Advances in Water Resources 160 (2022) 104075

Contents lists available at ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

Combining ground-based and remotely sensed snow data in a linear regression model for real-time estimation of snow water equivalent

Kehan Yang [a,b,*], Keith N. Musselman [b], Karl Rittger [b,c], Steven A. Margulis [d], Thomas H. Painter [e], Noah P. Molotch [a,b,f]

[a] Department of Geography, University of Colorado Boulder, Colorado, USA
[b] Institute of Arctic and Alpine Research, University of Colorado Boulder, Colorado, USA
[c] Earth Research Institute, University of California, Santa Barbara, CA, USA
[d] Department of Civil and Environmental Engineering, University of California Los Angeles, CA, USA
[e] Joint Institute for Regional Earth System Science and Engineering, University of California Los Angeles, CA, USA
[f] Jet Propulsion Laboratory, California Institute of Technology, CA, USA

ARTICLE INFO

Keywords:
Snow water equivalent
Snow-covered area
Real-time SWE modeling
Statistical-learning
Sierra Nevada snowpack
Remote Sensing

ABSTRACT

Effective water resources management in California relies substantially on real-time information of snow water equivalent (SWE) at basin scale and mountain ranges given that mountain snowpacks provide the primary water supply for the State. However, SWE estimation based solely on remote sensing, modeling, or ground observations does not meet contemporary operational requirements. In this context, this study develops a data-fusion framework that combines multi-source datasets including satellite-observed daily mean fractional snow-covered area (DMFSCA), snow pillow SWE measurements, physiographic data, and historical SWE patterns into a linear regression model (LRM) to improve SWE estimates in real-time. We test two LRMs: a baseline regression model (LRM-baseline) that uses physiographic data and historical SWE patterns as independent variables, and an FSCA-informed regression model (LRM-FSCA) that includes the DMFSCA from Moderate Resolution Imaging Spectroradiometer (MODIS) satellite imagery as an additional independent variable. By incorporating the satellite-observed DMFSCA, LRM-FSCA outperforms LRM-baseline with increased median R^2 from 0.54 to 0.60, and reduced median PBIAS of basin average SWE from 2.6% to 2.2% in the snow pillow SWE cross-validation. LRM-FSCA explains 87% of the variance in the snow course SWE measurements with 0.1% PBIAS, while LRM-baseline explains a lower 81% variance with 1.4% PBIAS, both of which show higher accuracy than SWE estimates from the two operational SWE datasets: the Snow Data Assimilation System (SNODAS, 73% and -2.4%, respectively) and Nationtional Water Model (NWM, 75% and -15.9%, respectively). Additionally, LRM-FSCA explains 85% of the median variance in the Airborne Snow Observatory SWE with -9.2% PBIAS, which is comparable to the LRM-baseline (86% and -11.3%, respectively) and considerably better than SNODAS (64% and 28.2%, respectively) and NWM (33% and -30.1%, respectively). This study shows a substantial model improvement by constraining the geographical and seasonal variation on snow-cover via satellite observation and highlights the values of using multi-source observations in real-time SWE estimation. The developed SWE estimation framework has crucial implications for effective water supply forecasting and management in California, where climate extremes (e.g., droughts and floods) require particularly skillful monitoring practices.

1. Introduction

Seasonal snow covers over 30% of the Earth's land surface and provides the water supply for approximately one-sixth of the global population (Dozier, 1989; Barnett et al., 2005). In the western United States, meltwater from seasonal snowpack contributes 50% to 80% of annual runoff (Stewart et al., 2004; Li et al., 2017). Snowmelt runoff recharges groundwater aquifers and surface water reservoirs, providing a vital water source for urban and agricultural areas (Barnett et al., 2005; Mote et al., 2005; Bales et al., 2006). Mountain snowpack strongly influences the timing and magnitude of streamflow (Hamlet et al., 2005; Lundquist et al., 2015), and thus accurate real-time information of the

* Corresponding author at: Department of Geography, Institute of Arctic and Alpine Research, University of Colorado Boulder, Colorado, USA.
E-mail address: Kehan.Yang@Colorado.EDU (K. Yang).

https://doi.org/10.1016/j.advwatres.2021.104075
Received 9 May 2021; Received in revised form 3 October 2021; Accepted 30 October 2021
Available online 2 November 2021

spatial variability of mountain snowpack is particularly important for effective water management and water supply forecasting (He et al., 2016). One traditional method to forecast water supply in California relies on *in-situ* observations of snow water equivalent (SWE), including an automatic snow pillow network providing real-time SWE measurements, and extensive manual snow courses taken at the beginning of each month during snow seasons (Pagano et al., 2004; He et al., 2016). However, these point-scale and transect-scale ground observations inadequately capture the spatial heterogeneity of snowpack distribution in mountainous terrain (Molotch and Bales, 2005; Bales et al., 2006). Recent studies indicate that spatial patterns of snow accumulation and melt are shifting due to climate change, further motivating improved understanding of the processes controlling streamflow generation (Nijssen et al., 2001; Barnett et al., 2005; Godsey et al., 2014; Mote et al., 2018). Therefore, to better support water management decision-making, it is imperative to have accurate real-time spatially distributed SWE information.

Many efforts have been made to estimate spatially distributed SWE in (near) real-time (i.e., with a one- or two-day processing lag time), yet all approaches have limitations when applied over large-scale mountainous regions (Dozier et al., 2016). For the *in-situ* SWE observations, the Natural Resource Conservation Service (NRCS) has operated the SNOw TELemetry (SNOTEL) network since the mid-1960s (Serreze et al., 1999). There are more than 800 SNOTEL sites located in the mountainous regions across the western United States, providing SWE measurements automatically every 15 minutes. Additionally, the California Legislature estabilished the California Cooperative Snow Surveys (CCSS) program in 1929 with alliances in more than 50 state, federal, and private agencies (Roos, 2004). They collect and analyze snow data from snow courses, and now there are more than 200 snow courses located throughout the Sierra Nevada. These ground-based SWE measurements have been interpolated into spatially distributed SWE based on their relationships with predictor variables (e.g., elevation, aspect, slope, solar radiation, wind, etc.) using spatial interpolation or regression methods (Carroll et al., 1999; Fassnacht et al., 2003; Fassnacht et al., 2012; Molotch et al., 2005). However, the accuracy of these statistcical methods is largly hindered by the high heterogeneity of SWE distribution over mountainous terrian, the observation desnity, and the representativeiness of the predict variables.

In addition to the interpolation or regression of *in-situ* SWE observations, remotely sensed passive microwave (PM) data have provided global SWE observations over the past three decades (Takala et al., 2017). Yet, applications of PM SWE data products are limited by many factors such as the mixed pixel problem at coarse spatial resolutions (e.g., 25 km), the saturation of the PM signal over deep snow, the many-to-one relationships between the PM signal, snow grain size, and SWE, and the impacts of forest cover (Durand and Margulis, 2007; Vander Jagt et al., 2013). While fusing ground-based snow depth observations could increase the spatial resolution of PM SWE data products (e.g., 5 km, Takala et al., 2017), it remains challenging to overcome the aforementioned issues, and thus PM-based SWE data products often exclude mountainous regions. Snow models, hydrological models, and land-surface models are also available for real-time SWE estimation. However, these models often show poor performance in SWE estimation partly because of the high uncertainty of precipitation model forcing which is particularly problematic in mountainous regions (Milly and Dunne, 2002; Adam et al., 2006). Other operational models, like the Snow Data Assimilation System (SNODAS) from the NOAA National Weather Service's National Operational Hydrologic Remote Sensing Center (NOHRSC) are usually available in (near) real-time (Carroll et al., 2001; Barrett, 2003). Clow et al. (2012) evaluated SNODAS with ground snow survey data in the Colorado Rocky Mountains and reported that SNODAS performed well in forested areas with 77% of the variance in SWE explained, but exhibited high uncertainties over alpine regions in which only 30% of the variance in SWE was explained because of high wind redistribution.

Despite the listed challenges, recent progress has been made in real-time SWE estimation by combining historical SWE estimates developed via reanalysis with statistical models (Schneider and Molotch, 2016; Bair et al., 2018; Zheng et al., 2018). In this context, reanalyses of SWE distribution using SWE reconstruction models and data assimilation techniques have been promising for SWE modeling over large-scale mountainous regions (Durand et al., 2008; Molotch, 2009; Guan et al., 2013; Girotto et al., 2014; Margulis et al., 2016; Rittger et al., 2016; Bair et al., 2016, 2018). These methods outperforms PM and SNODAS by the exploitation of satellite-based observations of snow cover depletion and/or snow albedo (Bair et al., 2019). Notwithstanding, these methods cannot be used for real-time SWE estimation given the retrospective calculation process. Therefore, to improve real-time SWE estimation, one effort has coupled the historical SWE from reconstruction models as an independent predictor for statistically-based SWE estimation (Schneider and Molotch, 2016; Bair et al., 2018; Zheng et al., 2018). Because snow distribution is often controlled by the interactions between meteorology and local constant topographic characteristics, snow patterns exhibit considerable consistency from year to year (Deems et al., 2008; Mendoza et al., 2020; Pflug and Lundquist, 2020). In other words, the historical SWE data usually contains relevant information regarding the SWE patterns of interest (e.g., real-time SWE). By incorporating this relevant pattern from the historical reconstructed SWE data, real-time SWE estimation has been substantially improved in some large mountain ranges, such as Upper Colorado River Basin (UCRB) (Schneider and Molotch, 2016) and a few basins in California Sierra Nevada (Zheng et al., 2018).

To evaluate modeled SWE accuracy, early studies have used many independent validation datasets including ground-based onservations and the Airborne Snow Observatory (ASO) SWE products. The ground-based SWE observations, particularly snow pillow and snow course data, represent the most accurate SWE information at point-scale (i.e., snow pillow) and transect-scale (i.e., snow course), and have thus been used as validation datasets in many previous research (Erickson et al., 2005; Franz et al., 2008; Raleigh and Lundquist, 2012; Margulis et al., 2015; Wrzesien et al., 2019). However, a limitation of these datasets is that they cannot fully capture the spatial variability of SWE over an entire watershed (Clark et al., 2011). Moreover, these sites are often located at easily accessible locations, further limiting their representativeness for evaluating spatially complex SWE estimates. Another commonly used validation dataset is based on the interpolation of intensive snow surveys (Molotch et al., 2005; Durand and Margulis, 2007; Guan et al., 2013; Meromy et al., 2013; Schneider and Molotch, 2016). Yet, the high logistical cost and feasibility of these surveys prevent its operational applications over larger areas.

The NASA ASO mission and now ASO Inc., which was created to transfer the technology for implementation around the globe, provides SWE estimates with unprecedented accuracy from year to year in several selected basins in the California Sierra Nevada and elsewhere in the western United States (Painter et al., 2016). The ASO aircraft carries a light detection and ranging (LiDAR) sensor and a visible through near-infrared imaging spectrometer to measure snow-on and snow-off

surface heights that are differenced to derive estimates of snow depth. By leveraging snow density calculated by a physically-based snow model (iSnobal) (Marks et al., 1999) constrained by *in-situ* measurements, ASO has derived very accurate SWE estimates at a high spatial resolution (50 m) (Painter et al., 2016). The detailed description of ASO SWE product and its uncertainty analysis is well documented in Painter et al. (2016). While the evaluation methods might be different, the magnitude of ASO SWE uncertainty is far less than that of the SWE errors from aforementioned SWE estimation studies, like SWE interpolation and/or regression (Fassnacht et al., 2003; Schneider and Molotch, 2016), reconstruction (Guan et al., 2013; Rittger et al., 2016) and data assimilation (Margulis et al., 2016). Thus, ASO SWE is increasingly used as ground 'truth' information to evaluate other SWE data products at large basin scales (Bair et al., 2016; Yang et al., 2018; Oaida et al., 2019). ASO SWE outperforms the traditional small-scale ground-based SWE observations by providing spatially complete SWE estimates over large watersheds at a high spatial resolution (Painter et al., 2016; Bormann et al., 2018). Yet, one limitation of using ASO SWE product in validation is that it only has a relatively short perioed (2013-present) compared to the snow pillow (mid-1960s) and snow survey (mid-1930s) datasets. Additionally, airborne campaigns are mostly conducted during the snow ablation period near and after peak SWE and for only a few selected watersheds, limiting our understanding of snow accumulation and SWE in time and space. Thus, using ASO data alone cannot comprehensively evaluate the SWE accuracy in the Sierra Nevada for watersheds and periods when the ASO data are not available.

The objectives of this study are to improve real-time SWE estimation over all basins in the California Sierra Nevada and to provide a comprehensive evaluation of the SWE estimates relative to other real-time SWE products currently produced operationally. Built on previous work, we developed a statistically-based data-fusion framework to estimate SWE in real-time, which combines ground snow pillow SWE measurements, physiographic data, satellite-observed daily mean fractional snow-covered area (DMFSCA), physically-based historical SWE patterns with a linear regression model (LRM). Previous studies have used remotely sensed FSCA data to inform hydrological models in many applications, where the SWE is often inferred from snow cover depletion curves (Liston, 1999; Gómez-Landesa and Rango, 2002; Andreadis and Lettenmaier, 2006; Clark et al., 2006; Niu and Yang, 2007). However, FSCA is not directly related to SWE and FSCA can remain constant as SWE changes from day to day rendering FSCA on a single day less useful in predicting SWE. Therefore, we introduce a new variable DMFSCA, which is calculated as the mean of time-series FSCA from the beginning of each water year to the date of simulation, as a predictor variable on SWE estimation. DMFSCA is sensitive to the magnitude of snow accumulation (i.e., increases DMFSCA) and ablation processes (i.e., decreased DMFSCA) as it describes the historical temporal variability of snow cover. For example, when pixels are completely covered by snow (i.e., FSCA equal 100%) for a modeling day, the DMFSCA value in a pixel with seasonal snowpack (e.g., high elevations with continuous snow accumulation) would be greater than in a pixel with an ephemeral snowpack (i.e., low elevations with frequent snowmelt), which are more likely to correspond with the SWE distribution than the FSCA values for these two pixels (both equal 100%). Using DMFSCA also has a low-pass filtering effect that reduces the influences of the noise in FSCA data for SWE modeling. Although DMFSCA does not provide direct SWE information, it is potentially useful as an explanatory variable in a statistical regression to explain the spatial distribution of SWE.

To examine the influence of satellite-derived DMFSCA for SWE estimation, we compared two LRMs in this study: a baseline regression model (LRM-baseline) in which the dependent variables included physiographic variables and historical SWE patterns from an SWE reanalysis dataset (Margulis et al., 2016), and an FSCA-informed SWE regression model (LRM-FSCA) in which the DMFSCA derived from MODIS satellite imagery (Rittger et al., 2020) was included as an additional independent variable. Two other SWE datasets were also included in comparison to provide a baseline of SWE accuracy in the operational applications: SWE estimates from SNODAS (Carroll et al., 2001; Barrett, 2003) and from the National Water Model (hereafter, NWM-SWE) retrospective runs version 1.2 operated by the National Oceanic and Atmospheric Administration (Gochis et al., 2018). We present the inter-model differences in terms of their accuracy, interannual variability, and spatial patterns. Section 2 describes the study area and Section 3 describes the datasets, the LRM framework, and the evaluation methods. The results are presented in Section 4, with the discussion and conclusion given in Sections 5 and 6, respectively.

2. Study area

The study area covers 20 major snow-dominated watersheds (Hydrological Unite Code 8, HUC8) in the Sierra Nevada California, with a total area of 49,409 km^2 (Fig. 1). Given ephemeral snowpacks variations below 1500 m elevation in the Sierra Nevada (Bales et al., 2006; Guan et al., 2013; Margulis et al., 2016; Rittger et al., 2016), the modeling domain is cutoff to areas above 1500 m, of which the average elevation is 2270 m ranging from ~1500 m to ~4410 m. The Sierra Nevada has a Mediterranean climate, with distinct dry summers and wet winters. The average annual temperature is approximately 10 °C and the total annual precipitation is approximately 830 mm estimated from Parameter-elevation Regressions on Independent Slopes Model (PRISM) 30-year normals (Daly et al., 1994). 80% of the annual precipitation occurs in the cold months from October to May mainly in the form of snow (Swain et al., 2016). Influenced by the prevailing west winds, the west-facing windward slopes of the Sierra Nevada capture the majority of the moisture from the Pacific Ocean, generating a large amount of orographic precipitation, while the eastern leeward slopes have more arid climate conditions. Based on National Land Cover Database 2011 (NLCD2011), evergreen needle-leaf forest (~38%) and shrub-lands (~36%) are the two major land cover types in the Sierra Nevada, with most of the forest cover on the western slopes of the Sierra, where the annual precipitation and temperature are both relatively high (Trujillo et al., 2012; Hansen et al., 2013).

The ASO mission (details on https://www.jpl.nasa.gov/missions/airborne-snow-observatory-aso/) began snow survey flights of the upper Tuolumne River Basin (TRB, Fig. 1) in 2013, and routinely provides SWE data at a 50 m resolution during the snowmelt seasons. The upper TRB is located on the west side of the Sierra Nevada with a total area of ~1200 km^2. The average elevation of the upper TRB is about 2680 m, with a wide elevation range from 1500 m to ~3970 m, a range similar to the entire Sierra Nevada. About 40% of the TRB is above treeline, which is proportionally much more alpine land area than the Sierra Nevada study area on the whole. Evergreen needle-leaf forest (~37%) and shrublands (~48%) are the two dominant land cover types in the upper TRB (Hansen et al., 2013). The average annual temperature and total precipitation are about 4 °C and 1260 mm, respectively (based on PRISM 30-year normals, Daly et al. (1994)). Snowmelt runoff from the TRB drains into the Hetch Hetchy reservoir, providing the primary water resource to the City of San Francisco and other Bay Area municipalities.

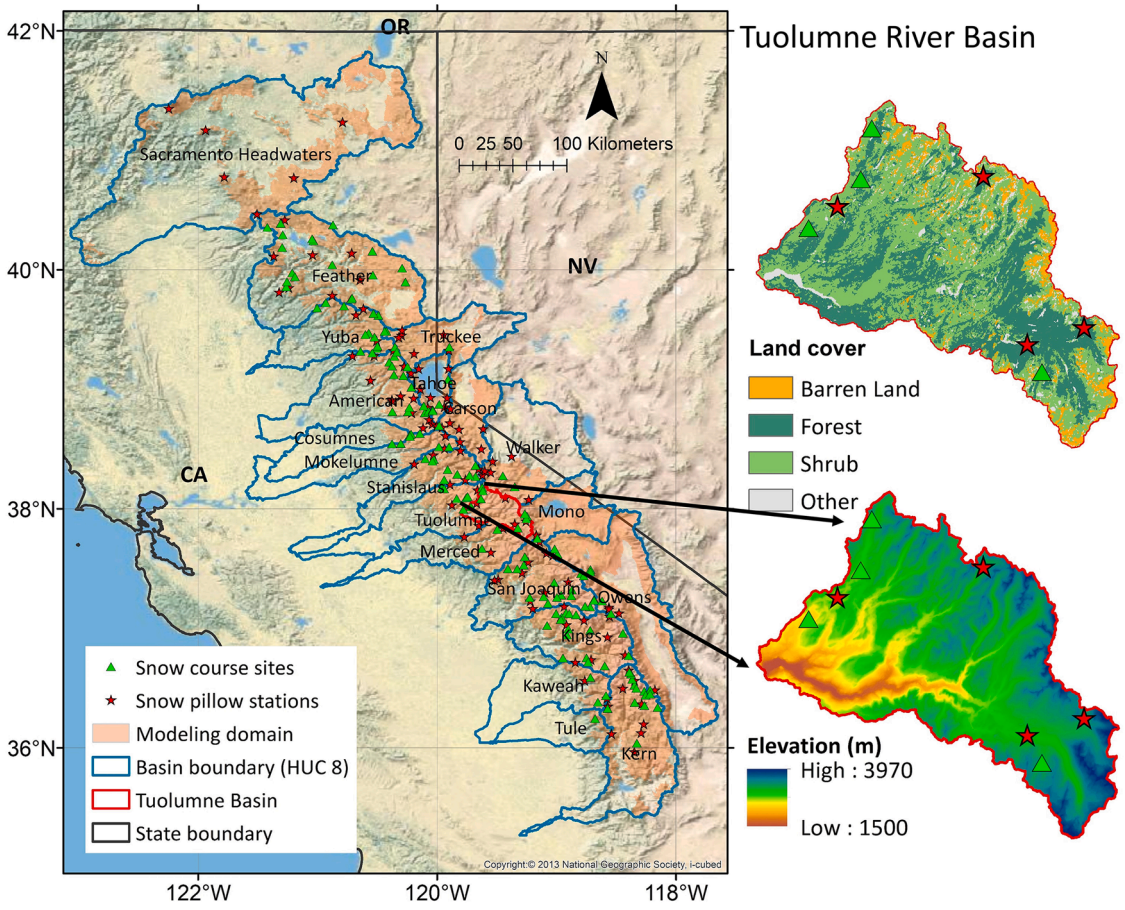

Fig. 1. Overview of the Sierra Nevada with the locations of snow pillows (red stars) and snow courses (green triangles) marked on the map. 20 watersheds (HUC8) are highlighted by the blue lines and the upper Tuolumne River Basin (TRB) highlighted by red lines is shown in the zoom views. The major land cover types in the upper TRB are displayed based on National Land Cover Dataset 2011 (NLCD2011). The study domain where the LRM runs is shaded in light orange (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

3. Methods and datasets

The LRM relies on the lasso and elastic-net regularized Generalized Linear Model (GLMNET, details on https://glmnet.stanford.edu/articles/glmnet.html) to approximate the relationship between the independent variables (listed in Table 1) and the dependent variable (i.e., SWE observations from ground snow pillow stations). Fig. 2 illustrates the schematic of the LRM with a detailed description in Sections 3.1 and 3.2. To examine the impacts of satellite-observed DMFSCA on model performance, we run two LRMs at a daily time step at 500-m

Table 1
Summary of the independent variables used in the LRMs and their importance on snow pillow SWE observations, with a detailed description of the methods used to calculate the physiographic variables. The importance of each variable is represented by the median of the squared Pearson-correlation coefficient (R^2) between FSCA-scaled snow pillow SWE (i.e., dependent variable) and each predictor variable for all simulations from the water year 2001 to 2019. Only significant correlations (p-value < 0.05) are summarized in this table.

Variable name	Description	Importance (median R^2)
Elevation	Derived from SRTM DEM	0.12
Latitude	Derived from SRTM DEM, the center of the grid	0.11
Longitude	Derived from SRTM DEM, the center of the grid	0.14
Northness	COSINE of the aspect calculated using SRTM DEM	0.09
Eastness	SINE of the aspect calculated using SRTM DEM	0.07
Slope	SINE of the slope calculated using SRTM DEM	0.08
Regional northness	COSINE of the regional aspect for every 4 × 4 km region	0.07
Regional eastness	SINE of the regional aspect for every 4 × 4 km region	0.08
Regional slope	SINE of the regional slope for every 4 × 4 km region	0.09
W/NW/SW distance to the ocean	Distance to the ocean calculated from west, northwest, and southwest	0.12/0.13/0.11
W/NW/SW barrier height	The elevation differences between the highest barrier and the grid in three directions of the ocean to the grid	0.16/0.14/0.18
W/NW/SW barrier distance	The distances from the highest barrier to the grid in three directions of the ocean to the grid	0.09/0.08/0.11
Historical SWE pattern	The historical SWE reanalysis with the closest SWE pattern compared with that on the target simulation day (details in Section 3.1.2)	0.50
Daily Mean FSCA (DMFSCA)	The mean value of daily fractional snow-covered area from the beginning of each water year (i.e., October 1st) to the date of the simulation	0.24

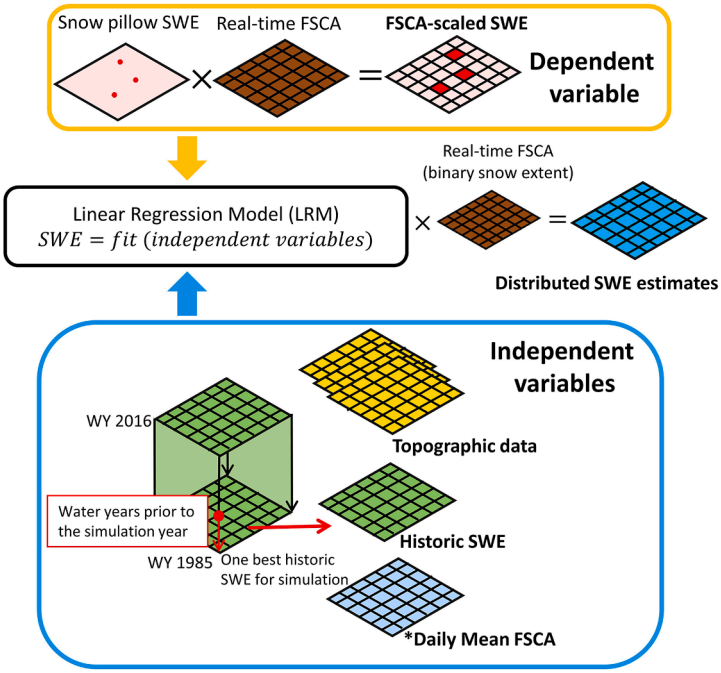

Fig. 2. Flowchart of the Linear Regression Model (LRM) for real-time SWE estimation. The dependent variable of the LRM is the snow pillow SWE observations, which is scaled by satellite-observed FSCA to the grid-cell scale. The independent variables (Table 1) of the LRM-baseline include 18 physiographic variables derived from the digital elevation model and one best-matched historical SWE reanalysis. The LRM-FSCA model includes the *Daily Mean FSCA (DMFSCA) as an additional independent variable. To mimic the real-time operational simulation, only one best-matched historical SWE from WYs before the simulation WY is used in the LRM. A binary snow extent derived from the FSCA data is used to determine the snow-covered areas for the final distributed SWE estimation.

spatial resolution (i.e., the same resolution as used by Schneider and Molotch (2016)) during the period when there is significant snow on the ground (i.e., from January to June) over the water year (WY; October 1–September 30) 2001 through 2019 across the modeling domain (Fig. 1): hereafter termed LRM-baseline and LRM-FSCA. Both simulations were evaluated through three methods including (1) cross-validation at snow pillow stations, (2) comparison against snow course measurements across the Sierra, and (3) comparison against the spatially distributed ASO SWE data in the upper TRB. The performance of SNODAS and the NWM-SWE were also included in the latter two evaluations (i.e., against snow course SWE and ASO SWE) to compare the performance of these complementary models relative to those developed herein. Given that snow pillow SWE observations were used in some of the presented models (i.e., LRMs and SNODAS), snow pillow SWE observations were not used for model comparison as they are not independent from model input data in these circumstances.

3.1. Data sources

3.1.1. Dependent variable

The SWE training samples for the LRM were acquired from 113 snow pillow stations across the study area (Fig. 1). The snow pillow stations over 1500 m span the elevation range from 1570 m to 3475 m with more sites located in the mid-elevations than lower and higher elevations (Fig. 3). The average elevation of snow pillow stations is about 2500 m and the average density is about 440 km^2 per site. While snow pillow stations may be inadequate to capture the high spatial variability of SWE across the Sierra, they are denser than the station networks in most mountainous regions globally, and thus the Sierra Nevada is a relatively ideal region for statistically-based SWE modeling. All snow pillow data were downloaded from the California Data Exchange Center (https://cdec.water.ca.gov/) and quality controlled. Because using the point-scale snow pillow SWE observations may not fully represent the average SWE estimates over the 500-m grid-cell (Molotch et al., 2005; Meromy et al., 2013), we used FSCA data to scale the point SWE observations to improve their representativeness at the 500 m grid scale (Schneider and Molotch, 2016).

3.1.2. Independent variables

Table 1 lists all the independent variables used in the two LRMs, including the 18 static physiographic variables and the dynamic historical SWE reanalysis and DMFSCA. Earlier studies of Bair et al. (2018); Fassnacht et al. (2003); Schneider and Molotch (2016) suggested that physiographic variables have significnat impacts on SWE distrubution, and thus we selected the listed physiographic variables (Table 1) for SWE modeling. All physiographic variables were derived from the Shuttle Radar Topography Mission 1 Arc-Second Global Digital Elevation Model (SRTM DEM), which was downloaded from

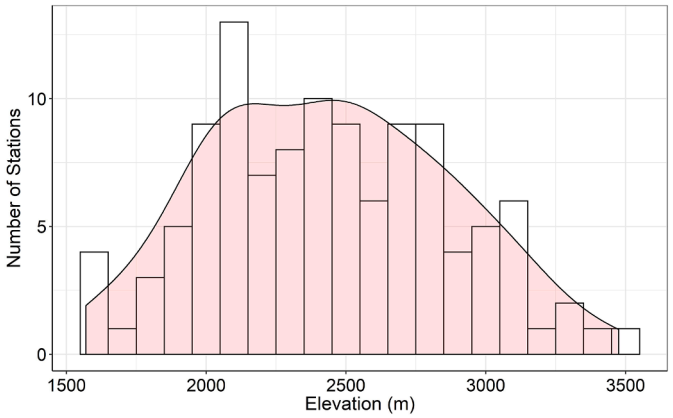

Fig. 3. The elevational distribution of snow pillow stations. Each bar represents the number of stations within every 100 m elevation band with the smoothed distribution shaded in light red (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

https://earthexplorer.usgs.gov/ at about 30-m spatial resolution and then resampled to 500-m spatial resolution using bilinear interpolation method.

The extent of SWE was determined by a gap-filled daily FSCA dataset, which was derived from the MODIS Snow Covered Area and Grain Size (MODSCAG) (Painter et al., 2009) at a ~463 m spatial resolution. MODSCAG shows better performance than the standard NASA snow cover product, MOD10A1 (Hall et al., 2002) when compared to FSCA from Landsat 5 and Landsat 7 during accumulation and melt (Rittger et al., 2013). Recent work has shown that MODSCAG has a bias of only 0.5% when compared to FSCA from Landsat 8 (Rittger et al., 2021). The MODSCAG FSCA was further refined to account for off-nadir viewing (Dozier et al., 2008), improved cloud detection and canopy adjustment (Rittger et al., 2020), and interpolated using a smoothing spline method (Dozier et al., 2008) creating spatially and temporally complete (STC) MODSCAG FSCA, hereby referred to as STC-MODSCAG. Data from WY 2021 is publicly available with a time lag of about two days and the historical dataset from 2001 to 2020 is available on request to early adopters (https://nsidc.org/snow-today/data_request). In addition to SWE reconstruction efforts previously noted, the FSCA data has been used for evaluating climate models (Wrzesien et al., 2015; Minder et al., 2016) and for validating regional scale models to better understand the sources of and deposition of dust and carbon (Sarangi et al., 2019, 2020).

We first resampled the STC-MODSCAG FSCA data from sinusoidal projection to the Universal Transverse Mercator (UTM) Zone 11 N with the 1983 North American Datum (NAD) (i.e., the original projection of the ASO SWE data) at 500 m resolution using the bilinear interpolation method. The resampled FSCA data were then used to scale snow pillow SWE observations as previously noted so that the point-scale SWE observations are more representative of the 500 m grid-cell mean SWE. Additionally, the independent variable, DMFSCA, was calculated based on daily STC-MODSCAG FSCA. Both STC-MODSCAG FSCA and DMFSCA were represented as a percentage (0–100%) value. Finally, we used the binary snow extent derived from this real-time STC-MODSCAG FSCA dataset to determine the snow-covered areas for the final modeled SWE estimates, in which the snow-free pixel was determined when the FSCA value equals 0.

The Sierra Nevada SWE reanalysis (SNSR, daily at about 100-m resolution from 1985 to 2016; data details on https://margulis-group.github.io/data/) (Margulis et al., 2015, 2016) was used as to provide analog historical SWE patterns (see Table 1, last variable) for the LRMs in real-time SWE estimation. The SNSR relies on a Land Surface Model (Simplified Simple Biosphere model, version 3) (Xue et al., 2003) coupled with a Snow Depletion Curve model (LSM-SDC) (Liston, 2004) to derive the prior SWE and FSCA estimates, in which the radiative and meteorological forcings are obtained from downscaled NLDAS-2 (Xia et al., 2012). The particle batch smoother (PBS) data assimilation (DA) scheme was then used to account for the uncertainty of the prior SWE estimate directly based on the FSCA retrieved from Landsat (i.e., Landsat 5 Thematic Mapper, Landsat 7 Enhanced Thematic Mapper, and Landsat 8 Operational Land Imager) (details in Margulis et al., 2015, 2016). SNSR provides the most accurate SWE estimates on average across the entire Sierra in a comparison with four other large-scale SWE datasets at various spatial resolutions (Yang et al., 2018). It also has a long historical modeling period (i.e., 32 years) relative to other datasets (Bair et al., 2016; Guan et al., 2013; Rittger et al., 2016), covering a large range of SWE conditions. Despite the good performance of SNSR, it is not applicable for real-time SWE estimation given its requirement of complete seasonal snow cover depletion information observed from satellites for the PBS DA process (Margulis et al., 2015, 2016). Thus, we selected the SNSR to provide historical SWE patterns for the LRM.

To select the historical SNSR that can best represent the spatial patterns of SWE on the target simulation date, we designed a two-step process to compare the snow pillow SWE patterns between the target simulation date and historical dates. First, we identified the historical dates that have similar numbers of non-zero SWE observations (e.g., the difference is $< 10\%$ in the study), which can reduce the impact of zero values on the linear regression fits in the next step. Secondly, the similarity of SWE patterns was quantified by the squared Pearson correlation coefficient (R^2) of the simple linear regression between snow pillow SWE on the two dates. For each date when the LRM is run, the historical SWE reanalysis with the highest R^2 was used as the best historical SWE distribution data. This two-step process assumes that the similarity in SWE patterns at stations is also true across the Sierra Nevada. To replicate how a real-time SWE estimation would produce, we only chose the historical SWE on the WYs before the current simulation WY even though the SNSR is available throughout 1986–2016, partially handicapping our approach by not using all of the historical data. All the historical SWE data were resampled from the native 100 m spatial resolution to 500 m using bilinear interpolation to match the extent and scale of the resampled STC-MODSCAG FSCA data.

To eliminate the influence of magnitude difference, all the independent variables (i.e., 18 physiographic variables, the DMFSCA, and the SNSR) were scaled to a mean of 0 and a standard deviation of 1. We used a mask of the maximum water extent in the Global Surface Water (GSW) dataset (Pekel et al., 2016) to exclude water bodies inside the study area that are often mapped (FSCA) or modeled (SWE) incorrectly. The maximum water extent data was derived from Landsat imagery that has a higher spatial resolution of 30 m for the 1984–2019 period. We resampled this dataset to the 500 m resolution using the nearest neighbor method which shows good performance in resampling categorical data.

3.2. Statistical model for SWE estimation

The GLMNET used in LRM outperformes the Generalized Linear Model (GLM) originally employed by Schneider and Molotch (2016) by optimally handling linearly related independent variables (e.g., the physiographic variables used by the LRM). To address multicollinearity among independent variables, the GLMNET uses an elastic-net approach to combine two regularization approaches: (1) shrinking the coefficients of correlated predictors to zero while keeping all independent variables (i.e., ridge regression), and (2) selecting one independent variable and discarding the others (i.e., lasso regression), resulting in a more stable and accurate model. The optimal hyperparameter was determined by a k-fold (i.e., 10 folds) cross-validation where the lowest mean absolute error among approximately one hundred model simulationswas chosen in the final model (details in Friedman et al., 2010; Simon et al., 2011, 2013; Tibshirani et al., 2012). Following Schneider and Molotch (2016), the Gaussian distribution with the identity link function was used to transform the dependent variable in the GLMNET. Given the requirement of the Gaussian distribution of SWE observations for the statistical LRM training, we only ran the model on the days with more than 40 snow pillow stations (i.e., about 35% of the total stations) reporting non-zero SWE values. We used 40 as an empirical threshold to stop running the LRM to ensure the representativeness of model training samples.

3.3. Two operational SWE datasets

SNODAS provides daily SWE estimates at 0600 Coordinated Universal Time (UTC) in (near) real-time for the contiguous United States at a 1000 m spatial resolution since WY 2004 (data are available at https://nsidc.org/data/g02158) (Barrett, 2003). SNODAS is a physically-based snow modeling and data assimilation system which ingests a variety of outputs from physically-based models, meteorological data, remote sensing data, and SNOTEL observations including snow depth and SWE (Carroll et al., 2006; Clow et al., 2012). The National Water Model (NWM) is an extension of the Weather Research and Forecasting Hydrological model (WRF-Hydro) coupled with the community Noah land surface model with multiparameterization options

(Noah-MP) to simulate the land surface processes, including SWE (Gochis et al., 2018). The NWM-SWE dataset used in this study is the output of the NWM retrospective version 1.2 (data are available at http://edc.occ-data.org/nwm/getdata/), which is a 25-year retrospective simulation from 1993 to 2017. NWM-SWE covers a large domain roughly from latitude 19 N to 58 N that includes the Continental United States, Canada, and Mexico at a spatial resolution of 1000 m. To assure data consistency with SNODAS, we used the NWM-SWE data at 0600 UTC in the comparison. For consistent data evaluation and inter-comparison, each SNODAS and NWM-SWE grid cell at 1000 m resolution was disaggregated evenly to four 500 m grid cells and then resampled to the NAD83/UTM zone 11 N using bilinear interpolation method.

3.4. Model evaluation

We used ground SWE observations and the ASO SWE product to provide a comprehensive validation for the four SWE models (i.e., LRM-baseline, LRM-FSCA, SNODAS, and NWM) at different scales. The ground observations (i.e., snow pillow and snow courses) provide the most accurate SWE estimates at a point-scale or transect-scale across the entire Sierra, while the ASO SWE data represents the most accurate basin-scale SWE estimates to the extent of our knowledge. Because snow pillow SWE observation is not an independent evaluation dataset for LRM and SNODAS, we only applied cross-validation at snow pillow sites to evaluate LRMs' performance in modeling SWE. The snow pillow cross-validation was conducted for all 2682 daily simulations (i.e., 141 days per year on average) for the 19-year modeling period (i.e., WYs 2001–2019). The SWE validation was conducted against a total of 6407 snow course station-years and 23 ASO flights (details in the following three sub-sections).

3.4.1. Cross-validation using snow pillow SWE observations

We used the Monte Carlo cross-validation (i.e., repeated random sub-sampling validation) method (Dubitzky et al., 2007) to evaluate the two LRMs' performance. For each simulation day, we randomly identified 10% of the 113 snow pillow observations (i.e., 11 samples) for testing with the remaining 90% of observations used for model training. Next, a simple linear regression model was fit between the modeled SWE estimates and the testing observation. We repeated this process 30 times in balance of computational cost and evaluation efficiency. The average values of the following four statistical metrics from all 30 simple linear regression models were used to describe the overall performance of each simulation: squared Pearson-correlation coefficient (R^2), root mean squared error (RMSE), relative mean absolute error (RMAE) (i.e., mean absolute error divided by the mean of FSCA-scaled snow pillow observations), and percent bias (PBIAS) of average SWE for modeled estimates against observations. The R^2 represents the correlation between modeled and observed SWE, and the RMSE represents the variability of the model residuals. The RMAE represents the relative errors of modeled SWE in the context of different SWE magnitudes. The last metric, PBIAS, represents the relative bias of average SWE over the entire study domain. The typical model performance is described by the distribution of each metric with the combination of the mean, median, and one standard deviation. These four metrics were also used in the subsequent validations using snow course measurements and the ASO SWE data.

3.4.2. Validation using snow course SWE measurements

There are 215 snow course sites located in the study area, with an average elevation of ~2400 m ranging from 1500 m to 3500 m. Each snow course consists of approximately 5–15 point-scale SWE measurements along an established transect using a calibrated Federal Snow Sampler (DeWalle and Rango, 2008). These data offer monthly SWE measurements near the first day of each month from January to May or June depending on the amount of snow remaining on the ground. To ensure the independence of the evaluation data from the training data (i.e., snow pillow SWE), we removed snow course sites located within a 2 km distance of snow pillows, leaving 148 snow course sites for evaluation. All the snow course data were downloaded from the California Data Exchange Center (http://cdec.water.ca.gov).

We extracted SWE estimates from the four SWE datasets at snow course sites and compared them to snow course SWE measurements. The spatial mismatch of the transect to 500 m pixels may cause uncertainties in this evaluation. To reduce the impact of this mismatch, we scaled snow course SWE measurements using FSCA data so that the observations can better represent the average SWE estimates at the 500 m resolution (Schneider and Molotch, 2016). Additionally, to reduce the impacts from resampling previously described, the SWE estimates at the eight neighboring grid-cells surrounding the target grid-cell (i.e., the snow course site located in the middle of the 3 × 3 grid-cell window) were also used in the evaluation. The SWE value with the smallest difference among the nine grid-cells was used in the calculation of the statistical metrics (Bair et al., 2016; Rittger et al., 2016; Margulis et al., 2016). Given the difference in data availability – LRMs from 2001 to 2019, SNODAS from 2004 to the present, NWM-SWE from 1993 to 2017 – we only compared the evaluation results in their overlap period from WY 2004 to WY 2017. For data consistency, only SWE estimates available for all datasets at snow course sites were compared. There remain 6407 station-years in total that contain coincident SWE estimates for the four SWE datasets. The same four statistical metrics (i.e., R^2, RMSE, RMAE, and PBIAS) were calculated to evaluate SWE accuracy at snow course sites.

3.4.3. Validation using ASO SWE data

Considering the overlap period (WYs 2004–2017) of the four SWE datasets and the ASO data availability (WY 2013 to present), we compared the ASO SWE validation for the four models throughout WYs 2013–2017 with a total of 19 days (i.e., 19 ASO flights). We also included the validation in WY 2018 and WY 2019 for LRM-baseline and LRM-FSCA with 4 additional flights to give a more comprehensive LRM model evaluation. Given the various spatial projections and resolutions of these datasets, we resampled all data into the same projection as the ASO SWE data (see Section 3.1.2) and 500 m spatial resolution. A 3 × 3 window (see Section 3.4.2) was also used in the validation against ASO SWE. The basin-wide PBIAS was calculated directly based on the resampled SWE data over the entire TRB, meaning no neighboring pixels were considered in the PBIAS calculation.

3.5. Inter-comparison of SWE datasets

The temporal overlap period of the four SWE datasets covers 14 WYs from 2004 to 2017 including the region's most recent snow drought period 2012–2016 and the wet years of 2011 and 2017. For each WY, we calculated the maximum snow water storage (hereafter, SWS) across the Sierra and the day of year (DOY) when the Sierra-wide SWS reaches the maximum. We introduce SWS (i.e., SWE multiples by the total area) in this comparison because it is a water volume metric that can reflect water availability for the study area directly, while SWE value only represents equivalent water depth. The maximum SWS over the entire Sierra was determined when the Sierra-wide average SWE reaches the peak value. In other words, the DOY of maximum SWS is the same as the DOY of maximum Sierra-wide average SWE. Additionally, to reveal the difference in maximum annual snow water availability at a pixel-scale among datasets, we compared the spatial distribution of 14-year average pixel-wise peak SWE and its changes as a function of elevation, where the pixel-wise peak SWE is defined as the maximum SWE estimates over the entire water year from October 1 to September 30 of the following year (Margulis et al., 2016).

4. Results

4.1. Evaluation of model performance

4.1.1. Sierra-wide cross-validation with snow pillow observations

LRM-FSCA exhibited notable improvements compared with LRM-baseline in the snow pillow cross-validation indicated by the median, mean, and standard deviation of all statistical metrics (Table 2). Specifically, LRM-FSCA explained 59±8% of the variance in snow pillow SWE, 6% higher than that for LRM-baseline (53±9%) (Table 2). The RMSE for LRM-FSCA was 17.4±8.7 cm, with an average reduction of 1.4 cm compared with the LRM-baseline (18.8±9.2 cm). LRM-FSCA also had a lower RMAE (51.3±160.0%) and basin-wide PBIAS (15.2±155.0%) than LRM-baseline, 56.0±204.0% and 17.2±200.0%, respectively. Both LRM-FSCA and LRM-baseline had high variation in RMAE and PBIAS values, but low variation in R^2 and RMSE values (Table 2). This can be explained by the high relative errors of a few simulations in very low SWE period when a small absolute difference could result in a great relative error.

The improvement in LRM-FSCA's performance was more significant during the primary snow accumulation periods (i.e., January-March) than during the snowmelt seasons (Fig. 4). The seasonal variability of model accuracy calculated in each month for LRM-FSCA and LRM-baseline exhibited similar patterns. LRM-FSCA exhibited much better performance than LRM-baseline in five of the six months we modeled, including January through May during the period of the largest volumes of melt. For the SWE estimates in June when snow remained primarily in high elevation regions, both LRM-FSCA and LRM-baseline showed similar low accuracy. This can be explained by the low representativeness of the training samples in June given that most stations are located at middle to low elevations and reported zero SWE. The R^2 for both LRMs exhibited little seasonal variability compared with the other three metrics, indicating that the spatial variability of the snow pillow observations was consistently captured by the model training over time. RMSE generally increased throughout the snow season. Increases in RMSE from January to April were consistent with the increased snow accumulation during that period. However, the continued increase in RMSE in May and June suggested a decreasing performance of the LRMs starting in May as snow accumulation was no longer high for this period. This general trend was also supported by the distribution of RMAE and basin-wide PBIAS (Fig. 4C and D), both of which indicated much better performance of LRMs in January through April than that in May and June.

4.1.2. Sierra-wide validation with snow course measurements

The SWE values of the two LRM models were relatively well aligned with snow course SWE values (Fig. 5A and B) whereas SNODAS and NWM-SWE values showed greater deviation from the 1:1 line (Fig. 5C and D). Adding the DMFSCA to the LRM largely reduced the uncertainty of SWE estimates over the snow course sites. LRM-FSCA outperformed the LRM-baseline with a 7% (0.06) increase in R^2, and 14% (2.5 cm), 20% (4.3%), and 93% (1.3%) reductions in RMSE, RMAE, and PBIAS, respectively. The overall PBIAS for LRM-FSCA and LRM-baseline were

Table 2
Performance summary of LRM-FSCA and LRM-baseline in the snow pillow cross-validation. For each simulation, a series of linear regression models were fit between FSCA-scaled snow-pillow SWE observations and modeled SWE estimates in the cross-validation to calculate the statistical metrics, and the mean values are summarized here. The bold numbers indicate the better performance.

Dataset	R^2			RMSE (cm)			RMAE (%)			PBIAS (%)		
	mean	Median	*SD	mean	median	*SD	mean	median	*SD	mean	median	*SD
LRM-FSCA	**0.59**	**0.60**	**0.08**	**17.4**	**15.9**	**8.7**	**51.3**	**30.2**	**160.0**	**15.2**	**2.2**	**155.0**
LRM-baseline	0.53	0.54	0.09	18.8	17.5	9.2	56.0	33.7	204.0	17.2	2.6	200.0

* SD represents standard deviation.

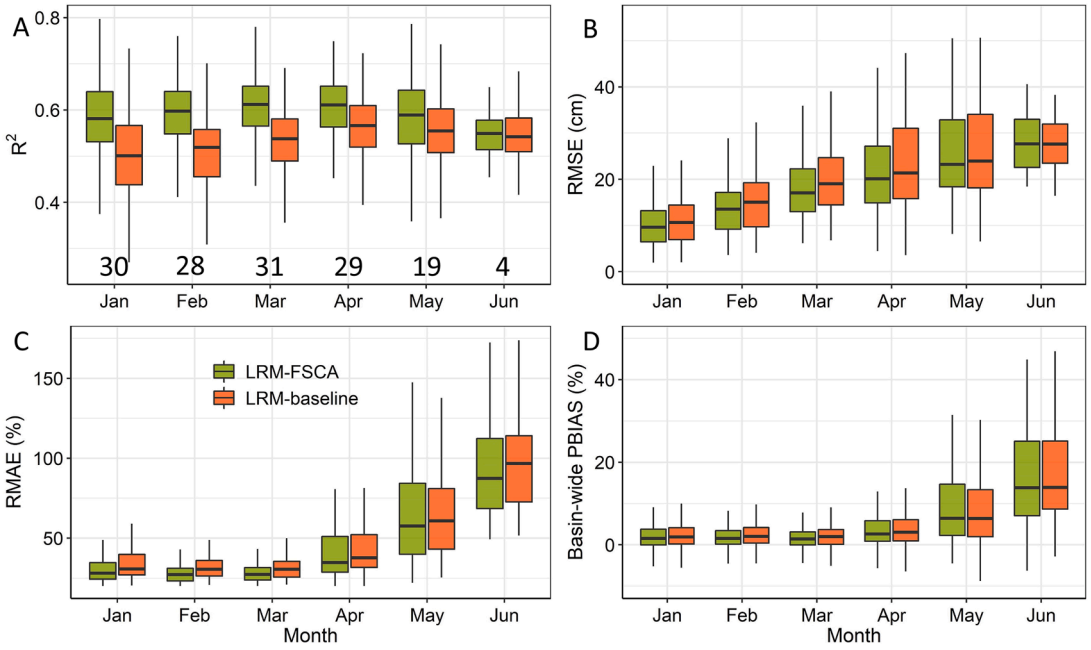

Fig. 4. Monthly snow pillow cross-validation results of LRM-FSCA and LRM-baseline. Panels A-D displays the monthly R^2, RMSE, RMAE, and basin-wide PBIAS for all daily simulations from WYs 2001–2019, respectively. The numbers labeled on figure A represent average numbers of model simulations for each month throughout WYs 2001–2019.

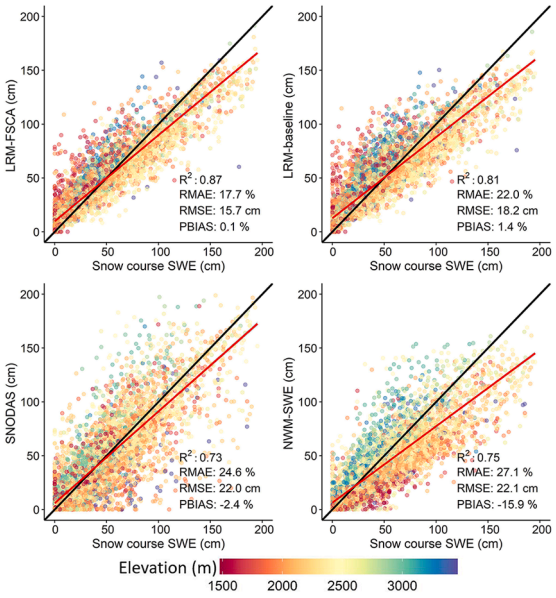

Fig. 5. Validation of the four SWE datasets at snow course sites. Each dot represents a pair of modeled SWE and FSCA-scaled snow course SWE. A total number of 6407 pairs were compared from the WYs 2004 to 2017 for the four SWE datasets. The black line represents a 1:1 line where gridded SWE equals to FSCA-scaled snow course SWE. A simple linear regression was fit (red line) between these two SWE values using the least-squares method, with four statistical metrics calculated and labeled on the bottom right corner of each panel (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

slightly positive, with the highest observed SWE values underestimated by both LRMs (see SWE values greater than 100 cm generally located below the 1:1 line in Fig. 5A and B). The overall performance of LRM-FSCA was notably better than SNODAS and NWM-SWE, showing both 29% (6.3 cm and 6.4 cm, respectively) lower RMSE, 28% (6.9%) and 35% (9.4%) lower RMAE, and 96% (2.3%) and 99% (15.8%) lower absolute PBIAS, respectively. SNODAS had slightly higher accuracy than NWM-SWE, with 9% (2.5%) and 85% (13.5%) lower RMAE and absolute PBIAS values, respectively, and similar RMSE values, while the R^2 for SNODAS (0.73) was slightly lower than that for NWM-SWE (0.75).

Because elevation is the main topographic control on SWE distribution, we also examined the relationship between elevation and SWE errors for each datasets (Fig. 5). The results showed that both LRMs underestimated SWE for the middle elevation snow course sites, but overestimated SWE for relatively high and low elevations (red and blue dots above 1:1 line in Fig. 5A and B). These SWE error patterns were possibly caused by a nonlinear relationship between SWE and elevation (e.g., a curve alike an upside-down "U" shape). No particular pattern was observed for SNODAS. NWM-SWE had a clear underestimation for low and middle elevations, but overestimation for relatively high elevations (around 3000 m). While the SWE estimation errors for each product had high variability, all datasets tended to overestimate SWE for very low SWE values but underestimated SWE for moderate to high SWE values (red lines in Fig. 5).

4.1.3. Basin-scale validation with ASO SWE data

The two LRMs also performed better than SNODAS and NWM-SWE when compared with the ASO data in the upper TRB during the overlapping period from WYs 2013–2017 with 19 ASO flights (Table 3 and Fig. 6). LRM-FSCA and LRM-baseline explained similar proportions of variance in ASO SWE with the same mean R^2 of 0.85, which was at least 31% (0.20) higher on average than the other two operational SWE datasets. The variance of ASO SWE explained by LRM-FCA, LRM-baseline, SNODAS, and NWM-SWE were 85±5%, 85±5%, 65±6%, and 32±18%, respectively (Table 3). LRM-FSCA, LRM-baseline, and SNODAS illustrated more robust SWE estimates than NWM-SWE given the relatively lower standard deviation of R^2 (Table 3).

The improvement of LRM-FSCA relative to the LRM-baseline was less evident over the TRB because the evaluation metrics showed cloase values (Table 3). LRM-FSCA had lower RMSE than LRM-baseline for 13 out of 19 days (Fig. 6B), but the median RMSE for LRM-FSCA (9.8 cm) was 26% (2.0 cm) higher than that for LRM-baseline (7.8 cm) for the period one(Table 3). The median RMSE for LRM-FSCA (13.3 cm) was 4% (0.5 cm) less than that for LRM-baseline (13.8 cm) in the period two,

Table 3
Summary of the validation results for the four SWE datasets using the ASO SWE data over the upper TRB for the 19 overlap days from WYs 2013 through 2017, and for two LRMs over the 23 overlap days from WYs 2013 through 2019. The bold numbers indicate the highest/higher accuracy for different SWE models in the two comparison periods, respectively.

Period	Dataset	R^2			RMSE (cm)			RMAE (%)			PBIAS		
		mean	median	SD	mean	median	SD	mean	median	SD	mean	median	SD
WYs 2013–2017	LRM-FSCA	**0.85**	0.85	0.05	**15.7**	9.8	17.2	**23.6**	19.8	12.5	-5.5	**-9.2**	27.7
	LRM-baseline	**0.85**	**0.86**	0.05	15.9	**7.8**	17.3	24.8	**19.3**	13.6	**-4.1**	-11.3	29.7
	SNODAS	0.65	0.64	0.06	18.9	14.8	11.6	35.1	26.7	20.7	35.4	28.2	41.9
	NWM	0.32	0.33	0.18	26.3	20.0	19.0	45.2	42.9	13.8	-28.8	-30.1	15.5
WYs 2013–2019	LRM-FSCA	**0.85**	0.85	0.05	**16.6**	**13.3**	15.7	**22.4**	**18.9**	11.7	-6.1	-9.0	25.3
	LRM-baseline	0.84	**0.86**	0.05	17.0	13.8	15.9	23.6	19.0	12.8	**-4.8**	**-7.4**	27.5

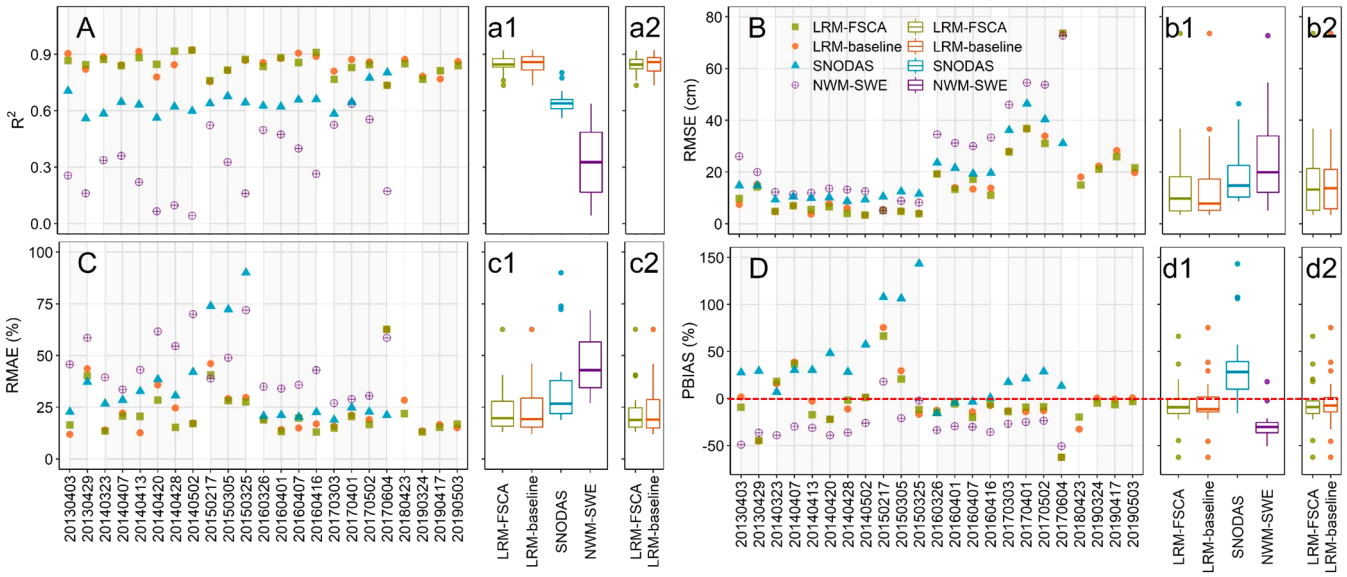

Fig. 6. Validation of the four SWE datasets using the ASO SWE data in the upper TRB. Panels A-D display the R^2, RMSE, RMAE, and PBIAS over the comparison period, respectively, with each water year shaded separately. Validation results were compared in two periods: WYs 2013–2017 when all four SWE datasets were available, and WYs 2013–2019 when only the two LRMs were available. Boxplots a1-d1 summarize the distribution of the statistical metrics in the first period with 19 overlap flights, while boxplots a2-d2 summarizethe statistical metrics for the two LRMs in the second period with 23 overlap flights.

further suggesting that LRM-FSCA and LRM-baseline had similar performance over the TRB during this validation period. In the daily model cross-validation with snow pillow SWE observations, we found that the improvement in LRM-FSCA's performance was more pronounced during snow accumulation periods than during the snowmelt seasons and LRM-FSCA only showed a little better performance than LRM-baseline in the late season. Given that ASO generally flies near or after peak SWE date (i.e., around April 1) and mostly in the late snowmelt season, it was not surprising to see that LRM-FSCA and LRM-baseline showed similar accuracy when validated with ASO SWE data.

The RMAE for LRM-FCA, LRM-baseline, SNODAS, and NWM-SWE were 23.6±12.5%, 24.8±13.6%, 35.1±20.7%, and 45.2±13.8%, respectively. LRM-baseline had the smallest PBIAS (-4.1±29.7%) compared with LRM-FSCA (-5.5±27.7%), SNODAS (35.4±41.9%), and NWM-SWE (-28.8±15.5%). Although NWM-SWE had a large PBIAS, the variability of its PBIAS was the smallest (15.5%), indicating it consistently underestimated SWE, which was also shown in Figure 6D with negative PBIAS values on 18 out of 19 dates.

To visualize spatial differences across SWE datasets not apparent by summary statistics in Fig. 6 or Table 3, we selected five days near peak SWE from WYs 2013 through 2017 to show the spatial distribution of SWE errors on one day in each WY. As with the summary statistics compared with the ASO SWE on a pixel-by-pixel basis, the magnitudes of SWE errors for LRM-FSCA and LRM-baseline were much smaller than that for SNODAS and NWM-SWE (Fig. 7). Both LRM-FSCA and LRM-baseline overestimate SWE in WYs 2014 and 2015, the extreme drought years, with a mixture of overestimation and underestimation in the normal and wet years of WYs 2013, 2016, and 2017. Additionally, the spatial distribution of SWE errors for the two LRMs revealed similar patterns. The LRMs tended to underestimate SWE in the northern high alpine regions of the TRB and overestimate SWE in the western low elevation valley regions. For the densely forested regions, mostly located in the southeastern of the TRB, the LRMs tended to overestimate SWE for all comparison years. Both the magnitude and the spatial pattern of SWE residuals for LRM-FSCA and LRM-baseline were identical, indicating that on average the two LRMs had similar accuracy in the upper TRB.

The spatial distribution of SNODAS errors exhibited greater variability and magnitude than the two LRMs, with a relatively large overestimation of SWE for most regions in WYs 2013, 2014, 2015, and 2017. In WY 2016, SNODAS overestimated SWE for the low elevation forested regions but underestimated SWE for high elevation alpine regions, exhibiting relatively identical patterns compared with the two LRMs. The SWE errors for NWM-SWE revealed rather similar spatial patterns under diverse snow conditions, with overestimation in some low elevation valleys and forests and underestimation at the medium to higher elevations. Overall, NWM-SWE underestimated SWE for most regions across the TRB.

4.2. *Spatial and temporal variability of SWE*

4.2.1. *Inter- and intra-annual variability of snow water storage*

To reveal the variability of snow water storage (SWS) across the Sierra Nevada, as well as the SWS differences among the four SWE datasets, we also compared the daily SWS through their 14-year overlap period (WYs 2004–2017) including the 2012–2016 California drought period with extremely low SWS (Fig. 8, gray shading). The two LRMs generally had higher SWS than SNODAS and NWM-SWE, particularly for the period close to the peak SWS date. Only subtle differences in SWS were observed among datasets in the early snow accumulation and late snowmelt seasons.

Looking more closely at the differences of maximum snow accumulation, the annual maximum SWS exhibited high interannual variability across all models (Fig. 9A and B) . The average maximum SWS for the driest year 2015 and wettest year 2011 derived from all the four SWE datasets were 3.4±0.6 gigaton (Gt) (i.e., 4.0 Gt for LRM-baseline, 3.7 Gt for LRM-FSCA, 3.2 Gt for SNODAS, and 2.7 Gt for NWM) and 38.6±7.7 Gt (i.e., 47.1 Gt for LRM-baseline, 42.5 Gt for LRM-FSCA, 34.6 Gt for SNODAS, and 30.0 Gt for NWM), respectively, with a 35.2 Gt SWS difference on average between these two years. The year-to-year changes of the maximum SWS for the four SWE datasets followed a similar pattern, but the magnitudes varied. LRM-baseline consistently estimated the highest SWS, followed by LRM-FSCA, SNODAS, and NWM-SWE, with the median maximum SWS of 25.0 Gt, 23.2 Gt, 16.0 Gt, and 13.8 Gt, respectively.

The maximum SWS difference (i.e., absolute value) among the datasets in any given year ranged from a low of 0.1 Gt in 2012 between LRM-FSCA and SNODAS to a maximum of 17.1 Gt in 2011 between LRM-baseline and NWM-SWE (Fig. 8). For each pair of datasets, the 14-

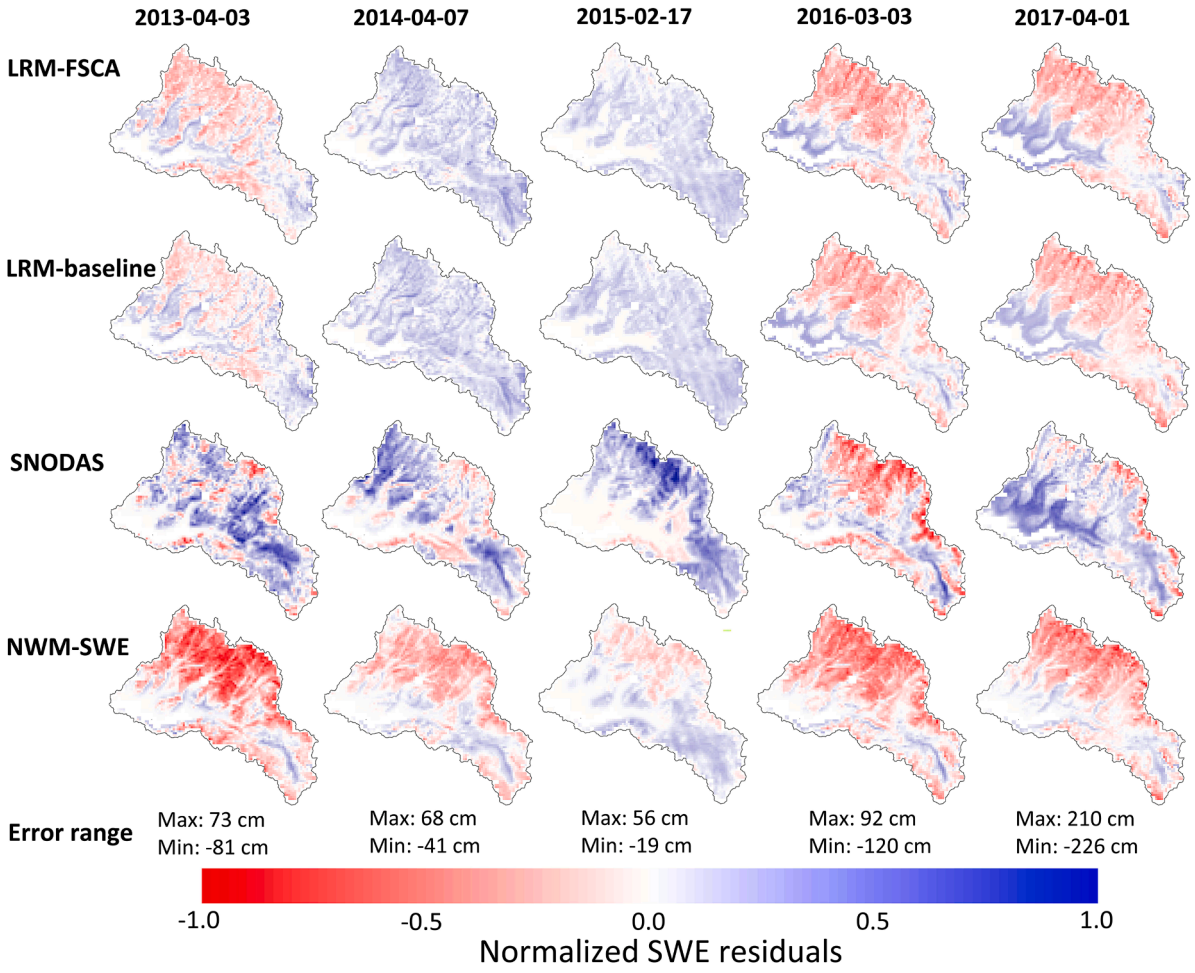

Fig. 7. Spatial distribution of SWE estimation errors for the four SWE datasets on one selected date near peak SWE date for each water year from 2013 through 2017. To fully display the inter-model differences in SWE errors, the SWE errors were normalized by the maximum absolute error on each comparison date of all four SWE datasets. The overall ranges of SWE errors are labeled on each date. The darker colors represent a higher magnitude of SWE errors.

year average SWS difference was the smallest between LRM-FSCA and LRM-baseline (1.2±1.2 Gt), ranging from 0.2 Gt to 4.6 Gt, and was largest between LRM-baseline and NWM-SWE (9.8±5.1 Gt), ranging from 1.3 Gt to 17.1 Gt.

The date of maximum SWS exhibited a very high inter-annual variability for all the four SWE datasets in the comparison period (Fig. 9C and D). LRM-FSCA and NWM-SWE generally peaked earlier over the entire Sierra, followed by SNODAS and LRM-baseline, with the median peaked dates of DOY of 70, 70, 73, and 74 (i.e., equivalent to March 11, +0, +3, and +4 days in a common year), respectively. The dates of maximum SWS were close in 11 out of 14 years among four datasets, with the exceptions of WYs 2010, 2013, and 2015. In WY 2010, a moderate SWE year, NWM-SWE peaked on March 14, 32 days earlier than SNODAS on April 15, while LRM-FSCA and LRM-baseline estimated peak SWS on April 8 and April 7, respectively. Notably, in WY 2013, an extremely dry year, both LRM-FSCA and NWM-SWE peaked in January, while LRM-baseline and SNODAS peaked in March. For dry WY 2015, only SNODAS peaked in March, while the other three datasets peaked in December or January. Although the peak SWS dates varied significantly in these two dry years, the SWS values between January to March did not change appreciably over time for all the four datasets in WY 2015 (Fig. 8), suggesting that the significant differences in the peak SWS date estimated from different datasets in dry years may only have a subtle influence on water volume forecasts.

4.2.2. *Spatial distribution of pixel-wise peak SWE*

LRM-FSCA and LRM-baseline generally had a similar spatial pattern of pixel-wise peak SWE, with only slightly lower SWE estimated by LRM-FSCA in relatively low elevations (<~2500 m) (Fig. 10) and on the eastern slopes of the Sierra (Fig. 11). Because the DMFSCA well represented the difference between low elevations and eastern Sierra with relatively ephemeral snowpacks versus high elevations with consistent snow covers, LRM-FSCA likely decreased SWE errors for low elevations and eastern Sierra where LRM-baseline tended to overestimate SWE. The SWE patterns for SNODAS and NWM-SWE were considerably different relative to the two LRMs, with much lower peak SWE at high elevations (>~3500 m) and low elevations (<~2200 m) (Fig. 10). Relative to the two LRMs, SNODAS exhibited notably greater SWE at around 2400–3300 m elevations. Interestingly, this relationship was reversed for the highest elevations above 3400 m where SNODAS SWE values were significantly lower than the two LRMs.

While the average pixel-wise peak SWE over the middle to high elevations exhibited distinct differences, the influence of these differences on the range-wide SWS was relatively modest given the significantly smaller alpine area compared with the area of the low elevations (<~2500 m) in the Sierra (red dots in Fig. 10). This also explained the high SWS difference among datasets near peak snow accumulation period when low elevation SWE was relatively high (Fig. 8). A small difference in low elevation SWE can result in relatively large differences in SWS given the large low elevation land area. Additionally, both

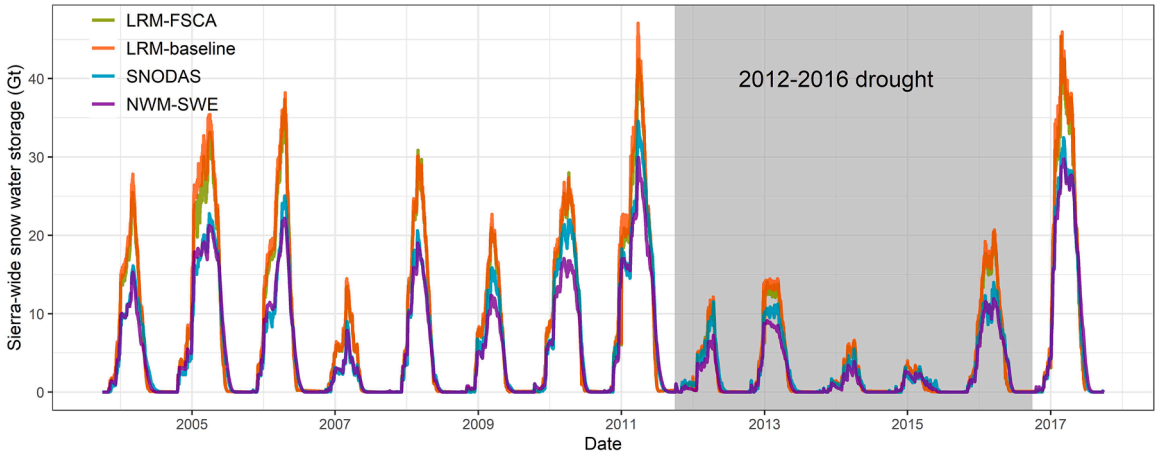

Fig. 8. 14-year interannual variability of SWS across the Sierra Nevada (WYs 2004-2017) . This comparison includes the extreme drought period from WYs 2012 to 2016 in California, which is shaded by gray.

SNODAS and NWM-SWE estimated much lower pixel-wise peak SWE than the LRMs for the eastern slopes of the Sierra and the low elevations of the western slopes of the Sierra (Fig. 11). NWM-SWE exhibited much lower SWE in majority regions compared with the other three datasets (Fig. 11).

5. Discussion

5.1. Comparison with existing SWE estimation methods

In this study, we combine ground observations with remotely sensed data in a linear regression model (LRM) for real-time SWE estimation in the Sierra Nevada Caliofornia. A key strength of the LRM is its simplicity compared with physically-based snow models (e.g., SWE reconstruction models, SWE data assimilation, and SNODAS). We present LRM-FSCA as a marked improvement on the LRM-baseline, the basis of which is established in Schneider and Molotch (2016). We include the new variable DMFSCA as an additional explanatory variable in the LRM given that DMFSCA describes the variability of snow cover over time that corresponds with the SWE accumulation and ablation processes. The results presented here indicate that including the satellite-observed DMFSCA as one additional predictor variable for the LRM has improved SWE estimation accuracy. This suggests that DMFSCA contains important information for real-time SWE modeling. Additionally, we show the performance of LRM-FSCA and LRM-baseline are considerably better than the operational SWE datasets (i.e., SNODAS and NWM-SWE). Hence, the real-time LRM SWE estimates have potentially crucial implications for practical water resource management. To give further context of the improved LRM-FSCA relative to other promising SWE estimation models represented in the literture, we compare our results with a few recently published studies in the text below.

Schneider and Molotch (2016) estimated SWE distribution using a similar method as LRM-baseline. They reported a mean R^2 value of 0.33 in snow pillow cross-validation, which is 44% (0.26) lower than the 0.59 reported here for LRM-FSCA, suggesting LRM-FSCA had a much improved performance. This is partly due to the higher density of the training samples (i.e., ~440 km^2 per snow pillow SWE observation for this study versus ~700 km^2 for Schneider and Molotch (2016)), and the higher accuracy of the historical SWE reanalysis data used in the Sierra Nevada than those used in the Rocky Mountains by Schneider and Molotch (2016) (Yang et al., 2018). The statistical model GLMNET used in this study also outperforms the original GLM model used by

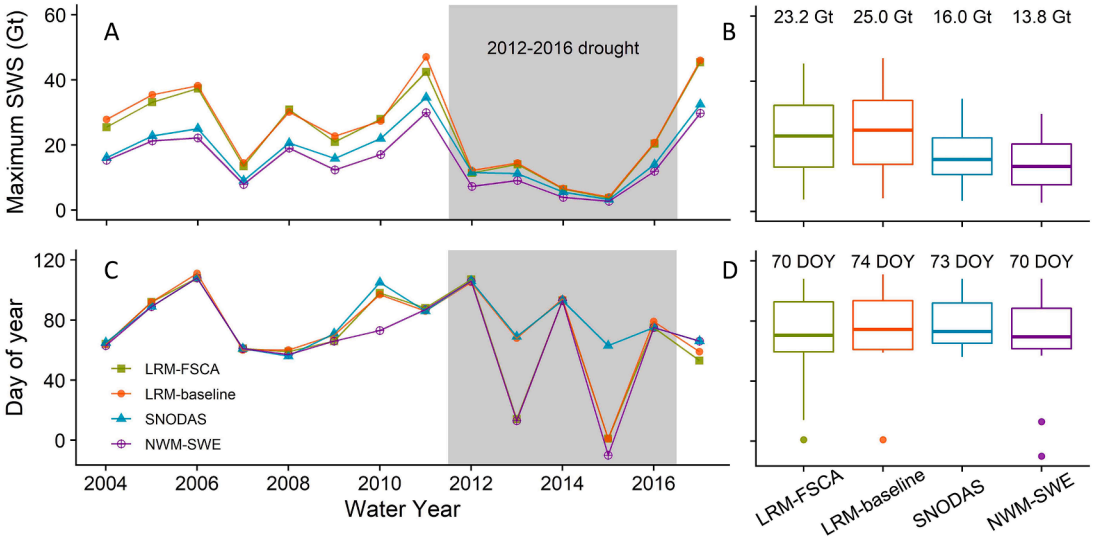

Fig. 9. Comparison of maximum SWS and the date of maximum SWS for the four SWE datasets across the Sierra. Subplots A and C show the 14-year time series of maximum SWS, and the date of maximum SWS is represented by the day of year (DOY). Their distributions are summarized in boxplots B and D, respectively. The median values of maximum SWS and the date of maximum SWS are labeled on the top of the boxplots.

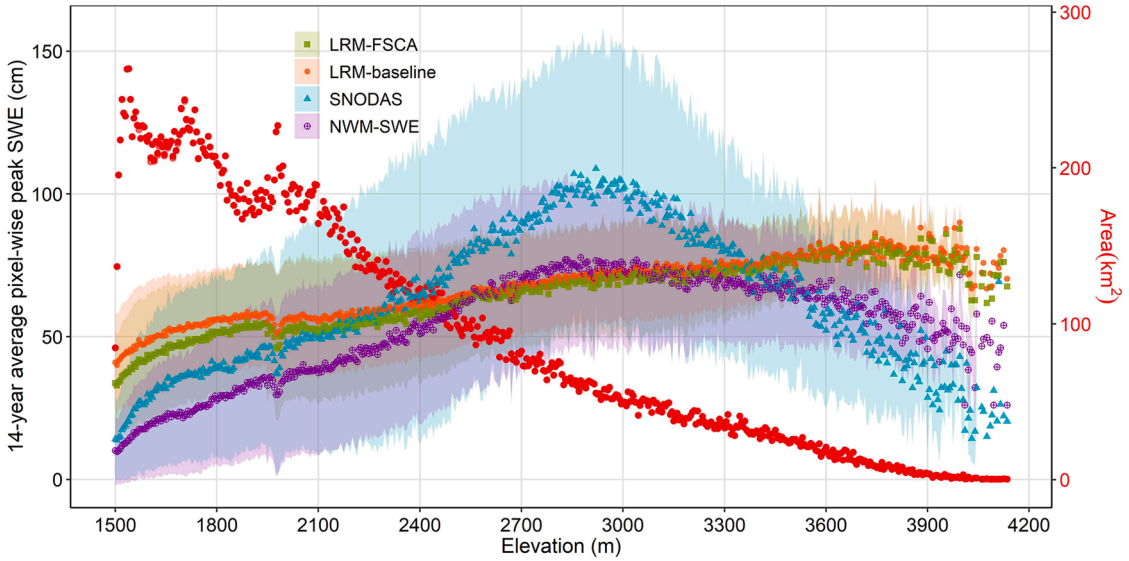

Fig. 10. The elevational distribution of 14-year average pixel-wise peak SWE for the four SWE datasets. Each point represents the average pixel-wise peak SWE value over a 10 m elevation interval in the 14-year comparison period (*y*-axis on the left), with the shaded regions represent one standard deviation. The red dots represent the area for each 10 m elevation band (*y*-axis on the right) (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

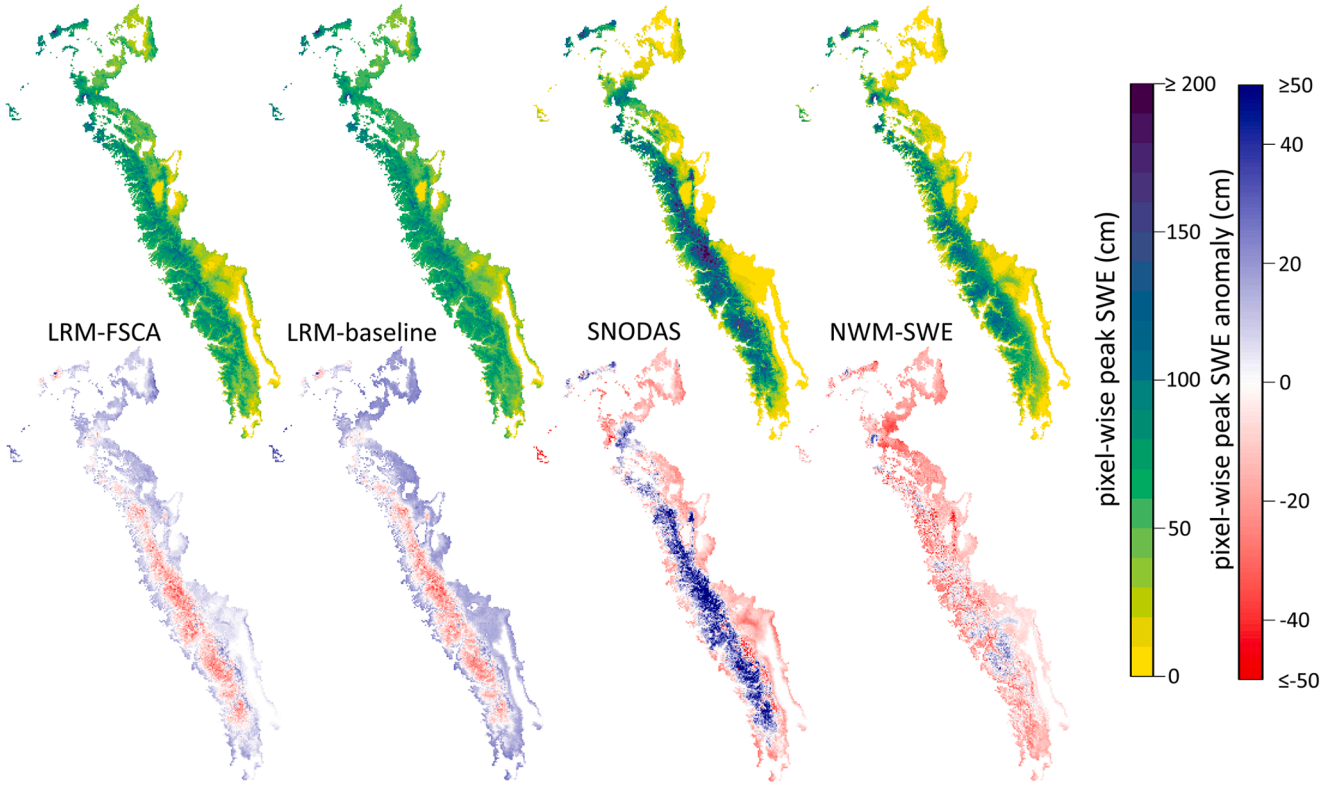

Fig. 11. The spatial distribution of 14-year average pixel-wise peak SWE for the four SWE datasets and their anomaly relative to the mean of the four SWE datasets during the overlap period of WYs 2004–2017 across the Sierra Nevada.

Schneider and Molotch (2016) in an environment with highly correlated independent variables given that GLMNET can adaptively choose a subset of the independent variables for a specific modeling date using the penalized maximum likelihood. Additionally, the differences in the patterns of winter snowfall precipitation between the maritime Sierra Nevada and the continental Rocky Mountains may also cause the differences in the model performance. The winter snowfall over the Sierra Nevada is strongly influenced by atmospheric rivers which often result in relatively uniform accumulation patterns across the Sierra Nevada (Ralph et al., 2004), while the snowfall over the larger Rocky Mountain domain studied by Schneider and Molotch (2016) is more complex spatially and is more significantly influenced by

land-atmosphere interactions other than orographic processes (Rasmussen et al., 2011).

Zheng et al. (2018) applied machine learning techniques to estimate SWE at the same 500-m resolution using observations from a dense wireless snow-depth sensor network and a physically-based SWE dataset (i.e., reconstructed SWE from Guan et al. (2013)). To get SWE at snow depth sensor locations, they applied the mean ratio of the co-located snow pillows and snow-depth sensors in the basin for average snow density. Even though the snow density shows much less spatial variability than snow depth (Painter et al., 2016; Sturm et al., 2010), using basin-average snow density may induce uncertainties for the final SWE estimation. Compared with our large-scale Sierra-wide SWE estimation, their SWE modeling is more explicit with a denser snow depth network (i.e., ~50 km^2 per sensor over specific watersheds inluding American Basin and Merced-Tuolumne Basin compared with our ~440 km^2 per site across the Sierra). They also evaluated their SWE estimates against the ASO SWE data in the upper TRB. Specifically, the average RMSE for their SWE estimation during the overlap period (i.e., April 1, 7, and 16, 2016) was 15 cm, which is almost the same as we estimated for the LRM-FSCA (14 cm) on the same dates. They also had an average R^2 value of 0.84, which is slightly lower than the reported 0.88 for LRM-FSCA on the same dates in our study. Although modeling SWE over the entire Sierra Nevada faces greater challenges given the significantly larger modeling domain and lower density of sampling sites, our SWE model has shown a similar performance compared with Zheng et al. (2018)'s SWE model.

LRM-FSCA also compared favorably well with SWE reconstruction models developed in the Sierra Nevada (Guan et al., 2013; Bair et al., 2016, 2018). Guan et al. (2013) reported an RMSE of 20.5 cm and 25.4 cm for their reconstructed SWE and SNODAS, respectively, when evaluated with 20 snow surveys in the Sierra Nevada, showing a 19% (4.9 cm) lower RMSE for their SWE reconstruction model than that for SNODAS. In this study, LRM-FSCA reported a 29% (6.3 cm) and 34% (5.0 cm) lower RMSE than SNODAS when compared with snow course SWE observations and ASO SWE, respectively. Although the evaluation datasets used by Guan et al. (2013) and this study are different, the relative improvements compared with SNODAS indicate that our LRM-FSCA is very likely to have a better performance than Guan et al. (2013)'s SWE reconstruction model. Bair et al. (2016) reported an average of 26.0% RMAE for their reconstructed SWE data when evaluated by ASO SWE data over WYs 2013 to 2015, while LRM-FSCA reported an average RMAE of 24.4% for the same period. Although our LRM-FSCA has a small negative bias (-1.9%), while the SWE estimation of Bair et al. (2016) had 0% bias, the developed LRM-FSCA suggests a comparable accuracy with Bair et al. (2016)'s reconstruction model.

5.2. Model limitations

One limitation of the developed SWE modeling framework is that it heavily relies on ground SWE observations, whereas some other models can be used in regions without ground observations (Guan et al., 2013; Margulis et al., 2016; Rittger et al., 2016; Bair et al., 2018). As a statistically-based model, the number and the distribution of the snow pillow stations have significant impacts on model performance. Hence, the LRM shows relatively better performance from January through April when more stations are covered by snow, providing a better representation of the overall SWE patterns across the Sierra. The model accuracy decreases in May and June as the number of snow pillows with zero SWE values increases.

In addition to the lower representativeness of snow pillow SWE observations in the late snowmelt season compared with those in the early snow season, the discrepancy due to incompatible spatial support of snow pillow SWE versus grid-based SWE estimates is another reason for the lower accuracy of the LRMs in the late snow season, although we used FSCA to scale the point snow pillow SWE observations to grid-cell in the model. Notwithstanding, the absolute error of the LRM is relatively low given that only a small region is covered by snow in the late snowmelt season (Fig. 10). The influence of this deterioration in model performance during the late snow season has little impact on the estimation of the total snow water storage or snow water availability given that both SWE and snow-covered regions (i.e., high elevation regions) are relatively low in the late snow season.

Furthermore, the spatial representativeness of the SWE training samples may have an impact on the LRM's accuracy. For example, snow pillow stations are typically located in small forest clearings and relatively flat locations (Meromy et al., 2013), there are less snow pillows on the eastern side of the Sierra Nevada, and the elevation distribution of the snow pillow (Fig. 3) differs significantly from the elevation distribution of the modeling domain across the Sierra (Fig. 10). Regions with less snow pillow coverage (e.g., high elevations in Fig. 5) are very likely to see a relatively lower accuracy in LRM SWE estimates.

The uncertainties in the historical SWE reanalysis data and the FSCA data are not considered in this study. The historical SWE reanalysis provides an analogous SWE pattern from a historical date for the target simulation date (Table 1). The basis for the value in using the historical SWE reanalysis is that SWE patterns on the simulation date often show similarity to the SWE from analog dates in the past. Hence, we rely on this analog SWE pattern to improve the statistically-based SWE estimation models (Schneider and Molotch, 2016). The accuracy of the historical SWE data has an important impact on the performance of the LRM given that the historical analog SWE distribution is the most informative predictor variable used in the LRM (Table 1). There are inherent uncertainties in choosing the best historical SWE reanalysis for the LRM. Herein we used snow pillow SWE observations on the target modeling date versus the historical date to inform the selection of the historical SWE date used in the LRM. Given the point-based nature of these snow pillow SWE observations, these data cannot fully capture the heterogeneity of the SWE distributions on the historical and target dates. However, for the historical SWE data selection, the snow pillow SWE observation is the only dataset we can rely on as the 'truth'.

The methods applied here are generally tranferable to other regions in which high-quality historical SWE data are available. For example, the SNSR SWE estimation model is generally applicable to different mountain locations across the world, but at the time of this writing the model has only been implemented in selected areas. Hence, high quality historical SWE data may not be available in other regions (e.g., the upper Colorado River Basin) and this may limit potential applications of the LRM based real-time SWE estimation method. In this context, other datasets with reasonably high accuracy SWE estimations, such as SWE recontructions from Guan et al. (2013), Rittger et al. (2016) and Bair et al. (2018) showing lower accuracy than SNSR but higher accuracy than statistically-based SWE regressions (Schneider and Molotch, 2016; Yang et al., 2018), could also be fed into the LRM. Future work is needed to investigate the influence of errors in different historical SWE data on the LRM SWE estimation accuracy.

In this study, FSCA data were used to determine snow extent, to scale the snow pillow and snow coure SWE observations to better represent SWE over a 500-m modeling grid-cell, and to calculate DMFSCA for the LRM-FSCA simulations. The daily gap-filled FSCA from STC-MODSCAG we used in this study is one of the most advanced products, although the uncertainty of this dataset is not negligible given the technical challenges in addressing issues like cloud cover contamination, canopy cover, and the effect of wide view angles of MODIS imagery particularly over forested regions (Dozier et al., 2008; Bair et al., 2019; Rittger et al., 2020).

To reduce the impacts of uncertainties in the historical SWE data and FSCA data, we scaled the historical SWE data and DMFSCA to mean 0 and standard deviation 1 so that only the patterns were ingested into the LRMs. Thus, the absolute values of the historical SWE data were not relevant to the LRM simulation on the target date. In other words, the performance of LRM will only be impacted by the relative SWE errors of the historical SWE data and the DMFSCA. As quantifying the

uncertainties in FSCA data is beyond the scope of this study, we do not discuss the impacts of the uncertainties in the FSCA-scale snow pillow SWE measurements (i.e., the model training samples) on the model performance.

6. Conclusion

This study offers evidence that satellite-derived DMFSCA contains significant information for statistically-based SWE estimation in real-time. Comparing two LRMs, LRM-FSCA has a better overall performance in SWE estimation than LRM-baseline. Both LRMs show high accuracy in the months January, February, and March, a slightly lower accuracy in April, and poor performance in May and June when there is less snow. Adding the DMFSCA improves LRM-FSCA median R^2 value from 0.54 to 0.60 and reduces the median PBIAS from 2.6% to 2.2% when validated by the snow pillow SWE observations in cross-validation. LRM-FSCA also shows a similar improvement in R^2 and PBIAS (from 0.81 to 0.87 and from 1.4% to 0.1%, respectively) in the snow course validation. The improvement of LRM-FSCA is less evident when validated by the ASO SWE data in the TRB. 85% of the ASO SWE variance is explained by the LRM-FSCA, with the median RMSE, RMAE, PBIAS values of 13.3 cm, 18.9% and -9.0%, which is comparable with LRM-baseline (0.86 R^2, 13.8 cm RMSE, 19.0% RMAE, –7.4% PBIAS) in the evaluation using ASO SWE data (i.e., 23 ASO flight covering the WYs 2013 through 2019) in the TRB. The SWE estimation errors for the LRMs exhibit similar patterns in the year-to-year comparison. The regions with higher errors are mostly found in the valleys, alpine regions, and dense forests. Both LRM-FSCA and LRM-baseline perform much better than SNODAS and NWM-SWE. NWM-SWE shows the lowest accuracy. The maximum snow water storage derived from all the four SWE datasets reveals a consistently high interannual variability, with a 35.2 Gt (139% of the 14-year mean) average difference between the driest year 2015 and wettest year 2011 in the 14-year comparison period, indicating an intense fluctuation of drought and flood risks in California. The developed LRM-FSCA with daily satellite-derived information improves real-time SWE estimation, and it will benefit the reservoir regulations, the management of agricultural water use, and the risk assessment of natural hazards like snow drought and rain on snow floods. Using multi-source observations including remote sensing, ground-based observations, and the statistical-learning model, the developed real-time SWE estimation model has important implications for effective water management and water supply forecasting in California.

CRediT authorship contribution statement

Kehan Yang: Conceptualization, Methodology, Software, Writing – original draft. **Keith N. Musselman:** Writing – review & editing. **Karl Rittger:** Resources, Writing – review & editing, Data curation, Validation. **Steven A. Margulis:** Resources, Writing – review & editing. **Thomas H. Painter:** Resources, Writing – review & editing. **Noah P. Molotch:** Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was supported by the NASA grants NNX17AF50G, 80NSSC17K0071, NNX17AF50G, 80NSSC18K0427 P00001, NA18OAR4590367. We also want to acknowledge the support from the Niwot Ridge LTER (NSF DEB 1637686) and NOAA RISA Western Water Assessment.

References

Adam, J.C., Clark, E.A., Lettenmaier, D.P., Wood, E.F., 2006. Correction of global precipitation products for orographic effects. J. Clim. 19 (1), 15–38. https://doi.org/10.1175/JCLI3604.1.

Andreadis, K.M., Lettenmaier, D.P., 2006. Assimilating remotely sensed snow observations into a macroscale hydrology model. Adv. Water Resour. 29 (6), 872–886. https://doi.org/10.1016/j.advwatres.2005.08.004.

Bair, E.H., Rittger, K., Davis, R.E., Painter, T.H., Dozier, J., 2016. Validating reconstruction of snow water equivalent in California's Sierra Nevada using measurements from the NASA Airborne Snow Observatory. Water Resour. Res. 52 (11), 8437–8460. https://doi.org/10.1002/2016WR018704.

Bair, E.H., Calfa, A.A., Rittger, K., Dozier, J., 2018. Using machine learning for real-time estimates of snow water equivalent in the watersheds of Afghanistan. Cryosphere 12 (5), 1579–1594. https://doi.org/10.5194/tc-12-1579-2018.

Bair, E.H., Rittger, K., Skiles, S.M., Dozier, J., 2019. An examination of snow albedo estimates from MODIS and their impact on snow water equivalent reconstruction. Water Resour. Res. 55 (9), 7826–7842. https://doi.org/10.1029/2019WR024810.

Bales, R.C., Molotch, N.P., Painter, T.H., Dettinger, M.D., Rice, R., Dozier, J., 2006. Mountain hydrology of the western United States. Water Resour. Res. 42 (8), W08432. https://doi.org/10.1029/2005WR004387.

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

Barrett, A.P. (2003). National operational hydrologic remote sensing center snow data assimilation system (SNODAS) products at NSIDC. NSIDC Spec. Rep. 11. 10.7265/N5TB14TC.

Bormann, K.J., Hedrick, A., Patterson, V.M., Deems, J.S., Marks, D., Painter, T.H., 2018. Per-pixel uncertainty for the Airborne Snow Observatory's SWE products. In: Proceedings of the American Geophysical Union Falling Meeting. Washington, D.C..

Carroll, S.S., Carroll, T.R., Poston, R.W., 1999. Spatial modeling and prediction of snow-water equivalent using ground-based, airborne, and satellite snow data. J. Geophys. Res. Atmos. 104 (D16), 19623–19629. https://doi.org/10.1029/1999JD900093.

Carroll, T., Cline, D., Fall, G., Nilsson, A., Li, L., Rost, A., 2001. NOHRSC operations and the simulation of snow cover properties for the coterminous U.S. In: Proceedings of the 69th Annual meeting of the Western Snow Conference, pp. 1–14.

Clark, M.P., Slater, A.G., Barrett, A.P., Hay, L.E., McCabe, G.J., Rajagopalan, B., Leavesley, G.H., 2006. Assimilation of snow covered area information into hydrologic and land-surface models. Adv. Water Resour. 29 (8), 1209–1221. https://doi.org/10.1016/j.advwatres.2005.10.001.

Clark, M.P., Hendrikx, J., Slater, A.G., Kavetski, D., Anderson, B., Cullen, N.J., et al., 2011. Representing spatial variability of snow water equivalent in hydrologic and land-surface models: a review. Water Resour. Res. 47 (7), W07539. https://doi.org/10.1029/2011WR010745.

Clow, D.W., Nanus, L., Verdin, K.L., Schmidt, J., 2012. Evaluation of SNODAS snow depth and snow water equivalent estimates for the Colorado Rocky Mountains, USA. Hydrol. Process. 26 (17), 2583–2591. https://doi.org/10.1002/hyp.9385.

Daly, C., Neilson, R.P., Phillips, D.L., 1994. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. J. Appl. Meteorol. 33 (2), 140–158. https://doi.org/10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2.

Deems, J.S., Fassnacht, S.R., Elder, K.J., 2008. Interannual consistency in fractal snow depth patterns at two colorado mountain sites. J. Hydrometeorol. 9 (5), 977–988. https://doi.org/10.1175/2008JHM901.1.

DeWalle, D.R., Rango, A., 2008. Principles of Snow Hydrology. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511535673.

Dozier, J., 1989. Spectral signature of alpine snow cover from the landsat thematic mapper. Remote Sens. Environ. 28 (4), 9–22. https://doi.org/10.1016/0034-4257(89)90101-6.

Dozier, J., Painter, T.H., Rittger, K., Frew, J.E., 2008. Time–space continuity of daily maps of fractional snow cover and albedo from MODIS. Adv. Water Resour. 31 (11), 1515–1526. https://doi.org/10.1016/j.advwatres.2008.08.011.

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

Dubitzky, W., Granzow, M., Berrar, D., 2007. Fundamentals of data mining in genomics and proteomics. Fundamentals of Data Mining in Genomics and Proteomics. https://doi.org/10.1007/978-0-387-47509-7.

Durand, M., Margulis, S.A., 2007. Correcting first-order errors in snow water equivalent estimates using a multifrequency, multiscale radiometric data assimilation scheme. J. Geophys. Res. Atmos. 112 (13), 1–16. https://doi.org/10.1029/2006JD008067.

Durand, M., Molotch, N.P., Margulis, S.A., 2008. A Bayesian approach to snow water equivalent reconstruction. J. Geophys. Res. Atmos. 113 (20), 1–15. https://doi.org/10.1029/2008JD009894.

Erickson, T.A., Williams, M.W., Winstral, A., 2005. Persistence of topographic controls on the spatial distribution of snow in rugged mountain terrain, Colorado, United States. Water Resour. Res. 41 (4), W04014. https://doi.org/10.1029/2003WR002973.

Fassnacht, S.R., Dressler, K.A., Bales, R.C., 2003. Snow water equivalent interpolation for the Colorado River Basin from snow telemetry (SNOTEL) data. Water Resour. Res. 39 (8), 1208. https://doi.org/10.1029/2002WR001512.

Fassnacht, S.R., Dressler, K.A., Hultstrand, D.M., Bales, R.C., Patterson, G., 2012. Temporal inconsistencies in coarse-scale snow water equivalent patterns: colorado River Basin snow telemetry-topography regressions. Pirineos 167 (0), 165–185. https://doi.org/10.3989/Pirineos.2012.167008.

Franz, K.J., Hogue, T.S., Sorooshian, S., 2008. Operational snow modeling: addressing the challenges of an energy balance model for National Weather Service forecasts. J. Hydrol. 360 (1–4), 48–66. https://doi.org/10.1016/j.jhydrol.2008.07.013 (Amst).
Friedman, J., Hastie, T., Tibshirani, R., 2010. Regularization paths for generalized linear models via coordinate descent. J. Stat. Softw. 33 (1), 1–22. https://doi.org/10.18637/jss.v033.i01.
Girotto, M., Cortés, G., Margulis, S.A., Durand, M., 2014. Examining spatial and temporal variability in snow water equivalent using a 27 year reanalysis: kern River watershed, Sierra Nevada. Water Resour. Res. 50 (8), 6713–6734. https://doi.org/10.1002/2014WR015346.
Gochis, D.J., Barlage, M., Dugger, A., FitzGerald, K., Karsten, L., McAllister, M., et al. (2018). The WRF-Hydro modeling system technical description, (Version 5.0). NCAR Technical Note, 107.
Godsey, S.E., Kirchner, J.W., Tague, C.L., 2014. Effects of changes in winter snowpacks on summer low flows: case studies in the Sierra Nevada, California, USA. Hydrol. Process. 28 (19), 5048–5064. https://doi.org/10.1002/hyp.9943.
Gómez-Landesa, E., Rango, A., 2002. Operational snowmelt runoff forecasting in the Spanish Pyrenees using the snowmelt runoff model. Hydrol. Process. 16 (8), 1583–1591. https://doi.org/10.1002/hyp.1022.
Guan, B., Molotch, N.P., Waliser, D.E., Jepsen, S.M., Painter, T.H., Dozier, J., 2013. Snow water equivalent in the Sierra Nevada: blending snow sensor observations with snowmelt model simulations. Water Resour. Res. 49 (8), 5029–5046. https://doi.org/10.1002/wrcr.20387.
Hall, D.K., Riggs, G.A., Salomonson, V.V., DiGirolamo, N.E., Bayr, K.J., 2002. MODIS snow-cover products. Remote Sens. Environ. 83 (1–2), 181–194. https://doi.org/10.1016/S0034-4257(02)00095-0.
Hamlet, A.F., Mote, P.W., Clark, M.P., Lettenmaier, D.P., 2005. Effects of temperature and precipitation variability on snowpack trends in the Western United States. J. Clim. 18 (21), 4545–4561. https://doi.org/10.1175/JCLI3538.1.
Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., et al., 2013. High-resolution global maps of 21st-century forest cover change. Science 342 (6160), 850–853. https://doi.org/10.1126/science.1244693.
He, M., Russo, M., Anderson, M., 2016. Predictability of seasonal streamflow in a changing climate in the Sierra Nevada. Climate 4 (4), 57. https://doi.org/10.3390/cli4040057.
Vander Jagt, B.J., Durand, M.T., Margulis, S.A., Kim, E.J., Molotch, N.P., 2013. The effect of spatial variability on the sensitivity of passive microwave measurements to snow water equivalent. Remote Sens. Environ. 136, 163–179. https://doi.org/10.1016/j.rse.2013.05.002.
Li, D., Wrzesien, M.L., Durand, M., Adam, J., Lettenmaier, D.P., 2017. How much runoff originates as snow in the western United States, and how will that change in the future? Geophys. Res. Lett. 44 (12), 6163–6172. https://doi.org/10.1002/2017GL073551.
Liston, G.E., 1999. Interrelationships among snow distribution, snowmelt, and snow cover depletion: implications for atmospheric, hydrologic, and ecologic modeling. J. Appl. Meteorol. 38 (10), 1474–1487. https://doi.org/10.1175/1520-0450(1999)038<1474:IASDSA>2.0.CO;2.
Liston, G.E., 2004. Representing subgrid snow cover heterogeneities in regional and global models. J. Clim. 17 (6), 1381–1397, 10.1175/1520-0442(2004)017≤1381:RSSCHI≥2.0.CO;2.
Lundquist, J.D., Hughes, M., Henn, B., Gutmann, E.D., Livneh, B., Dozier, J., Neiman, P., 2015. High-elevation precipitation patterns: using snow measurements to assess daily gridded datasets across the Sierra Nevada, California. J. Hydrometeorol. 16 (4), 1773–1792. https://doi.org/10.1175/JHM-d-15-0019.1.
Margulis, S.A., Girotto, M., Cortés, G., Durand, M., 2015. A particle batch smoother approach to snow water equivalent estimation. J. Hydrometeorol. 16 (4), 1752–1772. https://doi.org/10.1175/JHM-D-14-0177.1.
Margulis, S.A., Cortés, G., Girotto, M., Durand, M., 2016. A landsat-era sierra nevada snow reanalysis (1985–2015). J. Hydrometeorol. 17 (4), 1203–1221. https://doi.org/10.1175/JHM-D-15-0177.1.
Marks, D., Domingo, J., Susong, D., Link, T., Garen, D., 1999. A spatially distributed energy balance snowmelt model for application in mountain basins. Hydrol. Process. 13 (12–13), 1935–1959. https://doi.org/10.1002/(SICI)1099-1085(199909)13:12/13<1935::AID-HYP868>3.0.CO;2-C.
Mendoza, P.A., Musselman, K.N., Revuelto, J., Deems, J.S., López-Moreno, I., McPhee, J., 2020. Inter-annual and seasonal variability of snow depth scaling behavior in a sub-alpine catchment. Water Resour. Res. 1–17. https://doi.org/10.1029/2020wr027343.
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. Hydrol. Process. 27 (17), 2383–2400. https://doi.org/10.1002/hyp.9355.
Milly, P.C.D., Dunne, K.A., 2002. Macroscale water fluxes 1. Quantifying errors in the estimation of basin mean precipitation. Water Resour. Res. 38 (10) https://doi.org/10.1029/2001WR000759, 23-1-23–14.
Minder, J.R., Letcher, T.W., Skiles, S.M.K., 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. J. Geophys. Res. 121 (15), 9069–9088. https://doi.org/10.1002/2016JD024995.
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. Hydrol. Process. 23 (7), 1076–1089. https://doi.org/10.1002/hyp.7206.
Molotch, N.P., Bales, R.C., 2005. Scaling snow observations from the point to the grid element: implications for observation network design. Water Resour. Res. 41 (11), W11421. https://doi.org/10.1029/2005WR004229.
Molotch, N.P., Colee, M.T., Bales, R.C., Dozier, J., 2005. Estimating the spatial distribution of snow water equivalent in an alpine basin using binary regression tree models: the impact of digital elevation data and independent variable selection. Hydrol. Process. 19 (7), 1459–1479. https://doi.org/10.1002/hyp.5586.
Mote, P.W., Hamlet, A.F., Clark, M.P., Lettenmaier, D.P., 2005. Declining mountain snowpack in western north America. Bull. Am. Meteorol. Soc. 86 (1), 39–49. https://doi.org/10.1175/BAMS-86-1-39.
Mote, P.W., Li, S., Lettenmaier, D.P., Xiao, M., Engel, R., 2018. Dramatic declines in snowpack in the western US. NPJ Clim. Atmos. Sci. 1 (1) https://doi.org/10.1038/s41612-018-0012-1.
Nijssen, B., O'donnell, G.M., Hamlet, A.F., Lettenmaier, D.P., 2001. Hydrologic sensitivity of global rivers to climate change. Clim. Change 50, 143–175. https://doi.org/10.1023/A:1010616428763 (2001).
Niu, G.Y., Yang, Z.L, 2007. An observation-based formulation of snow cover fraction and its evaluation over large North American river basins. J. Geophys. Res. 112 (D21), D21101. https://doi.org/10.1029/2007JD008674.
Oaida, C.M., Reager, J.T., Andreadis, K.M., David, C.H., Levoe, S.R., Painter, T.H., et al., 2019. A high-resolution data assimilation framework for snow water equivalent estimation across the Western United States and validation with the Airborne Snow Observatory. J. Hydrometeorol. 20 (3), 357–378. https://doi.org/10.1175/JHM-D-18-0009.1.
Pagano, T., Garen, D., Sorooshian, S., 2004. Evaluation of official Western U.S. seasonal water supply outlooks, 1922–2002. J. Hydrometeorol. 5 (5), 896–909. https://doi.org/10.1175/1525-7541(2004)005<0896:eoowus>2.0.co;2.
Painter, T.H., Rittger, K., McKenzie, C., Slaughter, P., Davis, R.E., Dozier, J., 2009. Retrieval of subpixel snow covered area, grain size, and albedo from MODIS. Remote Sens. Environ. 113 (4), 868–879. https://doi.org/10.1016/j.rse.2009.01.001.
Painter, T.H., Berisford, D.F., Boardman, J.W., Bormann, K.J., Deems, J.S., Gehrke, F., et al., 2016. The Airborne Snow Observatory: fusion of scanning lidar, imaging spectrometer, and physically-based modeling for mapping snow water equivalent and snow albedo. Remote Sens. Environ. 184 (July), 139–152. https://doi.org/10.1016/j.rse.2016.06.018.
Pekel, J.F., et al., 2016. High-resolution mapping of global surface water and its long-term changes. Nature 540, 418–422. https://doi.org/10.1038/nature20584.
Pflug, J.M., Lundquist, J.D., 2020. Inferring distributed snow depth by leveraging snow pattern repeatability: investigation using 47 Lidar observations in the tuolumne watershed, Sierra Nevada, California. Water Resour. Res. 56 (9), 1–17. https://doi.org/10.1029/2020WR027243.
Raleigh, M.S., Lundquist, J.D., 2012. Comparing and combining SWE estimates from the SNOW-17 model using PRISM and SWE reconstruction. Water Resour. Res. 48 (1), W01506. https://doi.org/10.1029/2011WR010542.
Ralph, F.M., Neiman, P.J., Wick, G.A., 2004. Satellite and CALJET aircraft observations of atmospheric rivers over the Eastern North Pacific Ocean during the winter of 1997/98. Mon. Weather Rev. 132 (7), 1721–1745. https://doi.org/10.1175/1520-0493(2004)132<1721:SACAOO>2.0.CO;2.
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. J. Clim. 24 (12), 3015–3048. https://doi.org/10.1175/2010JCLI3985.1.
Rittger, K., Painter, T.H., Dozier, J., 2013. Assessment of methods for mapping snow cover from MODIS. Adv. Water Resour. 51, 367–380. https://doi.org/10.1016/j.advwatres.2012.03.002.
Rittger, K., Bair, E.H., Kahl, A., Dozier, J., 2016. Spatial estimates of snow water equivalent from reconstruction. Adv. Water Resour. 94, 345–363. https://doi.org/10.1016/j.advwatres.2016.05.015.
Rittger, K., Raleigh, M.S., Dozier, J., Hill, A.F., Lutz, J.A., Painter, T.H., 2020. Canopy adjustment and improved cloud detection for remotely sensed snow cover mapping. Water Resour. Res. 56 (6), 2019WR024914 https://doi.org/10.1029/2019WR024914.
Rittger, K., Bormann, K.J., Bair, E.H., Dozier, J., Painter, T.H., 2021. Evaluation of VIIRS and MODIS snow cover fraction in high-mountain asia using landsat 8. Front. Remote Sens. 2, 647154 https://doi.org/10.3389/frsen.2021.647154.
Roos, M., 2004. California's cooperative snow surveys program. In: Proceedings of the ICID 2nd Asian Regional Conference. Echuca–Moama, Australia. Citeseer.
Sarangi, C., Qian, Y., Rittger, K., Bormann, K.J., Liu, Y., Wang, H., et al., 2019. Impact of light-absorbing particles on snow albedo darkening and associated radiative forcing over high-mountain Asia: high-resolution WRF-Chem modeling and new satellite observations. Atmos. Chem. Phys. 19 (10), 7105–7128. https://doi.org/10.5194/acp-19-7105-2019.
Sarangi, C., Qian, Y., Rittger, K., Ruby Leung, L., Chand, D., Bormann, K.J., Painter, T.H., 2020. Dust dominates high-altitude snow darkening and melt over high-mountain Asia. Nat. Clim. Chang. 10 (11), 1045–1051. https://doi.org/10.1038/s41558-020-00909-3.
Schneider, D., Molotch, N.P., 2016. Real-time estimation of snow water equivalent in the Upper Colorado River Basin using MODIS-based SWE Reconstructions and SNOTEL data. Water Resour. Res. 52 (10), 7892–7910. https://doi.org/10.1002/2016WR019067.
Serreze, M.C., Clark, M.P., Armstrong, R.L., McGinnis, D.A., Pulwarty, R.S., 1999. Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data. Water Resour. Res. 35 (7), 2145–2160. https://doi.org/10.1029/1999WR900090.
Simon, N., Friedman, J., Hastie, T., Tibshirani, R., 2011. Regularization paths for Cox's proportional hazards model via coordinate descent. J. Stat. Softw. 39 (5), 1–13. https://doi.org/10.18637/jss.v039.i05.

Simon, N., Friedman, J., Hastie, T., 2013. A blockwise descent algorithm for group-penalized multiresponse and multinomial regression. Wiley Interdiscip. Rev. Comput. Stat. 1 (1), 128–129. https://doi.org/10.1002/wics.10.

Stewart, I.T., Cayan, D.R., Dettinger, M.D., 2004. Changes in snowmelt runoff timing in Western North America under a `business as usual' climate change scenario. Clim. Chang. 62 (1–3), 217–232. https://doi.org/10.1023/B:CLIM.0000013702.22656.e8.

Sturm, M., Taras, B., Liston, G.E., Derksen, C., Jonas, T., Lea, J., 2010. Estimating snow water equivalent using snow depth data and climate classes. J. Hydrometeorol. 11 (6), 1380–1394. https://doi.org/10.1175/2010JHM1202.1.

Swain, D.L., Horton, D.E., Singh, D., Diffenbaugh, N.S., 2016. Trends in atmospheric patterns conducive to seasonal precipitation and temperature extremes in California. Sci. Adv. 2 (4), e1501344 https://doi.org/10.1126/sciadv.1501344.

Takala, M., Ikonen, J., Luojus, K., Lemmetyinen, J., Metsamaki, S., Cohen, J., et al., 2017. New snow water equivalent processing system with improved resolution over Europe and its applications in hydrology. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10 (2), 428–436. https://doi.org/10.1109/JSTARS.2016.2586179.

Tibshirani, R., Bien, J., Friedman, J., Hastie, T., Simon, N., Taylor, J., Tibshirani, R.J., 2012. Strong rules for discarding predictors in lasso-type problems. J. R. Stat. Soc. Ser. B Stat. Methodol. 74 (2), 245–266. https://doi.org/10.1111/j.1467-9868.2011.01004.x.

Trujillo, E., Molotch, N.P., Goulden, M.L., Kelly, A.E., Bales, R.C., 2012. Elevation-dependent influence of snow accumulation on forest greening. Nat. Geosci. 5 (10), 705–709. https://doi.org/10.1038/ngeo1571.

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. Int. J. Climatol. 35 (9), 2472–2484. https://doi.org/10.1002/joc.4136.

Wrzesien, M.L., Pavelsky, T.M., Durand, M.T., Dozier, J., Lundquist, J.D., 2019. Characterizing biases in mountain snow accumulation from global data sets. Water Resour. Res. 55 (11), 9873–9891. https://doi.org/10.1029/2019WR025350.

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. J. Geophys. Res. Atmos. 117, D03110. https://doi.org/10.1029/2011JD016051.

Xue, Y., Sun, S., Kahan, D.S., Jiao, Y., 2003. Impact of parameterizations in snow physics and interface processes on the simulation of snow cover and runoff at several cold region sites. J. Geophys. Res. Atmos. 108 (D22), 2002JD003174 https://doi.org/10.1029/2002JD003174.

Yang, K., Musselman, K.N., Schneider, D., Painter, T.H., Margulis, S.A., Rittger, K., et al., 2018. An inter-comparison of five snow water equivalent estimation methods in the Sierra Nevada Mountains, California. In: Proceedings of the AGU Fall Meeting Abstracts (Vol. 2018, pp. C13I-1241).

Zheng, Z., Molotch, N.P., Oroza, C.A., Conklin, M.H., Bales, R.C., 2018. Spatial snow water equivalent estimation for mountainous areas using wireless-sensor networks and remote-sensing products. Remote Sens. Environ. 215 (March), 44–56. https://doi.org/10.1016/j.rse.2018.05.029.