Key Points:
- A new technique for estimating average snow water storage is presented for North American mountains
- Comparisons suggest that current global model estimates underestimate mountain snow by >50%
- Mountains hold a disproportionate amount of snow water storage, highlighting the need for better mountain representation in modeling

Supporting Information:
- Supporting Information S1

Correspondence to:
M. L. Wrzesien,
wrzesien.1@osu.edu

A New Estimate of North American Mountain Snow Accumulation From Regional Climate Model Simulations

Melissa L. Wrzesien[1,2] (iD), **Michael T. Durand**[1,2] (iD), **Tamlin M. Pavelsky**[3] (iD), **Sarah B. Kapnick**[4] (iD), **Yu Zhang**[1], **Junyi Guo**[1], **and C. K. Shum**[1,5] (iD)

[1]School of Earth Sciences, Ohio State University, Columbus, OH, USA, [2]Byrd Polar and Climate Research Center, Ohio State University, Columbus, OH, USA, [3]Department of Geological Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, [4]Geophysical Fluid Dynamics Laboratory, National Oceanic and Atmospheric Administration, Princeton, NJ, USA, [5]Institute of Geodesy and Geophysics, Chinese Academy of Sciences, Wuhan, China

Abstract Despite the importance of mountain snowpack to understanding the water and energy cycles in North America's montane regions, no reliable mountain snow climatology exists for the entire continent. We present a new estimate of mountain snow water equivalent (SWE) for North America from regional climate model simulations. Climatological peak SWE in North America mountains is 1,006 km^3, 2.94 times larger than previous estimates from reanalyses. By combining this mountain SWE value with the best available global product in nonmountain areas, we estimate peak North America SWE of 1,684 km^3, 55% greater than previous estimates. In our simulations, the date of maximum SWE varies widely by mountain range, from early March to mid-April. Though mountains comprise 24% of the continent's land area, we estimate that they contain ~60% of North American SWE. This new estimate is a suitable benchmark for continental- and global-scale water and energy budget studies.

1. Introduction

Mountain ranges act as natural water towers (Viviroli et al., 2007), storing winter snowfall and releasing water in warm months to provide crucial resources for ecosystems, agriculture, and other human enterprises. Despite the importance of snow, global snow storage estimates are highly uncertain, particularly for complex topography (Mudryk et al., 2015). Even in comparatively well observed mountain ranges, observation networks are not dense enough to provide realistic areal-average information about snowpack characteristics in mountains (Dozier et al., 2016). Satellite-based observations of snow storage often perform poorly in mountains due to high uncertainties in deep or wet snow and in dense vegetation (Lettenmaier et al., 2015; Takala et al., 2011; Vuyovich et al., 2014). Snow physics schemes within land surface models (LSMs) require accurate precipitation forcing to perform adequately, but mountain precipitation measurements are both sparse (Renwick, 2014) and prone to undercatch bias (Pan et al., 2003).

Global and regional climate models (GCMs and RCMs, respectively) couple land and atmospheric physics to simulate precipitation; indeed, GCM reanalysis products provide most existing global snow storage estimates (Mudryk et al., 2015). These reanalyses often underestimate snow storage in mountain areas (Broxton et al., 2016; Snauffer et al., 2016; Wrzesien et al., 2017). Existing data products show some consistency in identifying snow water equivalent (SWE) anomalies but diverge in the absolute values of snow storage (Mudryk et al., 2015). Recently, there has been a push to adapt RCMs to accurately simulate mountain snow (see Table S1 in the supporting information). Given their smaller spatial domains, RCMs can be run at finer spatial resolution, allowing for potentially more realistic simulation of orographic precipitation (e.g., Leung & Qian, 2003). RCMs can afford detailed regionally optimized physics; for example, precipitation microphysics schemes have been developed to better handle snowflake hydrometeor shape (Thompson et al., 2008) and simulate multiple hydrometeor types simultaneously (Molthan & Colle, 2012), leading to more accurate precipitation estimation (Liu et al., 2011). RCM LSMs have also been developed with snowpack dynamics at sub-GCM grid scales in mind, creating multilayer snowpack to improve snow ablation (Etchevers et al., 2004), multiple canopy scenarios for snow-related energy balance calculations (Niu et al., 2011), and targeted forest/canopy dynamics for improved snow accumulation (Rutter et al., 2009). With these advances, RCMs have been shown to produce reasonable snow simulations in mountains (see Table S1).

The goal of this paper is to describe a new estimate of North American mountain snow accumulation produced using the Weather Research and Forecasting (WRF; Skamarock et al., 2008) RCM. Throughout, we compare our estimate with the existing snow climatology developed by Mudryk et al. (2015), extracted over mountain areas. At a point or a grid cell, SWE is the equivalent depth of liquid water contained in a snowpack (presented in millimeters). Here we use "snow water storage" (SWS) to refer to SWE summed over all pixels within a mountain range (presented in cubic kilometers). Wrzesien et al. (2017) suggest that a "reasonable" estimate of range-wide SWS is accurate to the first significant figure, that is, ±50%. However, evaluating accuracy of continental-scale estimates is problematic since in situ point observations measure (and represent) a small fraction of an RCM pixel (Molotch & Bales, 2005; Nolin, 2012). Here for completeness, we compare to in situ observations; however, we acknowledge that this comparison cannot be considered an effective means of evaluating accuracy. Thus, we also compare to satellite gravimetry data sets, which measure the change in total terrestrial water storage (TWS; snow, soil moisture, lakes, etc.) across the entire continent; gravimetry is not of high enough spatial resolution to separate individual ranges. We stress that neither the in situ nor gravimetry comparisons are meant as model validation, but to evaluate whether the WRF estimate is reasonable. To create a comparable estimate of snow storage in North America, we merge our mountain snow climatology with the Mudryk et al. (2015) climatology values for nonmountain areas. Our optimization approach essentially combines computationally expensive higher-resolution model simulations where topography is complex and computationally cheap lower-resolution model simulations where topography is simpler.

We propose a novel solution to address the high computational cost associated with modeling a typical climatology. Instead of producing a single 30+ year model simulation, we propose a "representative climatology" constructed from one representative year in each mountain range. While range-wide SWS varies dramatically from year to year, we show, based on existing data products, that the summation of individual representative years for each mountain range likely gives similar estimates of characteristic snow storage as a true long-term average climatology for the entire North American mountain system.

2. Data and Methods

2.1. WRF

In this study, we simulate mountain snowpack with WRF version 3.6, coupled to the Noah LSM with multiparameterization options (Niu et al., 2011). Details of the physics options are described by Pavelsky et al. (2011) and are similar to other WRF mountain studies (Ikeda et al., 2010; Liu et al., 2016; Rasmussen et al., 2011). Simulations are one-way nested domains, where the outer and inner domains have a spatial resolution of 27 and 9 km, respectively. Outermost boundaries are forced by ERA-Interim (Dee et al., 2011), while the 9 km domain is forced by the 27 km domain. We use ERA-Interim as the forcing data set since it covers all of North America, unlike the North American Regional Reanalysis (NARR, Mesinger et al., 2006), and it has been used in other large-scale WRF studies (Liu et al., 2016). Previous studies suggest resolutions of <10 km are needed to capture realistic orographic precipitation processes in mountainous regions (Ikeda et al., 2010; Kapnick & Delworth, 2013; Minder et al., 2016; Pavelsky et al., 2011; Rasmussen et al., 2011; Wrzesien et al., 2017). We compare our 9 km results to 3 km model simulations for four mountain ranges (Brooks, Cascades, Sierra Nevada, and U.S. Rockies) and find that peak SWS from the 9 km simulation is +5.5, −8.2, −13.0, and −3.2%, respectively, of the 3 km peak (Figure S1).

Due to computational demands of simulating across the entire continent, we do not perform one continuous multiple-year model run, as previous studies have done (Liu et al., 2016). Instead, we introduce a new method of approximating the climatology by simulating separate model domains for a single representative year specific to each mountain range. A representative year for each range is selected by identifying years with near-average SWS from 2001 to 2014 in two different global and continental reanalyses (see Text S1 for details). While global modeled products are prone to significant bias in alpine regions because of course spatial resolution, they capture interannual variability relatively well and are suitable for identifying representative years (Mudryk et al., 2015; Wrzesien et al., 2017). We performed the WRF simulations on both the National Aeronautics and Space Administration Pleiades supercomputer and the National Science Foundation Extreme Science and Engineering Discovery Environment system. All WRF simulations required ~1.8 million core hours and months of real time. Details for each WRF domain are in Table S2 and Figure S2.
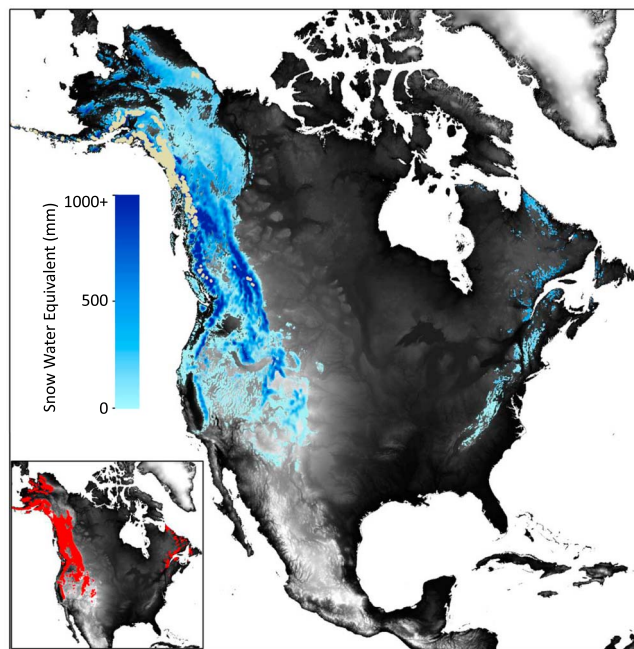
Figure 1. Peak snow water equivalent (SWE) for mountain areas in North America from separate representative year Weather Research and Forecasting simulations for each range. Tan areas indicate glaciers, for which Weather Research and Forecasting peak SWE values are not reported. Red areas in the inset map show the 30% of grid cells that comprise 75% of the SWE.

One potential disadvantage of our method is that we produce an estimate from sample years taken from 2001 to 2014 instead of a more traditional 30 year climatology. The tremendous computational cost of simulating all of North America at <10 km resolution prevented us from running WRF for more than 1 year for each mountain range. We performed an analysis to assess the suitability of our approach and found that for three independent global data products, the summation of SWS across representative years for each range is within ±7% of the 30 year average (see Text S1). We therefore conclude that our estimation methodology of simulating each mountain range for a single, representative year reasonably approximates a continental climatology at a fraction of the computational cost.

We chose WRF simulation domains such that they cover all mountains with seasonal snow in North America. The definition of "mountain" varies in the literature; here we used a definition from Kapos et al. (2000) based on elevation, slope, and local relief. We created a mask of North American mountains with seasonal snow at 1 km spatial resolution by intersecting the U.S. Geological Survey's Global 30 arc second elevation data set (https://lta.cr.usgs.gov/GTOPO30) with a map of seasonal snow cover derived from Moderate Resolution Imaging Spectroradiometer (Hall et al., 2006). We classified each mountain pixel as seasonal snow if it had an average duration of snow cover of greater than two consecutive months per year. More details are available in Text S2.

We exclude from analysis all grid cells classified as permanent snow/ice ("glaciers" hereafter) by the WRF land use scheme, which is based on the U.S. Geological Survey Global Land Cover Characterization (Anderson et al., 1976; Loveland et al., 2000). Thus, we focus the analysis on seasonally snow-covered regions—where accumulated snow melts completely during the warmer months—consistent with previous large-scale studies (e.g., Mudryk et al., 2015). Glaciers comprise approximately 2.9% of the mountain mask, with the majority of glaciers occurring in the Alaska and Coast Ranges (see Table S2 for glacier percentages and Figure 1 for glacier locations).

For each range, we present a map of peak SWE from the day of peak (DOP) SWS. Previous work suggests that range-wide and pixel-wise peak SWE values are comparable (Margulis et al., 2016). We also present maps of the DOP SWE, which are calculated pixel wise to capture how snow accumulation timing varies over and within each mountain range.

2.2. Comparison Data Sets and Methods
2.2.1. Automated Snow Pillows
We compare modeled SWE with ground observations from snow pillows. Daily SWE measurements are from the Snowpack Telemetry network (wcc.nrcs.usda.gov/snow), the California Department of Water Resources Data Exchange Center (cdec.water.ca.gov), the British Columbia Ministry of Environment (catalogue.data.gov.bc.ca/dataset/archived-automated-snow-station-data), and Alberta Environment and Parks (environment.alberta.ca/apps/basins/default.aspx). In total, there are 757 in situ points for comparison.

We match each ground station to its closest WRF grid cell in order to compare the SWE values. If multiple snow sensors fall within one 9 km WRF grid cell, we average together the observed SWE values. Due to differences in scale between a point measurement and a modeled grid cell (up to 6 orders of magnitude), we expect substantial scatter when comparing the model to snow pillows (Blöschl, 1999; Molotch & Bales, 2005; Nolin, 2012; see detailed explanation in Wrzesien et al., 2017). Indeed, differences in modeled and measured SWE can often be partially attributed to elevation differences between the station location and the grid cell average elevation (e.g., Pavelsky et al., 2011). Nonetheless, we include this comparison for completeness and because it is common practice in the literature (e.g., Arsenault et al., 2013; Cristea et al., 2014; De Lannoy et al., 2012; Leung & Qian, 2003; Rasmussen et al., 2011, 2014).

2.2.2. SNODAS

The Snow Data Assimilation System (SNODAS; Carroll et al., 2001; nsidc.org/data/G02158) is an operational, gridded product available over the continental United States at 30 arc second resolution (approximately 1 km at the equator) that provides daily estimates of SWE. We evaluate WRF SWE estimates by comparing maximum SWE values for each grid cell throughout the winter period from each product. For the comparison, we average together all SNODAS SWE values that lie within each WRF grid cell. The inclusion of SNODAS allows us to perform a more comprehensive comparison with WRF, as SNODAS, unlike snow pillows, is available continuously and at similar spatial scales as WRF.

2.2.3. CanSISE

We compare our SWE estimates to the Canadian Sea Ice and Snow Evolution Network (CanSISE) data ensemble, a 1° blended SWE product that covers the Northern Hemisphere from 1981 to 2010 (Mudryk & Derksen, 2017; Mudryk et al., 2015; available at nsidc.org/data/nsidc-0668). The CanSISE data product ("CanSISE" hereafter) estimates SWE by taking an ensemble mean approach to combine observations and model estimates of SWE from GlobSnow, ERA-Interim Land, Modern-Era Retrospective Analysis for Research and Applications, Crocus, and the Global Land Data Assimilation System. Full details validating continental snow (glacier zones are excluded) in CanSISE are described by Mudryk et al. (2015). We choose to not include higher-resolution CONUS data sets (e.g., North American Land Data Assimilation System, Mitchell et al., 2004; Xia, Mitchell, Ek, Cosgrove, et al., 2012; Xia, Mitchell, Ek, Sheffield, et al., 2012; NARR, Mesinger et al., 2006) since they do not cover the entire North American continent.

We construct a continent-wide snow accumulation estimate by combining the 30 year CanSISE SWE climatology estimates outside the mountains with our new WRF-derived mountain SWE estimates. Though coarse-resolution models do not produce reasonable SWE estimates in mountain areas, there is no evidence that CanSISE is biased in nonmountain areas (e.g., Snauffer et al., 2016).

2.2.4. Grace

In order to understand the role of SWE in the continental water storage budget, we compare our estimate of total North American SWE to the Gravity Recovery and Climate Experiment (GRACE; Tapley et al., 2004; Wahr et al., 2004; Syed et al., 2009, 2010) TWS anomaly observations for 2003 through 2010. Three different GRACE solutions are available: Center for Space Research at University of Texas, GeoforschungsZentrum Potsdam, and Jet Propulsion Laboratory. We use all three solutions in our TWS anomaly calculations to capture the mean and range of estimates.

For processing GRACE data, we use a decorrelation filter to diminish high-frequency geographically correlated errors (Duan et al., 2009) from the monthly geopotential solutions (the Level 2 product) and apply a correction of the glacial isostatic adjustment process using a model (Paulson et al., 2007); we also perform Gaussian smoothing and signal leakage correction along the coasts of North America (Guo et al., 2010; see Text S4 for additional discussion on GRACE processing). This processing results in the measurements of storage change (dS/dt), in millimeter per day, for the 8 year time series. We calculate the TWS volume by multiplying these changes by the study area, which in our case is North America excluding Greenland.

An additional consideration when using GRACE data is that any change in TWS will not be solely due to snow accumulation or melt. GRACE measures all changes in water mass, including soil moisture, lakes and reservoirs, and groundwater. Glacier mass changes, whether accumulation or ablation, will also be included in the long-term GRACE TWS estimates, unlike our WRF and CanSISE SWS estimates. Therefore, simultaneous increases in both the GRACE TWS anomaly and SWS do not mean that all changes are only due to snow accumulation.

3. Results and Discussion

Figure 1 shows the peak SWE for North American alpine regions. Across the entire continent, the average peak SWE for seasonally snow-covered mountain grid cells is 317 mm, with a standard deviation of 416 mm. The distribution of peak SWE values across all grid cells can be approximated by an exponential distribution ($f(x) = \lambda(e^{-\lambda x})$, where $\frac{1}{\lambda} = \mu = 316.7$), with values >4,000 mm in the Alaska Range and the Canadian Rocky Mountains. A small number of pixels dominate the total SWS, with 30.0% of the area containing 75.0% of the total SWE (see inset map in Figure 1). Neither latitude nor elevation explains the majority of the variance in peak SWE, whether considering all ranges together or ranges individually through multiple linear
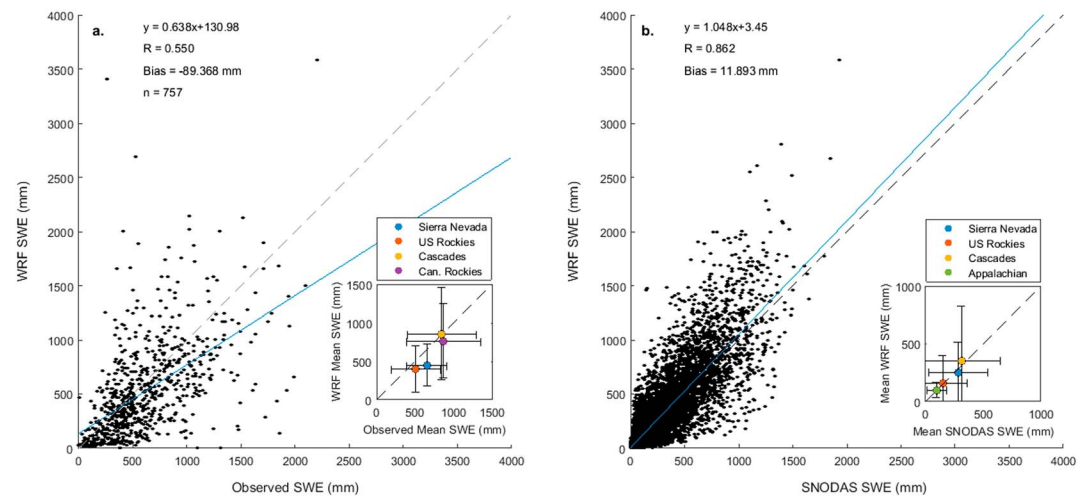
Figure 2. (a) Comparison of daily maximum snow water equivalent (SWE) from snow pillows with Weather Research and Forecasting (WRF) estimates. (a inset) Comparison of observed mean SWE and modeled mean SWE for the corresponding WRF grid cells for the four mountain ranges that have ≥20 snow pillows. Black lines indicate one standard deviation. (b) Comparison of WRF maximum SWE with maximum Snow Data Assimilation System (SNODAS) SWE for all mountainous grid cells in the continental United States. (b inset) Comparison of SNODAS mean SWE and WRF mean SWE for the four mountain ranges within the continental United States. Black lines indicate one standard deviation. For all figures, the dashed line is the one-to-one line and the blue line is the linear regression.

regression analysis. On average, latitude and elevation together explain ~30% of each mountain range's variation in peak SWE.

We also calculate the DOP SWE for each pixel (Figure S3) and the range-wide DOP SWS (Table S2). The pixel-wise mean DOP SWE is 24 March, with a standard deviation of 45 days (7 February to 8 May). DOP SWE occurs across a wide range of dates, generally from December to June (Figure S3). When averaged over the entire mountain range, the Brooks Range has the latest DOP SWE, and the Sierra Nevada has the earliest (Table S2). For all mountain pixels together, variance in DOP SWE is best explained by latitude ($R^2 = 0.27$), followed by grid cell peak SWE and elevation; more variance is explained ($R^2 = 0.42$) when all three variables are considered together.

Peak SWE and DOP SWE are somewhat correlated ($R^2 = 0.116$, $p < 0.01$), with higher elevations generally accumulating more snow and having a later DOP SWE. It is perhaps not surprising that elevation explains less of the variance in either peak SWE or DOP SWE, since once above tree line, snow accumulation tends to decrease due to wind effects. We also suggest that other variables, such as storm track and distance from moisture source, may have larger influences on where and when snow falls.

We compare WRF model estimates to 757 SWE observations from snow pillows located across western North America (Figure 2a) for representative years. WRF has a small negative bias, with a mean difference of −89 mm (14.7%) and an R^2 of 0.30. While useful as an initial assessment, comparing point-scale in situ measurements (~9 m^2 footprint) to grid cells many orders of magnitude larger (~81 km^2 footprint) is not a robust method for evaluation of gridded SWE in mountain regions (see Text S3 for additional reasoning for including snow pillow data in our comparisons). Figure 2a inset compares the average in situ and WRF peak SWE for the four ranges that have more than 20 snow pillows. Range-wide modeled and observed average SWE and standard deviation are similar for all four ranges.

Due to large differences in scale between snow pillows and WRF, we also compare WRF SWE to SNODAS estimates (Figure 2b). Compared to SNODAS, on a grid-cell-to-grid-cell basis WRF has a slight positive bias of 12 mm (~17%) and an R^2 of 0.74. At the range scale, WRF-SNODAS differences are −12%, 7%, 14%, and 1% for the Sierra Nevada, U.S. Rockies, Cascades, and Appalachians, respectively, with a mean bias of 2.5% (Figure 2b). The comparison to SNODAS provides a more robust estimate of range-wide uncertainty than does the snow pillow comparison, since SNODAS and WRF continuously cover the same spatial domain at comparable spatial resolution.
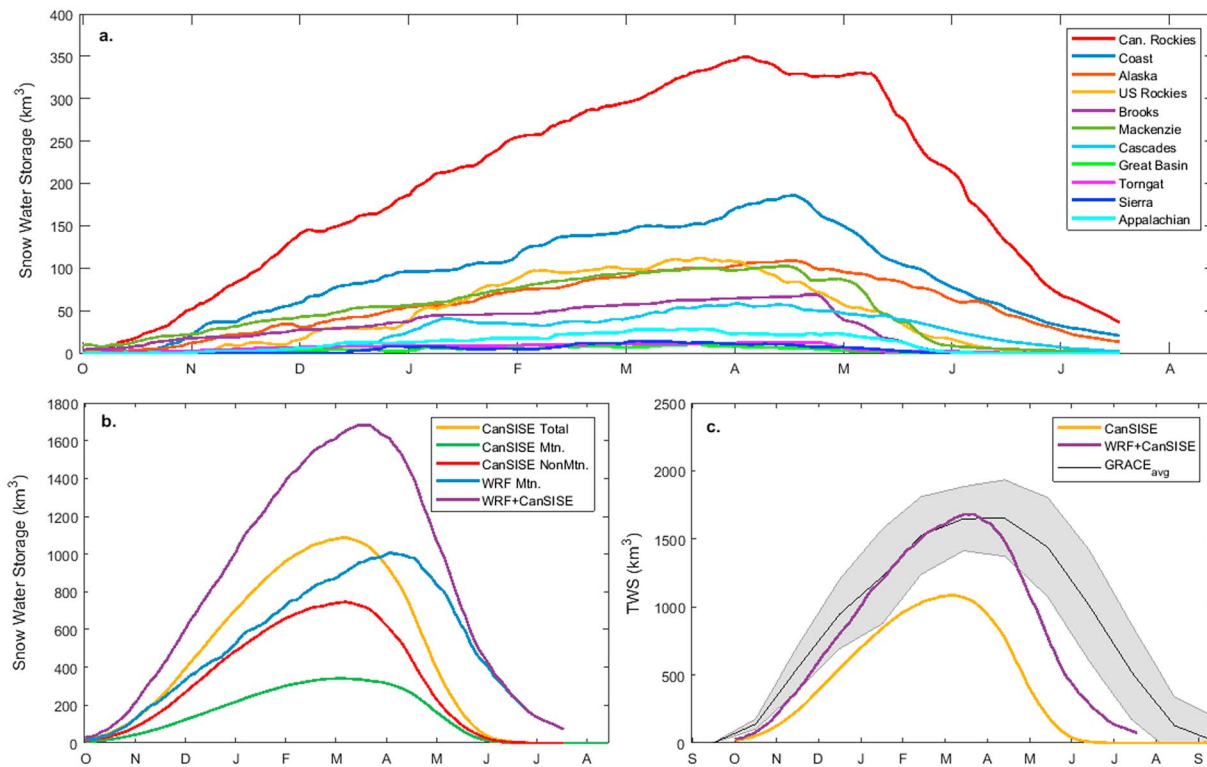
Figure 3. (a) Daily snow water storage (SWS) for each mountain range, in cubic kilometers. (b) Daily SWS (in cubic kilometers) for all of North America, as estimated by Weather Research and Forecasting (WRF) for mountains only (blue line), Canadian Sea Ice and Snow Evolution Network (CanSISE) for mountains only (green line), CanSISE for nonmountainous areas (red line), CanSISE total (yellow line), and WRF + CanSISE (purple line). (c) Comparison of continental water storage anomaly from Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage (TWS) (black line), CanSISE total snow accumulation (yellow line), and WRF + CanSISE (purple line). The shaded area is one standard deviation from the mean of the three GRACE solutions (Center for Space Research at University of Texas GeoforschungsZentrum Potsdam, and Jet Propulsion Laboratory).

Comparison of daily SWS time series, in cubic kilometers, in each mountain range (Figure 3a) shows that more than 50% of all mountain snow is located in the Coast and Canadian Rocky Mountains. Pacific coastal mountain ranges in western North America (Coast, Alaska, Cascades, and Sierra Nevada) have more snow per unit area (436 mm) at peak SWE than the western continental ranges (Brooks, Mackenzie, Canadian and U.S. Rockies, central Canada, and the Great Basin; 272 mm), although overall the continental ranges have more snow (644 km^3 for continental versus 367 km^3 for Pacific coast). The Canadian Rockies are an anomaly within the continental category, with the highest SWS and second highest normalized SWE (Table S2), despite being in the continental interior and separated from the Pacific moisture source by coastal mountains. Figure 3a also highlights that the DOP SWS differs from range to range, with more northern regions tending to have a later DOP SWS, in agreement with Figure S3. For example, the range-wide DOP SWS for the Alaska, Brooks, and Mackenzie Mountains is 17 April or later, while the Sierra Nevada, Great Basin, and U.S. Rocky Mountains reach peak SWS 22 March or earlier. Though we do not find a large correlation between DOP SWE and latitude in the pixel-wise analysis described above ($R^2 = 0.27$), when we compare range-scale DOP SWS and range-average latitude, we find a substantially stronger relationship ($R^2 = 0.88$, $p < 0.01$).

When combining the 11 individual mountain domains, the maximum SWS is 1,006 km^3 (Figure 3b). A linear regression model was used to estimate SWE over the masked-out glaciers, with latitude, longitude, and elevation as regression predictors (R^2 ranges from 0.16 to 0.36 with $p < 0.01$ for each of the five ranges with glacier cells). Inclusion of glacial regions would add an additional ~89 km^3 SWS to the mountain snow estimate.

The WRF peak mountain SWS estimate is 2.94 times larger than the CanSISE estimate of 342 km^3 (Figure 3b). Most striking, the WRF mountain estimate is nearly identical to the total continental storage in CanSISE (1,006 km^3 in WRF versus 1,087 km^3 in CanSISE), despite mountains covering ~25% of the total continent.

Previous studies indicated that coarse-resolution models like ERA-Interim, a component of CanSISE, fail to capture snow accumulation in mountain regions (Ikeda et al., 2010; Kapnick & Delworth, 2013; Pavelsky et al., 2011; Snauffer et al., 2016; Wrzesien et al., 2017), so the low estimate from CanSISE, produced at 1°, is not surprising.

Combining WRF mountain SWE with CanSISE nonmountain SWE produces a new North American snow accumulation estimate (WRF + CanSISE, hereafter), with a peak SWE of 1,684 km^3 (Figure 3b). Our new combined estimate is 55% larger than the CanSISE continent-wide estimate (1,087 km^3). Although mountains are approximately 25% of North American land area, 60% of all continental snow is located in the mountains based on WRF + CanSISE, compared to 31% from CanSISE alone. The original CanSISE estimate suggests that the proportion of North American snow stored in mountains is nearly equal to the proportion of mountain area, which implies that mountain snowpacks are not substantially deeper, on average, than nonmountain snowpacks. We have higher confidence in the WRF + CanSISE estimate because it is more consistent with the fact that mountains store substantially more snow per unit area than nonmountain areas (Mudryk et al., 2015; Snauffer et al., 2016; Sturm et al., 1995).

Finally, in order to understand the role of snow storage in the TWS budget, we compare the WRF + CanSISE SWS to the GRACE TWS anomaly (Figure 3c). TWS includes all variations in all components of water storage, not just snow, so the two are not directly comparable. We consider all of North America (minus Greenland), which resolves the issue of different spatial scales (9 km for WRF versus 1°, or ~300 km, for GRACE), since we integrate continental SWS from WRF + CanSISE and continental TWS from GRACE over the same region. The WRF + CanSISE SWS and the GRACE TWS anomalies are similar from October to mid-March but diverge thereafter. The maximum WRF + CanSISE SWS is similar to the maximum of the mean GRACE TWS estimate. This does not necessarily imply that WRF is too high but may suggest that much of the TWS increase during cold months is due to snow accumulation. The GRACE TWS estimate incorporates both positive and negative storage anomalies from non–snow-covered portions of the continent. Concurrent with the snow accumulation in northern and mountainous areas, some regions, such as the southeast United States, gain mass due to winter rainfall, while others, such as Mexico, lose mass due to winter drying, for example, from reductions in soil moisture; therefore, the positive and negative mass anomalies cancel out, leaving only the large snow signal. The late spring/summer lag between WRF + CanSISE SWS and GRACE TWS is consistent with the residence time on the continent for water stored as seasonal snow. As the snow changes phase, GRACE captures the movement of the snowmelt water through other storage reservoirs, while WRF + CanSISE SWS only shows the snow melting. Compared to other available estimates, the WRF SWS estimate therefore provides an upper bound on North American snow accumulation.

Here it is important to note that we do not use CanSISE, SNODAS, GRACE, or the snow pillows to validate WRF estimates. Rather, we compare WRF simulated SWE to other published estimates to assess whether the WRF values are reasonable. The challenge of working with gridded SWE data sets is that they are impossible to independently validate because there is no spatially continuous truth data set available. However, models like WRF and data sets like CanSISE provide important benchmark SWE estimates, and their consistency with other measures of SWE and the broader water cycle gives us confidence that they are reasonable. In the future, it may be possible to more completely validate these gridded data sets against spatially continuous observations from robustly ground-truthed airborne or satellite platforms. Though we cannot yet perform a full model validation, our results suggest that the estimates of North American snow storage presented here are more accurate than previous estimates and can better inform future hydrologic science.

4. Conclusions

We present a new estimate of SWS for North American mountains using the WRF RCM and a combined, continent-wide estimate of SWS by merging RCM results for the mountains with the CanSISE product everywhere else. Peak mountain SWS from our estimate is 2.94 times greater than the CanSISE mountain estimate. Comparison of WRF with CanSISE over the mountains shows that current global models often have a large negative bias in mountain snow, consistent with previous studies (e.g., Dutra et al., 2011; Kapnick et al., 2014; Snauffer et al., 2016; Wrzesien et al., 2017). This difference is substantial; incorrect estimates of snow accumulation lead to incorrect water budgets, which preclude accurate description of the water cycle itself (Rodell et al., 2015). Incorrect water budgets cause incorrect energy budgets, since the two are linked by

the changing phases of water (Stephens et al., 2012). Underestimation of snow not only impacts the understanding of the water budget from a purely scientific perspective but also influences efforts to quantify the economic importance of snow (Sturm et al., 2017).

References

Anderson, J. R., Hardy, E. E., Roach, J. T., & Witmer, R. E. (1976). *A land use and land cover classification system for use with remote sensor data. U.S. Geological Survey professional paper 964*. Reston, VA, USA: U.S. Geological Survey.

Arsenault, K. R., Houser, P. R., De Lannoy, G. J. M., & Dirmeyer, P. A. (2013). Impacts of snow cover fraction data assimilation on modeled energy and moisture budgets. *Journal of Geophysical Research: Atmospheres*, *118*, 7489–7504. https://doi.org/10.1002/jgrd.50542

Blöschl, G. (1999). Scaling issues in snow hydrology. *Hydrological Processes*, *13*(14-15), 2149–2175. https://doi.org/10.1002/(SICI) 1099-1085(199910)13:14/15%3C2149::AID-HYP847%3E3.0.CO;2-8

Broxton, P. D., Zeng, X., & Dawson, N. (2016). Why do global reanalyses and land data assimilation products underestimate snow water equivalent? *Journal of Hydrometeorology*, *17*(11), 2743–2761. https://doi.org/10.1175/JHM-D-16-0056.s1

Carroll, T., Cline, D., Fall, G., Nilsson, A., Li, L., & Rost, A. (2001). NOHRSC operations and the simulation of snow cover properties for the Conterminous U.S. In Proceedings of the 69th Annual Meeting of the Western Snow Conference, Sun Valley, ID, Western Snow Conference (14 pp.). Retrieved from www.westernsnowconference.org/sites/westernsnowconference.org/PDFs/2001Carroll.pdf

Cristea, N. C., Lundquist, J. D., Loheide, S. P., Lowry, C. S., & Moore, C. E. (2014). Modelling how vegetation cover affects climate change impacts on streamflow timing and magnitude in the snowmelt-dominated upper Tuolumne Basin, sierra Nevada. *Hydrological Processes*, *28*(12), 3896–3918. https://doi.org/10.1002/hyp.9909

De Lannoy, G. J. M., Reichle, R. H., Arsenault, K. R., Houser, P. R., Kumar, S., Verhoest, N. E. C., & Pauwels, V. R. N. (2012). Multiscale assimilation of advanced microwave scanning radiometer–EOS snow water equivalent and moderate resolution imaging spectroradiometer snow cover fraction observations in northern Colorado. *Water Resources Research*, *48*, W01522. https://doi.org/10.1029/2011WR010588

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/qj.828

Dozier, J., Bair, E. H., & Davis, R. E. (2016). Estimating the spatial distribution of snow water equivalent in the world's mountains. *WIREs Water*, *3*(3), 461–474. https://doi.org/10.1002/wat2.1140

Duan, X. J., Guo, J. Y., Shum, C. K., & van der Wal, W. (2009). On the postprocessing removal of correlated errors in GRACE temporal gravity field solutions. *Journal of Geodesy*, *83*(11), 1095–1106. https://doi.org/10.1007/s00190-009-0327-0

Dutra, E., Kotlarski, S., Viterbo, P., Balsamo, G., Miranda, P. M. A., Schär, C., et al. (2011). Snow cover sensitivity to horizontal resolution, parameterizations, and atmospheric forcing in a land surface model. *Journal of Geophysical Research*, *116*, D21109. https://doi.org/10.1029/2011JD016061

Etchevers, P., Martin, E., Brown, R., Fierz, C., Lejeune, Y., Bazile, E., et al. (2004). Validation of the energy budget of an alpine snowpack simulated by several snow models (SnowMIP project). *Annals of Glaciology*, *38*, 150–158. https://doi.org/10.3189/172756404781814825

Guo, J. Y., Duan, X. J., & Shum, C. K. (2010). Non-isotropic Gaussian smoothing and leakage reduction for determining mass changes Over land and ocean using GRACE data. *Geophysical Journal International*, *181*(1), 290–302. https://doi.org/10.1111/j.1365-246X.2010.04534.x

Hall, D. K., Salomonson, V. V., & Riggs, G. A. (2006). *MODIS/Terra snow cover monthly L3 global 0.05Deg CMG*. Version 5. MOD10CM. Boulder, CO: National Snow and Ice Data Center.

Ikeda, K., Rasmussen, R., Liu, C., Gochis, D., Yates, D., Chen, F., et al. (2010). Simulation of seasonal snowfall over Colorado. *Atmospheric Research*, *97*(4), 462–477. https://doi.org/10.1016/j.atmosres.2010.04.010

Kapnick, S. B., & Delworth, T. L. (2013). Controls of global snow under a changed climate. *Journal of Climate*, *26*(15), 5537–5562. https://doi.org/10.1175/JCLI-D-12-00528.1

Kapnick, S. B., Delworth, T. L., Ashfaq, M., Malyshev, S., & Milly, P. C. D. (2014). Snowfall less sensitive to warming in Karakoram than in Himalayas due to a unique seasonal cycle. *Nature Geoscience*, *7*(11), 834–840. https://doi.org/10.1038/ngeo2269

Kapos, V., Rhind, J., Edwards, M., Price, M. F., & Ravilious, C. (2000). Developing a map of the world's mountain forests. In M. F. Price, & N. Butt (Eds.), *Forests in sustainable mountain development: A state-of-knowledge report for 2000*, (pp. 4–9). Wallingford: CAB International.

Landerer, F. W., & Swenson, S. C. (2012). Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resources Research*, *48*, W04531. https://doi.org/10.1029/2011WR011453

Lettenmaier, D. P., Alsdorf, D., Dozier, J., Huffman, G. J., Pan, M., & Wood, E. F. (2015). Inroads of remote sensing into hydrologic science during the WRR era. *Water Resources Research*, *51*, 7309–7342. https://doi.org/10.1002/2015WR017616

Leung, L. R., & Qian, Y. (2003). The sensitivity of precipitation and snowpack simulations to model resolution via nesting in regions of complex terrain. *Journal of Hydrometeorology*, *4*(6), 1025–1043. https://doi.org/10.1175/1525-7541(2003)004%3C1025:TSOPAS%3E2.0.CO;2

Liu, C., Ikeda, K., Rasmussen, R., Barlage, M., Newman, A. J., Prein, A. F., et al. (2016). Continental-scale convection-permitting modeling of the current and future climate of North America. *Climate Dynamics*, *49*(1-2), 71–95. https://doi.org/10.1007/s00382-016-3327-9

Liu, C., Ikeda, K., Thompson, G., Rasmussen, R., & Dudhia, J. (2011). High-resolution simulations of wintertime precipitation in the Colorado headwaters region: Sensitivity to physics parameterizations. *Monthly Weather Review*, *139*(11), 3533–3553. https://doi.org/10.1175/MWR-D-11-00009.1

Long, D., Longuevergne, L., & Scanlon, B. R. (2015). Global analysis of approaches for deriving total water storage changes from GRACE satellites. *Water Resources Research*, *51*, 2574–2594. https://doi.org/10.1002/2014WR016853

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L., & Merchant, J. W. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, *21*(6-7), 1303–1330. https://doi.org/10.1080/014311600210191

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

Maussion, F., Scherer, D., Finkelnburg, R., Richters, J., Yang, W., & Yao, T. (2011). WRF simulation of a precipitation event over the Tibetan Plateau, China—An assessment using remote sensing and ground observations. *Hydrology and Earth System Sciences*, *15*(6), 1795–1817. https://doi.org/10.5194/hess-15-1795-2011

Meromy, L., Molotch, N. P., Link, T. E., Fassnacht, S. R., & Rice, R. (2013). Subgrid variability of snow water equivalent at operational snow stations in the western USA. *Hydrological Processes*, *27*(17), 2383–2400. https://doi.org/10.1002/hyp.9355

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., et al. (2006). North American regional reanalysis. *Bulletin of the American Meteorological Society*, *87*(3), 343–360. https://doi.org/10.1175/BAMS-87-3-343

Messerli, B., & Ives, J. D. (Eds) (1997). *Mountains of the world: A global priority*. New York and Carnforth: Parthenon.

Meybeck, M., Green, P., & Vörösmarty, C. (2001). A new typology for mountains and other relief classes: An application to global continental water resources and population distribution. *Mountain Research and Development*, *21*(1), 34–45. https://doi.org/10.1659/0276-4741(2001)021%5B0034,ANTFMA%5D2.0.CO;2

Minder, J. R., Letcher, T. W., & Skiles, S. M. (2016). An evaluation of high-resolution regional climate model simulations of snow cover and albedo over the Rocky Mountains, with implications for the simulated snow-albedo feedback. *Journal of Geophysical Research: Atmospheres*, *121*, 9069–9088. https://doi.org/10.1002/2016JD024995

Mitchell, K. E., Lohmann, D., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., et al. (2004). The multi-institution North American Land Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *Journal of Geophysical Research*, *109*, D07S90. https://doi.org/10.1029/2003JD003823

Molotch, N. P., & Bales, R. C. (2005). Scaling snow observations from the point to the grid element: Implications for observation network design. *Water Resources Research*, *41*, W11421. https://doi.org/10.1029/2005WR004229

Molthan, A. L., & Colle, B. A. (2012). Comparisons of single- and double-moment microphysics schemes in the simulation of a synoptic-scale snowfall event. *Monthly Weather Review*, *140*(9), 2982–3002. https://doi.org/10.1175/MWR-D-11-00292.1

Mudryk, L. R., & Derksen, C. (2017). *CanSISE observation-based ensemble of Northern Hemisphere terrestrial snow water equivalent, Version 2*. [North American Snow Water Equivalent]. Boulder, CO: NSIDC: National Snow and Ice Data Center. https://doi.org/10.5067/96ltniikJ7vd. [3 March 2017]

Mudryk, L. R., Derksen, C., Kushner, P. J., & Brown, R. (2015). Characterization of Northern Hemisphere snow water equivalent datasets, 1981–2010. *Journal of Climate*, *28*(20), 8037–8051. https://doi.org/10.1175/JCLI-D-15-0229.1

Niu, G. Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., et al. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research*, *116*, D12109. https://doi.org/10.1029/2010JD015139

Nolin, A. W. (2012). Perspectives on climate change, mountain hydrology, and water resources in the Oregon Cascades, USA. *Mountain Research and Development*, *32*(S1), S35–S46. https://doi.org/10.1659/MRD-JOURNAL-D-11-00038.S1

Pan, M., Sheffield, J., Wood, E. F., Mitchell, K. E., Houser, P. R., Schaake, J. C., et al. (2003). Snow process modeling in the North American Land Data Assimilation System (NLDAS): 2. Evaluation of model simulated snow water equivalent. *Journal of Geophysical Research*, *108*, 8850. https://doi.org/10.1029/2003JD003994

Paulson, A., Zhong, S., & Wahr, J. (2007). Inference of mantle viscosity From GRACE and relative sea level data. *Geophysical Journal International*, *171*(2), 497–508. https://doi.org/10.1111/j.1365-246X.2007.03556.x

Pavelsky, T. M., Kapnick, S., & Hall, A. (2011). Accumulation and melt dynamics of snowpack from a multiresolution regional climate model in the central Sierra Nevada, California. *Journal of Geophysical Research*, *116*, D16115. https://doi.org/10.1029/2010JD015479

Rasmussen, R., Ikeda, K., Liu, C., Gochis, D., Clark, M., Dai, A., et al. (2014). Climate change impacts on the water balance of the Colorado headwaters: High-resolution regional climate model simulations. *Journal of Hydrometeorology*, *15*(3), 1091–1116. https://doi.org/10.1175/JHM-D-13-0118.1

Rasmussen, R., Liu, C., Ikeda, K., Gochis, D., Yates, D., Chen, F., et al. (2011). High-resolution coupled climate runoff simulations of seasonal snowfall over Colorado: A process study of current and warmer climate. *Journal of Climate*, *24*(12), 3015–3048. https://doi.org/10.1175/2010JCLI3985.1

Renwick, J. (2014). MOUNTerrain: GEWEX mountainous terrain precipitation project. GEWEX News (Vol. 24, No. 4, pp. 5–6). Silver Spring, MD: International GEWEX Project Office. Retrieved from http://www.gewex.org/gewex-content/files_mf/1432213914Nov2014.pdf

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., et al. (2011). MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *Journal of Climate*, *24*(14), 3624–3648. https://doi.org/10.1175/JCLI-D-11-00015.1

Rodell, M., Beaudoing, H. K., L'Ecuyer, T. S., Olson, W. S., Famiglietti, J. S., Houser, P. R., et al. (2015). The observed state of the water cycle in the early twenty-first century. *Journal of Climate*, *28*(21), 8289–8318. https://doi.org/10.1175/JCLI-D-14-00555.1

Rodell, M., Famiglietti, J. S., Chen, J., Seneviratne, S. I., Viterbo, P., Holl, S., & Wilson, C. R. (2004). The Global Land Data Assimilation System. *Bulletin of the American Meteorological Society*, *85*(3), 381–394. https://doi.org/10.1175/BAMS-85-3-381

Rutter, N., Essery, R., Pomeroy, J., Altimir, N., Andreadis, K., Baker, I., et al. (2009). Evaluation of forest snow processes models (SnowMIP2). *Journal of Geophysical Research*, *114*, D06111. https://doi.org/10.1029/2008JD011063

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., et al. (2008). *A description of the advanced research WRF version 3. Technical Note 475*. Boulder, CO: National Center for Atmospheric Research.

Snauffer, A. M., Hsieh, W. W., & Cannon, A. J. (2016). Comparison of gridded snow water equivalent products with in situ measurements in British Columbia, Canada. *Journal of Hydrology*, *541*, 714–726. https://doi.org/10.1016/j.jhydrol.2016.07.027

Stephens, G. L., Li, J., Wild, M., Clayson, C. A., Loeb, N., Kato, S., & Andrews, T. (2012). An update on Earth's energy balance in light of the latest global observations. *Nature Geoscience*, *5*(10), 691–696. https://doi.org/10.1038/NGEO1580

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

Sturm, M., Holmgren, J., & Liston, G. E. (1995). A seasonal snow cover classification system for local to global applications. *Journal of Climate*, *8*(5), 1261–1283. https://doi.org/10.1175/1520-0442(1995)008%3C1261:ASSCCS%3E2.0.CO;2

Syed, T. H., Famiglietti, J. S., & Chambers, D. P. (2009). GRACE-based estimates of terrestrial freshwater discharge from basin to continental scales. *Journal of Hydrometeorology*, *10*(1), 22–40. https://doi.org/10.1175/2008JHM993.1

Syed, T. H., Famiglietti, J. S., Chambers, D. P., Willis, J. K., & Hilburn, K. (2010). Satellite-based global-ocean mass balance estimates of interannual variability and emerging trends in continental freshwater discharge. *Proceedings of the National Academy of Sciences*, *107*(42), 17,916–17,921. https://doi.org/10.1073/pnas.1003292107

Takala, M., Luojus, K., Pulliainen, J., Derksen, C., Lemmetyinen, J., Kärnä, J. P., et al. (2011). Estimating Northern Hemisphere snow water equivalent for climate research through assimilation of space-borne radiometer data and ground-based measurements. *Remote Sensing of Environment*, *115*(12), 3517–3529. https://doi.org/10.1016/j.rse.2011.08.014

Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F., & Watkins, M. M. (2004). GRACE measurements of mass variability in the Earth system. *Science*, *305*, 503–505. https://doi.org/10.1126/science.1099192

Thompson, G., Field, P. R., Rasmussen, R. M., & Hall, W. D. (2008). Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Monthly Weather Review*, *136*(12), 5095–5115. https://doi.org/10.1175/2008MWR2387.1

Viviroli, D., Dürr, H. H., Messerli, B., Meybeck, M., & Weingartner, R. (2007). Mountains of the world, water towers for humanity: Typology, mapping, and global significance. *Water Resources Research*, *43*, W07447. https://doi.org/10.1029/2006WR005653

Vuyovich, C. M., Henri, C. V., & Fern, D. (2014). Comparison of passive microwave and modeled estimates of total watershed SWE in the continental United States. *Water Resources Research*, *50*, 7206–7230. https://doi.org/10.1002/2013WR014956

Wahr, J., Swenson, S., Zlotnicki, V., & Velicogna, I. (2004). Time-variable gravity from GRACE: First results. *Geophysical Research Letters*, *31*, L11501. https://doi.org/10.1029/2004GL019779

Wrzesien, M. L., Durand, M. T., Pavelsky, T. M., Howat, I. M., Margulis, S. A., & Huning, L. S. (2017). Comparison of methods to estimate snow water equivalent at the mountain range scale: A case study of the California Sierra Nevada. *Journal of Hydrometeorology*, *18*(4), 1101–1119. https://doi.org/10.1175/JHM-D-16-0246.1

Wrzesien, M. L., Pavelsky, T. M., Kapnick, S. B., Durand, M. T., & Painter, T. H. (2015). Evaluation of snow cover fraction for regional climate simulations in the Sierra Nevada. *International Journal of Climatology*, *35*(9), 2472–2484. https://doi.org/10.1002/joc.4136

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., et al. (2012). Continental-scale water and energy flux analysis and validation for North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *Journal of Geophysical Research*, *117*, D03109. https://doi.org/10.1029/2011JD016048

Xia, Y., Mitchell, K., Ek, M., Cosgrove, B., Sheffield, J., Luo, L., et al. (2012). Continental-scale water and energy flux analysis and validation for North American Land Data Assimilation System project phase 2 (NLDAS-2): 2. Validation of model-simulated streamflow. *Journal of Geophysical Research*, *117*, D03110. https://doi.org/10.1029/2011JD016051