle panel) MR case. The relative differences in $\overline{<\sigma>_i}$ and $\overline{<CV>_i}$ are shown in the right panel. The basin-averaged differences are displayed at the bottom of the maps in the right panel.

As seen in Figure 3.13, the spatial distributions of $\overline{<\sigma>}$ and $\overline{<CV>}$ are highly variable. For both the baseline and MR cases, the high elevation areas accumulating large amounts of SWE (see Figure 3.7a) show $\overline{<\sigma>}$ on the order of 15 - 20 cm with a $\overline{<CV>}$ on the order of 10% 20%, while the lower parts of the UYRB have a $\overline{<\sigma>}$ around 5 cm or less,

with a $\overline{< CV >}$ higher than 60%. Regarding the relative difference between the MR and baseline cases (Figure 3.13a-b, right panel), no particular spatial pattern can be observed for $\overline{< \sigma >}$, with the exception of a few areas showing an underestimation on the order of 10%, bringing the basin-average difference to -1.06% or -1.8 cm. Figure 3.13b shows that the regions accumulating the most SWE with the lowest $\overline{< CV >}$ also have the lowest relative difference between the MR and baseline cases (white areas on the eastern and southern edges of the UYRB). Similar to the difference in $\overline{< \sigma >}$, the basin-average difference in $\overline{< CV >}$ of 0.47% is negligible.

3.4 Summary and conclusions

This study demonstrated the performance of a new MR terrain discretization approach in the context of a snow reanalysis framework using the assimilation of Landsat-derived *fsca* observations. The MR approach was shown to have an insignificant impact on the *fsca* observations assimilated and the reanalysis framework led to posterior SWE ensembles similar to the high-resolution 90 m baseline. The SWE reanalysis dataset generated with the MR approach matched the 90 m baseline ensemble median within 1 cm on average for peak SWE magnitude and within 1 day on average for timing of the accumulation and melt seasons. Most of the difference between the two approaches occurs in areas accumulating less than 15 cm of SWE, while areas accumulating more than that are estimated with a high degree of accuracy. In addition, the MR approach also preserved the SWE uncertainty, where the ensemble standard deviation and coefficient of variation showed differences on the order of -1% and 0.5% respectively. This study has demonstrated the feasibility of the MR approach in the context of a snow reanalysis framework, where the significant decrease in computational costs will allow much larger scale implementations of the SWE reanalysis over full mountain ranges, while preserving the accuracy of fine spatial resolution simulations.

CHAPTER 4

Joint assimilation of Landsat and MODIS-based fractional snow cover area products in a probabilistic SWE reanalysis framework

Relying on remote sensing information from Landsat only can be limiting given the presence of clouds, as well as the sensor lack of coverage over certain areas of the globe such as the High Mountain Asia. In order to implement a midlatitude SWE reanalysis dataset over seasonal montane snowpacks, the use of the complementary MODIS-based fractional snow cover area (*fsca*) product was investigated in the context of a joint data assimilation framework with Landsat derived *fsca* over three 1° by 1° tiles in the Sierra Nevada, Colorado River Basin, and High Mountain Asia for two representative water years. Due to its coarser spatial resolution and off-nadir looking nature, the MODIS sensor was found to overestimate forested fraction (*fveg*), and therefore underestimate *fsca* compared to Landsat products for the Sierra Nevada and Colorado River Basin tiles. The difference was exacerbated at the sensor high-viewing angles, which distorted the observation footprint and increased the obscuring effect of tall vegetation. Nadir-looking MODIS observations were found to be in good agreement with Landsat and were therefore successfully assimilated in the reanalysis framework. The magnitude of the posterior peak SWE (SWE_{peak}) generated by both sensors simultaneously was within a few cm of the posterior estimates generated using either Landsat or MODIS only. Improving the availability and quality of the measurements assimilated by including *fsca* observations from multiple sensors will facilitate the implementation of a high spatial resolution midlatitude SWE reanalysis dataset.

4.1 Background and motivation

Midlatitude montane snowpacks are a vital part of the global water and energy budget and their snow and ice melt runoff are providing fresh water to a significant part of the population (Barnett et al. [2005], Mankin et al. [2015]). However, the lack of good in-situ data (Serreze et al. [1999], Rohrer et al. [2013]), combined with the extreme complexity of the terrain (Baldo and Margulis [2017]) make the representation of snow water equivalent (SWE) difficult and generating a high spatial and temporal SWE dataset remains a challenge over these regions (Dozier et al. [2016]). One way to address this issue is to use ensemble-based data assimilation (DA) methods to overcome the lack of in-situ data and poorly characterized precipitation forcing in montane regions (Luo et al. [2003]). DA methods are currently the preferred method to estimate snow states, because assimilating remotely sensed observations offers spatially and temporally continuous estimates, while also providing a measure of their uncertainty (Clark et al. [2006], Andreadis and Lettenmaier [2006], Su et al. [2008], De Lannoy et al. [2010], Liu et al. [2013], Arsenault et al. [2013], Girotto et al. [2014b], Margulis et al. [2015], Kumar et al. [2015]).

The quality and availability of remote sensing images are crucial to generate a SWE reanalysis dataset that accurately represents the spatial variability of the seasonal montane snowpack. Due to their coarse spatial resolutions of 10+ km, observations acquired by passive microwave (PM) sensors such as AMSR-E (Tedesco et al. [2010]) cannot be solely considered in the DA framework, even though the temporal resolution, and the insensitivity to cloud cover could be useful. Other approaches using passive-active microwaves (Lemmetyinen et al. [2011]), or LiDAR altimetry (Deems et al. [2006] and Deems et al. [2013]) are promising in terms of directly retrieving snow depth observations at fine spatial resolutions, but the data is not commonly available and lacks spatial and temporal coverage. In contrast, visible and near infrared (VIS-NIR) reflectance observations can offer a global and continuous coverage at fine spatial resolutions ranging from 30 m for Landsat to ~500 m for MODIS and 1 km for AVHRR. Unlike PM observations, SWE cannot be directly estimated from VIS-NIR measurements, and a retrospective analysis of fractional snow cover area (*fsca*) has

to be performed instead.

Useful *fsca* observations were derived from Landsat by Cortés et al. [2014], and successfully assimilated by Margulis et al. [2016] and Cortés and Margulis [2017] to generate a 30 + year SWE reanalysis dataset at a resolution of 90 m over the Sierra Nevada and 180 m over the Andes respectively. The main advantage of using Landsat imagery is the length of its record (1985 - present), and its fine spatial resolution of 30 m with a nadir looking geometry. However, the presence of clouds and damaged Landsat 7 sensor since 2003, combined with its 16 days temporal resolution can significantly reduce the number of usable observations. In addition, gaps in the Landsat-era record exist in some regions of the globe (Kovalskyy and Roy [2013], Wulder et al. [2016]) during some years, which is especially problematic over the High Mountain Asia (HMA, Figure 1.1) region. In order to overcome Landsat limitations in the context of a global SWE reanalysis, the goal of this work is therefore to explore the simultaneous assimilation of both Landsat-derived *fsca* (Cortés et al. [2014]) and its complementary MODIS-derived MODSCAG *fsca* product by Painter et al. [2009]. In contrast to Landsat, MODIS has good coverage over midlatitude montane regions since 2001, with a daily temporal resolution that greatly increases the number of clear-sky observations. Its spatial resolution of ~500 m is not as fine as Landsat, but is still considered as high resolution, and MODIS-derived snow cover products were therefore previously used in many DA frameworks over montane regions (Su et al. [2008], De Lannoy et al. [2012], Chelamallu et al. [2014], Mir et al. [2015], Kumar et al. [2015]). One potential restriction when using MODIS-based observations is the effect of the sensor viewing geometry on SWE estimates (Dozier et al. [2008]), which has been under-studied in previous work. Given the scanning nature of the MODIS sensor, acquisitions can be made at zenith viewing angle varying from 0° to over 60°, which could be problematic over tall vegetated areas. As a result, the meaning of *fsca* is more complicated than in the Landsat case where it is simply a representation of snow depletion over time.

In this chapter, historical remote sensing *fsca* and *fveg* products from both the Landsat and MODIS sensors were compared and assimilated in a SWE reanalysis framework over

complex montane regions to address the following questions: 1) How does MODIS-derived *fsca* and *fveg* observations differ from Landsat over montane seasonal snowpack? 2) What are the implications of MODIS viewing geometry on the posterior SWE estimates? 3) Can both sensors be simultaneously assimilated in the SWE reanalysis framework?

The rest of the chapter is organized as follows: Section 4.2 illustrates the study area and the methodology used in this work. Section 4.3 compares Landsat and MODIS-based *fsca* and *fveg* products, and section 4.3.2 analyzes the impact of assimilating MODIS-based measurements in the context of a SWE reanalysis framework in order to answer the questions listed above. Finally, Section 4.5 summarizes the key points of this work.

4.2 Development of the SWE reanalysis framework at a global scale using both Landsat and MODIS-based observations

4.2.1 Application domain

In order to facilitate the global scale implementation of the SWE reanalysis framework, the montane areas highlighted in red in Figure 1.1 were divided into 1° by 1° tiles. The study areas shown in Figure 4.1 were chosen based on their complex physiography representative of montane snowpacks, which often results in complex spatial seasonal SWE patterns. Two test tiles were selected in the Western U.S. domain, N39W121 (39° North, 121° West) in the Sierra Nevada (SN), and N40W107 in the Colorado River Basin (CRB). One test tile, N35E75, was selected over the HMA domain. Two test water years (WYs) were selected from the overlap period between MODIS and landsat records (~2000 - present): WY 2003-04, and WY 2014-15.

The N39W121 tile is located in the Northwestern SN range, and Figure 4.1b shows that the region is relatively densely forested, with a mean forested fraction (*fveg*) of 34.7%, and a maximum of 80% in the western part of the tile. A significant portion of the region is below the snow/rain threshold of 1500 m and was therefore not represented. The highest portion of the SN is around 2500 m. In contrast, the N40W107 tile located in the CRB headwaters

is higher, with elevations ranging from 2000 m to 3500 m. The tile is slightly less forested than the SN, with a mean *fveg* of 20.2%, and a maximum around 60%. Regarding the HMA, the selected tile, N35E75, presents significantly different physiographic characteristics. The elevation is extremely variable and spans from 2000 m to 5500 m. As seen in Figure 4.1b, areas of the tile with elevations above 4500 m are covered by glaciers and were therefore excluded from the analysis. In contrast to the Western U.S. tiles that were densely forested, N35E75 forest cover is very sparse, with a mean *fveg* of 1.6% and a maximum of 10%.

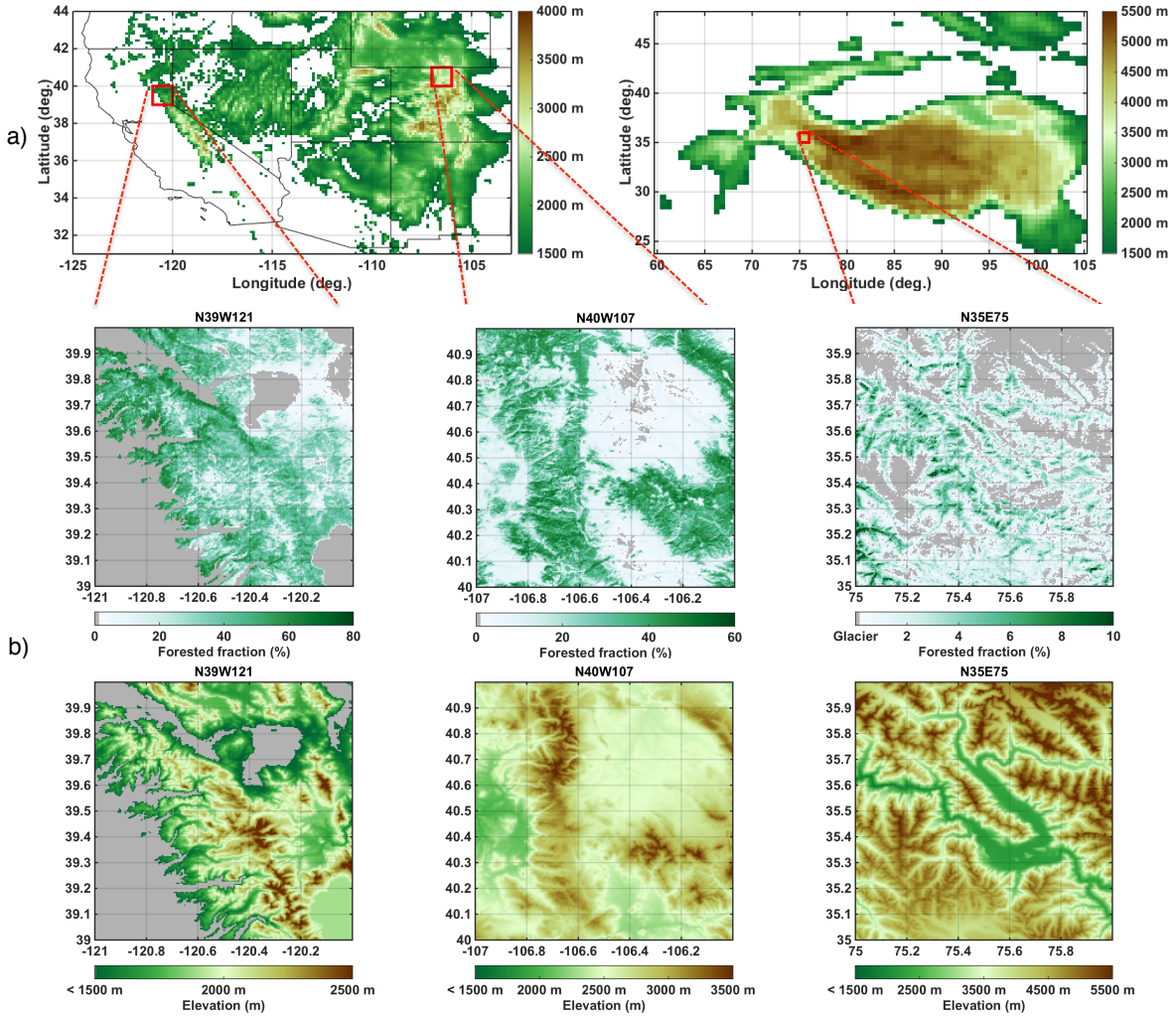

Figure 4.1: (a) Elevation maps of (left panel) the Western U.S. and (right panel) HMA regions, using using the 0.125° by 0.125° NLDAS-2 data for the Western U.S. and the 0.5° by 0.625° MERRA-2 data for the HMA. (b) (top panel) Forested fraction and (bottom panel) elevation maps at 480 m resolution for the three 1° by 1° tiles tested in the (left panel) Sierra Nevada (SN), (middle panel) Colorado River Basin (CRB) and (right panel) HMA. For all maps, only areas above 1500 m that are not glacierized are represented.

4.2.2 Static, meteorological and remote sensing data used

In contrast to the previous chapters, which focused exclusively on the conterminous United States, the inputs used for the three test tiles in this chapter draw from globally available datasets. Topographic data was based on the 30 m resolution digital elevation model from the Shuttle Radar Topography Mission (SRTM, Farr et al. [2007]), with gaps in coverage filled by the advanced spaceborne thermal emission and reflection (ASTER) DEM (JPL [2009]). Landcover was based on the 1 km information from the Advanced Very High Resolution Radiomater (AVHRR) global land cover classification database (Hansen et al. [2000]). Forest fraction was derived from the MODIS continuous vegetation field dataset (Sexton et al. [2013]). Since the method has been developed for seasonally snow-covered areas, glaciers and areas below 1500 m were excluded from the reanalysis (see Figure 4.1). The glacier mask used to identify pixels excluded from the reanalysis as seen for the HMA test tile in Figure 4.1b was extracted from the GLIMS glacier dataset (Kargel et al. [2014]). All static data was resampled to a baseline model resolution of 480 m in order to fit the resolution of the MODIS-based product as explained later in this section.

The hourly meteorological inputs were extracted from the 1/8 ° by 1/8 ° resolution NLDAS-2 dataset (Cosgrove et al. [2003], Xia et al. [2012]) for the SN and CRB tiles, and from the the 0.5° by 0.625° resolution MERRA-2 dataset (Gelaro et al. [2017]) for the HMA tile. Air temperature, specific humidity, surface pressure and shortwave and longwave radiative fluxes were downscaled probabilistically to the model resolution based on topographic corrections and uncertainty models using the methods outlined in Girotto et al. [2014b]. Zonal and meridional components of wind speed were downscaled following the approach of Liston and Elder [2006]. Precipitation was downscaled implicitly as part of the reanalysis step (Figure 4.2).

The *fsca* observations assimilated in the SWE reanalysis framework were derived from the Landsat and MODIS sensors. Useful Landsat data is available at a high resolution of 30 m since 1985 based on acquisitions from the Landsat 5 Thematic Mapper (TM), Landsat 7 Enhanced Thematic Mapper (ETM+) and Landsat 8 Operational Land Imager (OLI) sen-

sors. Images of *fsca* are derived using a spectral end-member unmixing approach developed by Cortés et al. [2014], and previously applied in Girotto et al. [2014b], Girotto et al. [2014a], Margulis et al. [2016], Cortés et al. [2016], and Cortés and Margulis [2017]. The landsat *fsca* data for WY 2003-04 are therefore a mix of Landsat 5 TM and Landsat 7 ETM+ images, and a mix of Landsat 7 ETM+ and Landsat 8 OLI for WY 2014-15. The MODIS-based *fsca* estimates, were extracted from the MODSCAG product (Painter et al. [2009]) from the MODIS sensor on board the Terra satellite available since 2001. The MODSCAG products were interpolated from their nominal resolution of ~463 m to a regular 480 m grid, and the 30 m landsat *fsca* images were aggregated to 480 m. As seen in Table 4.1, the SN tile benefits from a large number of available *fsca* observations from both Landsat and MODIS, with all MODIS-based acquisitions (both nadir and off-nadir-looking) being roughly five times as frequent as the Landsat ones. In the case of the CRB, the impact of cloud cover was more important than for the SN, and only half of the number of observations were available. Finally, the HMA tile Landsat and MODIS coverage was the smallest of the three tiles, with only 14 Landsat clear sky observations for WY 2003-04, making the additional 110 from MODIS an extremely useful source of *fsca* measurement.

Tile	WY 2003-04		WY 2014-15	
	MODSCAG	Landsat	MODSCAG	Landsat
SN	243	52	238	49
CRB	123	28	145	20
HMA	110	14	98	23

Table 4.1: Number of cloud-free *fsca* observations available over the SN, CRB and HMA test tiles for WYs 2003-04 and 2014-15. Landsat observations can come from different Landsat row/path, and multiple sensors (TM and ETM+ for WY 2003-04, and ETM+ and OLI for WY 2014-15).

4.2.3 Data assimilation framework

The SWE reanalysis framework used to assimilate both Landsat and MODSCAG products is the same Particle Batch Smoother (PBS; Margulis et al. [2015]) described in detail in Chapter 3, and is shown here schematically in Figure 4.2. The data assimilation (DA) framework merges historical remotely sensed *fsca* data with prior estimates of SWE from a model forced by uncertain inputs. The method is highly modular and can be used at variable spatial resolutions, with different meteorological input datasets (as shown in section 4.2.2), land surface models (LSMs), snow depletion curve (SDC) models, sources of *fsca* data, etc. (Figure 4.2). The LSM, and SDC used for this study are the same as the ones used in the previous chapters: The Simplified Simple Biosphere (SSiB) model developed by Xue et al. [1991], coupled with a three layer snow and atmosphere soil transfer (SAST) model (Sun and Xue [2001], Xue et al. [2003]) and the SDC from Liston [2004].

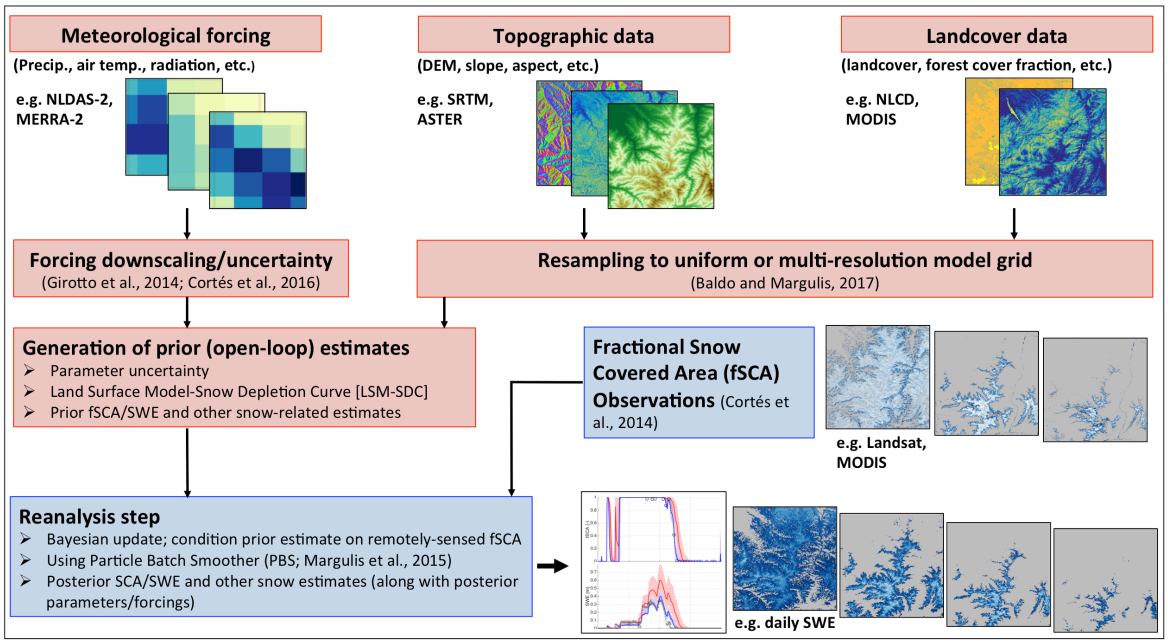

Figure 4.2: Reanalysis framework flow diagram. Red boxes represent the ensemble-based snow modeling system. Blue boxes represent data assimilation components.

Data assimilation systems require a measurement model that maps the model states to the measured variables. In the case of measuring *fsca*, each image pixel is composed of forested and non-forested fraction, and the forested part obscures what the sensor can see. The sensors are therefore assumed to only seeing snow over the non-forested fraction of each pixel. The following measurement model was therefore implemented in the PBS to estimate a bare-soil fraction (with the rest of the fraction being forest-covered) that provides a *fsca* prediction assumed to be what the sensors are seeing:

$$fsca_{predicted} = fsca_{bare} \times (1 - \overline{fveg_{winter}}) \tag{4.1}$$

where $\overline{fveg_{winter}}$ is the forested fraction measured by the sensor at nadir, averaged over winter (defined as $fsca > 0$), and $fsca_{bare}$ is the prior bare soil *fsca* predicted by the land surface model. In addition, *fsca* measurements from Landsat and MODSCAG are assigned the same upper limit $\overline{fveg_{winter}}$ as the prior predicted *fsca* in order to be consistent with the measurement model.Following the work of Margulis et al. [2016], a temporal average of nadir-looking *fveg* was chosen in order to reduce the effect of low-lying vegetation during the ablation season. The errors between the prior $fsca_{predicted}$ and the observation are then used by the following likelihood function to calculate the posterior weights that will be used to update each prior realization of snow states/fluxes (Margulis et al. [2015]):

$$w_r^+ = \frac{c_0}{N} p_v \left[fsca_{meas} - fsca_{predicted}^r, C_{v,meas} \right] \tag{4.2}$$

where c_0 is a normalization constant, N is the ensemble size (an ensemble of 100 realizations r was used in this work), and p_v is the probability function of the measurement error vector $fsca_{meas} - fsca_{predicted}^r$, constrained by the measurement error covariance $C_{v,meas}$. p_v assigns larger posterior weights w_r^+ for replicates with prior realizations closer to the observation. The measurement vector $fsca_{meas}$ can be composed of Landsat and/or MODSCAG observations depending on which sensor is assimilated. A seasonal screening is done to isolate

primarily observations during the ablation season.

4.3 Comparison of Landsat and MODSCAG *fsca* and *fveg* products

4.3.1 MODSCAG viewing geometry effects

In contrast to Landsat, which is nadir-looking, MODIS is a scanning sensor and can therefore acquire observations with significant off-nadir viewing angles up to 65° at the outer edge of the swath. As the viewing angle increases, the footprint of each pixel elongates, and snow can be obscured in the case of tall vegetation. Nadir-looking measurements should therefore have a stronger connection to the model state, and be assigned a lower measurement error covariance. In order to take into account the effect of MODIS viewing geometry in the data assimilation framework, the measurement error covariance $C_{v,MODSCAG}$ was derived as a function of the sensor viewing zenith angle, so that the error covariance is at its minimum at nadir (which corresponds to the Landsat measurement error covariance $C_{v,Landsat}$) and increases with the viewing angle.

In an attempt to create a space-time continuous/cloud-filled *fsca* dataset, Dozier et al. [2008] used a weighted least squares fitting of the raw MODSCAG data where the weighting coefficient $w(\theta_s)$ was given by:

$$w(\theta_s) = \frac{p^2 cos\theta_s}{p_{\parallel} p_{\perp}} \tag{4.3}$$

where p is the linear pixel dimension at nadir (463 m), θ_s is the sensor viewing angle, and $p_{\parallel}$ and $p_{\perp}$ are the along-track and cross-track pixel dimensions at a non-nadir scan angles and are given by Dozier et al. [2008] as follows:

$$
\begin{aligned}
p_{\|} &= \frac{p}{H}\left[(R+H)cos\theta - R\sqrt{1-sin^2\theta\left(\frac{R+H}{R}\right)^2}\right] \\
p_{\perp} &= R\Bigg\{sin_{-1}\left[\frac{R+H}{R}sin\left(\theta + tan^{-1}\frac{p}{2H}\right)\right] \\
&- sin^{-1}\left[\frac{R+H}{R}sin\left(\theta - tan^{-1}\frac{p}{2H}\right)\right] - 2tan^{-1}\frac{p}{2H}\Bigg\} \\
\theta &= sin^{-1}\left(\frac{R}{R+H}sin\theta_s\right)
\end{aligned}
\tag{4.4}
$$

where R is the radius of the earth (6371 km) and H is the orbit altitude above the surface (705 km). This weighting was done by Dozier et al. [2008] to treat nadir-looking observations as more trustworthy, and off-nadir as less trustworthy, which is what this work seeks to do as well.

Under assumptions of Gaussian measurement errors (which is consistent with the snow reanalysis) the weighting coefficient $w(\theta_s)$ can be interpreted as proportional to the inverse of the measurement error covariance $C_{v,MODSCAG}$. Since the error covariance (as a function of viewing geometry) is needed for the DA we borrow the prescribed weighting function to derive a new measurement error covariance function $C_{v,MODSCAG}(\theta_s)$:

$$
\begin{aligned}
C_{v,MODSCAG}(\theta_s) &= \frac{C_{v,MODSCAG}(nadir)}{w(\theta_s)} \\
C_{v,MODSCAG}(nadir) &= C_{v,Landsat} = 0.0225
\end{aligned}
\tag{4.5}
$$

The MODSCAG measurement error covariance at nadir ($C_{v,MODSCAG}(nadir)$) is assumed to be equal to that of Landsat ($C_{v,Landsat}$, which has been estimated at 0.0225, or a 15% measurement error standard deviation by Cortés et al. [2014].

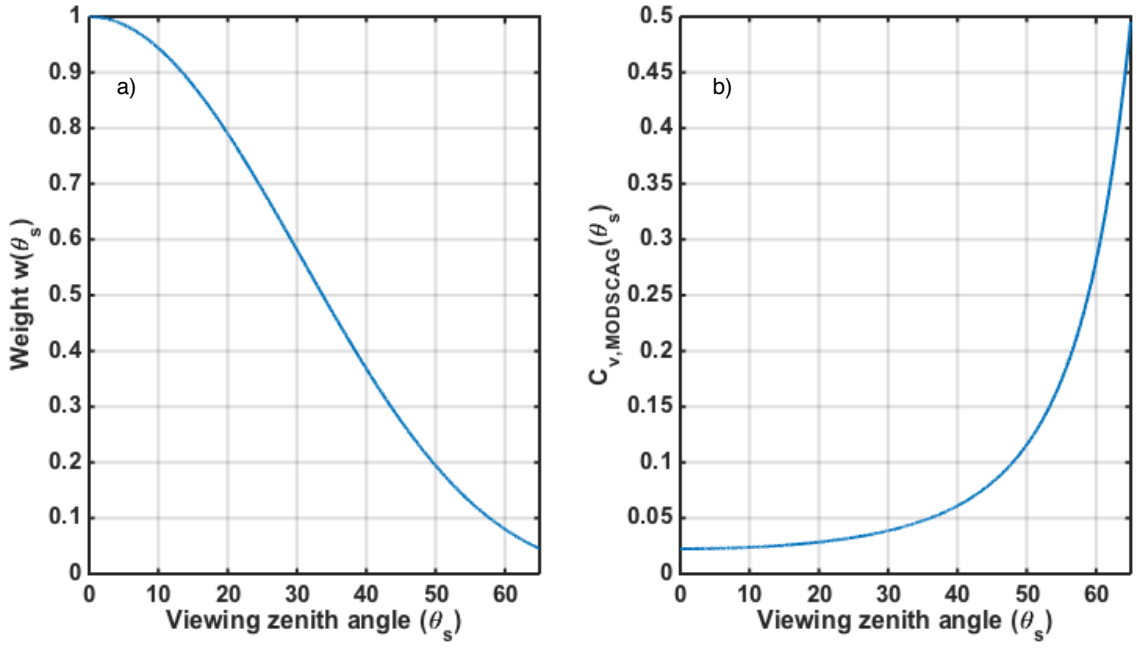

Figure 4.3: (a) Weighting function $w(\theta_s)$ and (b) MODSCAG measurement error covariance as a function of the sensor viewing angle θ_s.

As shown in Figure 4.3 (right panel), observations made at 35° would double the measurement error covariance from 0.025 at nadir to 0.05. In terms of error standard deviation, this translates into 15% at nadir and 22% at 35°. A viewing angle of 50° would multiply the nadir error covariance by 5, from 0.0225 to 0.11, leading to an error standard deviation of 33%. Off-nadir-looking measurements will therefore be assumed to have less information, and considered as less important in the DA compared to nadir.

4.3.2 Comparison of both sensors: Impact of MODSCAG viewing angle on *fveg* and *fsca*

The domain-averaged timeseries of nadir-looking (defined as having a viewing zenith angle of 12° or less) MODSCAG *fsca* and *fveg* show similar seasonal cycles and magnitude to Landsat for all three tiles and WYs tested as seen in Figure 4.4. As snow accumulates during the

winter, *fsca* increases and buries low-lying vegetation, which lowers *fveg* and leaves only the forest cover visible. Later in the season, *fsca* depletes and *fveg* increases as the low-lying vegetation resurfaces. Due to the lack of forest cover, the intra-annual dynamic of the HMA *fveg* is not as pronounced, and what is observed in Figure 4.4c is effectively the low-lying vegetation being covered by snow during the winter, and then reappearing during the ablation season. In contrast, the SN *fveg* showed no seasonality for WY 2014-15. This year was indeed particularly dry, and snow accumulated only over the highest areas of the tile, which are bare (Figure 4.1c), leaving the lower forested areas snow-free.

The difference with the most off-nadir-looking observations (defined as having a viewing zenith angle of 55° or more) is however noticeable, with off-nadir-looking *fveg* consistently overestimating nadir-looking MODSCAG and landsat observations for the SN and CRB tiles and both WYs tested (Figure 4.4a-b). Consequently, the higher *fveg* that amplifies the obscuring effect of snow cover leads to off-nadir-looking MODSCAG *fsca* lower than the nadir-looking measurements, especially during the ablation season (day of water year, or dowy 150 onward) of the wet WY 2003-04. As seen in Figure 4.4c, the HMA tile behaved differently than the Western U.S. tiles, due to its scarcely vegetated nature. The difference between nadir and off-nadir-looking MODSCAG observations is not as pronounced as in Figure 4.4a-b. The *fsca* and *fveg* measurements are comparable, which is consistent with the absence of tall vegetation over the tile. The MODSCAG *fsca* measurements are however consistently larger than Landsat, which suggests that the difference in nominal spatial resolution alone, regardless of viewing geometry can impact the *fsca* measured.

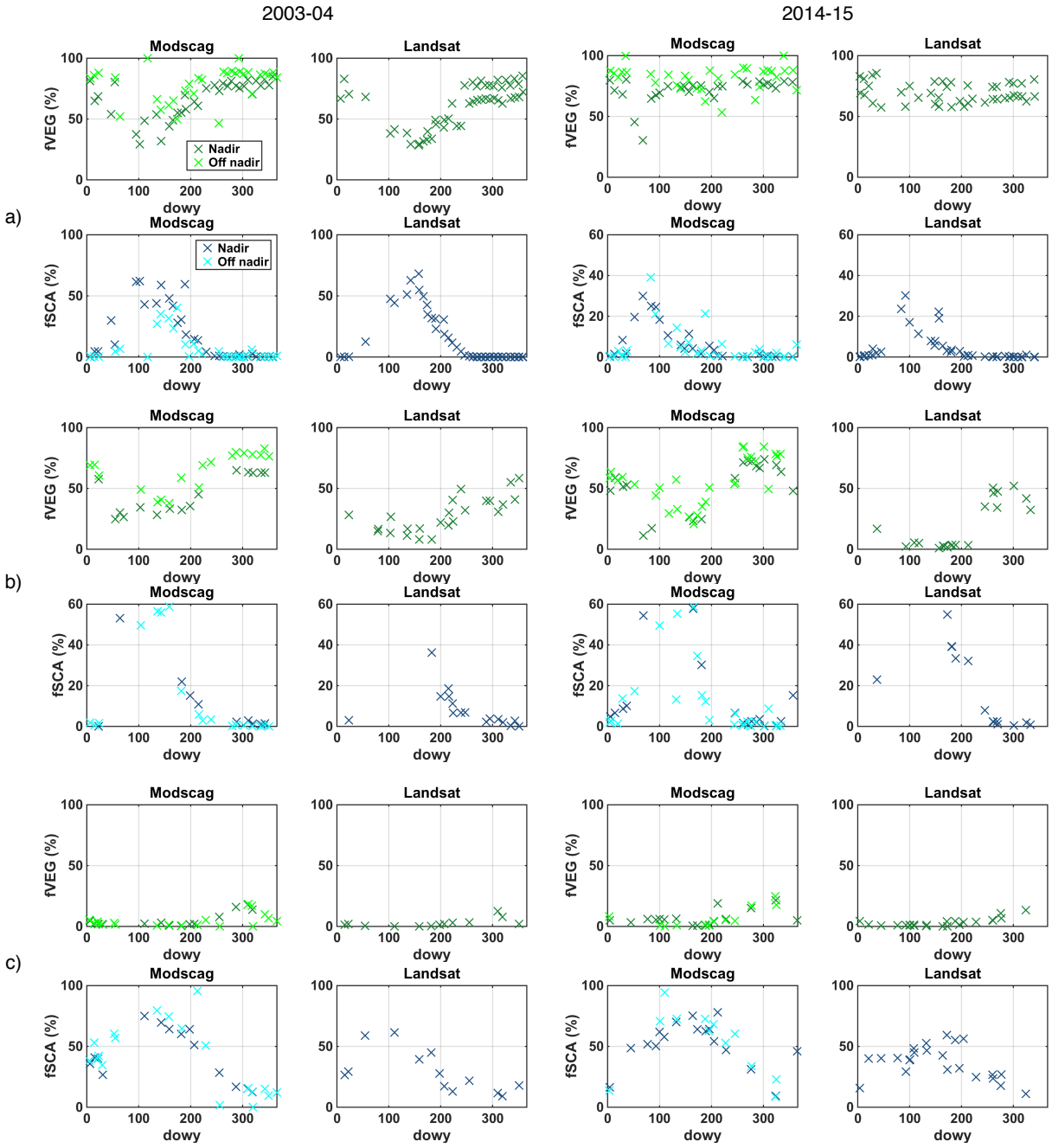

Figure 4.4: Domain-averaged timeseries of (top panels in green) *fveg* and (bottom panels in blue) *fsca* from MODSCAG and Landsat over the (a) SN, (b) CRB, and (c) HMA tiles for (left panel) WY 2003-04 and (right panel) WY 2014-15. MODSCAG nadir-looking observations are represented in dark colors, and off-nadir observations in light colors.

In order to better understand the impact of MODIS spatial resolution and viewing geometry on *fveg* and *fsca*, MODSCAG observations were extracted within one day of the Landsat acquisition for all three test tiles and WYs tested. Each Landsat image was compared to both a nadir ($w(\theta_s) > 0.90$, or within 12° of nadir) and a very off-nadir-looking ($w(\theta_s) < 0.10$, or with a viewing angle of 55° or more) MODSCAG observations. Since the MODSCAG algorithm is not sensitive enough to detect *fsca* below 15% (Painter et al. [2009]), this analysis focused exclusively on pixels where Landsat measured *fsca* of 15% or more to assess the impact of the sensor geometry only. All differences mentioned below have been calculated by subtracting Landsat from MODSCAG. A negative difference therefore means that MODSCAG is underestimating Landsat, and vice versa.

For densely forested areas such as in the SN and CRB tiles, the ability to see snow through tall vegetation can be significantly degraded at high viewing zenith angles as seen for two representative examples days in Figures 4.5 and 4.6. Even for nadir-looking acquisitions, Figure 4.5a and Figure 4.6a (top panel) clearly show that MODSCAG overestimated Landsat *fveg* over some areas of the tiles, despite having relatively unbiased difference distributions. This may be explained by the slight dissimilarity between the spectral endmember libraries used by MODSCAG and Landsat algorithms, as well as the difference in the sensors spatial resolution. Landsat observations were initially made at 30 m, and therefore capable of distinguishing areas of patchy forested cover that are classified as fully forested by MODSCAG at ~500m. The slight overestimation of *fveg* by nadir-looking MODSCAG therefore translated into a slight underestimation of *fsca* as seen in the top panels of Figure 4.5b and Figure 4.6b. The overestimation of *fveg* and underestimation of *fsca* was significantly aggravated for off-nadir-looking observations, with the *fveg* and *fsca* difference distributions respectively positively and negatively biased. The impact is especially visible over the SN tile, where snow almost completely disappeared at high viewing angles.

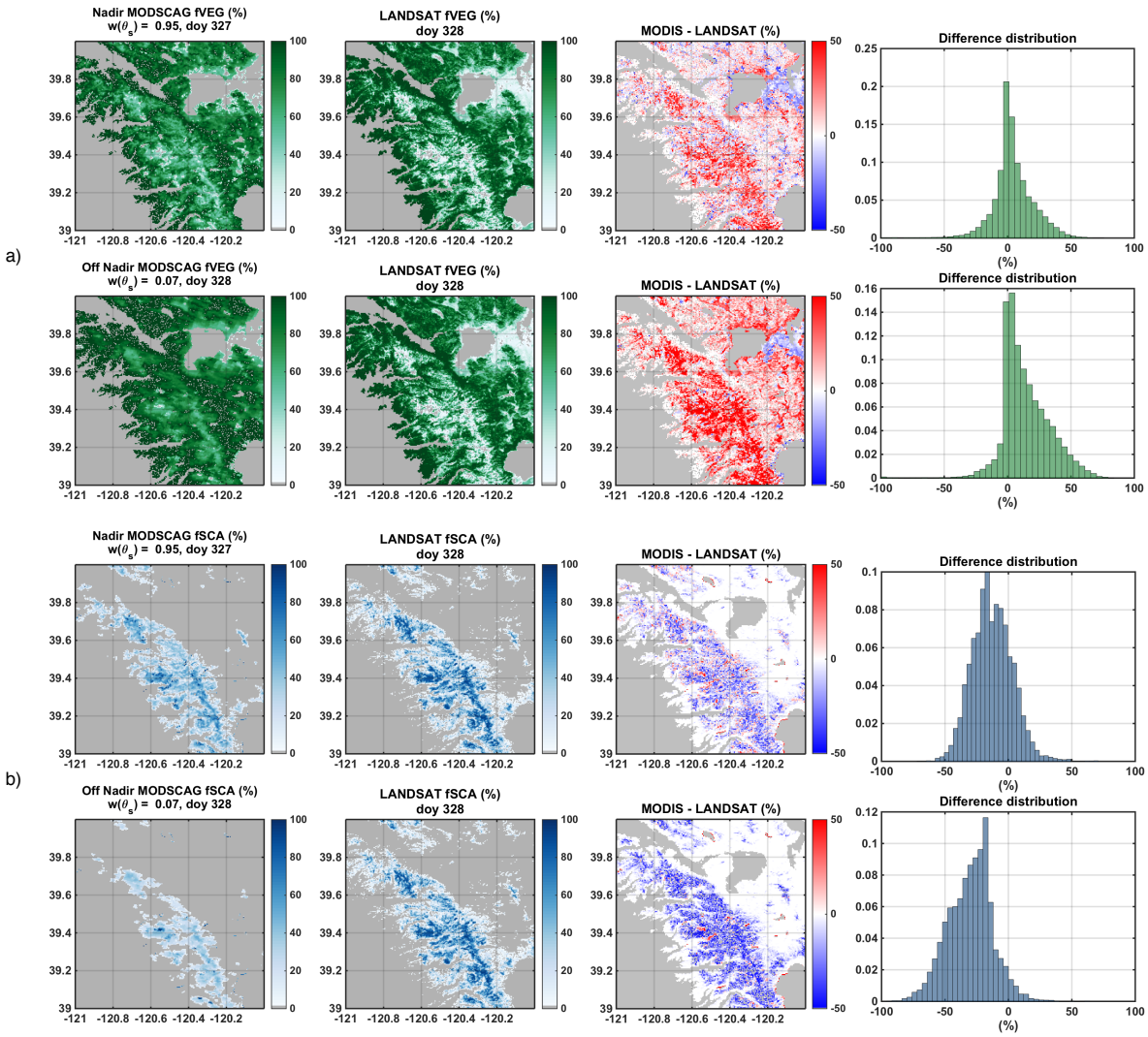

Figure 4.5: (a) maps of MODSCAG *fveg* for (top panel) nadir-looking acquisition and (bottom panel) off-nadir-looking acquisition around November 24, 2003 over the SN tile. The map of Landsat *fveg* follows the MODSCAG maps, and the relative difference maps as well as the distribution of the differences are displayed on the left panels. (b) maps of MODSCAG *fsca* for (top panel) nadir-looking acquisition and (bottom panel) off-nadir-looking acquisition around November 24, 2003 over the SN tile. The map of Landsat *fsca* follows the MODSCAG maps, and the difference maps as well as the distribution of the differences are displayed on the left panels. Only pixels with Landsat *fsca* larger than 15% were analyzed for the distributions shown in (a) and (b).

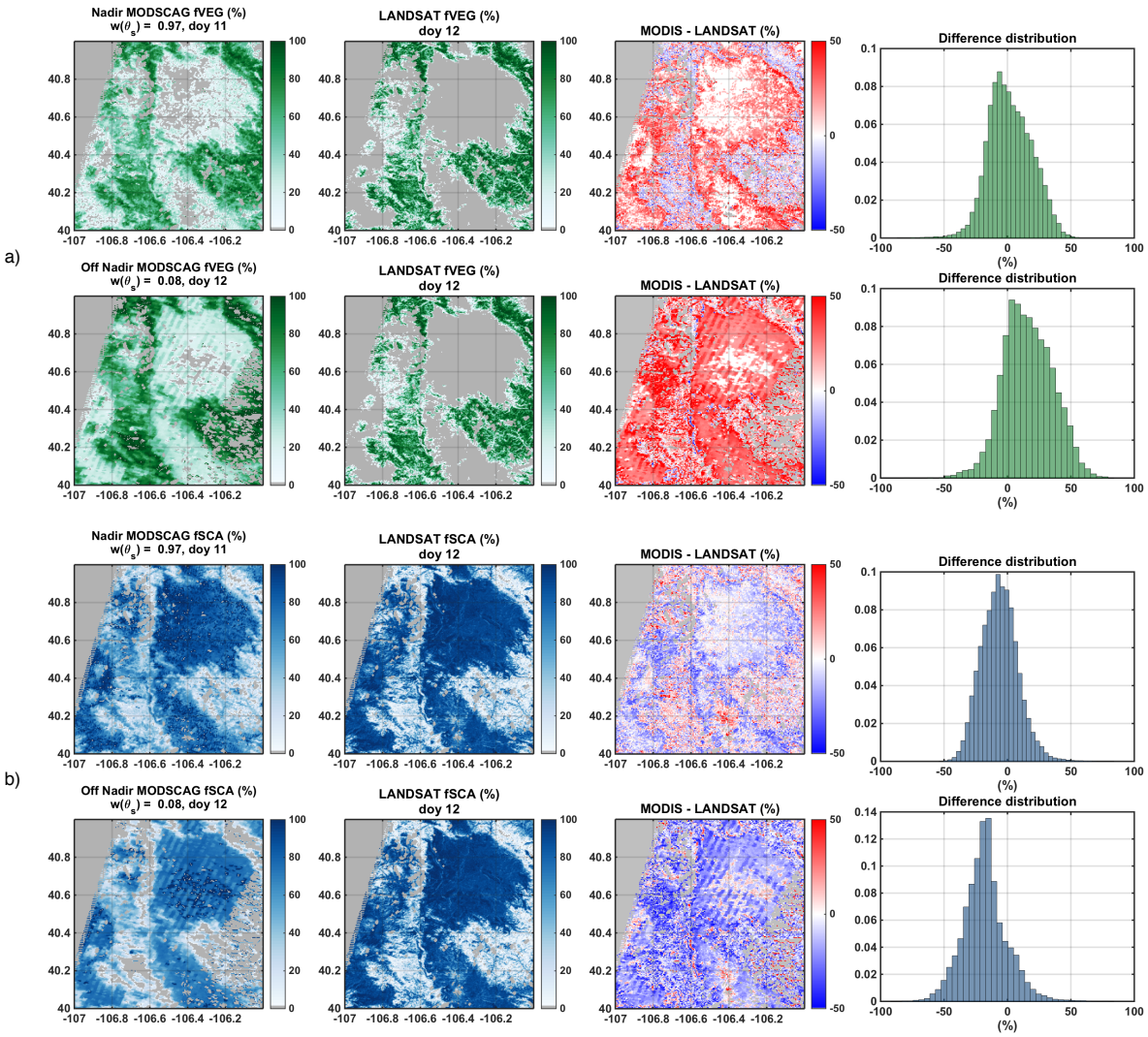

Figure 4.6: (a) maps of MODSCAG *fveg* for (top panel) nadir-looking acquisition and (bottom panel) off-nadir-looking acquisition around January 12, 2004 over the CRB tile. The map of Landsat *fveg* follows the MODSCAG maps, and the relative difference maps as well as the distribution of the differences are displayed on the left panels. (b) maps of MODSCAG *fsca* for (top panel) nadir-looking acquisition and (bottom panel) off-nadir-looking acquisition around January 12, 2004 over the CRB tile. The map of Landsat *fsca* follows the MODSCAG maps, and the difference maps as well as the distribution of the differences are displayed on the left panels. Only pixels with Landsat *fsca* larger than 15% were analyzed for the distributions shown in (a) and (b).

Regarding the HMA tile, the effects of MODSCAG coarser resolution and off-nadir-looking geometry are different given the lack of vegetation over the area. As seen in Figure 4.7a, almost no forested fractions were detected by either Landsat or MODSCAG. This means that the concealing of snow by tall vegetation was insignificant in this case. The slight overestimation of *fsca* by nadir-looking MODSCAG (Figure 4.7b, top panel) can be explained by the difference in spatial resolution. Similarly to the effect of the sensor resolution on *fveg* described previously for the SN and CRB tiles, Landsat is much more likely to detect bare soil pixels with patchy snow cover at 30 m, than MODIS at ~500m, leading to MODSCAG *fsca* slightly larger than Landsat. The overestimation of *fsca* is even more pronounced for the off-nadir-looking acquisition as seen in the bottom panel of Figure 4.7b. In addition to the effect of the coarser original resolution, the high viewing angles distort each pixel footprint, which can artificially increase *fsca* if no vegetation is present.

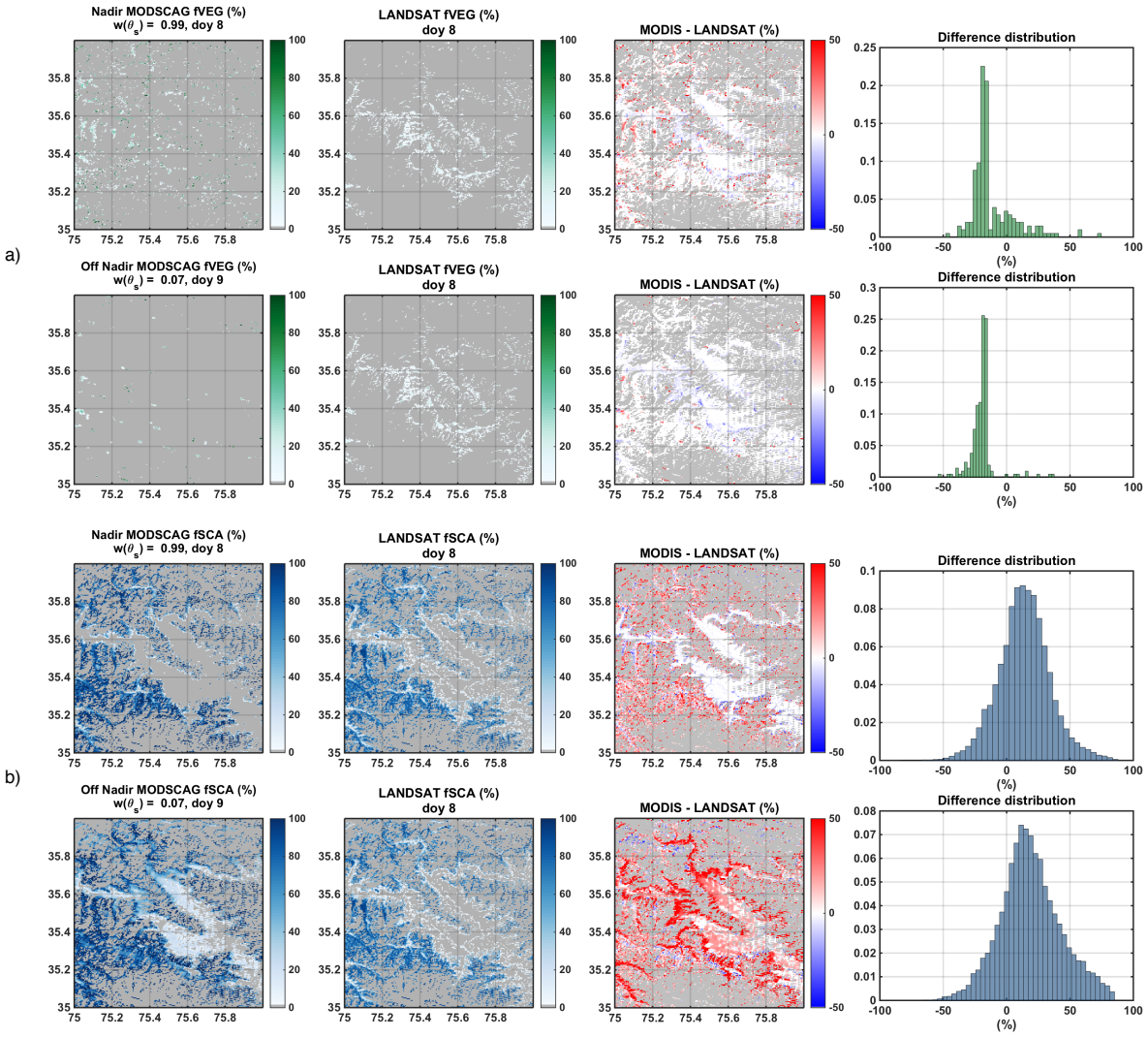

Figure 4.7: (a) maps of MODSCAG *fveg* for (top panel) nadir-looking acquisition and (bottom panel) off-nadir-looking acquisition around January 8, 2015 over the HMA tile. The map of Landsat *fveg* follows the MODSCAG maps, and the relative difference maps as well as the distribution of the differences are displayed on the left panels. (b) maps of MODSCAG *fsca* for (top panel) nadir-looking acquisition and (bottom panel) off-nadir-looking acquisition around January 8, 2015 over the HMA tile. The map of Landsat *fsca* follows the MODSCAG maps, and the difference maps as well as the distribution of the differences are displayed on the left panels. Only pixels with Landsat *fsca* larger than 15% were analyzed for the distributions shown in (a) and (b).

Figure 4.8 summarizes the differences between all common MODSCAG and Landsat observations made over snow covered pixels (with Landsat *fsca* of 15% or more). Figure 4.8a-b show that the difference distributions for nadir-looking MODSCAG *fveg* are clearly positively biased over both the SN and CRB tiles for the wet and dry WYs tested. The mean bias and spread of the difference distributions are similar, around 25%, and 22% respectively. Consistently, the difference distributions for nadir-looking MODSCAG *fsca* are clearly negatively biased, with means and standard deviations around -18% and -20% respectively. It is important to note that the nadir-looking angle threshold of 12° chosen might be playing a role in the differences observed, given that tall vegetation can obscure snow even at very small viewing angles. Similarly to what has been observed in Figures 4.5 and 4.6, the differences over the Western U.S. tiles are amplified by the high viewing zenith angle of the MODIS sensor. The spread of the distributions is similar to the nadir-looking cases but the means are significantly lower, especially for the SN tile where it is almost twice as large. Regarding the HMA tile, the differences in *fveg* are negligible due to the bare nature of the region. MODSCAG *fsca* performs better than for the forested tiles with only slightly positively biased difference distributions. For WY 2003-04, nadir and off-nadir acquisitions show a similar bias on the order of 5%, but the spread is larger for the off-nadir case. The distributions were more positively biased for WY 2014-15, which was slightly dryer, leading to a larger fraction of patchy snow pixels that were overestimated by MODSCAG.

In conclusion, MODIS coarser resolution of ~500 m, and scanning nature have the following impact on the *fsca* and *fveg* observed: 1) MODIS nadir-looking spatial resolution is not sensitive enough to capture areas of patchy snow or forest cover as accurately as Landsat, leading to an overestimation of *fveg* over densely forested regions, and an underestimation of *fsca* over bare soils. 2) The MODIS viewing geometry can emphasize the snow obscuring effect of tall vegetation, and thus underestimating *fsca* over forested areas. 3) The MODIS viewing geometry can distort the pixels footprint, which artificially increases *fveg* over forested areas and *fsca* over bare soil.

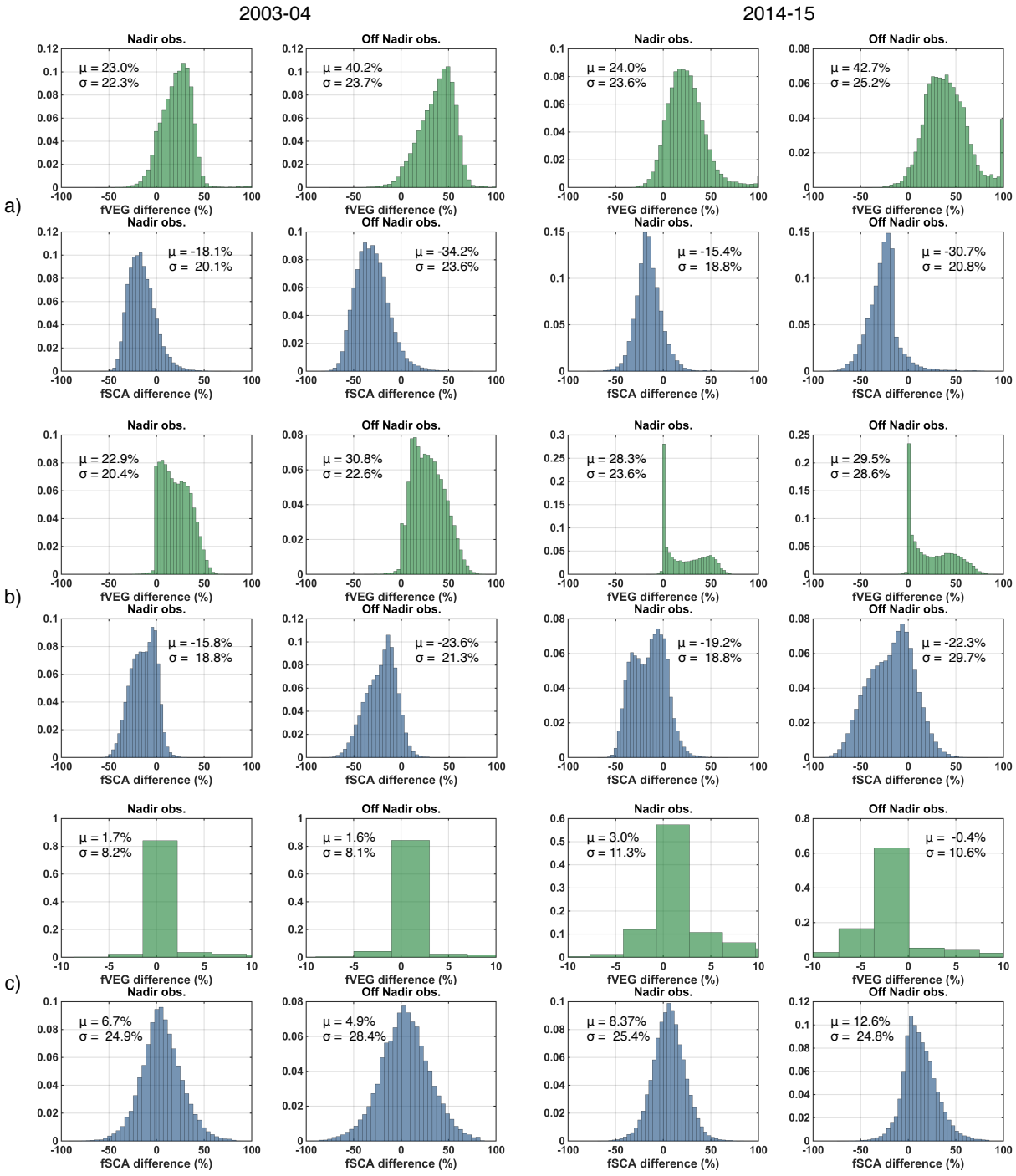

Figure 4.8: Distribution of the difference between Landsat and MODSCAG (green) *fveg* and (blue) *fsca* for the (a) SN, (b) CRB, and (c) HMA tiles. WY 2003-04 is on the left panel and WY 2014-15 is on the right panel. The mean (μ) and standard deviation (σ) are specified for each distribution.

4.4 Joint assimilation of Landsat and MODSCAG *fsca* products

4.4.1 Assimilation of *fsca* from Landsat, MODSCAG, and the combination of both

Three different assimilation setups were compared for WY 2003-04, over a forested tile in the CRB, and a bare tile in the HMA, in order to understand the joint impact of MODSCAG and Landsat on the SWE reanalysis. First, a baseline case assimilating only Landsat observations was performed. This setting was proven to yield extremely accurate SWE estimates over the CRB (Chapter 3 of this dissertation), but due to the lack of in-situ measurements over the HMA, a verification of the posterior SWE estimates over this tile was not feasible. Second, given the impact of MODIS zenith viewing angle on the observed *fsca* and *fveg* (section 4.3.1), only nadir-looking (defined as $w(\theta_s) > 0.90$, or within 12° of nadir) MODSCAG acquisitions were used in a MODSCAG-only case. Finally, a joint assimilation of Landsat and nadir-looking MODSCAG was performed.

Since all the observations assimilated are nadir (Landsat-only), or within 12° of nadir, the measurement error covariances $C_{v,MODSCAG}(12°)$ and $C_{v,Landsat}$ remain similar as seen in Figure 4.3, with the $C_{v,MODSCAG}(12°)$ only being slightly larger, making the MODSCAG observations with viewing angles larger than zero slightly less trustworthy than Landsat. In addition, the nadir-looking *fveg* used in the DA measurement model varies for each reanalysis. For the Landsat-only case, all winter Landsat *fveg* were averaged, and similarly all winter *fveg* were averaged for the nadir-looking MODSCAG-only case. The winter average of both *fveg* measurement vectors was used in the joint case. As seen in section 4.3.2, MODSCAG overestimates Landsat *fveg* over the forested Western U.S. tiles. The $\overline{fveg_{winter}}$ used in the measurement model (section 4.2.3) will therefore be the largest for the MODSCAG-only case, followed by the joint assimilation case, and then by the Landsat-only case. Consequently, the prior $fsca_{predicted}$ for the MODSCAG-only case will be the lowest for forested tiles such as the CRB. Regarding the HMA tile, most pixels are completely bare, and section 4.3.2 showed negligible differences between MODSCAG and Landsat *fveg*, which means that the

DA measurement model will perform similarly for each reanalysis case.

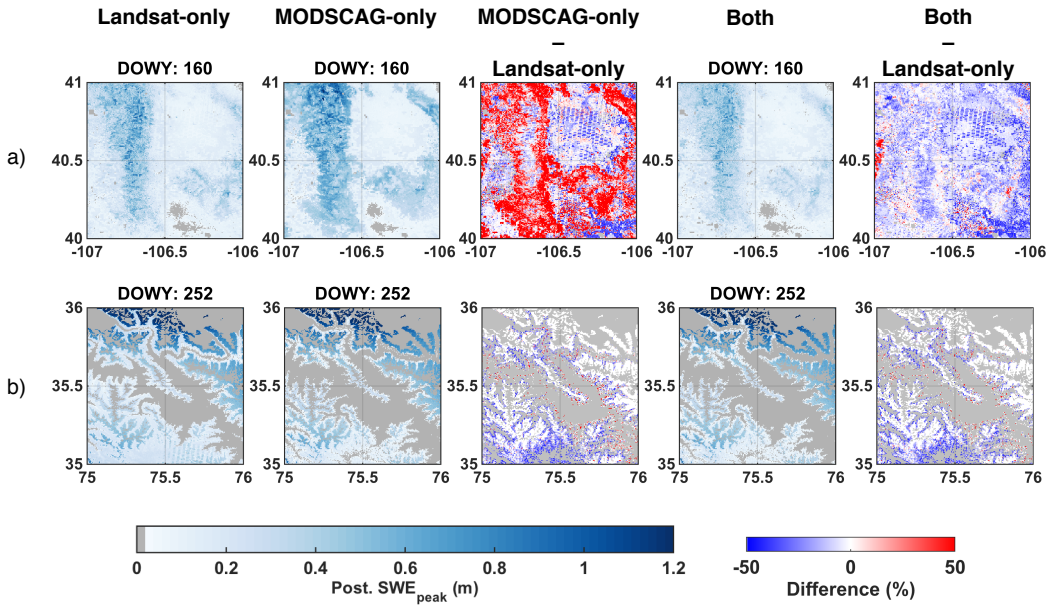

Figure 4.9: Maps of posterior peak SWE (SWE_{peak}) for the (a) CRB, and (b) HMA tiles for WY 2003-04. SWE_{peak} generated using Landsat-only, MODSCAG-only and both are represented in the first, second and fourth panels from the right respectively. The differences between the Landsat-only and MODSCAG-only cases are displayed in the third panel from the right, and the differences between the joint DA and the Landsat-only are on the fifth panel.

As seen in Figure 4.9, all three cases led to similar posterior peak SWE (SWE_{peak}) spatial distributions and timing. For the CRB, the joint assimilation case remained similar to the Landsat-only estimates, with underestimations on the order of 10% over areas accumulating significant amounts of SWE as seen in Table 4.2. In contrast, the MODSCAG-only case shows significantly larger SWE_{peak} estimates for the CRB tile (Figure 4.9a, and Table 4.2), with overestimations up to 50%. The HMA tile displays a different behavior, with Landsat-only estimates being larger than the other two (Figure 4.9b, and Table 4.2). However, given the limited amount of Landsat observations available for assimilation over the HMA tile for

that year (Table 4.1), the comparison with the Landsat-only case is not as meaningful for this particular case.

Tile	WY 2003-04	
	MODSCAG - Landsat	Both - Landsat
CRB	26.7%	-11.1%
HMA	-36.3%	-37.4%

Table 4.2: Domain-averaged relative differences of SWE_{peak} between the MODSCAG-only and Landsat-only cases, and between the assimilation of both MODSCAG and Landsat and the Landsat-only cases. Pixels with SWE_{peak} lower than 5 cm were excluded from the analysis.

The domain-averaged timeseries from Figure 4.10 show two different outcomes between the forested and bare tiles. As seen in Figure 4.10a, the MODSCAG-only posterior SWE ensemble median and interquartile range (IQR) are the largest, followed by the Landsat-only and then the joint estimates. The MODSCAG-only reanalysis therefore led to larger SWE, with more uncertainty, and the joint DA case led to smaller SWE, with less uncertainty than using Landsat-only. The magnitude of the differences are however small, on the order of 5 cm or less between each reanalysis case. Figure 4.10b shows that the Landsat-only case over the HMA tile led to the largest posterior ensemble SWE, with the joint DA posterior SWE being almost identical to the MODSCAG-only reanalysis. Given the small number of Landsat measurements available over the HMA tile for WY 2003-04, the similarity between the MODSCAG-only and joint DA cases was expected.

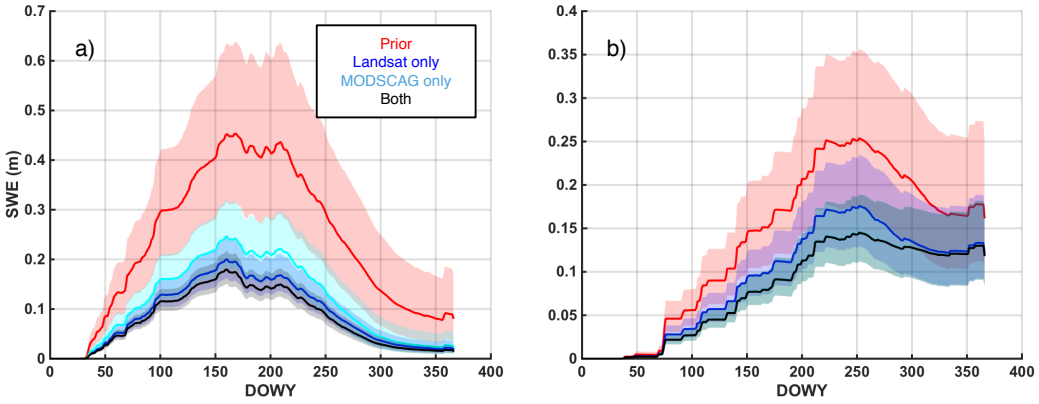

Figure 4.10: Domain-averaged timeseries of prior, landsat-based, MODSCAG-based and multi-sensor-based posterior SWE for the (a) CRB, and (b) HMA tiles for WY 2003-04. The ensemble medians are represented in solid lines, and the interquartiles ranges by shaded areas.

The impacts of the assimilation of MODSCAG observations over the CRB tile are further illustrated in Figure 4.11 over a pixel collocated with one snow pillow. First, the comparison of SWE timeseries with the in-situ data confirms the good performance of posterior estimates compared to the prior for all three reanalysis cases performed. It is however important to note that the in-situ SWE data from the SNOTEL stations are not necessarily representative of the 480 m by 480 m pixel, which may partly explain the differences observed in Figure 4.11b. As seen in Figure 4.11a, MODSCAG-only and the joint prior $fsca_{predicted}$ are lower than for the Landsat-only case, which is consistent with how the DA measurement model treats the three different sources of observation assimilated. The assimilation step using MODSCAG observations will therefore update the prior estimates by assigning a higher weight to the larger $fsca_{predicted}$ realizations that are closer to the measurements (Figure 4.11a). This effectively increases the posterior bare *fsca*, and consequently the mixed bare soil and forested posterior SWE estimates (Figure 4.11b). In addition to the impact in the measurement model, the high MODSCAG *fsca* detected late in the ablation season over the CRB tile in Figure 4.11a may be misinterpreted cloud cover, which can lead to a further

overestimation of posterior *fsca* and SWE. The increase of posterior bare *fsca* is mitigated in the joint assimilation case by Landsat measurements, leading to posterior SWE estimates in between the Landsat-only and MODSCAG-only cases as seen in the right panel in Figure 4.11a.

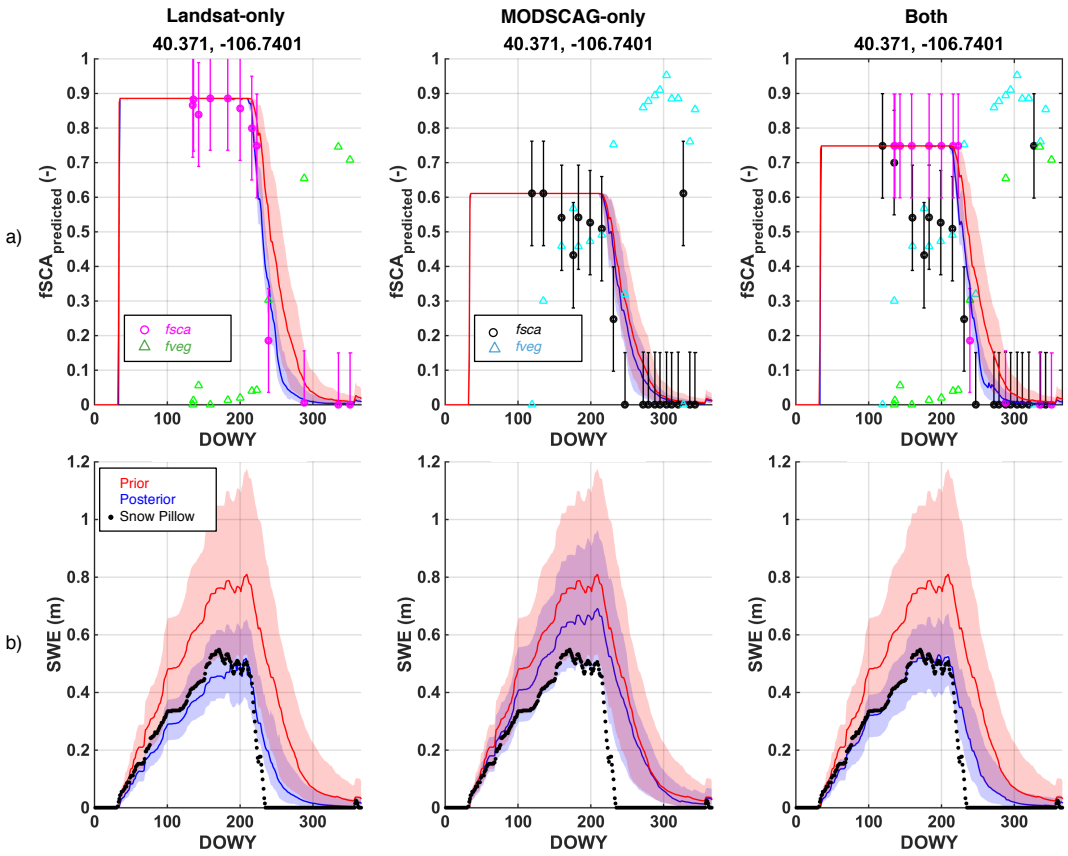

Figure 4.11: Timeseries of prior, landsat-based, MODSCAG-based and multi-sensor-based posterior (a) bare soil $fsca_{predicted}$ and the resulting (b) SWE (bare and forested mix) for a pixel located in the CRB tile for WY 2003-04. The ensemble medians are represented in solid lines, and the interquartile ranges by shaded areas. The uncertainty of of each *fsca* measurement is represented by an error bar. In-situ SWE measurements from the nearest snow pillow are also included.

In contrast to the CRB tile, which shared a similar number of nadir-looking MODSCAG, and Landsat measurements, the lack of Landsat coverage over the HMA means that the

assimilation step is mostly controlled by MODSCAG (Figure 4.10b). As seen in Figure 4.12a, only one Landsat measurement was available for assimilation over the pixel illustrated during WY 2003-04. Further, this observation occurred late during the ablation season, and low-lying vegetation increased the $\overline{fveg_{winter}}$ used in the measurement model, leading to smaller bare soil $fsca_{predicted}$ and a smaller update of the prior estimates. Ultimately the larger number of MODSCAG measurement available led to joint DA estimates being closer to the MODSCAG-only case.

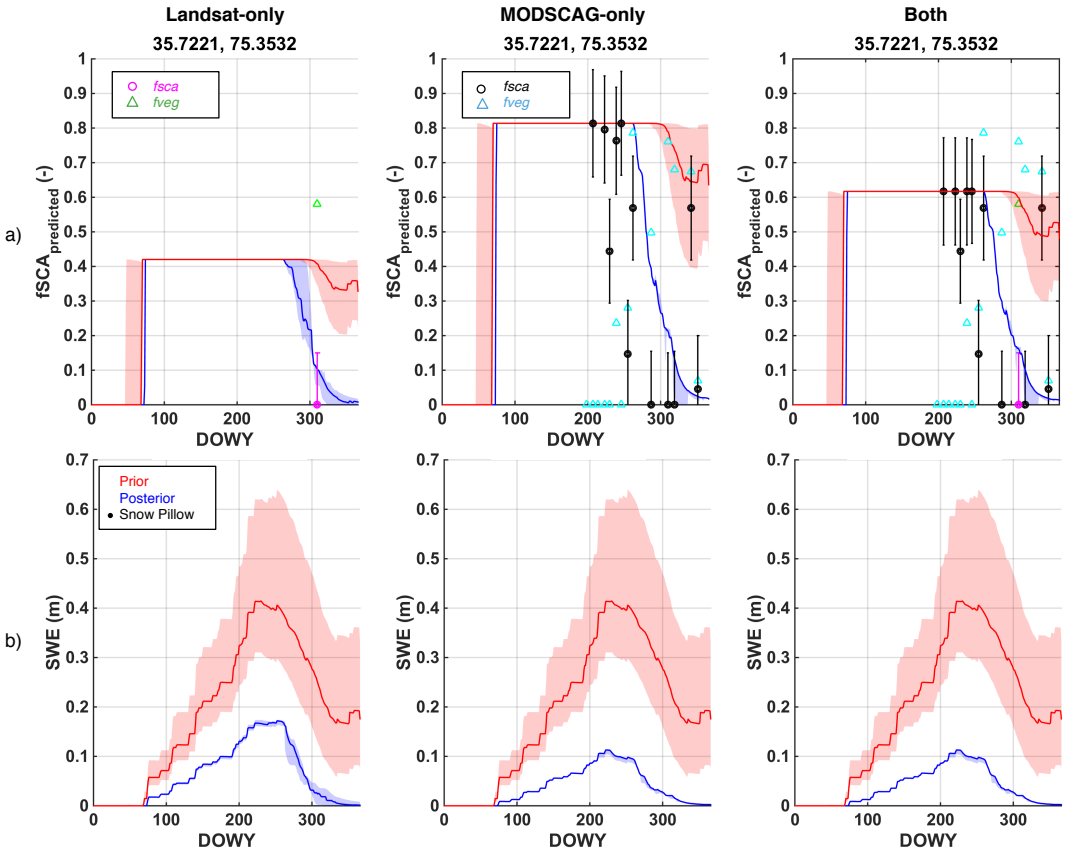

Figure 4.12: Timeseries of prior, landsat-based, MODSCAG-based and multi-sensor-based posterior (a) bare soil $fsca_{predicted}$ and the resulting (b) SWE (bare and forested mix) for a pixel located in the HMA tile for WY 2003-04. The ensemble medians are represented in solid lines, and the interquartile ranges by shaded areas. The uncertainty of of each *fsca* measurement is represented by an error bar.

4.4.2 Implementation of a different *fveg* in the measurement model: impact on posterior snow states over the forested CRB tile

As seen in Figures 4.11 and 4.12, using different sources of *fveg* to calculate $\overline{fveg_{winter}}$ in the measurement model may lead to significant differences between the Landsat-only and MODSCAG-only cases. Further, the goal of the measurement model is to take into account the effect of forested cover on the sensor ability to see snow, but as currently implemented $\overline{fveg_{winter}}$ might not be representative of forested cover due to snow interception by the canopy during the accumulation season, and low-lying *fveg* appearing during the ablation season. A new set of reanalyses were performed over the CRB tile and for WY 2003-04 using the static forest fraction $fveg_{static}$ (Sexton et al. [2013]) in the measurement model:

$$fsca_{predicted} = fsca_{bare} \times (1 - fveg_{static}) \tag{4.6}$$

In this configuration, $fveg_{static}$, and therefore $fsca_{predicted}$ are the same for the Landsat-only, MODSCAG-only, and joint DA cases. Since all observations are close to nadir, the $C_{v,meas}$ are also similar for all three reanalyses performed. In addition, *fsca* measurements from Landsat and MODSCAG were assigned the same upper limit $1 - fveg_{static}$ as the prior predicted *fsca* in order to be consistent with the measurement model. Finally, the MODSCAG cloud observations previously misclassified as high *fsca* during the late ablation season were corrected.

In contrast to Figure Figure 4.9a and Figure 4.10b, Figure 4.13 shows that assimilations using $fveg_{static}$ in the measurement model led to the MODSCAG-only posterior ensemble (both median and IQR) being closer and within 2 cm of the Landsat-only, and joint DA cases. This can be explained by the fact that $fveg_{static}$ is usually lower than the $\overline{fveg_{winter}}$ used previously in the MODSCAG-only reanalyses, which mitigates the effect of MODSCAG *fveg* on the update step, leading to lower posterior ensembles that better fit the observations.

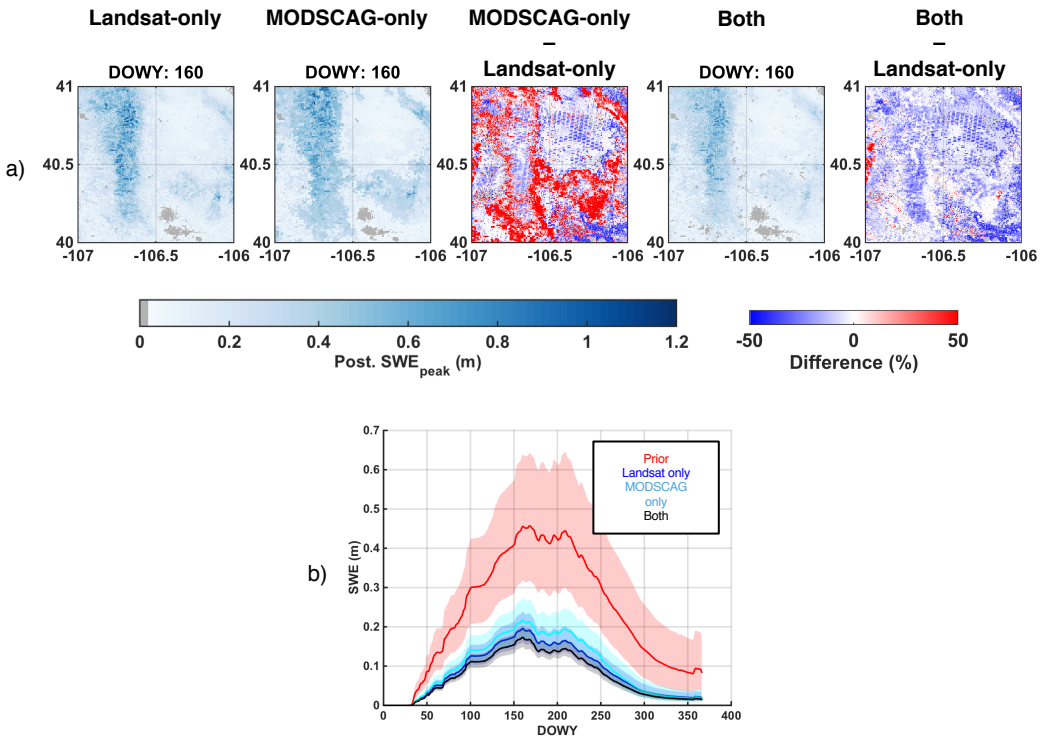

Figure 4.13: (a) Maps of posterior peak SWE (SWE_{peak}) for the CRB tile at time of peak SWE during WY 2003-04. SWE_{peak} generated using Landsat-only, MODSCAG-only and both are represented in the first, second and fourth panels from the right respectively. The differences between the Landsat-only and MODSCAG-only cases are displayed in the third panel from the right, and the differences between the joint DA and the Landsat-only are on the fifth panel. (b) Domain-averaged timeseries of prior, landsat-based, MODSCAG-based and multi-sensor-based posterior SWE for the CRB during WY 2003-04. The ensemble medians are represented in solid lines, and the interquartiles ranges by shaded areas.

Similarly, the pixel timeseries in Figure 4.14 show a slightly better agreement between the Landsat-only and MODSCAG-only posterior *fsca* and SWE due to the higher MODSCAG-only $fsca_{predicted}$ that led to smaller posterior SWE compared to Figure 4.11. In addition, all three reanalyses are consistent in terms of measurement model, and measurement error covariance, which means that the differences observed between the three posterior ensembles

are solely due to the differences in *fsca* measurements between Landsat and MODSCAG.

Other timeseries of pixels collocated with in-situ stations in the CRB tile also demonstrated improvements in fitting the posterior *fsca* with the remote sensing observations. Given the slight differences between MODSCAG and Landsat *fsca*, the domain-averaged MODSCAG-only and joint DA posterior estimates may differ from Landsat-only, but by a negligible margin of a few cm (Figure 4.13b). Using a static forest fraction in the measurement model may therefore better represent the impact of forested cover on the sensor ability to see snow at nadir, than a temporal average of vegetation fraction during the winter as done previously in section 4.4.1.

As mentioned previously, diagnosing the forest fraction from the winter timeseries of the remotely sensed *fveg* may not be appropriate due to the effects of canopy interception, and low-lying vegetation. Utilizing $\overline{fveg_{winter}}$ in the DA assimilation framework has however the advantage of being consistent with the observed *fsca*, and the inter-annual dynamic of the forest fraction. On the other end, $fveg_{static}$ might provide a better representation of the actual forest cover, but its static nature may not take into account the slowly-varying dynamics expected of forest changes.

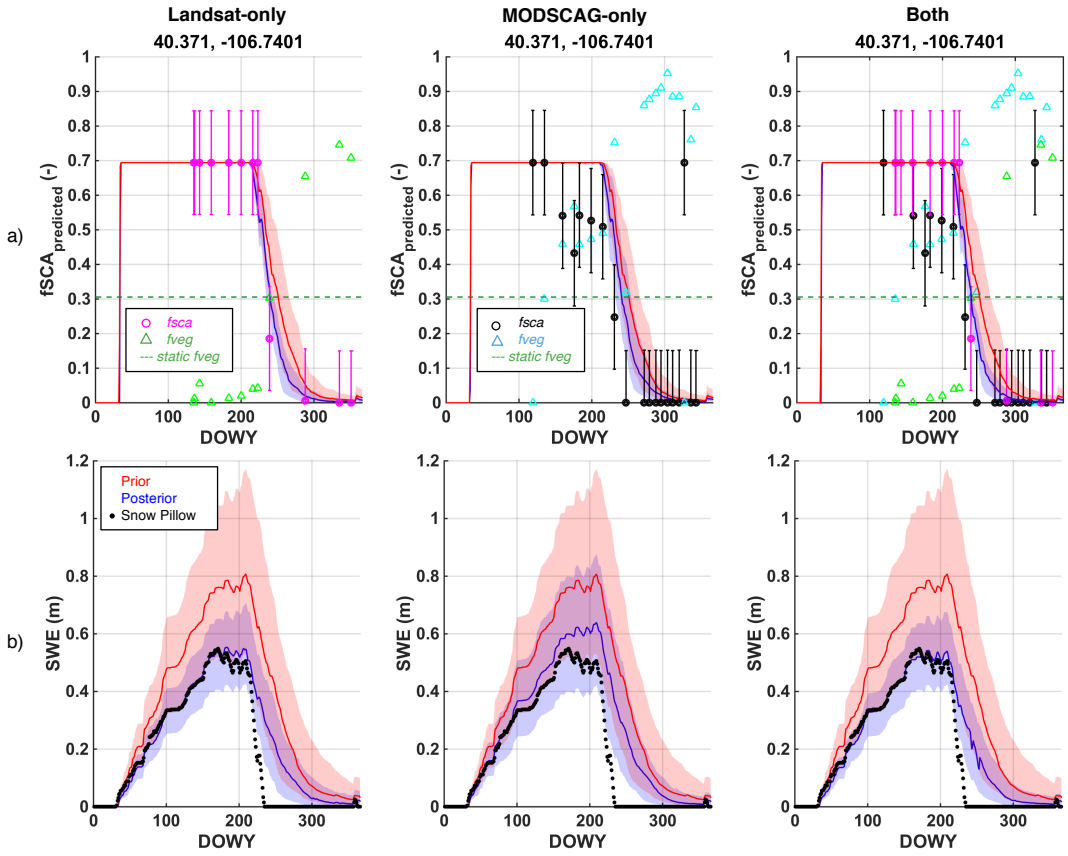

Figure 4.14: Timeseries of prior, landsat-based, MODSCAG-based and multi-sensor-based posterior (a) bare soil $fsca_{predicted}$ using $fveg_{static}$ and the resulting (b) SWE (bare and forested mix) for a pixel located in the CRB tile for WY 2003-04. The ensemble medians are represented in solid lines, and the interquartile ranges by shaded areas. The uncertainty of of each *fsca* measurement is represented by an error bar. The $fveg_{static}$ used in the measurement model is represented by the broken line. In-situ SWE measurements from the nearest snow pillow are also included.

4.5 Summary and preliminary conclusions

This work investigated the assimilation of the MODIS-based MODSCAG *fsca* product in complement to the Landsat-derived *fsca*, in order to increase the availability of remote sensing observations over midlatitude montane regions. MODSCAG was found to overestimate *fsca* over densely forested regions in the SN and CRB, especially at high zenith viewing angle, which increased the snow obscuring effect of tall vegetation. Over the scarcely forested HMA tile however, nadir-looking MODSCAG observations were in good agreement with Landsat, but off-nadir-looking measurements led to a distortion of the pixels footprints. The joint Landsat-MODSCAG data assimilation therefore only included close to nadir-looking MODSCAG *fsca*, and took into account the effect of the sensor geometry by increasing the measurement error with the viewing angle, and the effect of tall vegetation at nadir through the measurement model. The SWE reanalysis generated using both data streams was found to be identical to the Landsat-only SWE reanalysis in terms of peak SWE timing, and within a few cm in terms of peak SWE magnitude over the CRB tile, thus showing the potential usefulness of the approach over midlatitude montane ranges, especially over the Landsat-scarce HMA. The representation of forest fraction in the measurement model remains an open question that requires further analyses over more tiles and water years. In addition to MODSCAG *fsca* data, future work will include the assimilation of the MODDRFS dust forcing product (Painter et al. [2012]) in order to better represent the change in snow albedo due to dust loading. Finally, this joint data assimilation approach will be made compatible with the multi-scale nature of both sensors, as well as the multi-resolution approach described in Chapter 2 and Chapter 3.

CHAPTER 5

Conclusions

5.1 Summary and original contributions

This research developed different sets of useful tools to facilitate the large-scale implementation of a high resolution snow reanalysis framework over midlatitude montane ranges. First, a multi-resolution (MR) grid-based terrain discretization approach was developed in the context of sub 100 m snow modeling in Chapter 2. The main conclusions can be summarized as follows:

1. The physiographic complexity (CM) of the terrain can be constructed as a function of the standard deviations of elevation, northness index and forested fraction and serve as a proxy for snowmelt rate errors.

2. By focusing the use of fine and expensive spatial resolutions over areas of high CM exclusively, the MR approach was designed to mitigate the error of spatially distributed fine-resolution snow model applications.

3. The MR technique was successfully applied over a large test watershed in the Colorado River Basin and proved to be superior in terms of modeled snowmelt accuracy to the uniform grid-based terrain discretizations (at 180 m, 360 m, and 720 m), while increasing computational efficiency.

Such a technique was further proven to be robust in the context of a probabilistic data assimilation framework and was successfully applied over the same test watershed, to generate a MR SWE reanalysis dataset for the 30+ years Landsat-era. The main conclusions from Chapter 3 can be summarized as follows:

1. The MR approach was shown to have an insignificant impact on the *fsca* observations assimilated and the reanalysis framework led to posterior SWE ensembles similar to the high-resolution 90 m baseline, within 1 cm for peak SWE magnitude and 1 day for timing.

2. An important consequence of the MR approach is that most of the difference with the 90 m baseline occurred in areas accumulating less than 15 cm of SWE, while areas accumulating more than that are estimated with a high degree of accuracy.

3. In addition to estimating central tendencies with accuracy, the MR approach also preserved the SWE uncertainty, where the ensemble standard deviation and coefficient of variation showed differences on the order of -1% and 0.5% respectively with the 90 m baseline.

Finally, a large-scale implementation of the SWE reanalysis cannot be done by solely relying on Landsat *fsca* observations, which are limited by its low temporal resolution, cloud cover and coverage gap, especially over the HMA. The joint assimilation of Landsat and MODIS based *fsca* was therefore investigated in Chapter 4 and the main conclusions can be summarized as follows:

1. The impact of the MODIS viewing geometry cannot be ignored. Off-nadir-looking measurements of *fveg* were significantly overestimating the nadir-looking Landsat ones. Consequently, off-nadir MODIS *fsca* observations were underestimating Landsat over densely forested areas.

2. Over scarcely vegetated regions such as the HMA, MODIS performed much better, with differences with Landsat *fsca* smaller than 10%.

3. By restricting the assimilation of MODIS information to ~ nadir only, and assigning a measurement error function that increases with viewing angle, a joint assimilation of the *fsca* products from both sensors was successfully implemented over three large test tiles in the Western U.S. and HMA.

5.2 Potential for future work

The following improvements on the data assimilation are considered before performing a full-scale implementation over midlatitude montane regions:

1. Chapter 4 explored the joint assimilation at a uniform spatial resolution of 480 m, thus not fully leveraging the information contained in the 30 m Landsat images. The following step will therefore consists in making the simultaneous assimilation of multi-scale observations from Landsat and MODIS compatible with the multi-resolution approach described in Chapter 2.

2. The impact of dust on snow albedo can be significant over midlatitude montane regions, especially in the High Mountain Asia. Assimilating the MODDRFS dust forcing product from Painter et al. [2012] will therefore be implemented in the near future.

3. Biases in precipitation forcings are notably difficult to estimate given the lack of in-situ data over midlatitude montane regions. Large uninformed ensembles of posterior

precipitations will be generated for each mountain range in order to characterize MERRA-2 precipitation biases and uncertainty.

In continuity with the work of Margulis et al. [2016] and Cortés and Margulis [2017], the SWE reanalysis approach will then be integrated as part of a global assimilation system, thus providing a complete climatological understanding of spatial and temporal controls in montane seasonal SWE.

CHAPTER 6

Supporting Information

The following supporting information describes the methodology used to derive improved precipitation forcing using images from Landsat 5 TM sensor over the Upper Yampa River basin. Six SNOTEL stations and monthly manually sampled SWE from seven snow courses were used to quantify the gain in modeled SWE accuracy after correcting the raw NLDAS-2 precipitation.

The raw hourly NLDAS-2 precipitation forcing can be inaccurate (Luo et al. [2003]), and a pure forward modeling using nominal precipitation inputs led to significant modeled SWE errors when verified with continuous SWE measurements from six SNOTEL stations and monthly manually sampled SWE from seven snow courses. Such accumulation season errors will mask the impact of aggregation effects on the ablation season processes. As seen in Table 6.1, all three water years showed mean absolute errors ranging from 27 to 52 cm, and root mean square errors from 37 to 65 cm when compared to in-situ measurements. The mean errors are positive, which means that the simulation significantly underestimated SWE.

In order to provide more realistic snow water equivalent (SWE) realizations at high spatial resolutions, the downscaling of precipitation forcings were improved using fractional snow cover area (fSCA) images from Landsat 5 TM sensor via the reanalysis step of the probabilistic batch smoother (PBS) data assimilation framework developed by Margulis et al. [2015]. An initial PBS ensemble run was performed at 90 m to correct the precipitation scaling coefficients b using fSCA observations. These corrected precipitation scaling coefficients were then extracted and used to downscale the NLDAS-2 precipitation forcing to the model

scale in order to run the deterministic simulations for this study:

$$P_i^{90m} = b_i^{90m} \times P_i^{NLDAS} \tag{6.1}$$

where P_i^{90m} is the precipitation forcing for pixel i at 90 m, b_i^{90m} is the fSCA corrected precipitation scaling coefficient for pixel i at 90 m, and P_i^{NLDAS} is a weighted average of raw precipitation from the four NLDAS cells closest to pixel i. Using improved precipitation forcing to force the deterministic simulations therefore allowed a more accurate modeling of snow processes, while remaining a forward simulation without any data assimilation step.

As shown in Table 6.2, constraining the precipitation inputs to values that result in a snow cover simulation consistent with Landsat leads to more realistic SWE estimates. Baseline SWE for all three water years showed mean errors on the order of five centimeters, mean absolute errors around 10 cm, root mean square errors of 15 cm when compared to in-situ measurements. In-situ SWE measurements are usually made in easily accessible areas such as clearings and therefore could misrepresent a partially forested collocated 90 m pixel, which may partly explain the relatively small differences observed.

Difference (cm)	WY 1988 (normal melt)		WY 1993 (late melt)		WY 2004 (early melt)		All WYs	
	Pillows	Courses	Pillows	Courses	Pillows	Courses	Pillows	Courses
MD	52.2	38.8	NaN	45.4	NaN	27.3	52.2	36.6
MAD	52.2	38.8	NaN	45.4	NaN	27.5	52.2	36.7
RMSD	64.7	52.4	NaN	56.3	NaN	37.5	64.7	48.8

Table 6.1: Verification of the simulation using the raw NLDAS-2 precipitation forcing with In-Situ SWE observations from a network of six SNOTEL pillows and seven snow courses over the Upper-Yampa River Basin. The mean difference (MD), the mean absolute difference (MAD) and the root mean square difference (RMSD) are shown. An error of NaN was assigned when modeling led to no snow on the sensor collocated pixels.

Difference (cm)	WY 1988 (normal melt)		WY 1993 (late melt)		WY 2004 (early melt)		All WYs	
	Pillows	Courses	Pillows	Courses	Pillows	Courses	Pillows	Courses
MD	-0.7	2.1	5.3	5.6	-4.1	-5.3	-0.2	0.5
MAD	2.0	6.5	11.3	12.5	10.1	9.6	8.0	9.6
RMSD	2.1	8.8	15.8	16.6	14.4	12.5	12.6	13.0

Table 6.2: Verification of the baseline scenario using improved precipitation forcing with In-Situ SWE observations from a network of six SNOTEL pillows and seven snow courses over the Upper-Yampa River Basin. The mean difference (MD), the mean absolute difference (MAD) and the root mean square difference (RMSD) are shown.

Bibliography

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

Arsenault, K. R., Houser, P. R., De Lannoy, G. J. M., and Dirmeyer, P. A. (2013). Impacts of snow cover fraction data assimilation on modeled energy and moisture budgets. *Journal of Geophysical Research: Atmospheres*, 118(14):7489–7504.

Baldo, E. and Margulis, S. A. (2017). Implementation of a physiographic complexity-based multiresolution snow modeling scheme. *Water Resources Research*, 53(5):3680–3694.

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

Barrett, A. (2003). National operational hydrologic remote sensing center snow data assimilation system (snodas) products at nsidc. nsidc special report 11. boulder, co, usa. *National Snow and Ice Data Center. Digital media.*

Beven, K., Cloke, H., Pappenberger, F., Lamb, R., and Hunter, N. (2015). Hyperresolution information and hyperresolution ignorance in modelling the hydrology of the land surface. *Science China Earth Sciences*, 58(1):25–35.

Beven, K. J. and Kirby, M. J. (1979). A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin*, 24(1):43–69.

Bierkens, M. F. P., Bell, V. A., Burek, P., Chaney, N., Condon, L. E., David, C. H., de Roo, A., Dll, P., Drost, N., Famiglietti, J. S., Flrke, M., Gochis, D. J., Houser, P., Hut, R., Keune, J., Kollet, S., Maxwell, R. M., Reager, J. T., Samaniego, L., Sudicky, E., Sutanudjaja, E. H., van de Giesen, N., Winsemius, H., and Wood, E. F. (2015). Hyper-resolution global hydrological modelling: what is next? *Hydrological Processes*, 29(2):310–320. HYP-14-0850.

Bingham, N. H. and Fry, J. M. (2010). *Regression Linear Models in Statistics*. Springer.

Blöschl, G. (1999). Scaling issues in snow hydrology. *Hydrological Processes*, 13(14-15):2149–2175.

Broxton, P. D., Zeng, X., and Dawson, N. (2016). Why do global reanalyses and land data assimilation products underestimate snow water equivalent? *Journal of Hydrometeorology*, 17(11):2743–2761.

Chaney, N. W., Metcalfe, P., and Wood, E. F. (2016). Hydroblocks: a field-scale resolving land surface model for application over continental extents. *Hydrological Processes*, 30(20):3543–3559. HYP-15-0603.R1.

Chelamallu, H. P., Venkataraman, G., and Murti, M. (2014). Accuracy assessment of modis/terra snow cover product for parts of indian himalayas. *Geocarto International*, 29(6):592–608.

Chen, X., Kumar, M., Wang, R., Winstral, A., and Marks, D. (2016). Assessment of the timing of daily peak streamflow during the melt season in a snow-dominated watershed. *Journal of Hydrometeorology*, 17(8):2225–2244.

Christensen, N., Wood, A., Voisin, N., Lettenmaier, D., and Palmer, R. (2004). The effects of climate change on the hydrology and water resources of the colorado river basin. *Climatic Change*, 62(1-3):337–363.

Clark, M. P., Hendrikx, J., Slater, A. G., Kavetski, D., Anderson, B., Cullen, N. J., Kerr, T., Örn Hreinsson, E., and Woods, R. A. (2011). Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A review. *Water Resources Research*, 47(7):n/a–n/a.

Clark, M. P., Slater, A. G., Barrett, A. P., Hay, L. E., McCabe, G. J., Rajagopalan, B., and Leavesley, G. H. (2006). Assimilation of snow covered area information into hydrologic and land-surface models. *Advances in Water Resources*, 29(8):1209 – 1221.

Cline, D., Elder, K., and Bales, R. (1998). Scale effects in a distributed snow water equivalence and snowmelt model for mountain basins. *Hydrological Processes*, 12(10-11):1527–1536.

Cortés, G., Girotto, M., and Margulis, S. (2016). Snow process estimation over the extratropical andes using a data assimilation framework integrating merra data and landsat imagery. *Water Resources Research*, 52(4):2582–2600.

Cortés, G., Girotto, M., and Margulis, S. A. (2014). Analysis of sub-pixel snow and ice extent over the extratropical andes using spectral unmixing of historical landsat imagery. *Remote Sensing of Environment*, 141(0):64–78.

Cortés, G. and Margulis, S. (2017). Impacts of el niño and la niña on interannual snow accumulation in the andes: Results from a high-resolution 31 year reanalysis. *Geophysical Research Letters*, 44(13):6859–6867. 2017GL073826.

Cosgrove, B. A., Lohmann, D., Mitchell, K. E., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., Marshall, C., Sheffield, J., Duan, Q., Luo, L., Higgins, R. W., Pinker, R. T., Tarpley, J. D., and Meng, J. (2003). Real-time and retrospective forcing in the north american land data assimilation system (nldas) project. *Journal of Geophysical Research: Atmospheres*, 108(D22):n/a–n/a. 8842.

De Lannoy, G. J. M., Reichle, R. H., Arsenault, K. R., Houser, P. R., Kumar, S., Verhoest, N. E. C., and Pauwels, V. R. N. (2012). Multiscale assimilation of advanced microwave scanning radiometereos snow water equivalent and moderate resolution imaging spectroradiometer snow cover fraction observations in northern colorado. *Water Resources Research*, 48(1):n/a–n/a. W01522.

De Lannoy, G. J. M., Reichle, R. H., Houser, P. R., Arsenault, K. R., Verhoest, N. E. C., and Pauwels, V. R. N. (2010). Satellite-scale snow water equivalent assimilation into a high-resolution land surface model. *Journal of Hydrometeorology*, 11(2):352–369.

Deems, J. S., Fassnacht, S. R., and Elder, K. J. (2006). Fractal distribution of snow depth from lidar data. *Journal of Hydrometeorology*, 7(2):285–297.

Deems, J. S., Painter, T. H., and Finnegan, D. C. (2013). Lidar measurement of snow depth: a review. *Journal of Glaciology*, 59:467–479.

Derksen, C. and Brown, R. (2012). Spring snow cover extent reductions in the 20082012 period exceeding climate model projections. *Geophysical Research Letters*, 39(19):n/a–n/a. L19504.

Dozier, J., Bair, E. H., and Davis, R. E. (2016). Estimating the spatial distribution of snow water equivalent in the world's mountains. *Wiley Interdisciplinary Reviews: Water*, 3(3):461–474.

Dozier, J., Painter, T. H., Rittger, K., and Frew, J. E. (2008). Time–space continuity of daily maps of fractional snow cover and albedo from modis. *Advances in Water Resources*, 31(11):1515 – 1526. Hydrologic Remote Sensing.

Durand, M., Molotch, N. P., and Margulis, S. A. (2008). A bayesian approach to snow water equivalent reconstruction. *Journal of Geophysical Research: Atmospheres*, 113(D20):n/a–n/a.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., and Alsdorf, D. (2007). The shuttle radar topography mission. *Reviews of Geophysics*, 45(2):n/a–n/a. RG2004.

Fiddes, J. and Gruber, S. (2012). Toposub: a tool for efficient large area numerical modelling in complex topography at sub-grid scales. *Geoscientific Model Development*, 5(5):1245–1257.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper,

C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B. (2017). The modern-era retrospective analysis for research and applications, version 2 (merra-2). *Journal of Climate*, 30(14):5419–5454.

Girotto, M., Cortés, G., Margulis, S. A., and Durand, M. (2014a). Examining spatial and temporal variability in snow water equivalent using a 27 year reanalysis: Kern river watershed, sierra nevada. *Water Resources Research*, 50(8):6713–6734.

Girotto, M., Margulis, S. A., and Durand, M. (2014b). Probabilistic swe reanalysis as a generalization of deterministic swe reconstruction techniques. *Hydrological Processes*, 28(12):3875–3895.

Hansen, M. C., Defries, R. S., Townshend, J. R. G., and Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, 21(6-7):1331–1364.

Homer, C., Dewitz, J., Fry, J., Coan, M., Hossain, N., Larson, C., Herold, N., McKerrow, A., VanDriel, J. N., and Wickham, J. (2007). Completion of the 2001 national land cover database for the conterminous united states. 73(4):337–341.

Ivanov, V. Y., Vivoni, E. R., Bras, R. L., and Entekhabi, D. (2004). Catchment hydrologic response with a fully distributed triangulated irregular network model. *Water Resources Research*, 40(11):n/a–n/a. W11102.

Jepsen, S., Harmon, T., Meadows, M., and Hunsaker, C. (2016). Hydrogeologic influence on changes in snowmelt runoff with climate warming: Numerical experiments on a mid-elevation catchment in the sierra nevada, {USA}. *Journal of Hydrology*, 533:332 – 342.

JPL (2009). Aster global digital elevation model. *NASA JPL*.

Kargel, J. S., Leonard, G. J., Bishop, M. P., Kaab, A., and Raup, B. H. (2014). *Global Land Ice Measurements from Space*. Springer-Praxis.

Kovalskyy, V. and Roy, D. (2013). The global availability of landsat 5 tm and landsat 7 etm+ land surface observations and implications for global 30m landsat data product generation. *Remote Sensing of Environment*, 130(Supplement C):280 – 293.

Kumar, M., Marks, D., Dozier, J., Reba, M., and Winstral, A. (2013). Evaluation of distributed hydrologic impacts of temperature-index and energy-based snow models. 56.

Kumar, S. V., Peters-Lidard, C. D., Arsenault, K. R., Getirana, A., Mocko, D., and Liu, Y. (2015). Quantifying the added value of snow cover area observations in passive microwave snow depth data assimilation. *Journal of Hydrometeorology*, 16(4):1736–1741.

Lemmetyinen, J., Pulliainen, J., Arslan, A., Rautiainen, K., Vehvilainen, J., Wiesmann, A., Nagler, T., Rott, H., Davidson, M., Schuettemeyer, D., and Kern, M. (2011). Analysis of active and passive microwave observations from the nosrex campaign. In *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*, pages 2737–2740.

Liston, G. E. (2004). Representing subgrid snow cover heterogeneities in regional and global models. *Journal of Climate*, 17(6):1381–1397.

Liston, G. E. and Elder, K. (2006). A meteorological distribution system for high-resolution terrestrial modeling (micromet). *Journal of Hydrometeorology*, 7(2):217–234.

Liu, Y., Peters-Lidard, C. D., Kumar, S., Foster, J. L., Shaw, M., Tian, Y., and Fall, G. M. (2013). Assimilating satellite-based snow depth and snow cover products for improving snow predictions in alaska. *Advances in Water Resources*, 54(Supplement C):208 – 227.

Luo, L., Robock, A., Mitchell, K. E., Houser, P. R., Wood, E. F., Schaake, J. C., Lohmann, D., Cosgrove, B., Wen, F., Sheffield, J., Duan, Q., Higgins, R. W., Pinker, R. T., and Tarpley, J. D. (2003). Validation of the north american land data assimilation system (nl-

das) retrospective forcing over the southern great plains. *Journal of Geophysical Research: Atmospheres*, 108(D22):n/a–n/a. 8843.

Mankin, J. S., Viviroli, D., Singh, D., Hoekstra, A. Y., and Diffenbaugh, N. S. (2015). The potential for snow to supply human water demand in the present and future. *Environmental Research Letters*, 10(11):114016.

Margulis, S. A., Cortés, G., Girotto, M., and Durand, M. (2016). A landsat-era sierra nevada snow reanalysis (1985–2015). *Journal of Hydrometeorology*, 17(4):1203–1221.

Margulis, S. A., Girotto, M., Cortés, G., and Durand, M. (2015). A particle batch smoother approach to snow water equivalent estimation. *Journal of Hydrometeorology*, 16(4):1752–1772.

Marsh, C. B., Pomeroy, J. W., and Spiteri, R. J. (2012). Implications of mountain shading on calculating energy for snowmelt using unstructured triangular meshes. *Hydrological Processes*, 26(12):1767–1778.

Mascaro, G., Vivoni, E. R., and Méndez-Barroso, L. A. (2015). Hyperresolution hydrologic modeling in a regional watershed and its interpretation using empirical orthogonal functions. *Advances in Water Resources*, 83:190 – 206.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E. H., Ek, M. B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., and Shi, W. (2006). North american regional reanalysis. *Bulletin of the American Meteorological Society*, 87(3):343–360.

Mir, R. A., Jain, S. K., Saraf, A. K., and Goswami, A. (2015). Accuracy assessment and trend analysis of modis-derived data on snow-covered areas in the sutlej basin, western himalayas. *International Journal of Remote Sensing*, 36(15):3837–3858.

Molotch, N. P., Painter, T. H., Bales, R. C., and Dozier, J. (2004). Incorporating remotely-

sensed snow albedo into a spatially-distributed snowmelt model. *Geophysical Research Letters*, 31(3):n/a–n/a. L03501.

Newman, A. J., Clark, M. P., Winstral, A., Marks, D., and Seyfried, M. (2014). The use of similarity concepts to represent subgrid variability in land surface models: Case study in a snowmelt-dominated watershed. *Journal of Hydrometeorology*, 15(5):1717–1738.

Painter, T. H., Bryant, A. C., and Skiles, S. M. (2012). Radiative forcing by light absorbing impurities in snow from modis surface reflectance data. *Geophysical Research Letters*, 39(17):n/a–n/a. L17502.

Painter, T. H., Rittger, K., McKenzie, C., Slaughter, P., Davis, R. E., and Dozier, J. (2009). Retrieval of subpixel snow covered area, grain size, and albedo from modis. *Remote Sensing of Environment*, 113(4):868–879.

Pan, M., Sheffield, J., Wood, E. F., Mitchell, K. E., Houser, P. R., Schaake, J. C., Robock, A., Lohmann, D., Cosgrove, B., Duan, Q., Luo, L., Higgins, R. W., Pinker, R. T., and Tarpley, J. D. (2003). Snow process modeling in the north american land data assimilation system (nldas): 2. evaluation of model simulated snow water equivalent. *Journal of Geophysical Research: Atmospheres*, 108(D22):n/a–n/a.

Pomeroy, J. W., Gray, D. M., Brown, T., Hedstrom, N. R., Quinton, W. L., Granger, R. J., and Carey, S. K. (2007). The cold regions hydrological model: a platform for basing process representation and model structure on physical evidence. *Hydrological Processes*, 21(19):2650–2667.

Qu, Y. and Duffy, C. J. (2007). A semidiscrete finite volume formulation for multiprocess watershed simulation. *Water Resources Research*, 43(8):n/a–n/a. W08419.

Reichle, R. H., Koster, R. D., De Lannoy, G. J. M., Forman, B. A., Liu, Q., Mahanama, S. P. P., and Touré, A. (2011). Assessment and enhancement of merra land surface hydrology estimates. *Journal of Climate*, 24(24):6322–6338.

Revuelto, J., López-Moreno, J. I., Azorin-Molina, C., and Vicente-Serrano, S. M. (2014). Topographic control of snowpack distribution in a small catchment in the central spanish pyrenees: intra- and inter-annual persistence. *The Cryosphere*, 8(5):1989–2006.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin*, J. K., Walker, J. P., Lohmann, D., and Toll, D. (2004). The global land data assimilation system. *Bulletin of the American Meteorological Society*, 85(3):381–394.

Rohrer, M., Salzmann, N., Stoffel, M., and Kulkarni, A. (2013). Missing (in-situ) snow cover data hampers climate change and runoff studies in the greater himalayas. 468.

Serreze, M. C., Clark, M. P., Armstrong, R. L., McGinnis, D. A., and Pulwarty, R. S. (1999). Characteristics of the western united states snowpack from snowpack telemetry (snotel) data. *Water Resources Research*, 35(7):2145–2160.

Sexton, J. O., Song, X.-P., Feng, M., Noojipady, P., Anand, A., Huang, C., Kim, D.-H., Collins, K. M., Channan, S., DiMiceli, C., and Townshend, J. R. (2013). Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of modis vegetation continuous fields with lidar-based estimates of error. *International Journal of Digital Earth*, 6(5):427–448.

Sivapalan, M., Beven, K., and Wood, E. F. (1987). On hydrologic similarity: 2. a scaled model of storm runoff production. *Water Resources Research*, 23(12):2266–2278.

Su, H., Yang, Z.-L., Niu, G.-Y., and Dickinson, R. E. (2008). Enhancing the estimation of continental-scale snow water equivalent by assimilating modis snow cover with the ensemble kalman filter. *Journal of Geophysical Research: Atmospheres*, 113(D8):n/a–n/a. D08120.

Sun, S. and Xue, Y. (2001). Implementing a new snow scheme in simplified simple biosphere model. *Advances in Atmospheric Sciences*, 18(3):335–354.

Tedesco, M., Reichle, R., Low, A., Markus, T., and Foster, J. (2010). Dynamic approaches for snow depth retrieval from spaceborne microwave brightness temperature. *Geoscience and Remote Sensing, IEEE Transactions on*, 48(4):1955–1967.

Trujillo, E., Ramrez, J. A., and Elder, K. J. (2007). Topographic, meteorologic, and canopy controls on the scaling characteristics of the spatial distribution of snow depth fields. *Water Resources Research*, 43(7):n/a–n/a. W07409.

Tucker, G. E., Lancaster, S. T., Gasparini, N. M., Bras, R. L., and Rybarczyk, S. M. (2001). An object-oriented framework for distributed hydrologic and geomorphic modeling using triangulated irregular networks. *Computers Geosciences*, 27(8):959 – 973. Geocomputation Geosciences.

U. S. Geological Survey, W. R. D., Leavesley, G. H., Lichty, R. W., Troutman, B. M., and Saindon, L. G. (1983). Precipitation-runoff modeling system; user's manual. Technical report.

Vivoni, E., Ivanov, V., Bras, R., and Entekhabi, D. (2004). Generation of triangulated irregular networks based on hydrological similarity. *Journal of Hydrologic Engineering - ASCE*, 9(4):288–302.

Wang, R., Kumar, M., and Marks, D. (2013). Anomalous trend in soil evaporation in a semi-arid, snow-dominated watershed. *Advances in Water Resources*, 57(Supplement C):32 – 40.

Winstral, A., Marks, D., and Gurney, R. (2014). Assessing the sensitivities of a distributed snow model to forcing data resolution. *Journal of Hydrometeorology*, 15(4):1366–1383.

Wood, E. F., Roundy, J. K., Troy, T. J., van Beek, L. P. H., Bierkens, M. F. P., Blyth, E., de Roo, A., Dll, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaff, P. R., Kollet, S., Lehner, B., Lettenmaier, D. P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A., and Whitehead, P. (2011). Hyperresolution global land surface

modeling: Meeting a grand challenge for monitoring earth's terrestrial water. *Water Resources Research*, 47(5):n/a–n/a. W05301.

Wulder, M. A., White, J. C., Loveland, T. R., Woodcock, C. E., Belward, A. S., Cohen, W. B., Fosnight, E. A., Shaw, J., Masek, J. G., and Roy, D. P. (2016). The global landsat archive: Status, consolidation, and direction. *Remote Sensing of Environment*, 185(Supplement C):271 – 283. Landsat 8 Science Results.

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., Alonge, C., Wei, H., Meng, J., Livneh, B., Lettenmaier, D., Koren, V., Duan, Q., Mo, K., Fan, Y., and Mocko, D. (2012). Continental-scale water and energy flux analysis and validation for the north american land data assimilation system project phase 2 (nldas-2): 1. intercomparison and application of model products. *Journal of Geophysical Research: Atmospheres*, 117(D3):n/a–n/a. D03109.

Xue, Y., Sellers, P. J., Kinter, J. L., and Shukla, J. (1991). A simplified biosphere model for global climate studies. *Journal of Climate*, 4(3):345–364.

Xue, Y., Sun, S., Kahan, D. S., and Jiao, Y. (2003). Impact of parameterizations in snow physics and interface processes on the simulation of snow cover and runoff at several cold region sites. *Journal of Geophysical Research: Atmospheres*, 108(D22):n/a–n/a.

Zhou, Q., Lees, B., and Tang, G. (2008). *Advances in Digital Terrain Analysis*. Lecture Notes in Geoinformation and Cartography. Springer Berlin Heidelberg.