2013

Ensemble Modeling of SWE Distribution in the Bitterroot Mountains, Montana, USA

Frederick Kellner
The University of Montana

ENSEMBLE MODELING OF SWE DISTRIBUTION IN THE BITTERROOT MOUNTAINS, MONTANA, USA

By

FREDERICK KELLNER

B.S. Geoscience, University of Montana, Missoula, Montana, 2011

Thesis

presented in partial fulfillment of the requirements
for the degree of

Master of Science
Geosciences

The University of Montana
Missoula, MT

Official Graduation Date Summer 2013

Approved by:

Sandy Ross, Dean of The Graduate School
Graduate School

Joel Harper, Chair
Geosciences

Marco Maneta
Geosciences

Jesse Johnson
Computer Science

Kellner, Frederick, M.S, Summer 2013 Geosciences

Ensemble Modeling of SWE Distribution in the Bitterroot Mountains, Montana, USA

Chairperson: Joel Harper

The spatial distribution of snow remains poorly understood at the landscape scale, particularly at high elevation where snow can be under-represented by our current system of monitoring. The transferability of the processes and physiography that drive the spatial distribution of snow require further study. We apply an ensemble of three semi-independent snow models to the down-sloping side of the Bitterroot Mountains of Western Montana, which features an array of drainages of remarkably similar size and aspect. We modeled the snow water equivalent equal to the maximum snow accumulation plus any positive contributions to the snowpack during the melt season, for the years 2000-2010. The three models yield similar magnitudes and patterns of snow water equivalent distribution. We find that upwards of 70% of the snow water equivalent is found above the elevation of 1950 meters and this snow water equivalent is represented by a single SNOTEL station within our 1200 km^2 study area. The difference in snow water equivalent on north and south facing aspects within in individual drainages is found to be small. At lower elevations snow water equivalent increases with elevation, while above the elevation of 2000 meters snow water equivalent remains constant as elevation increases. The difference in snow water equivalent at lower and higher elevations within the study area is driven by snow accumulation processes that differ between valleys and valley sidewalls/ridgetops within the study area. The processes that control the spatial distribution of snow water equivalent within this study area are site specific and are likely not transferable to other regions.

Contents

List of Figures and Tables

1. Introduction

Total precipitation in mountainous areas of the Western US has been estimated to be comprised of 39% - 67% precipitation falling as snow [*Serreze et al.*, 1999].The hydrologic cycle of the Western US is influenced by the accumulation and melt of a seasonal snowpack. The mountain snowpack acts to store winter precipitation; the melting and release of this stored precipitation is a central component of runoff and stream flow. The release of this stored precipitation is important for the biological function of stream systems, and can affect the production of hydropower, agriculture and economies tied to recreation.

Projected increases in mean global surface temperature [*IPCC*, 2007] have created concern over the vulnerability of mountain snowpacks to increases in warming. Current literature has argued that the implications of a warmer planet are that there will be less accumulated snowpack, [e.g., *Mote et al.*, 2005; *Regonda et al.*, 2005], or that the timing of spring snowmelt, and stream runoff will occur earlier in the year [e.g., *McCabe and Clark*, 2005; *Stewart et al.*, 2005]. However, fundamental to understanding the impacts of natural variability and climate change on accumulated mountain snowcover, is knowledge of the quantity and spatial distribution of snow.

Perhaps the most fundamental problem in the realm of snow hydrology is understanding the spatial distribution of SWE at the landscape scale. This remains a critical problem for water resource managers and scientists as there exists neither a current system of on ground measurement or remote sensing tool to solve this problem. Our inability to answer to solve this problem leaves us with a poor understanding of the quantity and spatial and temporal distribution of snow at the landscape scale. This hinders our efforts to interpret the implications of past, present and future events related to snow accumulation and melt.

1.1. Snow Measurement

Perhaps the most complete estimate of snow water equivalent (SWE) in the mountains of the Western US comes from data gathered by the Natural Resources Conservation Service (NRCS). The NRCS began formal management of this program in 1935, which today consist of 900 snow course locations where snow is measured manually monthly to bimonthly and 750 SNOw TELemetry (SNOTEL) stations where snow is measured in real-time, in Alaska and the Western US.

Use of data from this SWE monitoring network for scientific investigation is hampered by the temporal and spatial resolution of these data. The intent of the NRCS snow monitoring network is to provide volumetric water supply forecasts through the statistical indexing of monitoring sites, which were located to provide ease of access and restriction of human disturbance [*Molotch and Bales*, 2005].

Even with over 1650 SWE measurement locations scattered throughout the West, the spatial resolution of these measurements is coarse. For example, the area above the elevation of 1400 m in Montana is equivalent to over 100,000 km^2, which is represented by 94 SNOTEL stations. This SNOTEL station to area ratio, equates to one SNOTEL station for every 1300 km^2 in Montana where a seasonal snowpack would accumulate.

In addition to these sites being spatially coarse, the record of measurement at these locations is relatively short. A few NRCS snow course sites have a record period that pre-dates the 1935 start date of the SWE monitoring program, while most of the 750 SNOTEL stations came into existence in 1978 or later [*Schaefer and Paetzold*, 2001]. Further, many stations have shortened or incomplete records due impart to closure or changes to monitoring sites.

Up-scaling of NRCS SWE data to the landscape scale is not only complicated by the spatial and temporal resolution of these data, but the representativeness of these point measurements relative to average SWE of surrounding basins and larger grid cells. SNOTEL measurement sites have been found to preferentially represent densely forested areas and under represent higher elevation terrain [*Molotch and Bales*, 2006]. The stationary nature of these sites leads to biases of SNOTEL SWE measurement relative to basin or grid cell SWE averages that are different during the accumulation season versus the ablation season [*Molotch and Bales*, 2005, *Neumann et al.*, 2006; *Meromy et al.*, 2012]. Further, a two year comparison of SWE measured at automated SNOTEL stations and SWE measured at manual snow courses that were cited in the same location as a SNOTEL station, found mean relative errors of 15% and 25% between the two measurements for each year of study. Because SNOTEL sites are biased in their representation of the surrounding basin or larger grid scales others have concluded that assimilation of data from the sites may decreases the accuracy of a landscape scale distributed mass and energy balance snow model [*Meromy et al.*, 2012].

1.2. Remote Sensing of Snow

There currently exist numerous remote sensing tools used to observe snow. Many of these are sensors such as MODIS, Thematic Mapper, and AVHRR, which are placed on airborne satellites. Products from these sensors often have adequate levels of temporal and spatial resolution, but only provide information as to the spatial extent of snow cover or in the case of MODIS the fractional snow cover within a sensor-viewing pixel. What these sensors fail to resolve is information about either snow depth or SWE on the landscape. Other methods using ground based or airborne Lidar, airborne passive microwave radiation, and airborne gamma radiation provide information about SWE on the landscape, however these technologies either don't work well beyond the plot scale or don't function well in the mountainous terrain of the Western US where SWE accumulates.

1.3. Modeling of SWE at the Landscape Scale

Without an adequate ground based or remote sensing network to understand the spatial distribution of SWE at the landscape scale, most efforts turn to modeling. Snow models can vary in simplicity from temperature index snowmelt models, to fully distributed mass and energy balance models of ever-increasing sophistication. Of these methods, neither is without its merits or faults. Mass and energy balance snow models are capable of providing fine temporal resolution, but require numerous inputs of meteorological and radiation data that is rarely gathered or are acquired at spatial scales that make the interpolation of variables to the landscape difficult. Even at the point scale where the collection of variables pertinent to the energy balance of a snowpack is convenient it has been shown that the energy exchange with the snow pack was under measured leading to a difficulty in closing the energy balance [*Helgason and Pomeroy*, 2012]. This is an important consideration when using an energy balance model as the inability to close the energy balance coupled with any errors associated with the parameterization and interpolation of radiation and meteorological data will be quickly compounded given the short time intervals in which calculations are performed in most energy balance models.

Conversely, temperature index snow models require a minimum of input data, and parameterization, while temperature, a driving input in these models is easily spatially interpolated relative to the many variables required by more sophisticated snow modeling methods. Temperature index models require a degree day factor or temperature snowmelt

coefficient, which is often empirically derived and dictates the temporal resolution of these models. Temporal resolution for temperature index models ranges from daily to hourly time steps but is usually not performed at sub hourly time steps, as model performance decreases with increases in temporal resolution [*Hock*, 2003]. The accuracy of temperature index coefficients affects model performance with the potential for the model to over or under predict snowmelt depending on the modeling time step. This is especially true if the modeling time period has snowmelt events in which energy exchange with the snow pack is not adequately represented by temperature such as rain on snow events. Subsequently, temperature index model performance is usually greatest when model coefficients are site specific and modeling is implemented over multiple model time steps.

At the catchment scale, temperature index model performance is similar to, or higher than energy balance models when used to simulate measured catchment run-off [*Hock*, 2003; *Zappa et al.*, 2003; *Debele et al.*, 2009]. However, others have shown that temperature index models can be sensitive to model coefficients, and may have poor performance modeling the spatio-temporal distribution of SWE in a small well instrumented catchment relative to an energy balance model [*Kumar et al.*, 2013]. Others have concluded that the sophistication of an energy balance model does not outweigh the results of a temperature index model because of the difficulty in acquiring and distributing meteorological and radiation variables required by energy balance models to the landscape scale [*Rango and Martinec*, 1996]. Because of these difficulties temperature index models are still widely found in current literature [*Molotch et al.*, 2004; *Molotch and Bales*, 2005; *Durand et al.*, 2008; *Molotch and Margulis*, 2008; *Gillan et al.*, 2010; *Rice et al.*, 2011].

A common method for modeling SWE at greater than the point scale is to use a reconstruction model, where daily snowmelt is determined from either an energy balance [e.g., *Cline et al.*, 1998] or temperature index model [e.g., *Martinec and Rango*, 1981], and daily snowmelt is summed over the course of the melt season. Because of their simplicity, temperature index models are frequently used as reconstruction SWE models for either the catchment or landscape scale [e.g., *Martinec and Rango*, 1981; *Durand et al.*, 2008; *Molotch and Margulis*, 2008; *Gillan et al.*, 2010; *Rice et al.*, 2011]. In instances where there is a lack of storm systems during the melt season, the modeled, integrated daily snowmelt of a reconstruction model can be considered a proxy for peak SWE.

Crucial to the implementation of a reconstruction model is knowledge of the spatial extent of snow cover which provides information as to when pixels within the model domain are snow free. As a result much of the previous research modeling SWE at the landscape scale using reconstruction models has focused on the effectiveness of these models when used with different available remotely sensed snow cover products [e.g., *Durand et al.*, 2008; *Molotch and Margulis*, 2008]. Efforts to compare reconstruction SWE model results with measured SWE that is interpolated using regression trees or with point scale SNOTEL SWE measurements, have produced maximum relative error of 55%-60% [*Durand et al.*, 2008], mean absolute error of 23% [*Molotch*, 2009] and mean relative difference 15%-18% [*Rice et al.*, 2011]. These comparisons show that the results of SWE modeling using a temperature index based reconstruction method are not unreasonable and reconstruction modeling is an effective method to model SWE at the landscape scale.

1.4. Current Understanding of the Spatial Distribution of SWE

Independent efforts to measure and model SWE at the catchment and landscape scale, have found that solar radiation [*Elder et al.*, 1991], slope [*Kerr et al.*, 2013], elevation [*Sexstone and Fassnacht*, 2013] and wind sheltering [*Erickson et al.*, 2005] were the primary controls on the spatial distribution of SWE for their respective study areas. The transferability of these physical processes and physiographic elements which control the spatial distribution of SWE to different study areas is still unclear and in many instances has not been studied.

Recent research indicates that a large percentage of total SWE occurs at high elevations and in most instances this high elevation SWE is poorly represented by our current systems of monitoring [*Gillan et al.*, 2010; *Rice et al.*, 2011]. When SWE is averaged along elevation intervals the gradients of SWE at higher elevations show variability between models used for the same study area [*Rittger et al.*, 2011], and inter-annual variability within a single study area [*Gillan et al.*, 2010]. The difference between the findings for the controls on the spatial distribution of SWE and the variability of higher elevation SWE gradients points to a need for further research into the distribution of SWE at a range of scales and global locations. Based on the findings of others we hypothesized that we would find a large percentage of total SWE at the higher elevations within our study site and the elevation gradients of this SWE would display greater variability relative to the SWE elevation gradients found at lower elevations. Further, our

study area is comprised of a series of west-east trending drainages of similar size with dominant north and south facing aspects. The nature of our study area allows us to test the control of aspect on the spatial distribution of snow, and to see if trends is SWE distribution found at for the entire study area are repeated within individual drainages within the study area.

This research seeks to further understand the spatial distribution of SWE at the landscape scale for a region of Western Montana. Rather than relying upon the output of a single model, we take an ensemble approach, using three different snow models. Our efforts are aimed at three objectives: 1) Characterize the differences and agreement between the three models used in this research, and combine these results into an ensemble model output; 2) Characterize the spatial distribution of SWE and the processes and physiographic elements, such as aspect and elevation, that control this distribution within our study area; 3) Determine the SWE elevation gradient at higher elevations within our study area where SWE measurement is under-represented.

2. Methods

2.1. Study Area and Model Domain

The study area encompasses 1200 km^2 within the Bitterroot Mountain Range in West Central Montana (Figure 1). This portion of the Bitterroot Range is unique in that all of its watersheds trend in a west-east direction. Subsequently, the valley sidewalls of these drainages are largely north and south facing with over 70% of total surface area in the study area being comprised of north and south facing aspects (Figure 2).

A 30 m X 30 m resolution US Geological Survey (USGS), digital elevation model (DEM) at the spatial extent of the model domain (Figure 1), was used to represent elevation and calculate slope and aspect within the study area. This DEM was resampled to a resolution of 500 m X 500 m, and 1 km X 1 km to create two separate DEMs for modeling that match the spatial resolution of MODIS snow cover products and SNODAS model output used in this research. To isolate the study area and the west-east trending watersheds within it, we used the ESRI ArcGIS Watershed Delineation Tool with the model DEMs. Pour points for these watershed delineations were placed in the lowest elevation pixel at the location where watershed sidewalls meet the Bitterroot valley floor.

Use of MODIS snow cover products and SWE volume calculations requires knowing the surface area for pixels within each model domain. Rather than simply use the surface area

resulting from the spatial resolution of pixels in each model, we created a new surface area field that accounts for the slope of each pixel.

2.2. Models

Three separate models are used in this research to provide for inter-model comparison, and to produce a landscape scale SWE product from an ensemble of these models. These models differ from each other in how each model represents the physics of snow accumulation and melt. Because of this, the models differ in their temporal estimation of SWE. Two of the models used are SWE reconstruction models, while the third model is a forward SWE model (Figure 3). This is an important distinction as a forward model simulates SWE depth at a given time interval from the beginning of the accumulation season through the end of the melt season. In contrast, a reconstruction model integrates SWE melt over the course of the melt season. This integration results in an approximation of peak SWE, plus any positive contributions to SWE accumulation during the melt season. Herein, we refer to modeled peak SWE plus any positive contributions to SWE accumulation as PSWEP.

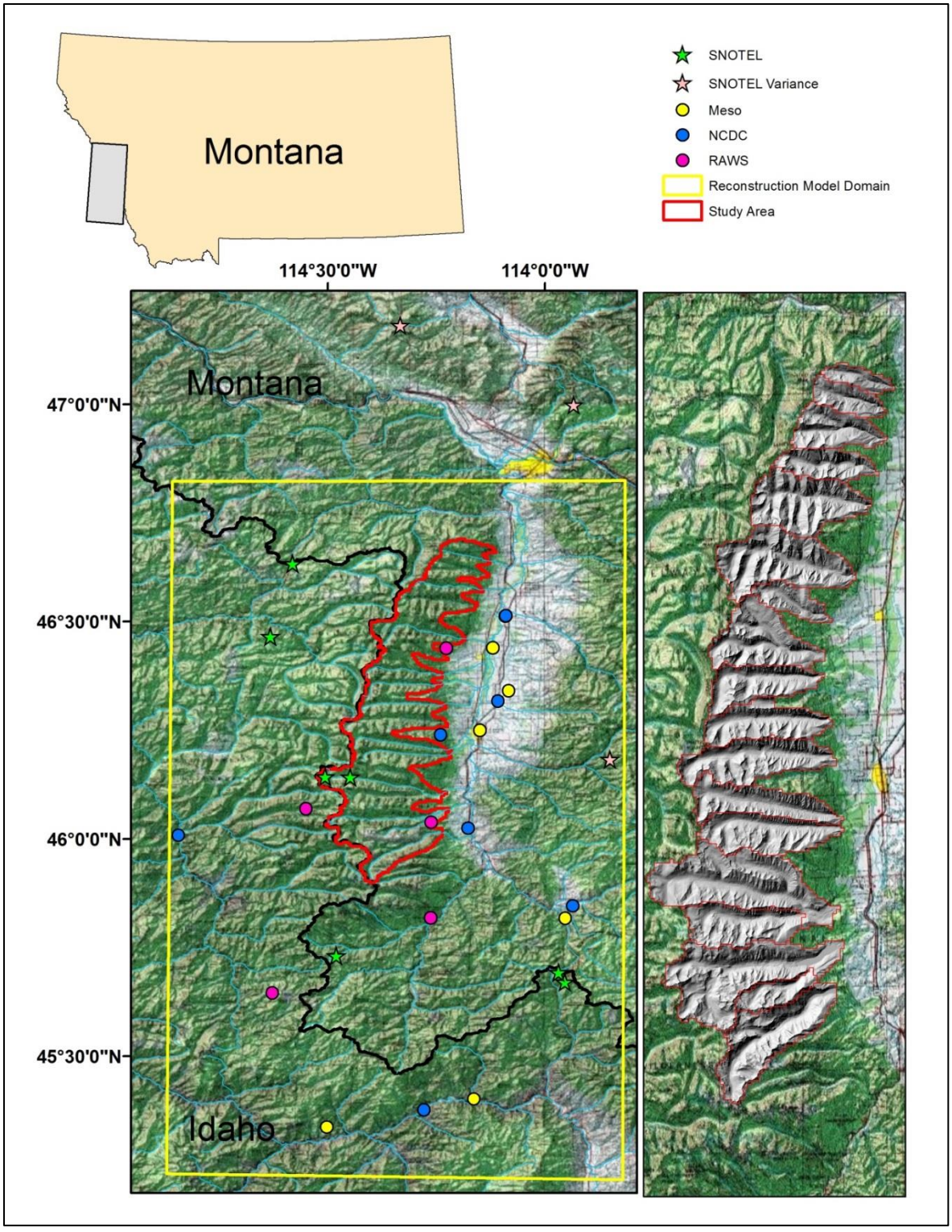

Figure 1: Study area located in the Bitterroot Range of Montana. Right hand panel is a hillshade showing the west-east trending nature of the drainages within the study area.

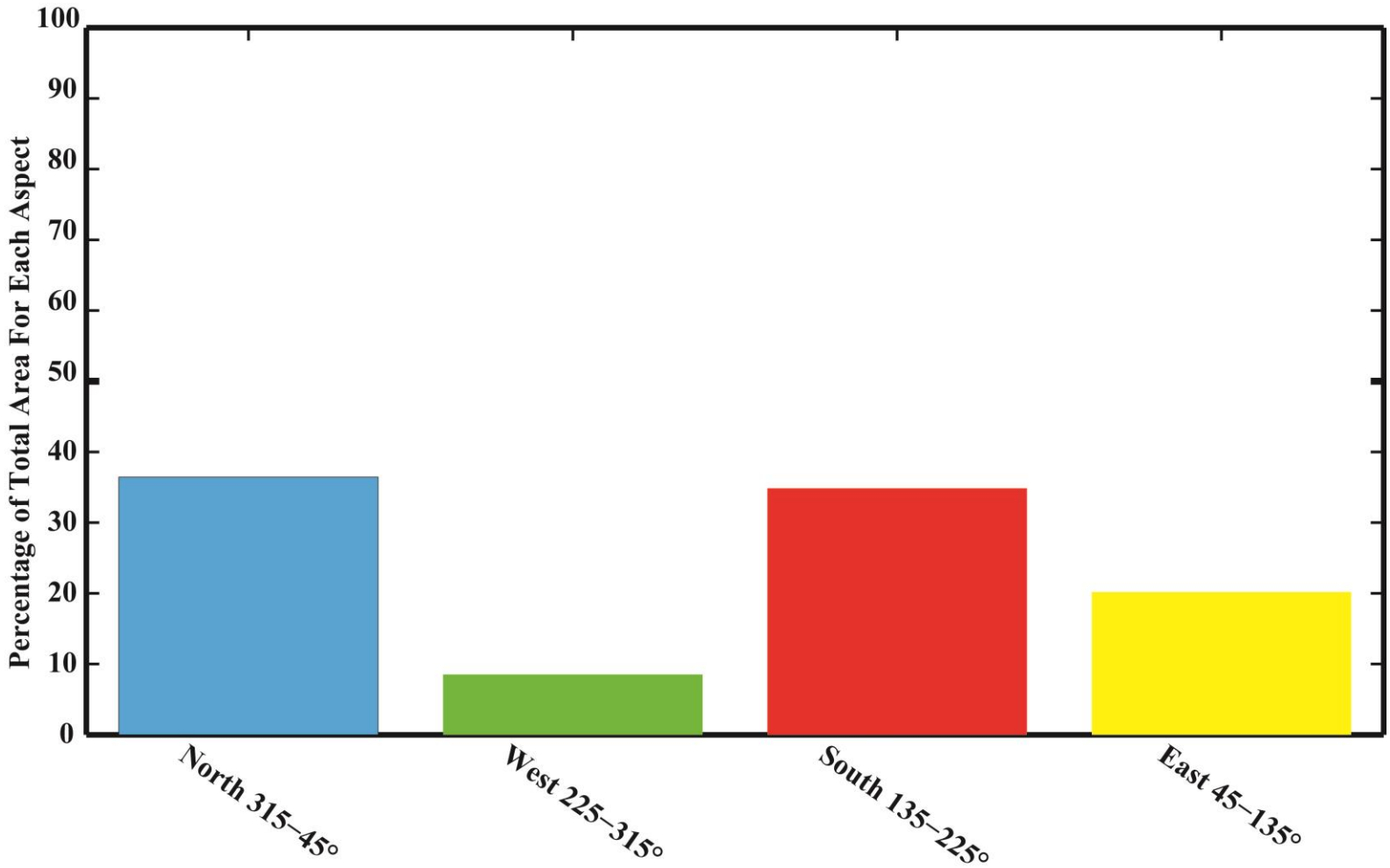

Figure 2: Percentage of area by aspect relative to total area with the study site.

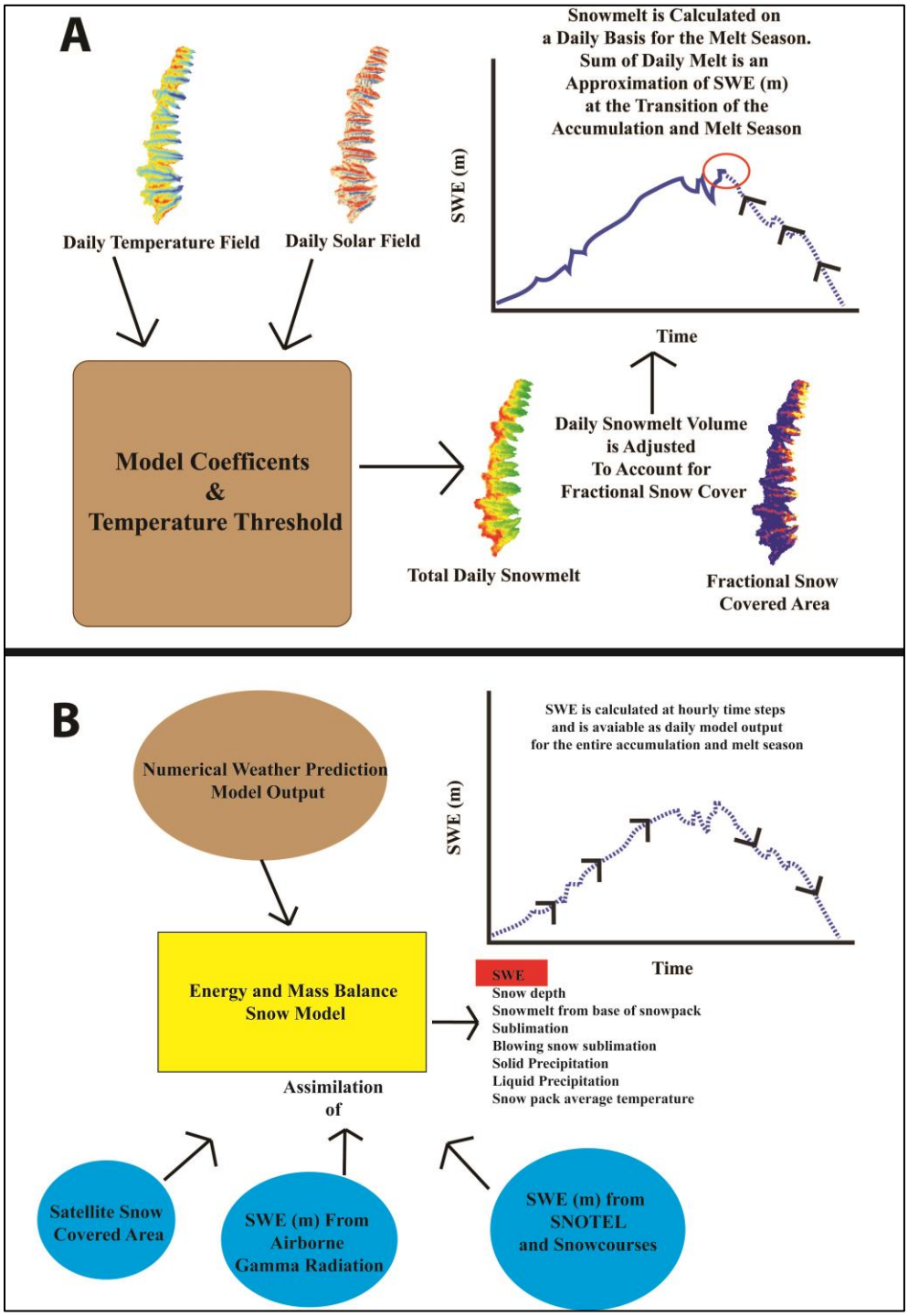

Figure 3: Conceptualization of reconstruction and forward snow models. Panel A represents a snowmelt reconstruction model and panel b represents a forward snow model.

2.3. MODIS Snow Cover Products

Using a SWE reconstruction model requires information about when pixels within the model domain are snow covered. Typically, this is determined from a remotely sensed image or snow cover product. Using these data will often dictate the spatial resolution of the model. Here we use the Moderate Resolution Imaging Direction Spectroradiometer (MODIS), daily fractional snow covered area (FSCA) and eight day snow cover products [*Hall et al.,* 2006a, *Hall et al.,* 2000b]. We downloaded these MODIS Collection 5 products, produced from the Terra satellite for the years 2000-2010. Snow cover products were left in their original resolution of 500 m X 500 m, however, we used the USGS MODIS Reprojection Tool, to mosaic, clip and reproject these data from their native sinusoidal projection to UTM Zone 11 projection.

2.3.1. MODIS Cloud Fill - Because the MODIS fractional snow covered area (FSCA) product is remotely sensed by airborne satellites, cloud cover during the time in which MODIS images are gathered presents a problem for use with a SWE reconstruction model. An algorithm was developed to fill in daily MODIS FSCA pixels obscured by cloud cover within the study area. The first step of this algorithm is to determine the percentage of cloud obscured pixels within the study area. When the percentage of pixels in the study area obscured by cloud cover is greater than 90% (Table 1), the next previous FSCA image was used in place of that FSCA image.

In addition to the daily FSCA product, MODIS produces an eight day, maximum snow cover extent product. This is a binary product where pixels that have had no observed snow during an eight day period receive a zero value and pixels which have observed snow receive a value of one. From this eight day snow cover product a daily no snow mask was created and was applied to the daily FSCA pixels within the study area.

The eight day maximum snow cover extent was also used to create daily inputs to the cloud fill algorithm that are related to the change in snow covered area with respect to elevation. To create these daily inputs the maximum snow cover extent was multiplied by the slope corrected surface area of each pixel in the model domain and then binned at 25 meter elevation intervals. For each 25 meter elevation bin, the area of snow covered pixels was divided by the total area within the elevation bin. This percentage of snow covered area was then plotted versus

its corresponding elevation bin. From these plots, cut off points were selected that represent low and high elevations where the change in the percentage of snow covered area with elevation is essentially constant (Figure 4). In some instances high and low cut off points did not exist and none were selected. Two elevation gradient points were selected from these plots. Between these elevation gradient points snow covered area is increasing with elevation within the model domain (Figure 4).

These input points where then used within the study area to fill in any pixels that remained cloud obscured after the application of the no snow mask. The mean FSCA was calculated for pixels within the study area that are less than or equal to the low cut off point. The calculated mean FSCA value was then applied to all cloud obscured pixels within the study area that are less than or equal to this low elevation cut off point. The same method was applied to the cloud obscured pixels with elevations greater than or equal to the high cut off point. If no cut off points were selected, all cloud obscured pixels within the study area were filled in with the following method.

The two elevation gradient points selected are used to determine within the study area the change in FSCA with respect to elevation. Between these two gradient points, mean FSCA was calculated at 25 meter elevation interval within the study area. A linear regression of mean FSCA and elevation between these two elevations was performed. The slope derived from this linear regression was used to fill any remaining cloud obscured pixels within the study area using an inverse distance weighted spatial interpolation based on the Linear Lapse Rate Adjustment (LLRA) method [*Dodson and Marks,* 1997].

2.3.2. Missing MODIS FSCA Images - Occasionally, MODIS FSCA products have missing days (Table 1) because of sensor errors and other problems. To create a FSCA image for missing days we used the difference in FSCA between the next available and previous available FSCA image. This difference in FSCA was divided the number of missing days and added to the day previous the missing FSCA image to create a new FSCA image. This creates linearly increasing or decreasing FSCA for the day before the missing image the newly created image and the image that occurs after the newly created image.

Table 1: Number of missing daily MODIS FSCA images and number of daily MODIS FSCA images with 90% or more of pixels obscured by cloud cover.

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Missing MODIS Images	9	18	12	0	1	0	0	0	1	0	0
MODIS Images Cloud Cover > 90%	39	65	65	64	62	72	67	66	68	64	59

2.4. Ground Based SWE Reconstruction Model

A Ground Based SWE reconstruction model (GB-reconstruction model) was developed using measured surface air temperature, calculated potential clear sky solar radiation and the cloud filled MODIS fractional snow covered area product. This SWE reconstruction model is a temperature index model that has been modified to index both temperature and solar radiation to calculate daily snowmelt from a snowpack. This modified temperature index model is similar to those used by others [*Molotch et al.*, 2004; *Durand et al.*, 2008; *Gillan et al.*, 2010].

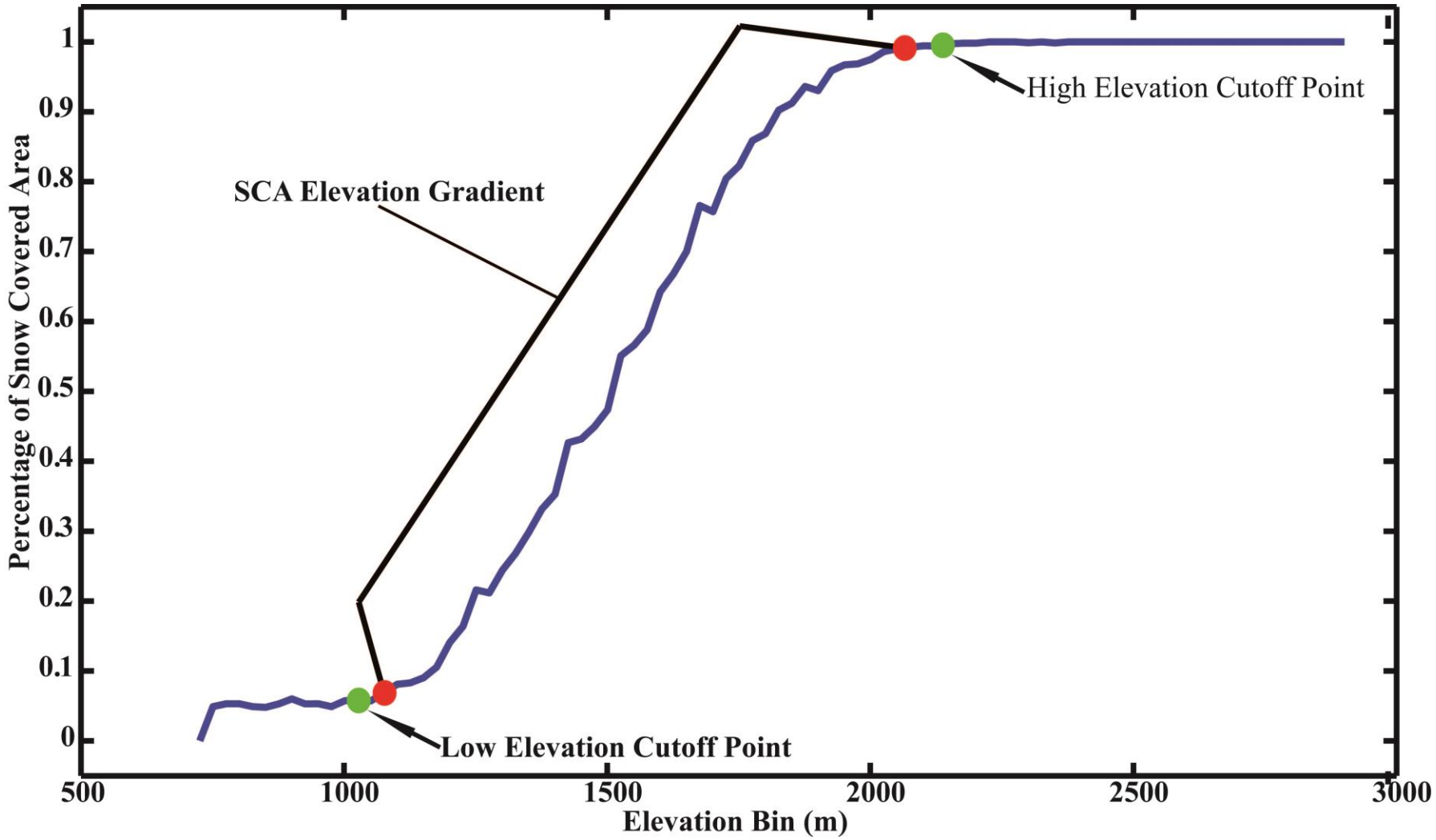

Figure 4: Example of high and low cutoff point and elevation gradient points used in MODIS cloud fill algorithm. Green dots represent high and low cutoffs. Red dots represent elevations for calculation of FSCA gradients.

Table 2: List of meteorological stations used in creation of GB-reconstruction model daily temperature fields, GB and WRF-reconstruction model temperature and solar radiation coefficients and WRF model average temperature Bayesian optimal interpolation.
*** Indicates stations used for creation of reconstruction model coefficients. † Indicates stations used in Bayesian optimal interpolation.**

Station Name	Station Type	Station Elevation (m)	Latitude	Longitude
Bitterroot Bell	Meso	1015	46°26'00"	114°07'00"
Corvallis	Meso	1096	46°19'00"	114°04'00"
Daly Creek†	SNOTEL	1762	46°11'00"	113°51'00"
Darby	Meso	1182	46°01'00"	114°10'00"
Hamilton	Meso	1089	46°15'00"	114°09'00"
Hamilton	NCDC	1076	46°14'00"	114°14'00"
Hells Half Acre	RAWS	2468	45°14'00"	144°37'00"
Indianola	Meso	1075	45°24'02"	114°09'47"
Lolo Pass*†	SNOTEL	1597	46°38'00"	114°35'00'
Little Rock Creek	RAWS	1678	46°02'00"	114°15'00"
Lolo Pass	SNOTEL	1597	46°38'00"	114°35'00'
Moose Creek*	SNOTEL	1889	45°40'00"	113°57'00"
Nez Perce*	SNOTEL	1722	45°44'00"	114°29'00"
Saddle Mountain*	SNOTEL	2420	45°42'00"	113°58'00"
Savage Pass†	SNOTEL	1880	46°28'00"	114°38'00"
Selway Lodge	NCDC	786	46°00'00"	114°50'00"
Shoup	NCDC	1036	45°22'00"	114°16'00"
Skalkaho†	SNOTEL	2210	46°15'00"	113°46'00"
Skull Gulch	Meso	1791	45°20'00"	114°30'00"
Sleeping Woman†	SNOTEL	1875	47°20'00"	114°20'00"
Smith Creek	RAWS	1722	46°26'00"	114°13'00"
Stevensville	NCDC	1028	45°54'00"	113°44'00"
Stuart Peak†	SNOTEL	2256	47°00'00"	113°56'00"
Sula	Meso	1392	45°49'00"	113°57'00"
Sula 3	NCDC	1364	45°50'00"	113°56'00"
Twelvemile Crk* †	SNOTEL	1706	45°49'00"	113°57'00"
Twin Lakes*†	SNOTEL	1950	46°09'00"	114°30'00"
West Fork	RAWS	1584	45°49'00"	114°15'00"
Western Ag. Ctr	NCDC	1096	46°19'00"	114°06'00"

2.4.1. GB-Reconstruction Model: Temperature Interpolation - Measured average daily surface air temperature from 25 different meteorological stations (Table 2) from within and outside of the study area (Figure 1), were used to create daily temperature fields for the years 2000-2010. Temperature data that was error flagged by providers was removed from the temperature record. A daily temperature lapse rate for the model domain was calculated through linear regression of station elevation and average surface temperature. Temperature was interpolated across the study area using average surface temperature and temperature lapse rates to create daily average surface temperature fields using an inverse distance weighted interpolation with the LLRA method [*Dodson and Marks*, 1997]. The LLRA method uses the calculated lapse rate and station elevation to change average surface temperature to its sea level equivalent. This removes topographic bias from the temperature before interpolation. Temperature is then interpolated to each pixel across the model domain using inverse distance weighting and returned to its correct spatial elevation using the calculated daily lapse rate and elevation of each pixel from the model DEM.

2.4.2. GB-Reconstruction Model: Potential Clear Sky Solar Radiation -To create a daily solar radiation index we calculated clear sky potential solar radiation [*Hock*, 1999] at the beginning of each hour for all hours of the day. Calculation of clear sky potential solar radiation incorporates slope and aspect of pixels, while elevations from the model DEM were used to find shadowing of pixels within the study area during hourly radiation calculations. Pixels that were determined to be shadowed during the time of radiation calculation received zero values. We then summed these hourly radiation values to create a bulk daily radiation value.

2.4.3. GB-Reconstruction Model: Coefficients and Critical Values - Our modified temperature index model uses average daily surface air temperature, bulk daily potential clear sky solar radiation and coefficients for these two variables, while incorporating FSCA to estimate daily SWE melt. Daily SWE melt was calculated for each pixel within the study area at a daily time step.

$$M = \left\{\left(\left((T * \alpha) + (I * \beta)\right)\right) * FSCA\right\} \qquad T \geq T_c \ \& \ FSCA \geq FSCA_C$$

$$M = 0 \qquad T < T_c \ or \ FSCA < FSCA_C$$

Where M is the daily SWE (m), T is daily average surface temperature, I is the bulk daily potential clear sky solar radiation, α is the temperature coefficient $(m\ (℃\ d\ )^{-1}\ )$ and β is the solar radiation coefficient $\left(m\ (\frac{W}{m^2}\ d)^{-1}\ \right)$. No SWE melt is produced for a pixel for that day if either the temperature or FSCA or were below their critical values, which are $T_c = 1$ ℃ and $FSCA_C = 10\%$.

Temperature and solar radiation coefficients were calculated using measured SWE at 6 SNOTEL stations (Table 2) from within and outside of the study area. To calculate the coefficients the yearly ten day period of largest SWE melt from each station was found. This was performed for the years 2000-2010. This SWE melt was paired with the corresponding recorded average surface temperature at the SNOTEL station and the calculated potential clear sky solar radiation for the pixel in the model domain in which the SNOTEL station occurs. Any instances where measured SWE depth increased, or the measured average surface temperature was below 1 °C were removed from this record. A multiple linear regression was performed on these data. The coefficients resulting from this regression are $0.0019\ m\ \ (℃\ d\ )^{-1}$ for α and $1.3\ x\ 10^{-6}\ m\ (\frac{W}{m^2}\ d)^{-1}$ for β.

2.4.4. GB-Reconstruction Model Modeling Time Period and Melt Season - Modeling was performed for the time period of 2000-2010. For each year the model was run during the melt season which was established as day 60 -200 of the year, except in the year 2000. In the year 2000 the model was run from day 70-200, as no suitable MODIS FSCA image was available until day 70 of that year. Our melt season time period of day 60-200 of the year is similar to that used by others [*Molotch*, 2009]. SWE melt for each day during the period of 60-200 was summed to produce modeled annual PSWEP.

2.5. WRF SWE Reconstruction Model

A second SWE reconstruction model was developed that is similar to the GB-reconstruction model, but uses the output of the Weather and Research Forecasting (WRF) model to force SWE melt. WRF is a regional climate model that dynamically downscales the output of a global climate model (GCM). The Advanced Research WRF model version 3.2 [*Skamarock and Klemp,* 2008], was used for GCM downscaling. The GCM output that was downscaled for this research is the National Center for Environmental Prediction Global

Forecasting System Final (NCEP GFS-FNL). NCEP GFS-FNL has 1.0° X 1.0° resolution, with six hour temporal resolution. WRF downscaling of NCEP GFS-FNL was performed for the time period 2000-2010. After dynamical downscaling of NCEP GFS-FNL, the final WRF output is 4 km x 4 km resolution with hourly temporal resolution.

While forcings between the WRF SWE reconstruction model (WRF-reconstruction model) and GB-reconstruction model differ, the same coefficients, critical values, model time periods, melt seasons and daily cloud filled MODIS FSCA were used for both models. Key differences in the forcings of these two models are the distribution of temperature and solar radiation affected by cloud cover. The spatial resolution of WRF model output requires further downscaling to the resolution of the reconstruction model. We used WRF pixel centroids to distribute temperature across the model domain. These pixel centroids are evenly spaced and distributed across the model domain and represent both mountainous and valley floor temperatures across the reconstruction model domain. In contrast, temperature stations used in the GB-reconstruction model are primarily located in non-mountainous terrain and often clustered near communities. Unlike the GB-reconstruction model, the WRF-reconstruction model accounts for cloud cover in its calculation of shortwave radiation by applying a cloud cover factor to the calculated potential clear sky solar radiation of the GB-reconstruction model using shortwave radiation model output from WRF.

For use in the WRF-reconstruction model, WRF surface temperature and shortwave wave solar radiation were converted to daily resolution through averaging of hourly model output. These model output were then clipped to the spatial extent of the reconstruction model domain.

WRF GCM downscaling was performed in two batches for the time periods 2000-2006 and 2007-2010. Because of a discrepancy between a georeferencing file used for these two WRF runs, the location of pixel centroids differs between the two WRF model runs. A bilinear spatial interpolation script written for the National Center for Atmospheric Research Command Language was used to align the pixel centroids of the 2000-2006 WRF output to the pixel centroids of the 2007-2010 WRF output.

2.5.1. WRF-Reconstruction Model: Surface Temperature - We observed that the WRF daily average surface temperature at higher elevations within and outside the study area were cooler than those measured at meteorological stations. This cold bias resulted in poor performance of the WRF-reconstruction model. Temperature bias is not uncommon when using

regional climate models (RCM). Bias can be introduced through either the GCM model output being used to drive the RCM, or from the RCM itself [*Gao et al.*, 2011]. To address this cold temperature bias and the poor performance of the WRF-reconstruction model we performed a Bayesian optimal interpolation [*Wikle and Berliner*, 2007] on the WRF daily average surface temperature model output. The interpolation was performed using the average surface temperature data from 10 SNOTEL stations (Table 2) located within and outside of the model domain (Figure 1) for the years 2000-2010.

After the Bayesian interpolation was performed on the WRF daily average surface temperatures, these temperatures were further downscaled to the spatial resolution of the reconstruction model at 500 m X 500 m. To further downscale the temperature to this resolution a daily temperature lapse rate was calculated through linear regression using the daily average surface temperature and elevation of each WRF pixel. The centroid of each WRF pixel was overlaid on the reconstruction model DEM, and the WRF average surface temperature was adjusted using the calculated daily lapse rate and the difference in elevation between the WRF pixel and the DEM. These adjusted temperatures, and the daily temperature lapse rates were then used to create daily temperature fields using an inverse distance weighted interpolation with the LLRA method [*Dodson and Marks*, 1997].

2.5.2. WRF-Reconstruction Model: Solar Radiation - WRF shortwave solar radiation was not downscaled and directly input into the WRF-reconstruction model; rather, it was used to apply a cloud cover factor to the potential clear sky solar radiation calculated for the GB-reconstruction model. To create this cloud cover factor the potential clear sky solar radiation was calculated for the model domain at 4 km X 4 km WRF resolution, using the same method employed by the GB-reconstruction model. Since the WRF radiation values represent the solar radiation averaged over the day, these values were multiplied by 24 to produce a daily bulk shortwave radiation value. These values and their pixel centroids were added to ESRI ArcGIS as X Y data and a nearest neighbor interpolation was performed to create a 4 km X 4 km field with the same spatial extent as the WRF-reconstruction model domain. The percentage change was calculated between the potential clear sky solar radiation and the WRF nearest neighbor interpolated shortwave solar radiation for each pixel within the model domain. This percentage change was then applied to each GB-reconstruction model radiation pixel within the study area

that was either within or intersected by the larger WRF pixel to create the radiation fields for the WRF-reconstruction model.

2.5.3. WRF-Reconstruction Model: Coefficients and Thresholds - For the WRF-reconstruction model we used the same threshold values as those used in the GB-reconstruction model, but calculated new temperature and solar radiation coefficients. Temperature and solar radiation coefficients were calculated using measured SWE at 6 SNOTEL stations (Table 2) from within and outside of the study area. To calculate the coefficients the yearly ten day period of largest SWE melt from each station was found. This was performed for the years 2000-2010. This SWE melt was paired with the corresponding WRF-reconstruction model downscaled average surface temperature and solar radiation values from the pixel in the model domain in which the SNOTEL station occurs. The coefficients resulting from this regression are $0.0017\ m\ \ (℃\ d\ )^{-1}$ for α and $2.1\ x\ 10^{-6}\ m\ (\frac{W}{m^2}\ d)^{-1}$ for β.

2.6. SNODAS Model

The SNOw Data Assimilation Systems (SNODAS) is a hydrologic product produced by the National Weather Service's, National Operational Hydrologic Remote Sensing Center (NOHRSC). This product is the output of a physics based, spatially distributed mass and energy balance snow model, which assimilates remotely sensed snow cover and ground measured SWE data [*Barrett*, 2003]. The downloadable SNODAS product is a suite of model outputs related to snow cover and SWE.

SNODAS is available at a daily temporal resolution and 1 km X 1 km spatial resolution for the contiguous United States from 2003 through present. SNODAS model output was downloaded for the years 2004-2010 [*NOHRSC*, 2004]. SWE depth was extracted from the suite of SNODAS model outputs, clipped and left in its native geographic coordinate system during analysis.

Because SNODAS is a forward SWE model, we wanted the results from SNODAS to be similar to those our reconstruction models. To accomplish this we used the daily SNODAS SWE results to determine the change in SWE for each day of day 60-200 melt seasons for each year of modeling. We summed the absolute value of negative daily SWE change over the time period of

day 60-200 of the year for each pixel in the study area, which resulted in a reconstructed SWE for the study area from the SNODAS results.

2.7. Ensemble Model

To create an ensemble PSWEP product from the three models used in this research we summed and averaged the PWSEP results from each model for each year in which model output was available for all three models. Because the reconstruction models have a different spatial resolution than the SNODAS model we upscaled the reconstruction model PSWEP values to the resolution of the SNODAS model. To upscale the reconstruction models we converted the pixels values to point using the pixel centroid. These point values were unprojected and using an inverse distance weighting interpolation were rescaled to the same spatial resolution and extent as the SNODAS model. The SNODAS model PSWEP fields had pixels which contains zero values for SWE (m) for many of the years of model output in different locations throughout the study area. Because the SNODAS model PSWEP is the result of summing multiple days of SWE change we believe these zero values to be an error in the SNODAS model output. To account for these zero values when summing and averaging the model values for the Ensemble model we ignored the SNODAS zero values and used the average of the two reconstruction models for the Ensemble model values. No weighting was applied to any of the models when averaging the values for the Ensemble model. We decided against using weighting when averaging as we lacked a full understanding of the bias or error associated with any of the models. Since we lacked this information, any weighting we applied could be either erroneous or arbitrary.

2.8. Model Uncertainty and Sensitivity

The reconstructive nature of the models used in this research, modeling for a time period prior to the start of research and the landscape scale of these models makes model validation difficult. Because of these difficulties and lacking any significant data with which to validate the models we instead try to address model uncertainty and sensitivity through several different means.

2.9. GB-Reconstruction Model Uncertainty

One source of uncertainty in the GB-reconstruction model results is the interpolation of point scale temperature. We assume that the measured point scale temperature is representative of the temperature of a 500 m X 500 m pixel within the model domain in which the station is located. At lower elevation flat sites this assumption is likely accurate, however, at higher elevation sites within the model domain which have a greater topographic variability, this assumption may introduce error into the model results.

To further understand the effect of this assumption on our modeling results we calculated the standard deviation of elevation within each 500 m X 500 m grid cell in which the meteorological station was located using a 30 m x 30 m DEM. We adjusted the measured average temperature for each meteorological station using the station elevation plus or minus its corresponding standard deviation of elevation and the calculated daily temperature lapse rates for the model domain. These adjusted temperatures were then interpolated across the model domain using the LLRA method. We re-ran the GB-reconstruction model using these standard deviation temperature fields, while using the same MODIS and solar radiation fields and model coefficients.

2.9.1. GB-Reconstruction Model Jackknifing

- Uncertainty in the GB-Reconstruction model results can also be introduced through the inclusion or exclusion of meteorological stations. The stations likely to have the largest impact on reconstruction model results are SNOTEL stations as these have the most complete temperature record and are used in the calculation of model coefficients. We used a process of jackknifing where we removed SNOTEL temperatures used in temperature field creation, and the calculation of model coefficients as another way to understand GB-reconstruction model uncertainty.

2.10. Reconstruction Model Sensitivity

Sensitivity of the reconstruction models to changes in the model coefficients was determined by randomly selecting 50 coefficient values from two arrays of 100 temperature and solar radiation coefficient values. Temperature and solar radiation coefficient values were incrementally increased and decreased by 1% intervals up to 50% plus or minus of the original

coefficient values. The reconstruction models were then re-run using these randomly selected 50 different coefficient values.

2.11. SNODAS Model Sensitivity and Uncertainty

SNODAS Model sensitivity and uncertainty information is not provided by NOHRSC. However, recent comparison of measured mean SWE (m) and modeled SWE (m) from SNODAS for sites in Colorado found root mean square error (RMSE) between 5 -12 (cm) [*Clow et al.,* 2012]. The range in these RMSE values is driven by location and land cover type. These RSME values were highest in open non-forested sites while forest covered areas had the lowest RSME values.

These validation results while encouraging are not without their difficulties. *Clow et al.,* [2012] sampled 45 different areas which represented SNODAS model pixels and the selection of these sites was not random and they avoided pixels which had slopes greater than 30°. Further, the points sampled within each grid cell were not random, and 33 point scale SWE measurements were averaged to represent the mean SWE of that pixel to compare with SNODAS. There is uncertainty in how representative the sample of mean point scale measurements is of the surrounding 1 km X 1 km grid cell, along with the sample of 45 different SNODAS grid cells with slopes of less than 30°, in terms of overall SNODAS performance. Because of this uncertainty, the RMSE values reported by *Clow et al.,* [2012], have the potential to be higher or lower than what is reported.

The assimilation of ground based data by the SNODAS model allows for improvement in model results by computing differences between model results and the ground measurement and using these differences to rerun the model and improve SWE estimates [*Barrett,* 2003]. Presumably, where there are more observation stations estimates of SWE will improve. Within our study area which encompasses over 1200 km^2 there are only two SNOTEL stations of which SNODAS would be ingesting data from.

3. Results

3.1. Model Performance

A benchmark used in the assessment of modeling results is comparison between model output and point scale SWE recorded at Natural Resources Conservation Service SNOTEL

stations. Point scale SWE measurements are highly variable in space because of accumulation an melt processes, wind redistribution, interactions with vegetation, and topography [*Elder et al.*, 1991; *Deems*, 2006]. High resolution Lidar measurement of SWE in a catchment of less than 1 km^2 found SWE (m) to vary between 0- 9 (m) [*Grünewald et al.*, 2010]. Because of this variability, point scale SWE measurement may not be representative of larger grid scales [*Molotch and Bales*, 2005; *Rice and Bales*, 2010]. Consequently, here we use point scale SWE (m) measurement from SNOTEL stations to provide a first order model assessment, as opposed to a full model validation. Two SNOTEL stations, Twin Lakes at 1950 m elevation and Twelvemile Creek at 1706 m elevation are located within our study area. We compare to model output the absolute value of negative daily SWE change over the time period of day 60-200 from each of these SNOTEL stations.

All three models capture the annual variability in SNOTEL PSWEP for the modeling time period at both stations (Figure 5). The SNODAS pixel SWE (m) values are both greater and less than the point scale SWE (m) at the higher elevation station, with values ranging from 70%-225% of the SNOTEL value (Table 3). For every year of modeling the reconstruction models have pixel SWE (m) values that are less than the point scale SNOTEL SWE (m) value at the higher elevation station. The largest reconstruction model pixel SWE value is 83% of the point scale SNOTEL value at the higher elevation station. With the exception of a few years, the pixel SWE (m) values are greater than the point scale SNOTEL value at the lower elevation station. At this SNOTEL station SWE (m) values for all three models range from 86%-200% of the point scale SWE (m) measurement.

The years 2005 and 2009 produce the smallest and largest SWE (m^3) for the GB-reconstruction model, respectively (Table 4). The lowest and highest percent of average snowpack for the Bitterroot Basin listed by the NRCS Montana State Basin Outlook Reports for March, April and May 1st for the years 2000-2010, are 2005 and 2008 respectively. NRCS basin reports are available at http://www.wcc.nrcs.usda.gov/. The year of highest SWE (m^3) for both reconstruction models is 2009 and ranks 5th 4th and 3rd for March, April and May 1st for the NCRS reports. The Bitterroot Basin in these NRCS reports, while encompassing the study area is much larger than the study area and includes the Sapphire Mountains. For the SNODAS modeling time period of 2004-2010 the largest and smallest SWE (m^3) are in 2005 and 2008 respectively. For the WRF-reconstruction model the year of smallest SWE (m^3) is 2002. NRCS

Bitterroot Basin percentage of average snowpack for March, April and May 1st in 2002 are the 4th 5th and 7th highest values for the 11 years modeled.

3.2. Reconstruction Models: Uncertainty and Sensitivity

3.2.1. GB-Reconstruction Model Jackknifing - Through the process of jackknifing where SNOTEL stations are removed from the creation of temperature fields and model coefficient calculations, the root mean square difference (RMSD) between the model results and jackknifed model results vary from 0.00-0.13 (m) depending on the year of modeling and stations removed (Table 5). The largest RMSD values are produced when the Nez Perce station is removed. Removal of both the SNOTEL stations within the study area, Twin Lakes and Twelvemile, produces the smallest RMSD values.

Table 3: Percentage of modeled SWE (m) relative to point scale observed SWE (m) at Twin Lakes and Twelvemile Creek SNOTEL stations within the study area.

	Percentage of Modeled SWE Relative to Twin Lakes SNOTEL SWE				Percentage of Modeled SWE Relative to Twelvemile SNOTEL SWE			
Year	**GB Model**	**WRF Model**	**SNODAS Model**	**Ensemble Model**	**GB Model**	**WRF Model**	**SNODAS Model**	**Ensemble Model**
2000	74%	76%	-	-	142%	138%	-	-
2001	77%	68%	-	-	179%	165%	-	-
2002	64%	26%	-	-	157%	86%	-	-
2003	55%	49%	-	-	125%	121%	-	-
2004	81%	83%	115%	98%	140%	144%	137%	115%
2005	77%	83%	225%	127%	168%	148%	200%	153%
2006	63%	60%	155%	99%	134%	138%	193%	122%
2007	57%	48%	70%	60%	155%	139%	137%	135%
2008	65%	66%	101%	76%	86%	93%	129%	101%
2009	82%	80%	82%	85%	163%	165%	133%	122%
2010	80%	79%	86%	88%	199%	191%	121%	128%
Mean	70%	65%	119%	91%	150%	139%	150%	125%
Standard Deviation	9.4%	17%	47%	19%	29%	29%	30%	15%

Table 4: Annual SWE (m^3) for all models.

Year	GB Total SWE (m^3)	WRF Total SWE (m^3)	SNODAS Total SWE (m^3)	Ensemble Total SWE(m^3)
2000	6.0×10^8	6.8×10^8	-	-
2001	4.8×10^8	4.8×10^8	-	-
2002	7.2×10^8	3.3×10^8	-	-
2003	5.5×10^8	5.7×10^8	-	-
2004	5.7×10^8	6.2×10^8	5.9×10^8	5.8×10^8
2005	4.3×10^8	4.5×10^8	3.4×10^8	4.1×10^8
2006	6.1×10^8	6.0×10^8	8.2×10^8	6.9×10^8
2007	5.6×10^8	5.2×10^8	5.2×10^8	5.8×10^8
2008	6.9×10^8	7.3×10^8	9.0×10^8	7.7×10^8
2009	7.7×10^8	7.8×10^8	7.5×10^8	7.8×10^8
2010	5.7×10^8	5.5×10^8	4.3×10^8	5.3×10^8
Mean	6.0×10^8	5.7×10^8	6.2×10^8	6.2×10^8
Standard Deviation	9.5×10^7	1.2×10^8	1.9×10^8	1.3×10^8

Table 5: RMSD between all pixel SWE (m) values for the GB-reconstruction model and jackknifed GB-reconstruction model results.

Year	Lolo	Moose Creek	Nez Perce	Saddle Mtn	Twelvemile	Twin Lakes	Twin Lakes and Twelvemile
2000	0.08	0.09	0.09	0.02	0.06	0.06	0.00
2001	0.02	0.03	0.04	0.02	0.05	0.05	0.01
2002	0.03	0.04	0.06	0.04	0.08	0.08	0.01
2003	0.03	0.03	0.05	0.04	0.06	0.06	0.01
2004	0.04	0.05	0.06	0.02	0.07	0.06	0.01
2005	0.08	0.08	0.09	0.05	0.05	0.06	0.01
2006	0.09	0.10	0.12	0.07	0.08	0.07	0.01
2007	0.02	0.03	0.04	0.04	0.05	0.04	0.01
2008	0.10	0.11	0.13	0.05	0.07	0.05	0.01
2009	0.11	0.11	0.13	0.07	0.08	0.06	0.02
2010	0.09	0.10	0.12	0.06	0.06	0.06	0.01
Mean	6.3×10^{-2}	7×10^{-2}	8.5×10^{-2}	4.4×10^{-2}	6.5×10^{-2}	6×10^{-2}	1×10^{-2}
Standard Deviation	3.3×10^{-2}	3.2×10^{-2}	3.4×10^{-2}	1.8×10^{-2}	1.2×10^{-2}	1.0×10^{-2}	4.3×10^{-3}

Table 6: RMSD between all pixel SWE (m) values for the GB-reconstruction model and GB-reconstruction model standard deviation of elevation adjusted temperature results.

Years	GB Minus One Standard Deviation	GB Plus One Standard Deviation
2000	2.2×10^{-2}	1.2×10^{-5}
2001	1.6×10^{-2}	9.8×10^{-6}
2002	2.7×10^{-2}	1.3×10^{-5}
2003	2.2×10^{-2}	1.4×10^{-5}
2004	2.3×10^{-2}	1.3×10^{-5}
2005	1.8×10^{-2}	8.9×10^{-6}
2006	2.2×10^{-2}	1.3×10^{-5}
2006	1.6×10^{-2}	7.5×10^{-6}
2007	2.2×10^{-2}	1.3×10^{-5}
2008	2.5×10^{-2}	1.7×10^{-5}
2009	2.3×10^{-2}	1.6×10^{-5}
2010	2.2×10^{-2}	1.2×10^{-5}
Mean	2.2×10^{-2}	1.2×10^{-5}
Standard Deviation	3.2×10^{-3}	2.6×10^{-6}

3.2.2. GB-Reconstruction Model: Standard Deviation of Elevation Adjusted Temperatures - The standard deviation of elevation (SDE) adjusted temperatures of the GB-reconstruction model have a larger RMSD of individual pixel values when the temperature is adjusted by a subtracting a SDE (Table 6). Adjusting the temperatures by adding a standard deviation of elevation results in RMSD that are 3-4 orders of magnitude smaller than the RMSD when the temperatures are adjusted by subtracting a SDE.

3.2.3. Reconstruction Models: Sensitivity to Coefficients - When the sign of the percentage change in the reconstruction models' coefficients is the same, the mean percentage change in SWE varies by a similar amount as the percentage change in the coefficient values (Figure 6). The mean percentage change in SWE is reduced when the signs of the percentage change to the model coefficients are opposite of each other. A plus or minus 50% change in coefficient values of the same sign for the GB-reconstruction model produces a RMSD of individual pixel values ranging from 0.18 - 0.33 (m) depending on the year of modeling. For the WRF-reconstruction model these RMSD values range from 0.11 - 0.24 (m) depending on the year of modeling.

3.3. Inter-model Comparison

3.3.1. SWE Magnitudes and Variability - For all models, the mean of all pixel SWE (m) values is within the same order of magnitude (Table 7). The SNODAS model has the widest range of mean SWE (m) values. However, the maximum individual pixel SWE (m) value for the SNODAS model is an order of magnitude larger than the reconstruction models in 2005. In 2010, the maximum individual pixel SWE (m) value for the reconstruction models is an order of magnitude larger than the SNODAS model. The SNODAS model has the highest standard deviation of individual pixel SWE (m) values, while reconstruction models have the same value (Table 7). When individual pixel SWE (m) values are binned at 50 meter intervals, the variability of the pixel SWE (m) values within an elevation interval is greatest between the elevations of 1750-2100 m within the study area (Figure 7) for all three models. For the reconstruction models the individual interval of highest variability is 1800/1850 m. For the SNODAS model this interval is 1950/2000 m.

Table 7: Annual mean SWE (m) and standard deviation of all pixel SWE values for all models.

Year	GB Mean SWE (m)	WRF Mean SWE (m)	SNODAS Mean SWE (m)	Ensemble Mean SWE (m)	GB Standard Deviation of SWE (m)	WRF Standard Deviation of SWE (m)	SNODAS Standard Deviation of SWE (m)	Ensemble Standard Deviation Of SWE (m)
2000	0.51	0.58	-	-	0.20	0.23	-	-
2001	0.42	0.41	-	-	0.15	0.15	-	-
2002	0.62	0.28	-	-	0.18	0.11	-	-
2003	0.48	0.49	-	-	0.18	0.20	-	-
2004	0.49	0.53	0.49	0.49	0.23	0.26	0.24	0.19
2005	0.37	0.39	0.28	0.34	0.17	0.18	0.25	0.17
2006	0.53	0.52	0.68	0.57	0.20	0.21	0.35	0.23
2007	0.48	0.44	0.43	0.46	0.22	0.21	0.21	0.21
2008	0.60	0.63	0.74	0.64	0.22	0.23	0.30	0.23
2009	0.66	0.67	0.62	0.65	0.24	0.24	0.25	0.23
2010	0.49	0.48	0.35	0.44	0.18	0.17	0.19	0.17
Mean	0.51	0.49	0.47	0.51	0.20	0.20	0.25	0.20
Standard Deviation	8.2×10^{-2}	0.11	0.16	0.11	-	-	-	-

3.3.2. SWE vs. Elevation - A correlation coefficient was calculated to quantify the linear dependence of SWE (m) and elevation for all individual pixels. The correlation coefficients are greatest for the reconstruction models in 2005 (Table 8). The smallest correlation coefficients occur in 2000 and 2010 for the GB and WRF-reconstruction models respectively. In 2002 the correlation coefficient is negative for the WRF-reconstruction model. For the SNODAS model the largest and smallest correlation coefficients are in 2009 and 2004, 2005 respectively. The linear dependence of SWE and elevation observed for all individual pixel SWE (m) values is maintained when SWE and elevation are isolated to only north and south facing pixels (Table 8).

3.3.3. Mean SWE vs. Elevation - When SWE (m) is averaged over 50 meter intervals of elevation all three models show a similar pattern of linearly increasing SWE with increasing elevation at lower elevations. Between the elevations of 1900 -2600 m, SWE (m) remains relatively constant or decreases with increases in elevation (Figure 8). After 2600 m elevation the model results diverge with the reconstruction models showing a decrease in SWE with increases in elevation, and the SNODAS model having constant, decreasing or increasing SWE with increases in elevation.

3.3.4. SWE Volume Distribution - Surface area and SWE (m^3) was binned and averaged at 50 m intervals. Mean SWE (m^3) distribution is closely related to the distribution of mean surface area within the study area (Figure 9; Figure 10). For the reconstruction models the greatest mean SWE (m^3) and surface area occur at the elevation bin of 2200/2250 m, except for the WRF model in 2002. For the WRF model in 2002, the greatest mean SWE (m^3) occurs at the elevation bin of 1950/2000 m. For the SNODAS model, the elevation bin that has the greatest mean surface area occurs at 2200/2250 m. However, the elevation bin which has the greatest mean SWE (m^3) is more variable than the reconstruction models and occurs between the elevations of 1950-2250 m.

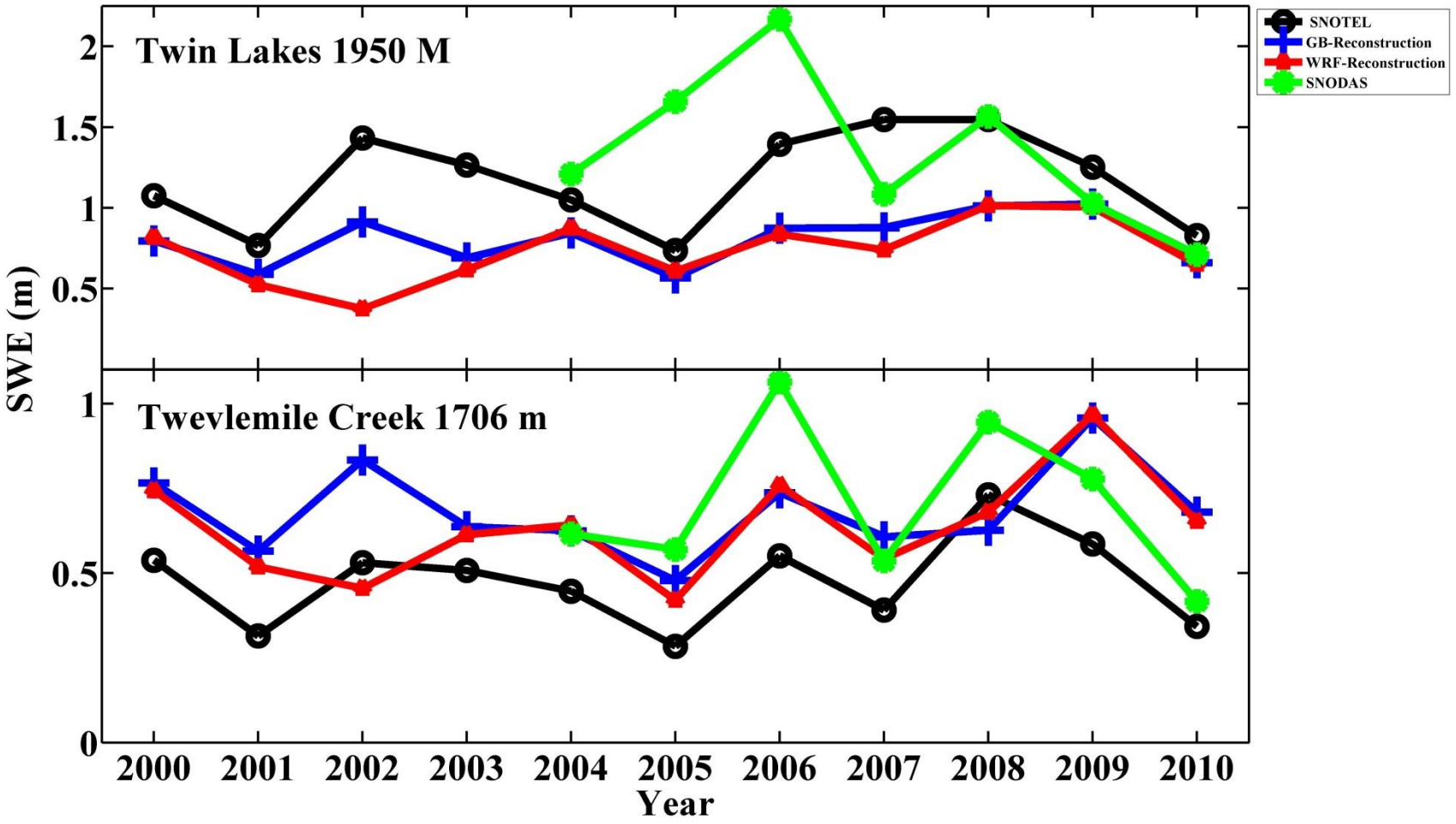

Figure 5: Comparison of point scale SWE (m) from SNOTEL stations and model pixels.

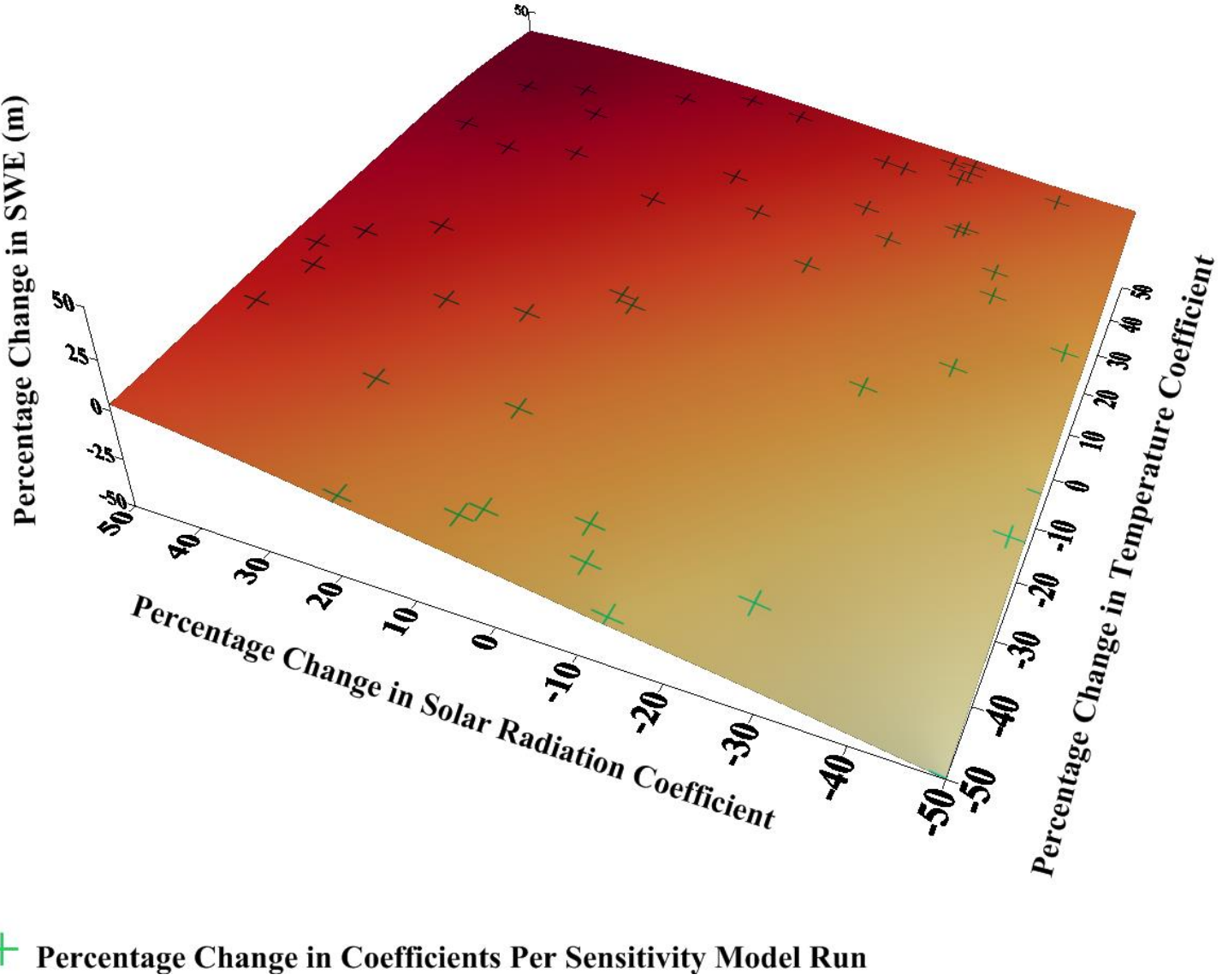

Figure 6: Sensitivity of mean SWE (m) from Reconstruction models to changes in temperature and solar radiation model coefficients

Table 8: Annual correlation coefficient of SWE and elevation for all individual model pixels and north and south facing model pixels.

Year	All GB Pixels	N Pixels	S Pixels	All WRF Pixels	N Pixels	S Pixels	All SNODAS Pixels	N Pixels	S Pixels	All Ensemble Pixels	N Pixels	S Pixels
2000	0.35	0.34	0.37	0.47	0.47	0.46	-	-	-	-	-	-
2001	0.43	0.42	0.44	0.26	0.28	0.22	-	-	-	-	-	-
2002	0.40	0.39	0.37	-0.23	-0.25	-0.27	-	-	-	-	-	-
2003	0.48	0.49	0.45	0.40	0.41	0.36	-	-	-	-	-	-
2004	0.59	0.60	0.63	0.60	0.61	0.63	0.25	0.27	0.18	0.61	0.31	0.26
2005	0.63	0.66	0.63	0.61	0.63	0.60	0.25	0.22	0.22	0.58	0.28	0.30
2006	0.45	0.51	0.37	0.31	0.35	0.21	0.29	0.25	0.20	0.47	0.33	0.24
2007	0.60	0.63	0.57	0.52	0.55	0.48	0.45	0.42	0.39	0.61	0.42	0.39
2008	0.53	0.55	0.50	0.40	0.41	0.34	0.51	0.47	0.44	0.59	0.49	0.45
2009	0.51	0.51	0.48	0.36	0.36	0.32	0.62	0.60	0.56	0.61	0.62	0.58
2010	0.39	0.42	0.37	0.17	0.20	0.12	0.55	0.53	0.49	0.54	0.53	0.49
Mean	0.49	0.50	0.47	0.35	0.37	0.32	0.42	0.40	0.36	0.57	0.42	0.39
Standard Deviation	9×10^{-2}	0.10	0.10	0.22	0.23	0.24	0.14	0.14	0.14	5.0×10^{-2}	0.12	0.12

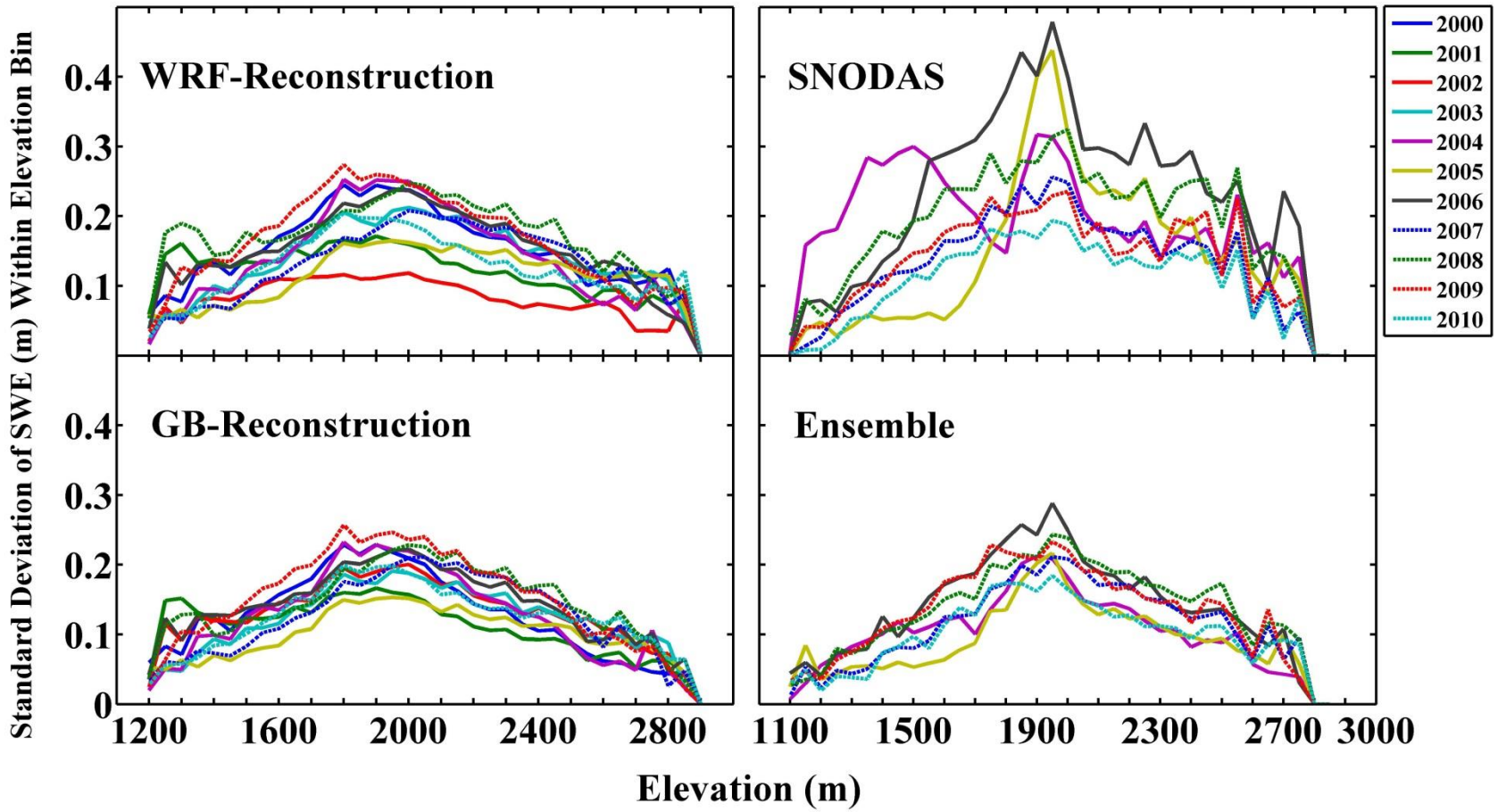

Figure 7: Standard deviation of individual pixel SWE (m) values within 50 meter elevation intervals. Elevations on X axis represent an elevation bin of that elevation plus 50 meters.

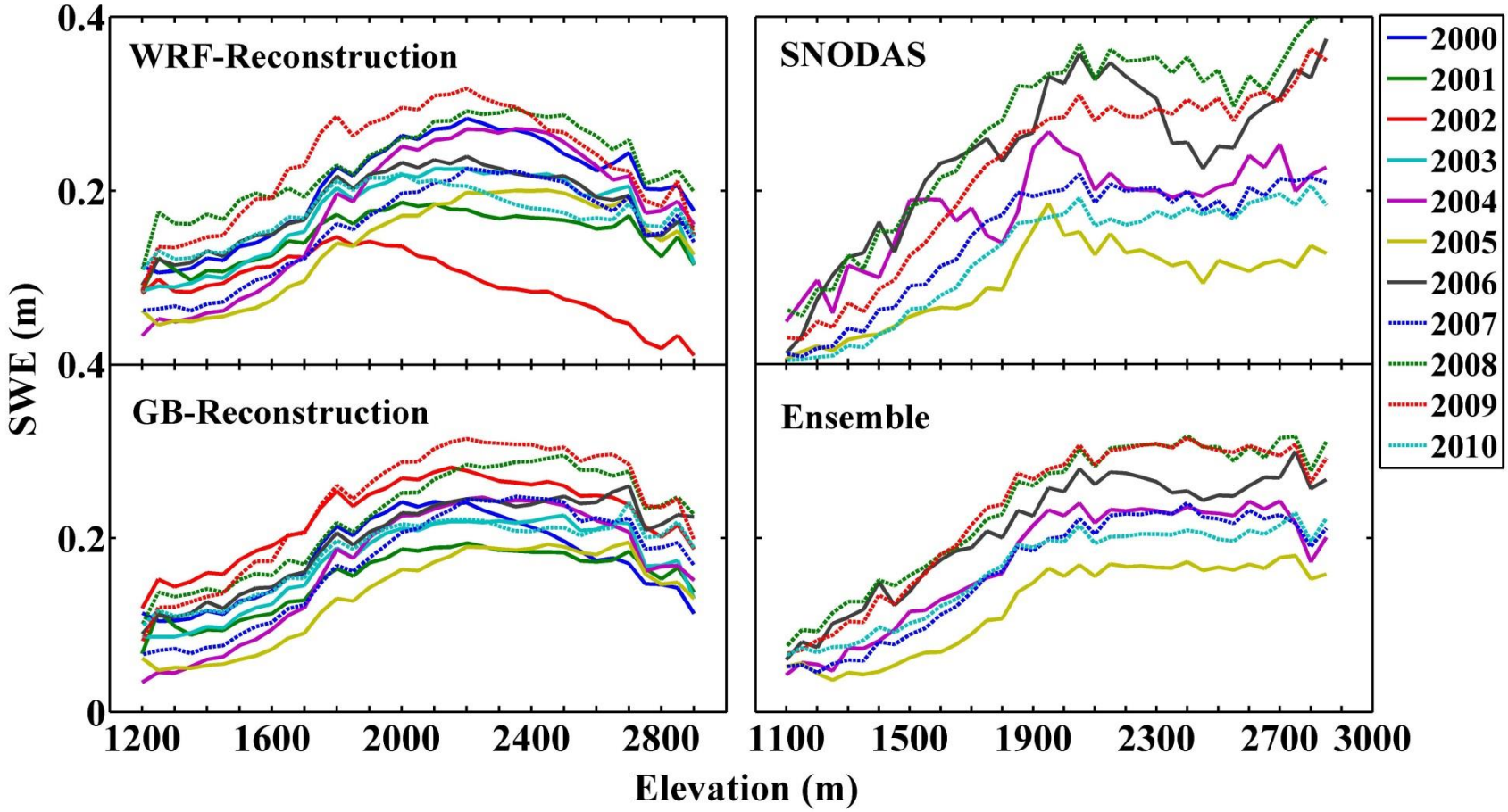

Figure 8: Mean of individual pixel SWE (m) values within 50 meter elevation intervals. Elevations on X axis represent an elevation bin of that elevation plus 50 meters.

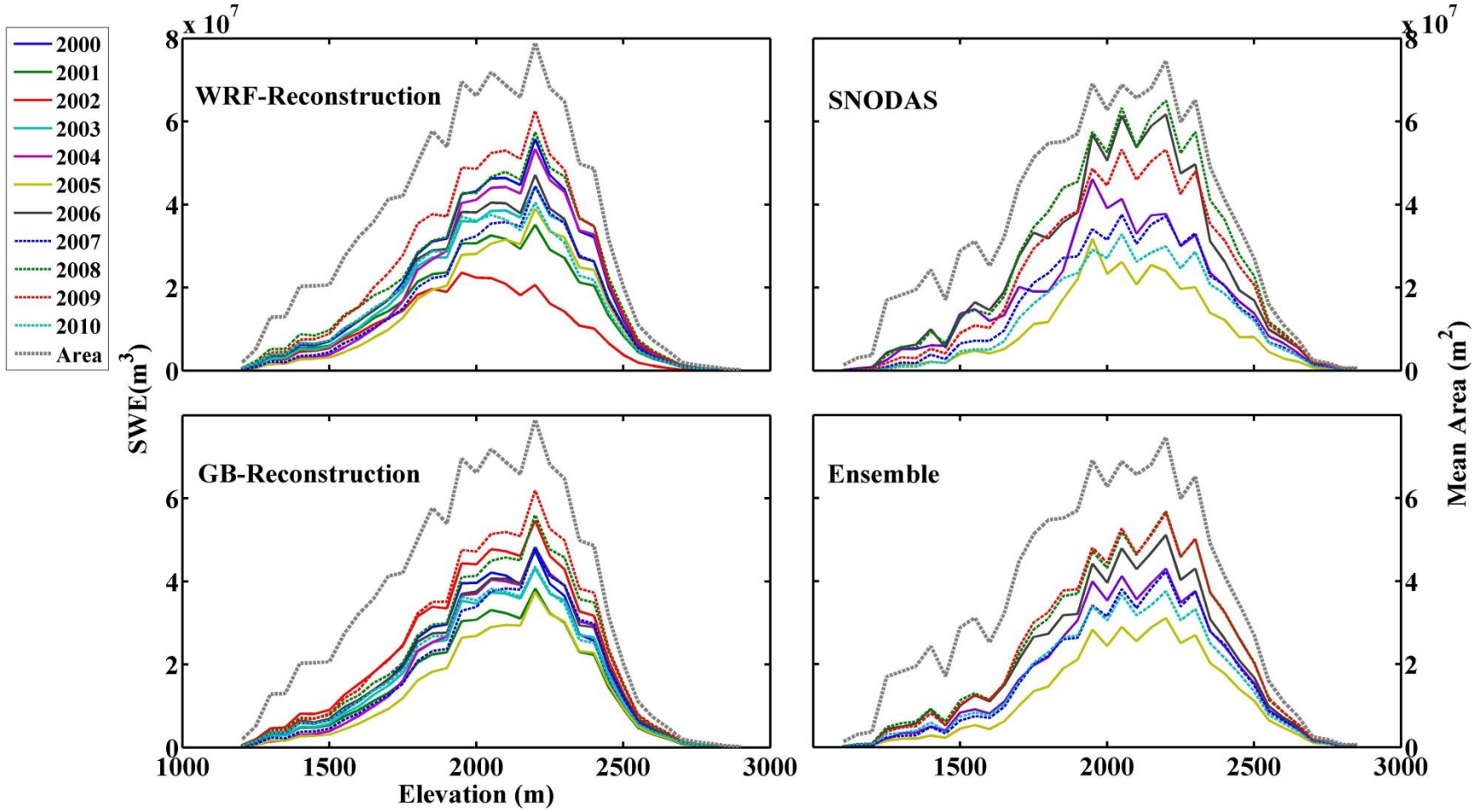

Figure 9: Mean SWE (m^3) and mean surface area within 50 meter elevation intervals. Elevations on X axis represent an elevation bin of that elevation plus 50 meters.

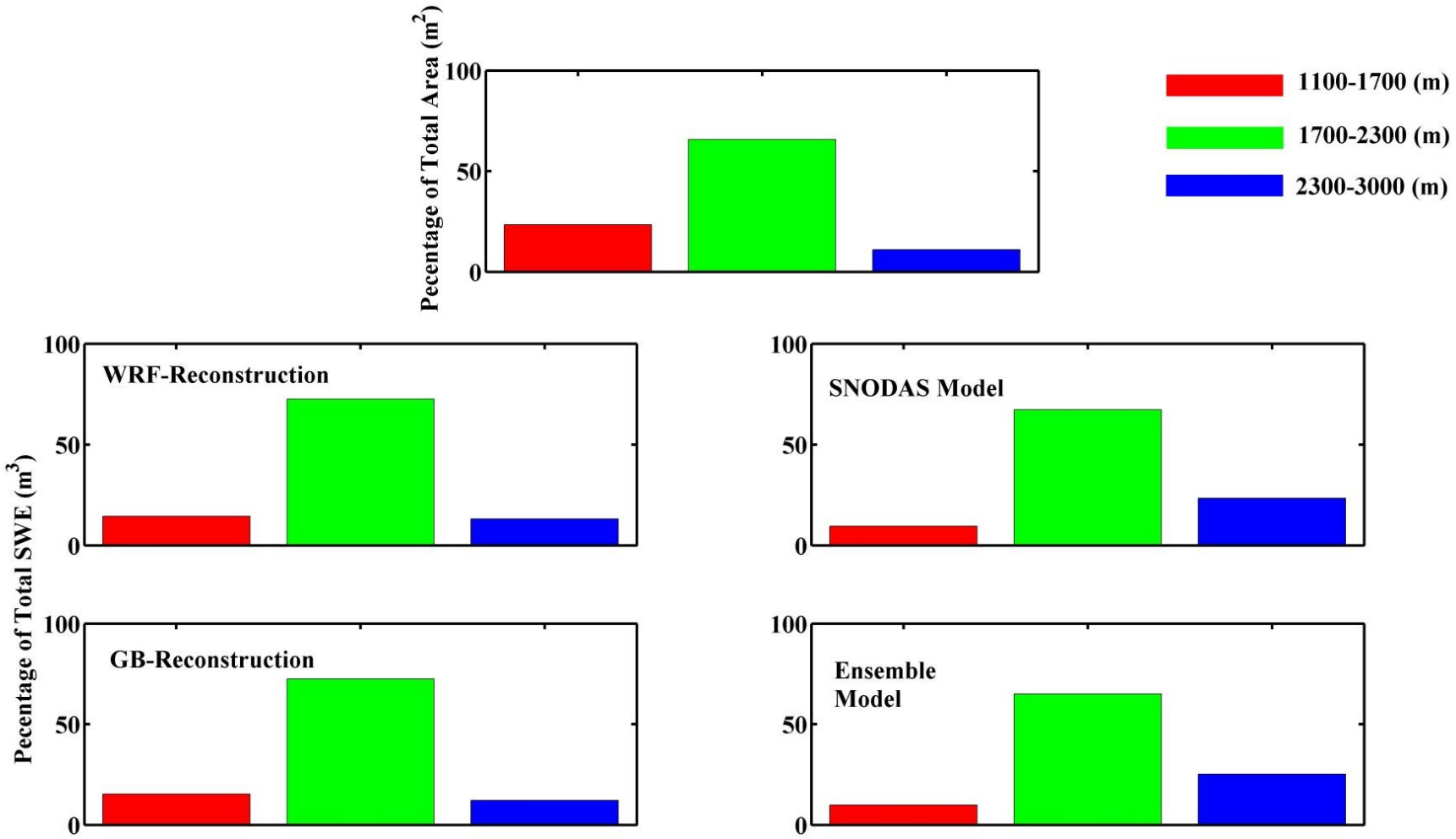

Figure 10: Percentage of total SWE (m^3) for three elevation intervals within the study area.

3.3.5. Bulk SWE: North and South Facing Pixels - In a comparison of the difference in bulk SWE (m), (SWE volume divided by area) for all north (315°-45°) and south (135°-225°) facing pixels within individual drainages for the study area, the bulk SWE (m) is greater on north facing drainage sidewalls (Figure 11) for all three models. Of the 486 bulk SWE (m) difference calculations made for drainages in all three models for all years, 16% of these calculations resulted in bulk SWE (m) on south facing drainage side walls that was greater than or equal to the bulk SWE (m) on northing slopes. For the GB and WRF reconstruction models the bulk SWE (m) on south facing drainage walls is greater than or equal to north facing drainage sidewalls in 20% and 12% of the bulk SWE difference calculations, respectively. For the SNODAS model this number is 24%. When comparing bulk SWE on south facing drainage sidewalls, south facing bulk SWE ranges from 49% – 135% of north facing bulk SWE (Figure 12). When only considering bulk SWE (m) difference calculations where the bulk value is larger for north facing slopes the annual average of these values for all drainages ranges from 6.4 – 7.5 cm depending on the model and year of modeling. When only considering bulk SWE (m) that is larger on north facing slopes and is expressed as a percent of south facing bulk SWE (m) divided by north facing bulk SWE (m) the minimum of these percentages range from ranges from 34% - 57% depending on the model and the drainage.

3.4. Ensemble Model Performance and Results

Because the Ensemble model is an average of our three models, the previously discussed results for the individual models are very similar for the Ensemble. This is particularly true of the patterns of mean SWE (m) versus elevation (Figure 8) and the distribution of SWE (m^3) within the study area (Figure 9).

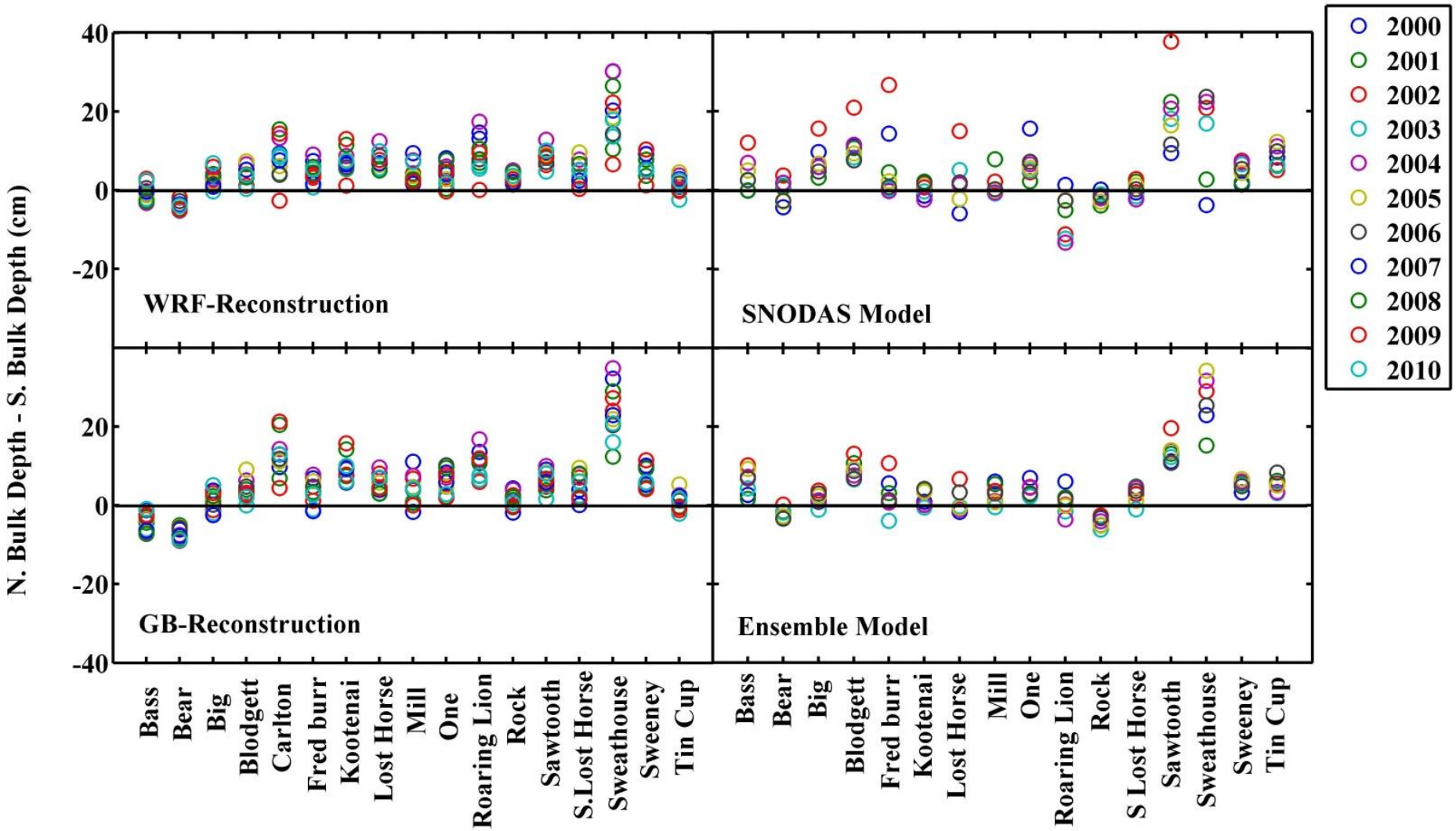

Figure 11: Difference in bulk SWE (m) between north and south facing drainage sidewalls.

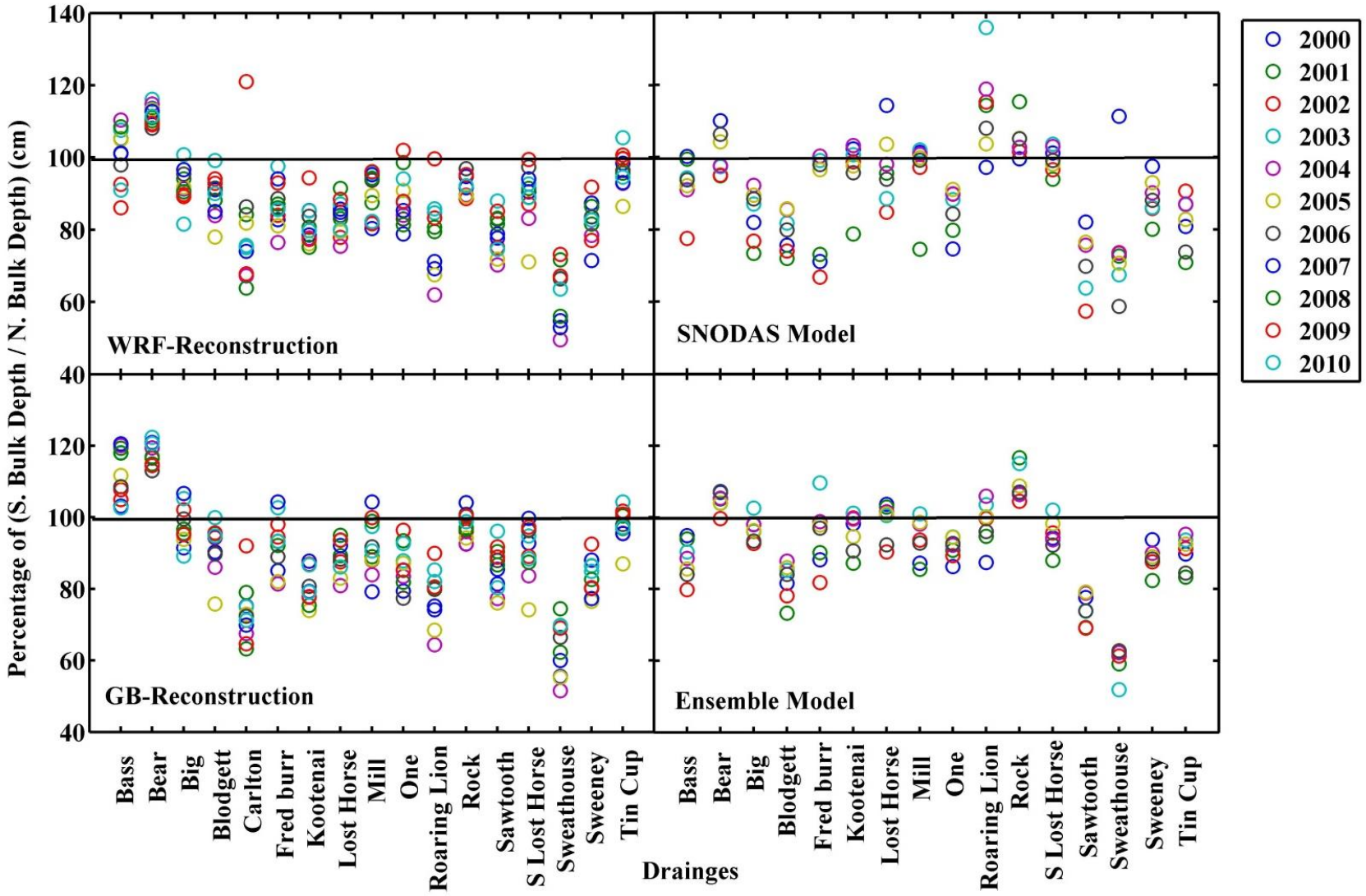

Figure 12: Percentage of south facing bulk SWE (m) relative to north facing bulk SWE (m).

4. Discussion

4.1. Model Performance

4.1.1. Reconstruction Models: Performance and Comparison - We expected that the GB and WRF-reconstruction models would have similar performance because both models use the same cloud filled MODIS FSCA images as well as temperature and FSCA threshold values. Additionally, the temperature output of the WRF model underwent a Bayesian optimal interpolation that used the variance of average temperature from some of the meteorological stations used to create temperature fields for the GB-reconstruction model. Further, there is similarity between the solar radiation values used in the two models. The WRF-reconstruction model solar radiation values are always equal to or smaller than those used in the GB-reconstruction model because of the cloud cover factor determined from the WRF model solar radiation output that was applied to the GB-reconstruction model potential clear sky solar radiation values.

Because of these similarities between the reconstruction models, the RMSD for individual pixel SWE (m) values between the two reconstruction models, ranges from 7.2×10^{-3} - 0.33 (m) depending on the year of modeling. If the year 2002 is excluded, the average of the RMSD values between the two reconstruction models is 2.4×10^{-2} (m). In addition to the small RMSD values between the two reconstruction models, the trends in the distribution of SWE (m^3) (Figure 9; Figure 10) and SWE (m) averaged over elevation intervals (Figure 8), are very similar as well.

4.1.2. WRF-Reconstruction Model Performance: Year 2002 Results - For the WRF and GB-reconstruction models, 2002 was the year of lowest and second lowest annual mean melt season temperature, respectively. For the WRF-reconstruction model this resulted in particularly low PSWEP, which produced a negative linear dependence of SWE (m) and elevation, something not observed in any of the other years of modeling for any of the models. With the exception of 2002, the absolute difference in the annual mean melt season temperature between the WRF and GB-reconstruction models is 0.42 °C or less. In 2002, the GB-reconstruction model annual mean melt season temperature is 4.3 °C warmer than that of the WRF-reconstruction model. The highest mean annual melt season temperature for the WRF-reconstruction model of 7.3 °C occurred in 2007 and is 22 times larger than the temperature of 0.33 °C modeled for 2002.

These cold temperatures are restricted to the melt season of 2002. The Bayesian optimal interpolated temperatures of the WRF model output for all of 2002 have maximum, minimum and mean temperatures that are comparable with those of the other years of modeling. These cold melt season temperatures are the result of the Bayesian optimal interpolation as melt season temperatures of such a small magnitude are not in the melt season temperatures of the WRF model output before the Bayesian optimal interpolation.

4.1.3. Reconstruction and SNODAS Model Performance and Comparison - We are encouraged by the similar performance of the three models used in this research. While the SNODAS model has the greatest variability (Table 7) of the three models, the mean individual pixel SWE (m) values and total SWE (m^3) are within the same order of magnitude. The mean individual pixel SWE (m) values for the SNODAS model range from 71%-130% of the reconstruction models values, and SNODAS total SWE (m^3) values range from 92%-175% of the reconstruction values.

Because of the spatial variability of SWE, much of the variability observed in the individual pixel SWE (m) values of the SNODAS model results are probably representative of the on ground conditions within the study area. However, for all years of modeling and scattered through all elevations of the study area there are a number of SNODAS model pixels which have values of zero SWE (m). Since we are summing the absolute value of negative daily SWE change over the time period of day 60-200 of the year for each pixel in the study area, the probability of these pixels being SWE free, particularly at higher elevations is very low and we believe that these zero SWE (m) values represent an error in the SNODAS model output. Further, analysis of snowmelt runoff at the base of the snow pack, a diagnostic model output variable from SNODAS does not have any of these zero values.

While the agreement between these three models is encouraging, this model agreement could be occurring for reasons that are unrelated. However, because these models all represent the physics of snowmelt differently and ingest different forcing data it seems more likely that the model agreement is the result of the models correctly capturing the physical processes that control and effect the spatial distribution of SWE. This is not to say that any of the individual model or Ensemble model results are a replication of SWE on the landscape during our modeling

period, but that the patterns and distributions of SWE resulting from our modeling are reflective of the on ground conditions and the processes which drive the spatial distribution of SWE.

4.2. Model Sensitivity and Error

Lacking a data set of ground measured SWE for our study area that is representative of our model pixel resolution or that is similar in nature to the PSWEP from each model makes error analysis of any of these results difficult. Previous efforts to understand the error associated with reconstruction models that are similar to those implemented in this research have used several different approaches. These include, comparison of modeled SWE with a sample of measured point scale SWE [e.g., *Rice et al.*, 2011], measured point scale SWE that has been interpolated to the landscape scale using a regression tree [e.g., *Durand et al.*, 2008;] or averaging a sample of measured point scale SWE to represent the mean SWE value at the resolution of a select number of model pixels [e.g., *Clow et al.*, 2012; *Molotch*, 2009]. While these methods provide some basis for comparison, they are not actually quantitative model validation experiments. Both methods require on the ground measurement of SWE, which are limited in scope due to the cost and danger associated with on ground snow measurement. Collection of SWE measurements to represent an individual model pixel must be numerous as point scale SWE measurement is usually not representative of the larger surrounding grid scale [*Molotch and Bales*, 2005; *Rice and Bales*, 2010]. Further, on ground SWE measurement must be performed during a short time period to represent a snap shot in time of on ground SWE, because accumulation, ablation, and wind redistribution processes can quickly change the results of on ground catchment or landscape scale SWE measurement. Because of these issues, the maximum relative error of 55%-60% [*Durand et al.*, 2008], mean absolute error of 23% [*Molotch*, 2009] and mean relative difference 15%-18% [*Rice et al.*, 2011] observed by others when comparing modeled and measured SWE could be higher or lower than these reported values.

In the reconstruction models the limited number of model forcing inputs and parameterization help to limit error propagation in the results of these models. Further, because of the small magnitude of both the temperature and solar radiation coefficients used in the reconstruction models, errors associated with either the temperature or solar radiation forcing data would result in a small propagation of error within the results. The GB-reconstruction model

was most sensitive to the removal of the Nez Perce SNOTEL station from creation of daily temperature fields and calculation of model coefficients. The annual RMSD values for the Nez Perce station range from $4.1x10^{-2}$ - 0.13 (m) (Table 5). When these annual RMSD values are expressed as a percentage of the annual mean of individual pixel SWE (m) values for the GB-reconstruction model (Table 7) these percentages range from 10%-24%. Sensitivity testing of the WRF-reconstruction model was more limited because we were further downscaling the output of a model. However, because of the small RMSD values between the two reconstruction models and the similarities in the sensitivities of the models to changes in model coefficients (Figure 6), the error associated with either reconstruction model are likely within the same order of magnitude.

The comparison of SNODAS model pixels with average SWE measurement from corresponding on the ground locations for a study area in Colorado [*Clow et al.*, 2012] produced a RSME of 5-12 (cm). Location and vegetation cover accounted for the range of RMSE values observed. The largest RMSE values occurred in higher elevation locations that had limited or non-existent tree cover. *Clow et al.*, [2012] attributes these higher RSME values to wind redistribution which the SNODAS model accounts for poorly. The tree-line or above tree-line sites in the Bitterroot are lower in elevation than those of the *Clow et al.*, [2012] study and are likely less affected by wind, which may produce lower RMSE values for our study area.

While we cannot put an absolute value on the error of any of the models used in this research, we have shown that the mean SWE (m) and total SWE (m^3) of our 3 semi-independent models all produce reasonable distributions of SWE which have similar patterns and are on average within 21% each other. Further the sensitivity of our reconstruction models to large changes in model coefficients or jackknifing of temperature fields produces RMSD values of 0.33 (m) or less. When we consider model validation attempts for similar reconstruction models and SNODAS we find that the range of errors reported from their limited validation attempts are in the same order of magnitude as individual pixel SWE (m) values produced by their models. Perhaps the best estimate we can make of model errors is from examination of SNODAS model error. Validation of SNODAS found RSME values of 5-12 (cm) [*Clow et al.*, 2012], however, we believe that given the method employed to validate the SNODAS model there is potential for these values to be larger or smaller than what is reported. With this in mind given the similar performance of the reconstruction models and SNODAS we have concluded that any error

associated with our modeling results is either within the same order of magnitude or at the most one order of magnitude larger than any of the individual, or mean SWE (m) and total SWE (m^3) values.

4.4. Ensemble Model

Because we lack the ability to determine superior performance between our three models we believe the average of these three models to be the best estimate of SWE distribution for our study area. However, because of the differences in resolution between these models and our need to understand if SWE distribution is a function of model performance we refer to all model results in the following discussions.

4.5. Landscape SWE Distribution: Patterns and Processes

4.5.1. Percentages of Total SWE (m^3) – The distribution of SWE (m^3) closely matches the distribution of surface area (Figure 9). The partitioning of SWE (m^3) into broad elevation intervals reveals that the largest percentages of SWE (m^3) and area are found between the elevations of 1700-2300 m elevation (Figure 10). Within the study area the highest elevation SNOTEL station is Twin Lakes located at 1950 m elevations. When looking at the percentage of total SWE (m^3) above this elevation the average of all of the models for all years of modeling is 70%. Our study area was selected to include the west-east trending drainages that feature north and south facing sidewalls and didn't encompass the entire portion of Central Bitterroot Range. To the south of the study area while still located in the Central Bitterroot Range the Saddle Mountain SNOTEL station sits at 2420 meters elevation. The percentage of total SWE (m^3) located above this station is on average 10% for all of the models. This illustrates the finding of others that large quantities of SWE on the landscape are unrepresented by our current system of measurement [*Gillan et al.*, 2010; *Rice et al.*, 2011]. For our study area up to 70% of total SWE (m^3) is represented by a single SNOTEL station.

4.5.2. Bulk SWE North and South Facing Drainage Sidewalls – For the majority of bulk SWE (m) calculations between north and south facing drainage sidewalls, north facing drainage sidewalls have a greater bulk SWE (m) than south facing drainage sidewalls (Figure 11). The mean bulk SWE difference for instances where only north bulk SWE (m) is greater than south, ranges from 6.4 – 7.8 (cm) for all of the individual drainages. The smallest south facing bulk

SWE (m) value is 49% of the corresponding north facing bulk SWE (m) value. These differences while small in some instances are not trivial and likely fall outside of the bounds of error for our models. However, because the difference in these bulk SWE (m) values are small we performed a statistical analysis using a Kruskal-Wallis test to test for a statistical difference in the population of annual bulk SWE (m) for north and south facing drainage sidewalls for each drainage within the study area. For the 17 drainages within the model domain, six were statistically different for the GB-reconstruction model and four were statistically different for the WRF-reconstruction model at the 95% confidence interval (Table 9). For the SNODAS model the results for only one of the drainages is statistically different.

The differences between these bulk SWE (m) values could be caused by sublimation or depositional processes. It is unlikely that more snow is preferentially deposited on north facing slopes during precipitation events particularly at the catchment scale, but prevailing winds and differences in vegetation on north and south facing slopes within the study area could account for the differences in these bulks SWE (m) values. Rates of sublimation vary depending on the site, but losses of 15% [*Hood et al.*, 1999] and 20-32% [*MacDonald et al.*, 2010] of cumulative snowfall are not uncommon. *Hood et al.*, [1999] found that sublimation occurred primarily during the accumulation season, and condensation to the snowpack was more likely during the melt season with a total of 2% of cumulative snowfall being added back to the snowpack through condensation. If preferential sublimation of snow is responsible for the differences in the bulk SWE (m) it is occurring during the accumulation season as all of our model results implicitly account for the loss of SWE from sublimation. The calculation of model coefficients for the reconstruction models and the PSWEP values from SNODAS, are based on the daily change in SWE. This change in SWE represents a loss of mass from both the production of snowmelt and sublimation. If sublimation was missing from these results then the results of bulk SWE (m) calculations for north and south facing slopes could underrepresent the total snow accumulation on these slopes and either slope could have more bulk SWE (m) than what has been calculated. Since our results represent the loss of mass from both sublimation and snowmelt, then similar amounts of snow accumulate on these north and south facing drainage sidewalls as the models are producing similar amounts of total snowmelt. When this is considered with the lack statistical difference in bulk SWE (m) calculations for individual drainages and the small magnitude of these bulk SWE difference calculations, our interpretation is that within the scale and error of

these models there is little to no difference between total accumulated snow on north and south facing drainage sidewalls in our study area.

4.5.3. SWE Elevation Gradients - When SWE (m) is averaged at 50 m intervals across the study area, the patterns of the SWE (m) elevation gradients for all three models are very similar from year to year and appear to be scaled by the climate of that year (Figure 8). Plotting the mean of these SWE (m) elevation gradients over the average surface area for the study area shows that elevations of maximum SWE depth corresponds closely with the elevations of maximum surface area (Figure 13), and as mean surface area decreases SWE (m) remains constant. The mean SWE (m) elevation gradient for all of the models (Figure 13), show a distinct pattern of SWE (m) increasing linearly as elevation increases at lower elevations within the study area until approximately 2000 m elevation (Figure 13). After this elevation zone, the SWE gradient switches from linearly increasing with increases in elevation to flattening out and becoming constant as elevation increases.

When we observed these flattened out SWE (m) elevation gradients we wondered if these were a reflection of the on ground SWE conditions within the study area or if the observed SWE (m) elevations gradients were an artifact of modeling. If the solar radiation and temperature values above the elevations of 2000 m used by the models were too low, daily melt produced by the models would be too low and may produce the observed flattened SWE (m) elevation gradients. This is an important consideration as a large portion of surface area within the study area is between the elevations of 2000 and 2400 m (Figure 13). If these SWE (m) elevation gradients are incorrect and are supposed continue to increase linearly above the elevations of 2000 m within the study area, a large volume of SWE is missing from our model results.

In our reconstruction models we are distributing temperature throughout the model domain using either observed temperatures or those from the WRF model pixels and temperature lapse rates determined from the elevation and temperature from meteorological stations and the elevation and temperature WRF model pixel values. Temperatures above 2000 m in the reconstruction models could be colder that what is naturally occurring from either a limited number of observations stations or WRF model pixels above the elevations of 2000 m or from the lapse rates used to distribute these temperatures across the model domain. These lapse rates could be producing temperatures at higher elevations that are too cold because they are not

accurately capturing cold air drainage within the Bitterroot Mountains that would cause temperature inversions resulting in warmer temperatures at higher elevations. In the GB-reconstruction model only two of the meteorological stations used to distribute temperature across the model domain are above the elevation of 2000 m (Table 2). Within the WRF-reconstruction model 56 of the 74 WRF model pixel centroids used to distribute temperature in the WRF-reconstruction model study area have an elevation above 2000 m.

When we considered further if these flattened SWE (m) elevation gradients were an artifact of modeling we found some evidence to the contrary. First, we found this pattern in all of the model results (Figure 8; Figure 13). The SNODAS model does display more variability than the reconstruction models but there still exists a pattern of flattened SWE (m) elevation gradients above the elevation of 2000 m (Figure 8). Further, the SNODAS model uses different data and methods to distribute temperature within its domain. Second, all of the models ingest some remotely sensed snow covered area product to determine whether pixels within the model domain were snow covered. The SNODAS model ingests AVHRR, NOAA-15 satellite data to determine aerial snow cover extent and we have not downloaded these data for a direct comparison. However, when the melt out dates of individual MODIS FSCA pixels used in the reconstruction models are averaged over 50 m elevation intervals within the study area, the patterns observed are similar to those when SWE (m) is averaged over elevation (Figure 14). Like the SWE (m) elevation gradients of the reconstruction models, the change in average pixel melt out dates becomes more constant at higher elevations within the study area.

The results of our average melt out date gradients are contrary to those found by others for a portion of the Sierra Nevada Range in California [*Rice et al.*, 2011]. Our results show melt out dates that are not largely different above the elevations of 2100 m, whereas *Rice et al*., [2011] found that each elevation interval of 300 m, melted out 2-3 weeks later than the successive lower elevation interval. *Rice et al*., [2011] used a different method to determine FSCA using MODIS satellite imagery, and their study area has a maximum elevation that is 1000 m higher than our study area.

Lastly, when we looked at the trends in SWE (m) for three SNOTEL stations within and outside of the study area we found a similar pattern of SWE (m) increasing at lower elevations within the study area and decreasing at higher elevation (Figure 15).

With some confidence that the SWE (m) elevation gradients observed within the study area were not an artifact of the modeling, we sought some explanation of physical processes that could be driving these patterns. When considering the physical processes driving the observed SWE (m) elevation gradients within the study area it is important to understand the distribution of elevation within the study area, particularly above 2000 m as this is the elevation where the SWE (m) elevation gradients change. In Figure 16, the elevations above 2000 m within the study area highlighted in blue. When the area above the elevation of 2000 m is highlighted, it becomes apparent that at lower elevations where the SWE (m) elevation gradient is more linear is occurring within the drainage valleys within the study area. Valley sidewalls and ridge top comprise the areas above 2000 meters in the study area. This is further exemplified by elevation transects that are plotted for a ridge and valley within the study area (Figure 17).

Subsequently, these elevation transects show that the difference between the SWE (m) elevation gradients above and below of 2000 m within the study area are a function of SWE accumulation processes in the valley bottoms and valley sidewalls/ridgetops of the drainages within the study area. In the valley bottoms, SWE is a function of elevation which is what is expected from orographic precipitation processes. However, along ridgetops and valley sidewalls within the study area SWE and elevation are not strongly correlated which indicates that orographic precipitation does not strongly control the spatial distribution of SWE in these locations.

Interpreting the processes driving these flattened out SWE (m) elevation gradients along ridgetops and valley sidewalls within the study area is more difficult. It is possible that within the study area above the elevation of 2000 m precipitation has reached its maximum and will not increase with elevation, as associated with orographic precipitation resulting in the observed flattened out SWE (m) elevation gradients. Further, processes such as wind and avalanching occurring above the elevations of 2000 m where tree cover is more sparse could result in preferential redistribution of SWE within the study area resulting in the observed SWE (m) elevation gradients. These processes are likely secondary to precipitation patterns, indicating a need for further research of orographic precipitation processes particularly on the lee side of mountain ranges.

Modeling efforts for the Sierra Nevada Mountains of California that used a model of interpolated SWE measurement, SNODAS and a reconstruction model found differences in the

SWE (m) elevation gradients between the three models they employed [*Rittger* et al., 2011]. Similar to our results they found that the interpolation SWE model and SNODAS had SWE elevation gradients values that increased with increases in elevation for the lower elevations within their study area up to the elevation of 3000 (m). After this elevation, the SWE interpolation and SNODAS model, SWE (m) elevation gradients decreased as elevation increased. However, the SWE (m) elevation gradient of their reconstruction model was linearly increasing across all elevations within their study area.

Our SWE elevation gradients also differ from those of *Gillan et al.,* [2010]. In their study area the pattern of individual SWE (m) elevation gradients were the same across the lower and mid elevations within their study area and are scaled by the climate of the modeling year. At higher elevations, the patterns of SWE elevation gradients had much greater inter-annual variability. This is counter to our results, where SWE (m) elevation gradient patterns at higher elevations are the similar from year to year.

The differences between the SWE elevation gradients for the individual model results of *Rittger et al*., [2011], at higher elevations and the difference between the SWE elevation gradients of *Rittger et al*., [2011] and *Gillan et al.,* [2010] compared to our own results points to a need for a greater understanding of high elevation snow throughout the Western US. These higher elevations are currently underrepresented by our present systems of on ground SWE measurement, have been show to hold large percentages of total basin SWE [*Gillan et al.,* 2010; *Rice et al.*, 2011], and will likely be less affected by variability and changes to future climate.

4.5.5. Physiography and SWE Distribution - In an examination of some of the physiographic controls on the spatial distribution of SWE for our study area there is not a single physiographic element that strongly controls the spatial distribution of SWE. At the lower elevations within valley bottoms within our study area SWE elevation gradients would indicate that orographic processes and thus elevation are controlling the spatial distribution of SWE. Others have found that elevation was important in predicting the spatial distribution of SWE throughout their study area [*Sexstone and Fassnacht*, 2013]. However, at higher elevations on ridge tops and valley sidewalls within our study area SWE elevation gradients flatten out. the higher elevations is not strongly controlling the distribution of SWE, and throughout the study area, other physiographic elements such as aspect and slope appear to be at best secondary

controls on the spatial distribution of SWE. These results are not surprising given that previous research on the spatial distribution of SWE have found different physiographic elements to be important when considering spatial SWE distribution [e.g., *Elder et al.*, 1991; *Kerr et al.*, 2013; *Sexstone and Fassnacht*, 2013]. It is likely that the physical processes which drive the spatial distribution of SWE in many instances are site or region specific. This should give pause to those implementing methods of SWE interpolation or snow model parameterization that are largely based on physiography. Lastly, these results show that the implementation of SWE monitoring system will be most effective if sensors are place across a variety of physiographic locations and elevations.

Table 9 : Number drainages where the bulk SWE on north and south facing drainage sidewalls is statistically different.

GB-Reconstruction	WRF-Reconstruction	SNODAS	Ensemble
6	4	1	2

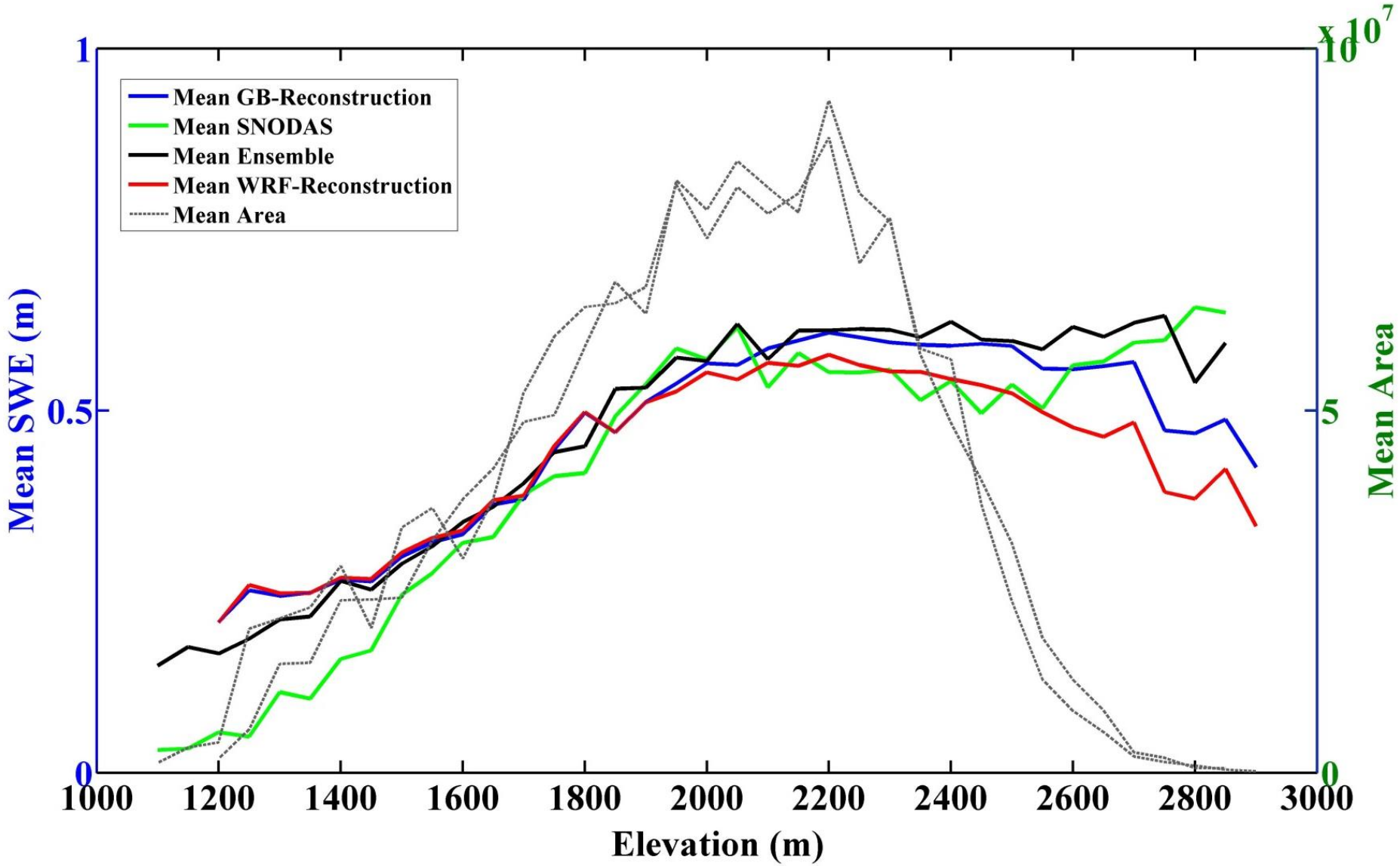

Figure 13: Mean of SWE elevation gradients for all years of modeling and mean surface area for all models.

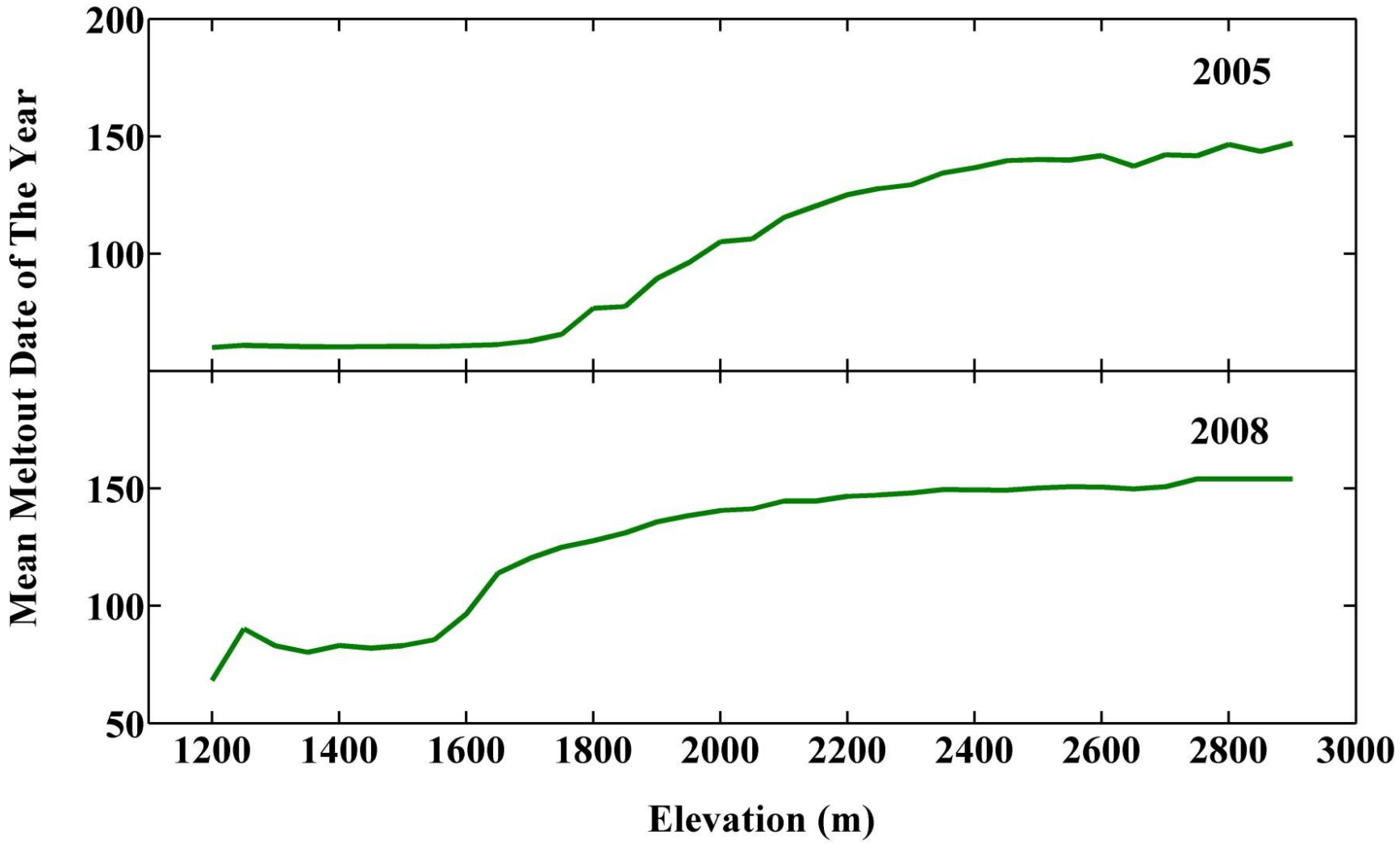

Figure 14: Mean meltout dates from cloud filled MODIS FSCA, for all north and south facing pixels, for 2005 and 2008 the driest and wettest years of modeling.

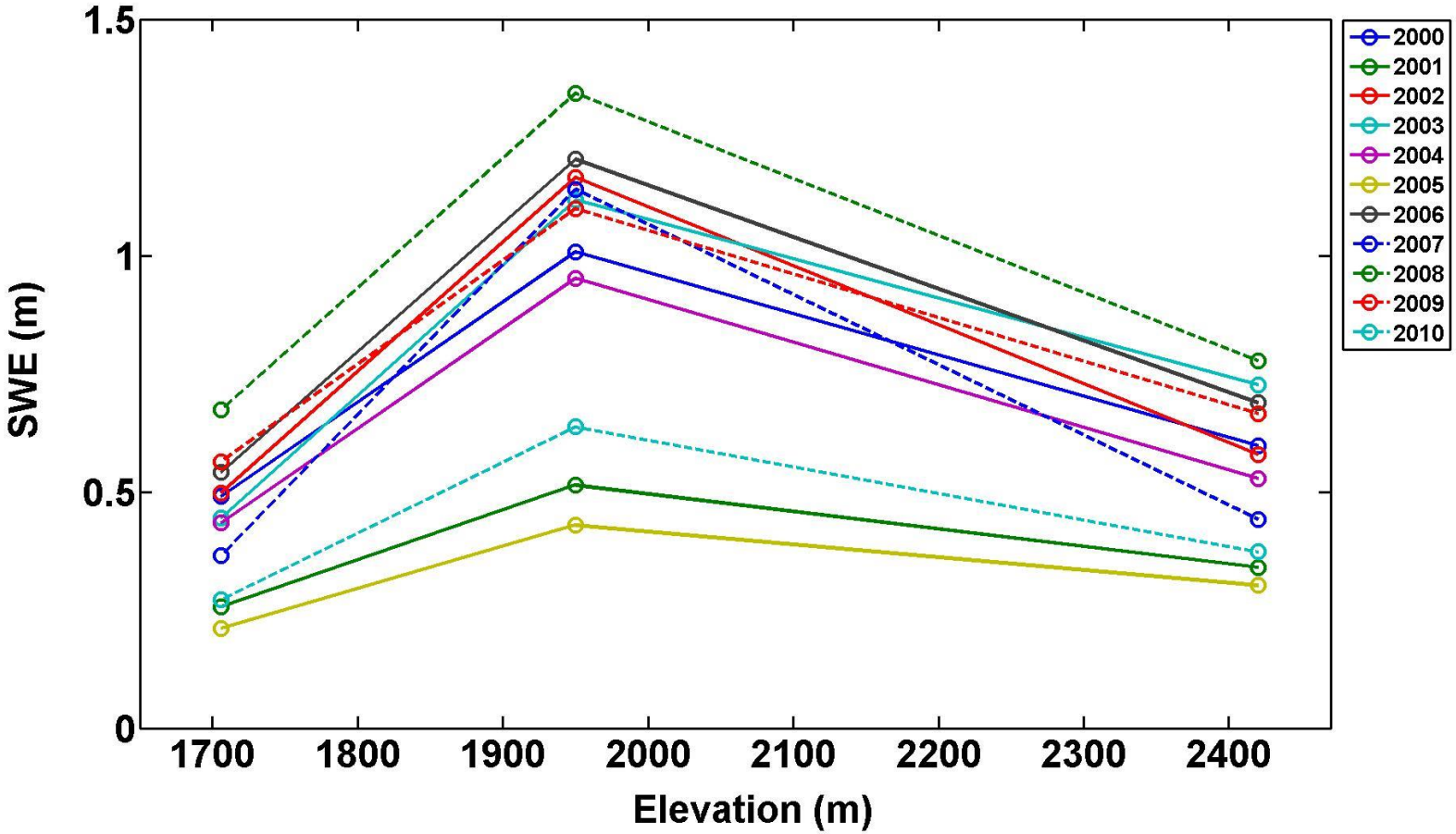

Figure 15: Trends in SWE (m) for SNOTEL stations from within and outside of the study area for the years 2000 - 2010. The station at 1706 meters and 1950 meters are the Twelvemile Creek and Twin Lakes SNOTEL stations located within the study area. The station at 2420 meters is the Saddle Mountain SNOTEL station which is located within the Central Bitterroot Mountains but outside of the study area.

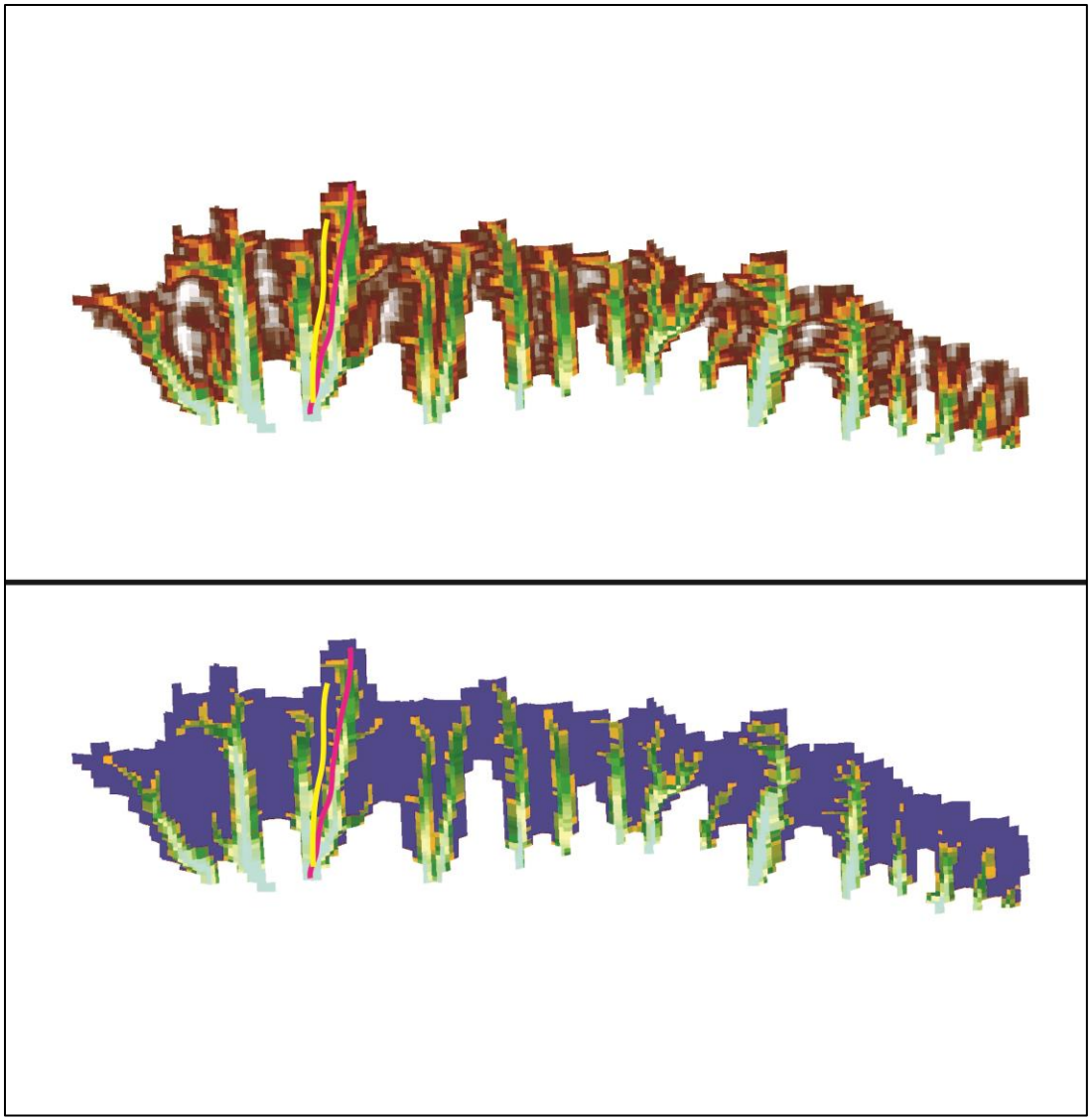

Figure 16: Elevation within the study area, with elevations above 2000 meters highlighted in blue in lower panel of figure. Yellow line is elevation profile on a ridge top. Maroon line is an elevation profile in a valley bottom.

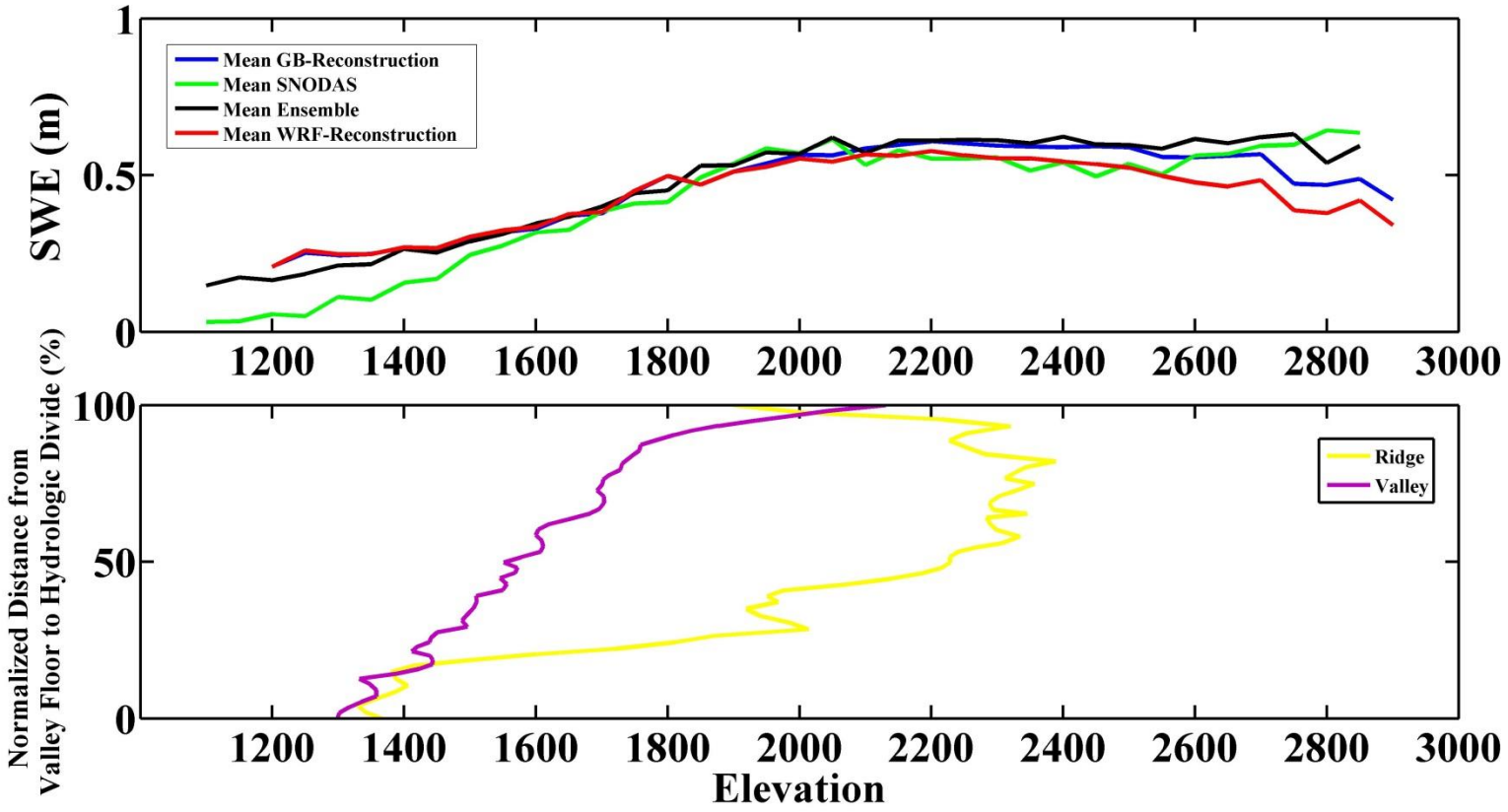

Figure 17: Elevation transects sub-plotted with SWE (m) elevation gradients within the study area. Valley transect is the maroon line and ridge transect is yellow line.

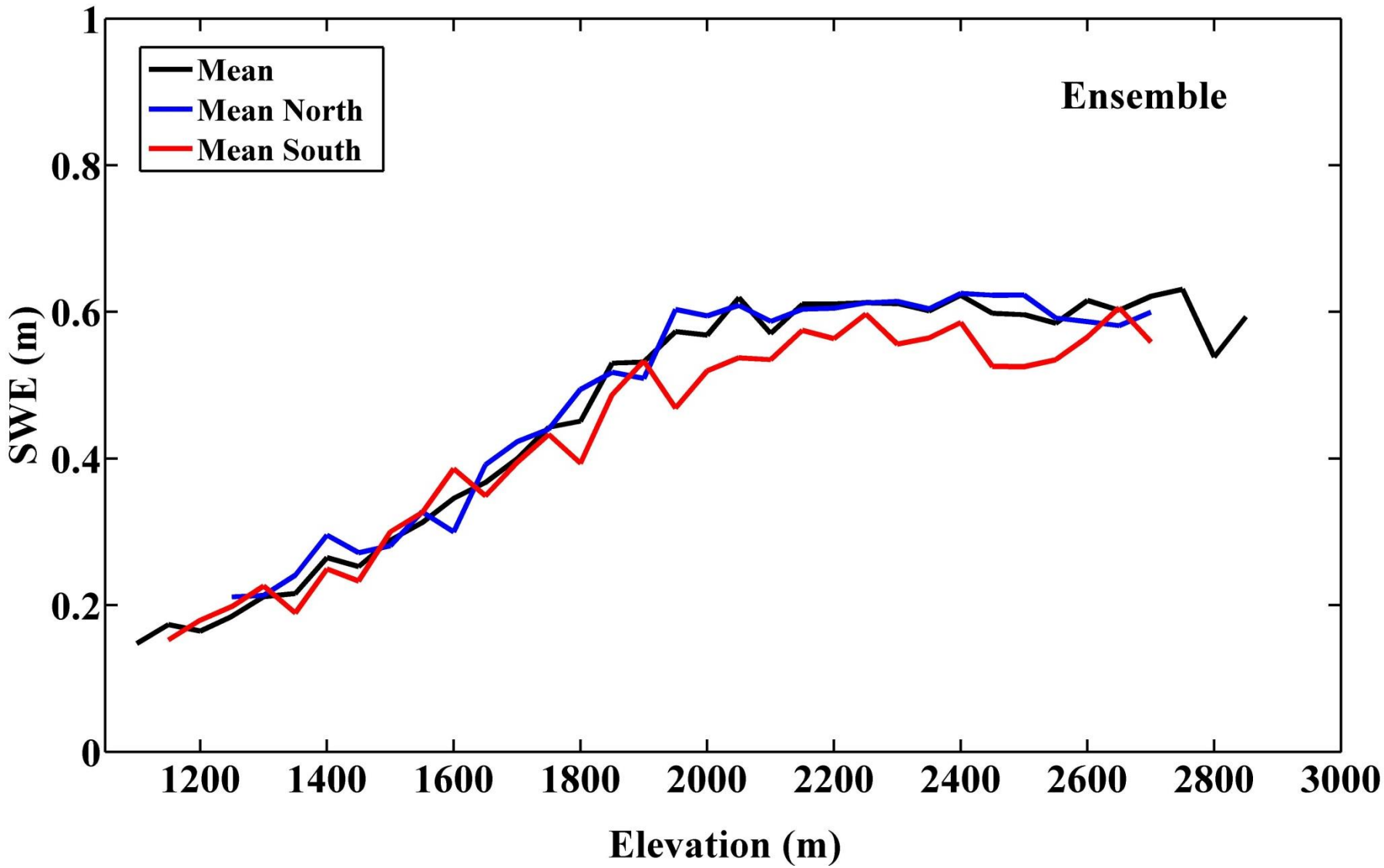

Figure 18: Mean SWE (m) elevation gradients for all north and south facing pixels, for the Ensemble model.

5. Conclusions

The use of three semi-independent snowmelt and snow models produces mean SWE (m) and total SWE (m^3) values that are within the same order of magnitude for a study area within the Central Bitterroot Mountain Range in Montana. Upwards of 70% of total SWE (m^3) is above the elevation of 1950 m elevation within our study area. This volume of snow is currently represented by a single SNOTEL station. Analysis of bulk SWE (m) on north and south facing drainage sidewalls indicates that the difference in snow accumulation on these slopes is minimal within this study area. SWE (m) elevation gradients within this study area increase linearly at lower elevations and flatten out and become constant at higher elevations. The difference between these SWE (m) elevation gradients at higher and lower elevations is driven by snow accumulation process that are occurring in drainage valleys, versus valley sidewalls and ridgetops. In drainage valleys SWE is a function of elevation indicating orographic precipitation is controlling SWE distribution. At higher elevations on valley sidewalls and ridgetops, where SWE (m) gradients remain constant with elevation, other precipitation processes and secondary redistribution processes such as wind and avalanching are likely controlling the spatial distribution of SWE. These SWE (m) elevation gradients indicate that SWE monitoring should occurring across a variety of physiographic locations and elevations.

These results are for a region on the lee side of mountainous terrain. Subsequently, processes at work within our study area are likely different from those for other regions or areas found on the windward side of a mountain range. Therefore, these results may not be representative of other regions and point to a need for further study of the spatial distribution of SWE at a variety of scales and global locations.

6. References Cited

Barrett, A. (2003), National operational hydrologic remote sensing center snow data assimilation system (SNODAS) products at NSIDC. *National Snow and Ice Data Center, Cooperative Institute for Research in Environmental Sciences, 2003*.

Cline, D. W., R. C. Bales, and J. Dozier (1998), Estimating the spatial distribution of snow in mountain basins using remote sensing and energy balance modeling a modeling spatial distribution of snow water equivalence basin at the a function the ecology, *Water Resources*, *34*(5), 1275–1285.

Clow, D., L. Nanus, K. Verdin, and J. Schmidt (2012), Evaluation of SNODAS snow depth and snow water equivalent estimates for the Colorado Rocky Mountains, USA, *Hydrological Processes*, doi:10.1002/hyp.

Debele, B., R. Srinivasan, and a. K. Gosain (2009), Comparison of Process-Based and Temperature-Index Snowmelt Modeling in SWAT, *Water Resources Management*, *24*(6), 1065–1088, doi:10.1007/s11269-009-9486-2.

Deems, J. (2006), Fractal distribution of snow depth from LiDAR data, *Journal of Hydrometeorology*, 285–297.

Dodson, R., and D. Marks (1997), Daily air temperature interpolated at high spatial resolution over a large mountainous region, *Climate Research*, *8*, 1–20.

Durand, M., N. P. Molotch, and S. a. Margulis (2008), Merging complementary remote sensing datasets in the context of snow water equivalent reconstruction, *Remote Sensing of Environment*, *112*(3), 1212–1225, doi:10.1016/j.rse.2007.08.010.

Elder, K., J. Dozier, and J. Michaelsen (1991), Snow Accumulation and Distribution in an Alpine Watershed, *Water Resources Research*, *27*(7), 1541–1552, doi:10.1029/91WR00506.

Erickson, T. a., M. W. Williams, and A. Winstral (2005), Persistence of topographic controls on the spatial distribution of snow in rugged mountain terrain, Colorado, United States, *Water Resources Research*, *41*(4), n/a–n/a, doi:10.1029/2003WR002973.

Gao, Y., J. a. Vano, C. Zhu, and D. P. Lettenmaier (2011), Evaluating climate change over the Colorado River basin using regional climate models, *Journal of Geophysical Research*, *116*(D13), D13104, doi:10.1029/2010JD015278.

Gillan, B. J., J. T. Harper, and J. N. Moore (2010), Timing of present and future snowmelt from high elevations in northwest Montana, *Water Resources Research*, *46*(1), 1–13, doi:10.1029/2009WR007861.

Grünewald, T., M. Schirmer, R. Mott, and M. Lehning (2010), Spatial and temporal variability of snow depth and SWE in a small mountain catchment, *The Cryosphere Discussions*, *4*(1), 1–30, doi:10.5194/tcd-4-1-2010.

Hall, D. K., G. A. Riggs and V. V. Salomonson (2006a), MODIS/Terra snow cover Daily L3 global 500m grid V005, 2000-2010, http://nsidc.org/data.mod10a2v5.html, Natl Snow and Ice Data Cent., Boulder, Colo. (Updated daily/weekly.)

Hall, D. K., G. A. Riggs and V. V. Salomonson (2006b), MODIS/Terra snow cover 8-Day L3 global 500m grid V005, 2000-2010, http://nsidc.org/data.mod10a2v5.html, Natl Snow and Ice Data Cent., Boulder, Colo. (Updated daily/weekly.)

Helgason, W., and J. Pomeroy (2012), Problems Closing the Energy Balance over a Homogeneous Snow Cover during Midwinter, *Journal of Hydrometeorology*, *13*(2), 557–572, doi:10.1175/JHM-D-11-0135.1.

Hock, R. (1999), A distributed temperature-index ice- and snowmlet model including potential direct solar radiation. *Journal of Glaciology, 45*(149), 101-111.

Hock, R. (2003), Temperature index melt modelling in mountain areas, *Journal of Hydrology*, *282*(1-4), 104–115, doi:10.1016/S0022-1694(03)00257-9.

Hood, E., M. Williams, and D. Cline (1999), Sublimation from a seasonal snowpack at a continental, mid-latitude alpine site, *Hydrological Processes*, *13*, 1781–1797.

Kerr, T., M. Clark, J. Hendrikx, and B. Anderson (2013), Snow distribution in a steep mid-latitude alpine catchment, *Advances in Water Resources*, *55*, 17–24, doi:10.1016/j.advwatres.2012.12.010.

Kumar, M., D. Marks, J. Dozier, M. Reba, and A. Winstral (2013), Evaluation of distributed hydrologic impacts of temperature-index and energy-based snow models, *Advances in Water Resources*, doi:10.1016/j.advwatres.2013.03.006.

MacDonald, M. K., J. W. Pomeroy, and a. Pietroniro (2010), On the importance of sublimation to an alpine snow mass balance in the Canadian Rocky Mountains, *Hydrology and Earth System Sciences*, *14*(7), 1401–1415, doi:10.5194/hess-14-1401-2010.

Martinec, J., and A. Rango (1981), Areal Distribution of Snow Water Equivalent Evaluated by Snow Cover Monitoring, *Water Resources Research*, *17*(5), 1480–1488.

McCabe, G., and M. Clark (2005), Trends and variability in snowmelt runoff in the western United States, *Journal of Hydrometeorology*, 476–483.

Meromy, L., N. Molotch, E. Timothy, S. Fassnacht, and R. Rice (2012), Subgrid variability of snow water equivalent at operational snow stations in the western USA, *Hydrological Processes*, doi:10.1002/hyp.

Molotch, N. P. (2009), Reconstructing snow water equivalent in the Rio Grande headwaters using remotely sensed snow cover data and a spatially distributed snowmelt model, *Hydrological Processes*, *1089*(January), 1076–1089, doi:10.1002/hyp.

Molotch, N. P., and R. C. Bales (2005), Scaling snow observations from the point to the grid element: Implications for observation network design, *Water Resources Research*, *41*(11), 1–16, doi:10.1029/2005WR004229.

Molotch, N. P., and R. C. Bales (2006), SNOTEL representativeness in the Rio Grande headwaters on the basis of physiographics and remotely sensed snow cover persistence, *Hydrological Processes*, *20*(4), 723–739, doi:10.1002/hyp.6128.

Molotch, N. P., and S. a. Margulis (2008), Estimating the distribution of snow water equivalent using remotely sensed snow cover data and a spatially distributed snowmelt model: A multi-resolution, multi-sensor comparison, *Advances in Water Resources*, *31*(11), 1503–1514, doi:10.1016/j.advwatres.2008.07.017.

Molotch, N. P., T. H. Painter, R. C. Bales, and J. Dozier (2004), Incorporating remotely-sensed snow albedo into a spatially-distributed snowmelt model, *Geophysical Research Letters*, *31*(3), L03501, doi:10.1029/2003GL019063.

Mote, P. W., A. F. Hamlet, M. P. Clark, and D. P. Lettenmaier (2005), Declining Mountain Snowpack in Western North America*, *Bulletin of the American Meteorological Society*, *86*(1), 39–49, doi:10.1175/BAMS-86-1-39.

Neumann, N. N., C. Derksen, C. Smith, and B. Goodison (2006), Characterizing local scale snow cover using point measurements during the winter season Characterizing Local Scale Snow Cover Using Point Measurements During the Winter Season, *Atmoshpere-Ocean*, *44*(3), 257–269.

National Operation Hydrologic Remote Sensing Center. 2004 *Snow Data Assimilation System (SNODAS) Data Products at NSIDC*, 2004 to 2010. Boulder, Colorado USA: National Snow and Ice Data Center. http://dx.doi.org/10.7265/N5TB14TC

Rango, A., and J. Martinec (1996), Revisiting the Degree-Day Method for Snowmelt Computations, *Water Resources Bulletin*, *31*(4).

Regonda, S., B. Rajagopalan, M. P. Clark, and J. Pitlick (2005), Seasonal cycle shifts in hydroclimatology over the western United States, *Journal of Climate*, 372–385.

Rice, R., and R. C. Bales (2010), Embedded-sensor network design for snow cover measurements around snow pillow and snow course sites in the Sierra Nevada of California, *Water Resources Research*, *46*(3), 1–13, doi:10.1029/2008WR007318.

Rice, R., R. C. Bales, T. H. Painter, and J. Dozier (2011), Snow water equivalent along elevation gradients in the Merced and Tuolumne River basins of the Sierra Nevada, *Water Resources Research*, *47*(8), W08515, doi:10.1029/2010WR009278.

Rittger, K., A. Kahl, and J. Dozier (2011), Topographic distribution of snow water equivalent in the sierra nevada *Proceedings of the 79th Annual Western Snow Conference*.

Schaefer, G., and R. Paetzold (2001), SNOTEL (SNOwpack TELemetry) and SCAN (soil climate analysis network), in *Proceedings of an International Workshop on Automated Weather Stations for Applications in Agriculture and Water Resources Management.*

Serreze, M. C., M. P. Clark, R. L. Armstrong, D. a. McGinnis, and R. S. Pulwarty (1999), Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data, *Water Resources Research*, *35*(7), 2145–2160, doi:10.1029/1999WR900090.

Sexstone, G. a., and S. R. Fassnacht (2013), What drives basin scale spatial variability of snow water equivalent during two extreme years?, *The Cryosphere Discussions*, *7*(3), 2943–2977, doi:10.5194/tcd-7-2943-2013.)

Skamarock, W. C., and J. B. Klemp (2008), A time-split nonhydrostatic atmospheric model for weather research and forecasting applications, *Journal of Computational Physics*, *227*(7), 3465–3485, doi:10.1016/j.jcp.2007.01.037.

Solomon, S., D. Qin, M. Manning, M. Marquis, K, Averyt, M.M.B. Tignor, H.L. Miller Jr., Z. Chen. (2007) Technical Summary. In: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovermental Panel on Climate Change.* Cambridge University Press, Cambride, United Kingdom and New York, NY, USA.

Stewart, I., D. Cayan, and M. Dettinger (2005), Changes toward earlier streamflow timing across western North America, *Journal of Climate*, 1136–1156.

Wikle, C. K., and L. M. Berliner (2007), A Bayesian tutorial for data assimilation, *Physica D: Nonlinear Phenomena*, *230*(1-2), 1–16, doi:10.1016/j.physd.2006.09.017.

Zappa, M., F. Pos, P. Warmerdams, and J. Gurtz (2003), Seasonal water balance of an Alpine catchment as evaluated by different methods for spatially distributed snowmelt modelling, *Nordic Hydrology*, *34*(3), 179–202.

7. Appendix

7.1. SNODAS PSWEP Methods

To use the results of the SNODAS model in the manner of a reconstruction SWE model we tried several methods before selecting our final method of determining PSWEP from the SNODAS model output. The first method we employed was to determine the day of peak SWE for the entire study area. From the date of peak SWE for the entire study area we added any positive contributions in SWE that occurred after this date until complete melt out of the model pixel. We found the results of this method to be very similar those of the PSWEP that we report in our results (Figure 19A). The year 2006 shows the largest difference between these two different calculations. Differences between these two methods are driven by melt that occurs before the date of peak SWE. In years where either the date of peak SWE is close to day 60 of the year or when little to no melt occurs between day 60 of the year and the date of peak SWE the difference between the results is very small.

In a separate analysis we looked at the SNODAS model output of water from the base of the snowpack (WBS). In the analysis of WBS from SNODAS, we summed all of the WBS for day 60-200 of the year for each pixel in the study area (Figure 20A). In this comparison the WBSand PSWEP reported in the results, the values are very similar. The WBS values can be larger or smaller than the PSWEP values that we report in the results. The difference between these values could be explained by the loss of SWE from sublimation that would be present in the PSWEP reported in our results. In some instances the loss of mass from WBS and sublimation would be present in our PSWEP results and these values would be larger than the WBS values. Why the WBS values would be larger than our PSWEP values such as in 2006 is unclear.

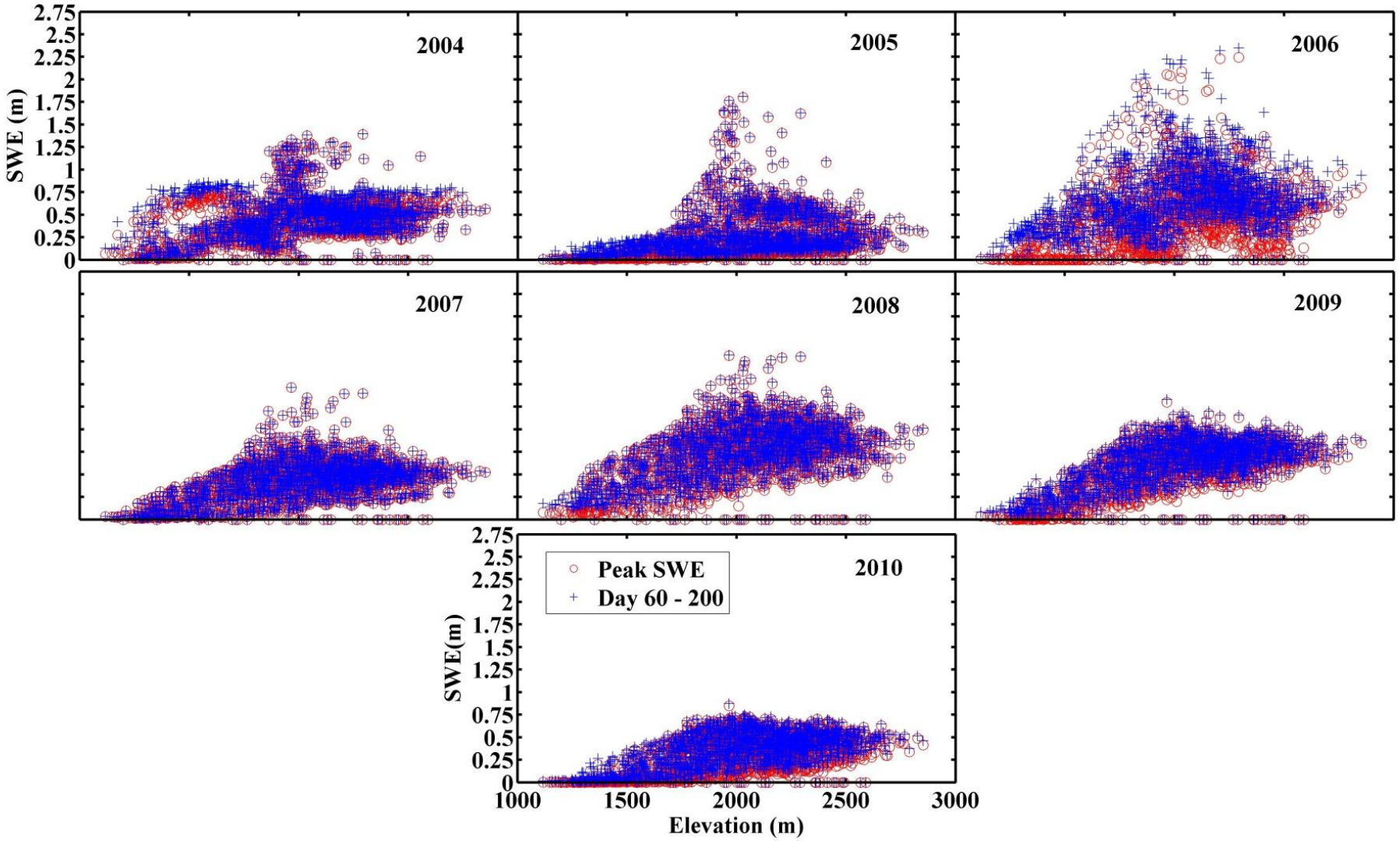

Figure 19A: Comparison of SNODAS PSWEP calculated using two different methods.

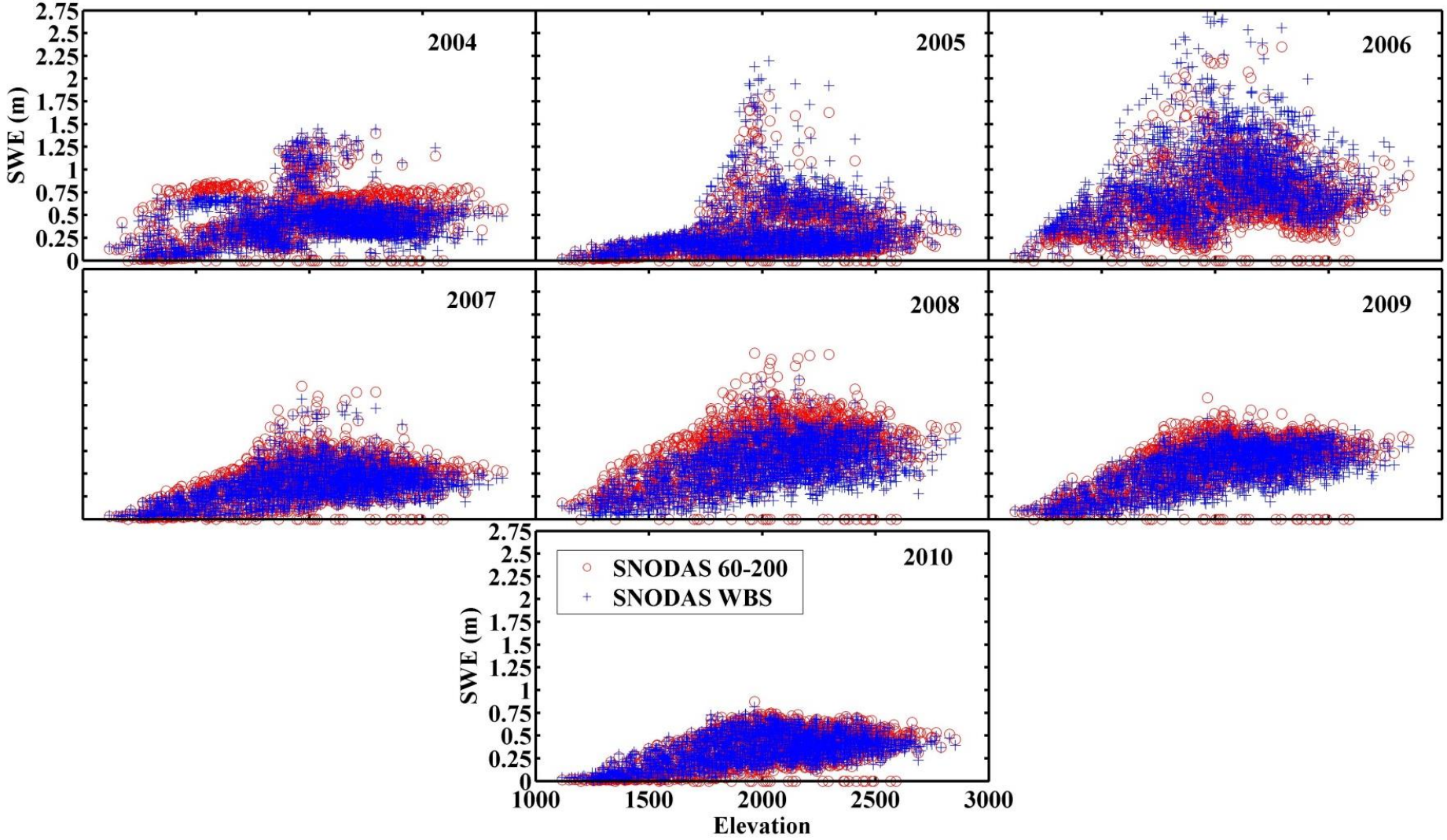

Figure 20A: Comparison of SNODAS PSWEP reported in results and the SNODAS model output variable, water from the bottom of the snowpack summed for day 60 - 200 of the year.

7.2. WRF Cold Temperature Bias

The initial run of the WRF-reconstruction model using the downscaled WRF model temperature output had poor model performance producing PSWEP values that we felt were too low when compared with the SNODAS and GB-reconstruction models. We determined that the temperatures modeled by WRF were colder than observed temperatures within the study area and were producing low daily snowmelt values. Subsequently, we performed a Bayesian optimal interpolation on the WRF temperature model output using the variance of temperature for SNOTEL stations located within the study area. Here we provide a comparison WRF pixels located closest to the Twin Lakes and Twelvemile Creek SNOTEL stations located within the study area, with the WRF temperature previous to the Bayesian optimal interpolation (Figure 21A; Figure 22A). The WRF temperatures are colder than the observed temperatures particularly during the early part of the melt season. The elevation of the Twin Lakes and Twelvemile Creek SNOTELs are 1950 m and 1706 m respectively. The elevations of the corresponding WRF pixels are 2048 m and 2086 m. The mean annual difference in temperature between these WRF simulated temperatures ranges from 4.1 °C - 6.1 °C for the Twin Lakes site and 4.4 °C – 5.9 °C for the Twelvemile Creek site depending on the year of modeling. While the difference in elevation between the WRF pixels and SNOTEL stations could cause some of the difference in temperature between the observed and modeled, the difference in these temperature values is still larger than can be explained by a temperature lapse rate and the elevations differences alone.

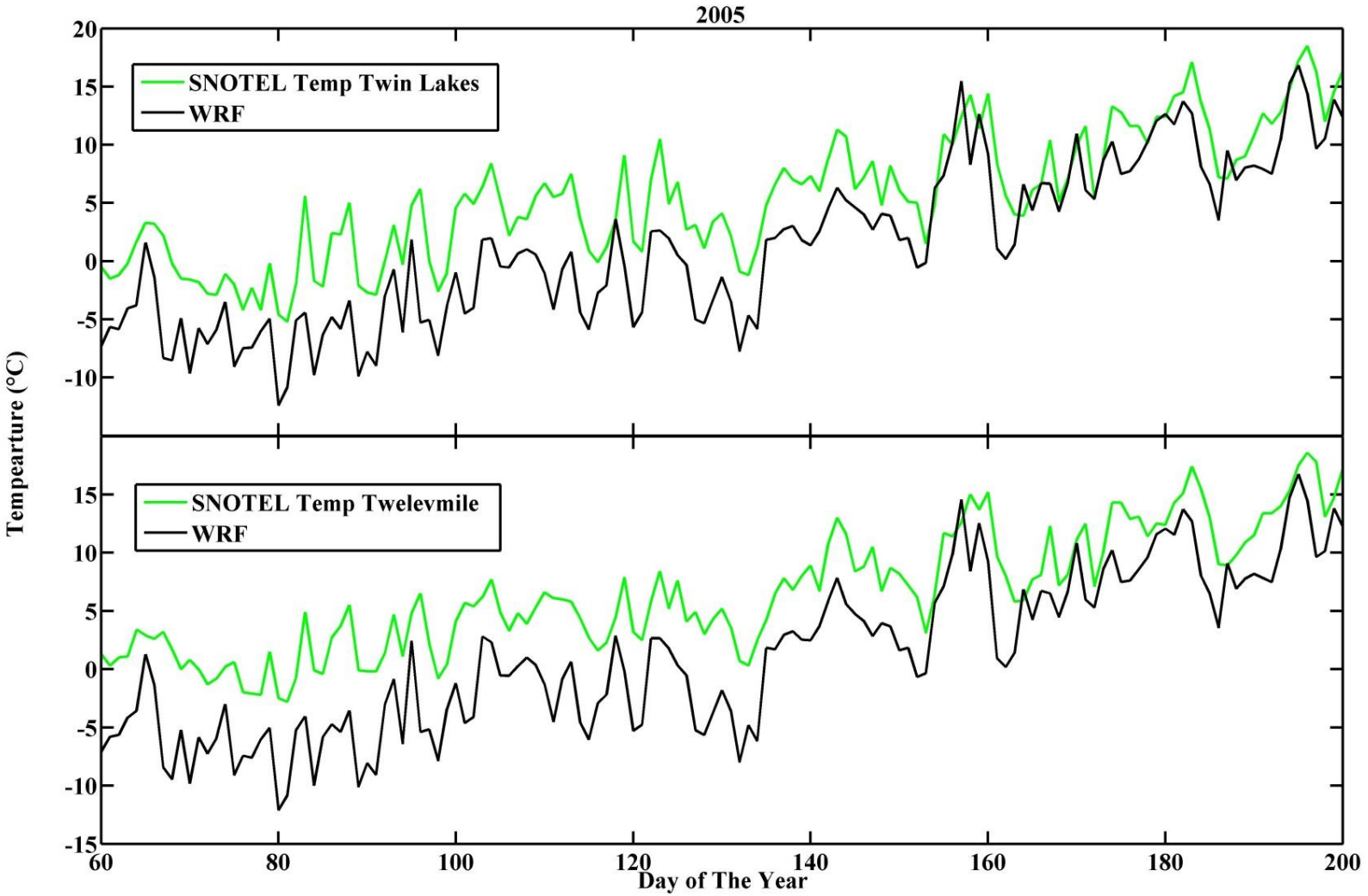

Figure 21A: Comparison of temperature recorded at Twin Lakes and Twelvemile SNOTEL stations and the simulated temperatures from the nearest WRF model pixels, for the year 2005, the driest year of modeling

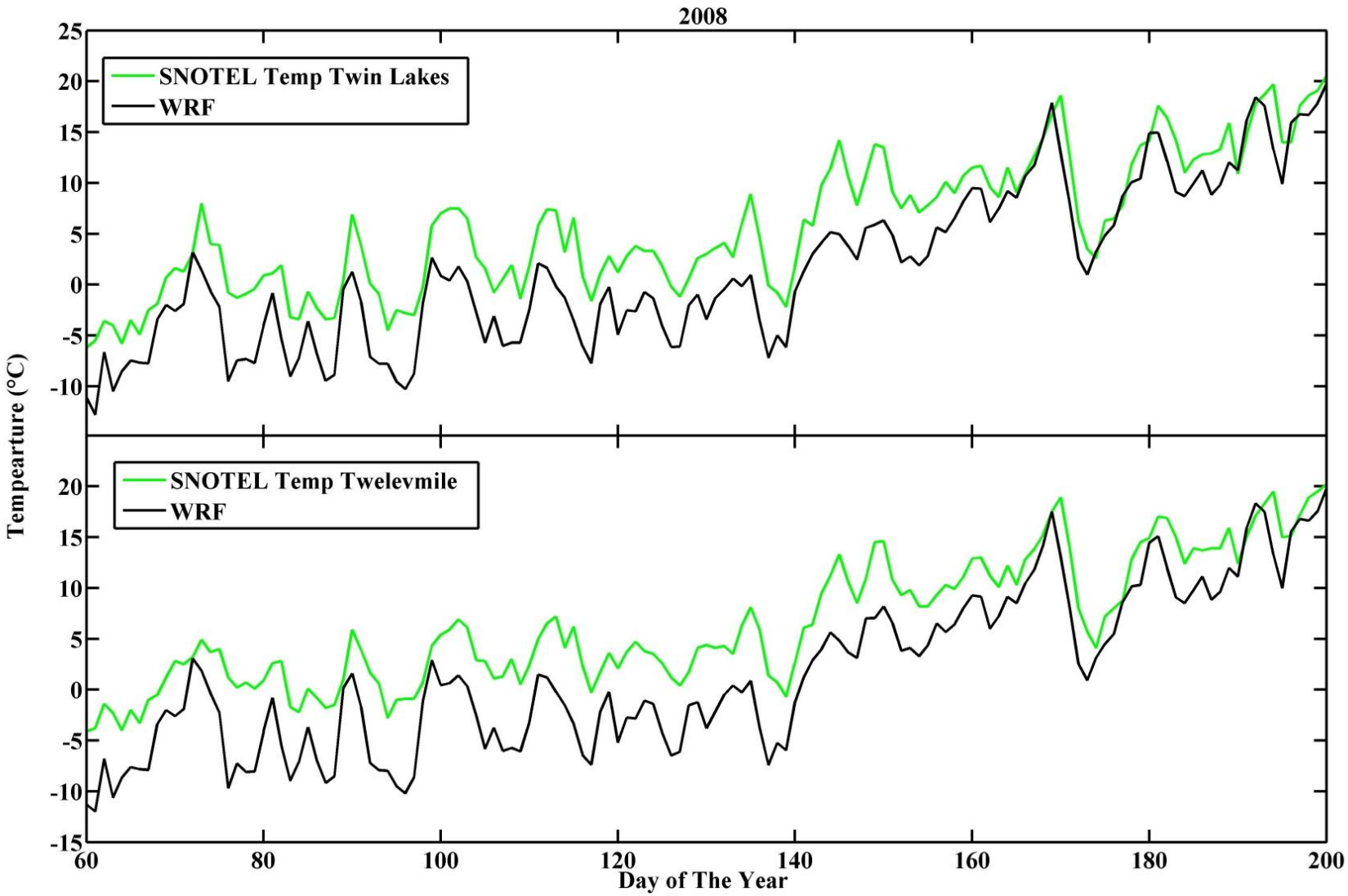

Figure 22A: Comparison of temperature recorded at Twin Lakes and Twelvemile SNOTEL stations and the simulated temperatures from the nearest WRF model pixels, for the year 2008, the wettest year of modeling.

7.3. Slope and SWE

A major component of understanding the spatial distribution of SWE was to observe the control of physiographic elements on the spatial distribution of SWE as way to further understand physical processes. One of the physiographic elements we analyzed was the linear dependence of SWE and slope. Slope is capable of driving SWE distribution in areas where avalanche activity is significant. In these instances, more SWE would be accumulating on lower angle slopes after being preferentially removed from higher angle slopes through avalanching. Our analysis showed that the linear dependence of SWE and slope is not of great importance within our study area (Figure 23A).

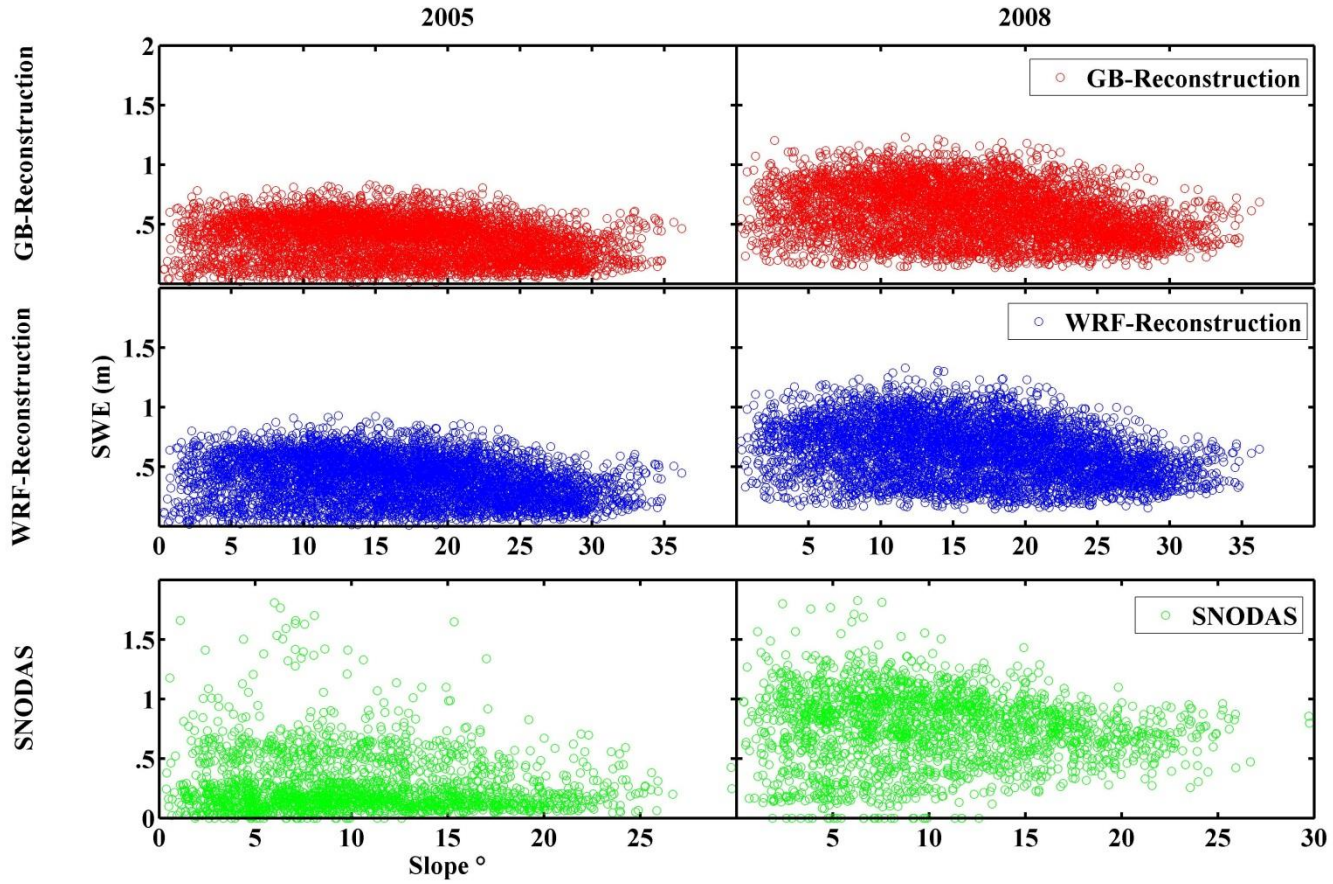

Figure 23A: Comparison of the linear dependence of SWE and slope for the years 2005 and 2008.