Global snow water equivalent product derived from machine learning model trained with in situ measurement data

Jungho Seo[1,2], Mahdi Panahi[1], JiHyun Kim[1], Sayed M. Bateni[2,3], Yeonjoo Kim[1]

[1]Civil and Environmental Engineering, Yonsei University, Seoul, 03722, South Korea
5   [2]Civil, Environmental and Construction Engineering, and Water Resources Research Center, University of Hawaii at Manoa, Honolulu, Hawaii, 96822, USA
[3]UNESCO-UNISA African Chair in Nanoscience and Nanotechnology College of Graduate Studies, University of South Africa, Muckleneuk Ridge, Pretoria 392, South Africa

10   *Correspondence to*: Yeonjoo Kim (yeonjoo.kim@yonsei.ac.kr)

Abstract. Snow water equivalent (SWE) quantifies the volume of water stored in snowpacks and therefore critically attributes to the timing and amount of water discharged into groundwater sources and rivers. The SWE has been estimated using various methods, including in situ measurements, remote sensing, and physics-based models. However, each of these methods present certain limitations, including high costs, low spatiotemporal resolution, and uncertainty in model representation and parameter

15   calibration. To address these challenges, in this study, we developed a machine learning-based daily global gridded SWE (SWEML) product with a spatial resolution of 0.25º, covering the period from 1980 through 2020. To develop this product, we first applied the k-means clustering algorithm using topographical and climatic variables to classify global in situ SWE measurements into 13 clusters. Subsequently, we adopted the random forest algorithm to correlate daily in situ SWE measurements ($n = 11,653$) with meteorological forcing and terrain attributes. We compared SWEML with other SWE datasets,

20   including the GlobSnow dataset from the European Space Agency, the Global Land Data Assimilation System dataset, and SWE estimates from the Advanced Microwave Scanning Radiometer for the Earth Observation System. The overall root mean square error (RMSE) was 10.80 mm, and the overall bias was -6.89 mm globally, in particular, with high accuracy with Pearson correlation coefficient, R, of 0.99 and RMSE of 16.88 mm in mountainous and high-elevation areas, such as the Rocky Mountains in the U.S. Furthermore, both snow accumulation during winter and snow melting during spring, were well depicted

25   in the SWEML, which is only possible with a high-temporal-resolution product. Overall, the daily gap-free global SWEML product introduced in this study can significantly contribute to water resource management efforts in snow-dominant regions

and provide a robust reference for data assimilation in global-scale land surface modeling. The SWEML is available at

https://doi.org/10.5281/zenodo.14195794 (Seo et al., 2024).

1. Introduction

30

Snow is an essential component of the hydrological cycle of Earth, affecting the volume of water flowing into rivers, lakes, and streams and providing a supply of water for millions of people (Barnett et al., 2005; Viviroli et al., 2011; Musselman et al., 2021). Snow can be characterized by its depth, cover, and snow water equivalent (SWE), which quantifies the amount of water stored in the snowpack. Accurate SWE estimations are essential for various purposes, including water resource

35   management (Musselman et al., 2021), global water cycle modeling (Kraft et al., 2022), and the prediction of landslides (Matsuyama et al., 2021), flood (Meshyk et al., 2021), and drought prediction (Dierauer et al., 2019). Despite being recognized as an essential climate variable by the Global Climate Observing System (GCOS) of the World Meteorological Organization, SWE remains an under-observed component of the global water cycle (Mudryk et al., 2015; Pulliainen et al., 2020; Tsang et al., 2022).

40   To date, numerous ground-based observation networks have been deployed to measure the SWE, including the Snow Climate Analysis Network (SCAN), SNOwpack TELemetry Network (SNOTEL), Russian Snow Survey Data (RSSD), and Norwegian Water Resources and Energy Directorate (NVE). However, these ground-based measurements are expensive, labor-intensive, limited to point scale data, and insufficient to cover the vast extent of the Northern Hemisphere (Tsang et al., 2022; Fang et al., 2022). As viable alternatives, passive remote sensing observations can provide long-term global-scale SWE data (Luojus

45   et al., 2021; Tsang., 2022). However, SWE estimates based on remote sensing have some inherent limitations depending on the sensors mounted onto each instrument (Frei et al., 2012; Fang et al., 2022). For instance, SWE retrievals from passive microwave instruments, such as the Special Sensor Microwave/Imager and the Advanced Microwave Scanning Radiometer for the Earth Observation System (AMSRE), exhibit uncertainties under deep-snow conditions (i.e., SWE >150 mm) and are particularly sensitive to forest cover (Frei et al., 2012; Mortimer et al., 2020). Hence, significant efforts have been devoted to

50   leveraging radar data across different bands (e.g., C, P, Ku, and X) to acquire long-term spatially distributed SWE data (Yueh et al., 2021). This use of active data has yielded promising results, but their implementation remains challenging and demands advanced technical developments (Merkouriadi et al., 2021; Tsang et al., 2022).

In addition to passive remote sensing observations, physically based models also provide seamless global-scale SWE estimates (Kim et al., 2021), but there can be substantial differences between their estimates owing to uncertainties in model

55 parameterizations, simplistic model assumptions, and structural model errors (Snauffer et al., 2016; Lundquist et al., 2019; Wrzesien et al., 2019; Menard et al., 2021). To address these challenges, Pulliainen et al. (2006) and Takala et al. (2011) combined ground synoptic weather measurements with space-borne passive microwave data within a Bayesian non-linear iterative assimilation approach to produce the GlobSnow v3.0 dataset. Despite certain limitations, such as overestimations and the exclusion of mountain areas, GlobSnow v3.0 is still the reliable SWE product at a continental scale across the Northern

60 Hemisphere (Takala et al., 2011; Luojous et al., 2021). Each of these existing datasets has its own limitations, and thus it remains necessary to develop an accurate SWE product that can be used reliably by the scientific community (Bokhorst et al., 2016; Mortimer et al., 2020).

Machine learning (ML) has been used as a viable alternative approach that merges independent variables from various sources to simulate the SWE (Broxton et al., 2019; Shao et al., 2022; Cui et al., 2023). Furthermore, these ML approaches have been

65 demonstrated to effectively learn the complex relationships between meteorological variables (i.e., predictors) and the SWE (Fang et al., 2019; Cui et al., 2023). Consequently, numerous studies have utilized ML for SWE estimations. For instance, Broxton et al. (2019) adopted an artificial neural network to combine snow density measurements with lidar snow depth data, thereby creating high-resolution SWE maps. Furthermore, Shao et al. (2022) applied the ridge regression ML approach for SWE estimation in regions above 45°N, while Cui et al. (2023) employed a long short-term memory algorithm for SWE

70 estimation in Sierra Nevada, California. As illustrated in these studies, ML algorithms can provide offer more accurate SWE estimates compared to traditional physics- and remote-sensing-based approaches (Bair et al. 2018; Broxton et al., 2019; Shao et al., 2022; Cui et al., 2023). Furthermore, ML approaches can successfully predict or simulate various hydrometeorological processes, including runoff (e.g., Ghiggi et al., 2019), streamflow (e.g., Cheng et al., 2020; Cho and Kim, 2022), soil moisture (e.g., O and Orth, 2021), and precipitation (e.g., Ravuri et al., 2021; Choi and Kim, 2022), along with their social impacts (e.g.,

75 Dodangeh et al., 2020; Seo and Kim, 2023) using diverse meteorological and geographical datasets from local to global scales.

In this study, we develop a new global-scale SWE product covering the period from 1980 through 2020 using the random forest (RF) algorithm. Notably, the RF algorithm uses meteorological and terrain attributes as predictor variables and relates them to daily SWE measurements as dependent variables. To develop the SWE estimation tool, we collected 11,653 daily in situ SWE measurements from eight sources during 1980–2020. Subsequently, mini-batch k-means (MBK) clustering was

80   adopted to regionalize these measurements. Thereafter, we constructed SWEML – a global daily SWE product covering all land areas except Antarctica, using RF models trained for each cluster. SWEML was evaluated with that of reference datasets, and it demonstrated an improved performance, especially at high elevations and the ability to reflect monthly trends more precisely. A schematic overview of this study is provided in Fig. 1.

2. Data and methods

85   ### 2.1. SWE data

2.1.1. In situ measurements and reanalysis data

    In situ daily SWE measurements were collected from eight different sources spread across the Northern Hemisphere (Table 1 and Fig. 2a). The RSSD provides point-scale SWE data across Russia (Bulygina et al., 2011), while the Hemispheric-Scale Snow Course (HSSC) data contain SWE measurements from Russia, Finland, Canada, and some parts of Eastern Europe

90   (Brönnimann et al., 2018; Brown et al., 2019; Pulliainen et al., 2020). Meanwhile, the Global Historical Climatology Network (GHCN) provides SWE measurement data observed by numerous ground stations spread across approximately 179 countries worldwide and has been subject to rigorous quality checks (Menne et al., 2012). In this study, we used GHCN SWE data exclusively for the U.S. owing to temporal limitations. The SNOTEL and SCAN data include an extensive network of automated sub-daily SWE observations, encompassing a diverse range of snow climates and characteristics across the western

95   U.S., including high-latitude tundra and taiga regions, the Rocky Mountains, and the maritime Cascades Mountain range (Serreze et al., 1999; USDA NRCS, 2022b, 2022c). Meanwhile, the California Cooperator Snow Sensors (CCSS) provide sub-daily snowpack data collected from remote mountain locations in the western U.S. (Osterhuber, 2014; USDA NRCS, 2022a). The NVE provides real-time quality-checked SWE data (NVE, 2023), and the GCOS includes SWE data collected from various global sources. Among these diverse sources, we utilized GCOS SWE data from Switzerland (Marty, 2020). Furthermore, we

100   utilized the SWE reanalysis data with a spatial resolution of 0.25° in China, including the in situ measurements and field survey course by the National Meteorological Information Centre of the China Meteorology Administration (Jiang et al., 2022). To ensure data quality, we excluded data with fewer than three months of observations and removed outliers not subjected to data correction by the related agency. Using the coordinates of the selected point-scale SWE measurements, we obtained

meteorological data corresponding to each SWE measurement. The SWE measurements and their corresponding

105 meteorological data are used as outputs and inputs for the RF algorithm. In total, we collected SWE data at 11,653 grid points,

with most of these data covering a period of more than 20 years, with a consistent elevation distribution (Fig. 2b).

In-situ SWE measurements collected by various agencies and countries are inevitably inconsistent because different

measurement methods are used. Furthermore, point-scale data potentially lack spatial representativeness regarding its

surroundings (e.g., high-elevation regions) owing to high heterogeneity. To address these issues, in this study, we adjusted the

110 raw in situ SWE measurements within individual grid cells to align them with the mean and variance of the corresponding

SWE data sourced from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5-Land. This allowed us

to consistently use SWE measurements from different sites and time periods, while maintaining the daily temporal variations.

Notably, point-scale SWE measurements are particularly effective at capturing temporal variations and physical snow

processes (Kinar and Pomeroy, 2015; Zhang and Ma, 2018), thus point-scale data generally provide sufficient information to

115 understand snow dynamics at the grid scale.

2.1.2. Reference datasets for validation

In this study, we employed datasets provided by the Global Land Data Assimilation System (GLDAS, Rodell et al., 2004),

GlobSnow (ESAGB, Luojus et al., 2021), and AMSRE (Kelly, 2009; Imaoka et al., 2010; Tedesco and Jeyaratnam, 2019) for

result validation. Notably, the GLDAS, based on a land surface model, provides precipitation, temperature, runoff, soil

120 moisture, and snow data with spatial resolutions of 0.25° and 1.0° and temporal resolutions of 3 hours, 1 day, and 1 month

(Rodell et al., 2004). For this study, we used the daily SWE dataset from the GLDAS Catchment Land Surface Model.

Meanwhile, the ESAGB dataset includes SWE and snow-cover data for the Northern Hemisphere above a latitude of 35° N,

released by the European Space Agency (Luojus et al., 2021). The SWE estimates in this dataset are based on the combination

of Canadian Meteorological Center daily snow depth analysis data (Walker et al., 2011), ground-based weather station data,

125 and satellite-based microwave radiometer observations. However, the ESAGB dataset exhibits limited spatial coverage as it

excludes regions with very-high-elevation mountainous areas or grids below 35°N. AMSRE is a satellite-based dataset that

includes data collected by microwave scanning radiometers on the Aqua satellite of the National Aeronautics and Space

Administration (NASA) Earth Observation System and the GCOM-W1 satellite of the Japan Aerospace Exploration Agency

(JAXA). The AMSRE from NASA (AMSR-E) provides global daily SWE data from 19 June, 2002, to 3 October, 2011, and

130   the AMSRE from JAXA (AMSR2) has been providing global daily SWE data from 2 July, 2012, to the present. For this study,

we utilized SWE data with a spatial resolution of 0.25° (approximately 25 km) and a daily temporal resolution. Although the

AMSRE satellite began recording observations in 2002, there are gaps during 2011 and 2012. Therefore, we used the grids

available across all datasets to ensure consistency in the spatiotemporal extent of the reference data. We used each SWE dataset

for validation based on the spatiotemporal extent of AMSRE and ESAGB, focusing on grids above 35° N and the period from

135   2013 through 2017. In addition, we considered only the observed dates based on the rotation cycle of AMSRE. Consequently,

we evaluated the accuracy of SWEML using 1,706,825 data points from 6,133 sites between 2013 and 2017. We calculated

the error to evaluate the accuracy by comparing SWEML and reference datasets with in situ measurements. The root mean

square error (RMSE), mean absolute error (MAE), Pearson correlation coefficient (R) and bias were adopted to assess the

accuracy of the SWE datasets. The equations of the accuracy evaluation metric are described as follows:

140   $$RMSE = \sqrt{\frac{\sum_{i=0}^{n}(r_i - m_i)^2}{n}},$$ (1)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|r_i - m_i|,$$ (2)

$$R = \frac{\sum(r_i - \bar{r})(m_i - \bar{m})}{\sqrt{\sum(r_i - \bar{r})^2 \sum(m_i - \bar{m})^2}},$$ (3)

$$Bias = r_i - m_i,$$ (4)

where n is the number of samples, r_i is the SWE value of each dataset, m_i is the in situ measurements, $\bar{r}$ and $\bar{m}$ are the average

145   of SWE dataset and in situ measurements, respectively.

2.2. Ancillary data

2.2.1. Topographical data

Five terrain characteristics including vegetation type, elevation, and the angle, anisotropy, and slope of the orography were

collected from multiple sources (Table 2) and used as inputs to the RF algorithm owing to the significant influence of these

150   characteristics on the SWE. Vegetation type data were sourced from GLDAS Dominant Vegetation Type Data Version 2. It

affects the melting and accumulation of snow (Friedl et al., 2002; Varhola et al., 2010; Snauffer et al., 2016; Luojus et al.,

2021) owing to its shading effects, such as canopy interception and changes in the radiation balance, that can affect ablation

by up to 70% and accumulation by up to 40% compared to nearby open areas (Varhola et al., 2010). Elevation is an important

predictor of the SWE that has been widely used in previous studies (Kim et al., 2021; Fang et al., 2022; Shao et al., 2022).

155 Higher elevations are typically associated with lower temperatures, leading to an earlier onset of the snowing season and

delayed melting compared to lower elevations (Xie et al., 2020). Elevation also interacts with land cover to influence snow

accumulation, with more pronounced differences in melting observed between clear-cut and forested areas at higher elevations

(Trujillo et al., 2012; Tennant et al., 2017). For this study, we utilized the mean and standard deviation of the digital elevation

data from ETOPO1 (Amante and Eaken, 2009). For orography, which is the study of the terrain characteristics of mountainous

160 areas, we utilized slope, aspect, and anisotropy data sourced from ERA5 (Harrison et al., 2022; Hersbach et al., 2023). Notably,

slope, aspect, and anisotropy have a strong relationship with solar radiation and blowing wind, which affects SWE in

mountainous regions (Tennant et al., 2017; Niedzielski et al., 2019). Consequently, numerous studies have adopted these

variables as SWE predictors (Niedzielski et al., 2019; Fang et al., 2022; Cui et al., 2023).

2.2.2. Meteorological data

165 We utilized hourly meteorological forcing data with a spatial resolution of 0.25° from the ERA5 reanalysis dataset, provided

by the ECMWF (Hersbach et al., 2020). We converted the temporal scale of meteorological forcing data from hourly to daily

for our use. Notably, ERA5 merges various types of data, such as ground-based synoptic station data, radar precipitation

information, and satellite radiance via the 4D-Var data assimilation, to provide comprehensive and accurate reanalysis datasets

(Hersbach et al., 2020). It incorporates more observation data and advanced modeling techniques compared to its predecessor,

170 ERA-Interim, thus it offers improved forcing data (Albergel et al., 2018). Furthermore, the ERA5 dataset provides a consistent

data structure that is suitable for the derivation of long-term global-scale data. The direct use of observation data, such as

satellite data across a global scale, can pose challenges due to spatiotemporal differences between the variables of interest.

Therefore, the SWE data constructed using our approach can be employed to evaluate the accuracy of SWE data originating

from other approaches. ERA5 is also open source and provides the latest data, suggesting that our products can be updated

175 regularly as needed.

In this study, we trained the RF model using five meteorological forcings (land surface temperature, precipitation, relative

humidity, downward solar radiation, and precipitation type), snow depth, and the antecedent precipitation index (API) (Table

2). Notably, the number of years and months were also included to account for annual and seasonal variations, respectively.

We selected input forcing variables based on a literature review, and correlation analysis. Correlation analysis in Fig. 3 showed

180   that each input variable exhibited weak correlations with other input variables, suggesting that they could be considered

independent. Among the selected meteorological forcings, snow depth exhibits the strongest negative correlation with land

surface temperature, while relative humidity has the strongest positive correlation with precipitation type. Relative humidity,

in particular, influences the temperature threshold for rain–snow partitioning (Ye et al., 2013; Jennings et al., 2018; Wang et

al., 2019). Precipitation is a fundamental factor controlling the SWE (Raleigh et al., 2015; Broxton et al., 2016b; Cho et al.,

185   2022), with the SWE being sensitive to both the amount of precipitation (Schirmer et al., 2011; Broxton et al., 2016a) and

rain–snow partitioning (Harder and Pomeroy, 2014; Jennings and Molotch, 2019). Hence, the precipitation type (rain or snow)

on land (Hersbach et al., 2023) and relative humidity were considered as inputs in this study to address rain–snow partitioning.

Snow depth directly affects the SWE and was, therefore, adopted as an input variable (Hill et al., 2019). Meanwhile, land

surface temperature and downward solar radiation affect the phase transitions of snow, influencing its melting and freezing

190   behaviors upon contact with the land surface; hence, these variables significantly impact the SWE (Anderton et al., 2004; Lapo

et al., 2015). Finally, SWE is influenced by the precipitation in prior days; hence, we utilized the API, which represents the

accumulation of precipitation, as an input variable. Notably, the API is calculated as the weighted sum of daily precipitation

over previous days (Kohler and Linsley, 1951; Singh, 1989) and is defined as follows:

$$API = \sum_{n=-1}^{-i} P_n k^{-n},$$ (5)

195   where P_n is the precipitation on the nth day, i denotes the number of previous days, and k is the decay constant. Typically, i

is set to 5 days or more (Ali et al., 2010; Hong et al., 2010), while k, representing an empirical parameter, is set to values

ranging from 0.8 to 1.0 (Zêzere et al., 2005; Ma et al., 2014; Dingman, 2015; Li et al., 2021a). In this study, k was varied from

0.8 to 1 in increments of 0.05, while i was varied from 1 to 30 in increments of 1. Finally, we selected i and k values that led

to the highest correlation between the API and in situ SWE. In this study, i and k of values were 23 and 1, respectively, yielded

200   the highest correlation coefficient between API and SWE.

2.3. Regionalization

SWE presents regional variability based on meteorological and geographical characteristics (Mizukami et al., 2011; Herderson et al., 2018; Pulliainen et al., 2020). Hence, regionalization can be useful for analyzing the SWE dynamics, especially when SWE data are collected in different continents with diverse meteorological and geographical attributes. In this study, we regionalized the training and testing data (Fig. 4) using the MBK clustering method, a modification of the traditional k-means clustering algorithm (Sculley, 2010). Notably, the traditional k-means approach is an unsupervised ML algorithm that assigns each data point to the nearest cluster and iterates this process until convergence, thereby dividing the dataset into distinct, non-overlapping clusters (Li et al., 2021b; Hu et al., 2023). However, this approach can be computationally intensive and inefficient owing to the repeated use of every data point in each iteration. The MBK clustering method addresses these issues by adopting a probabilistic approach for each iteration, thus reducing computational costs, particularly for extensive datasets (Biswas et al., 2023; Ikotun et al., 2023).

To determine the optimal number of clusters, we adopted three techniques: the elbow method (Thorndike, 1953), silhouette coefficient (Rousseeuw, 1987), and gap statistic (Tibshirani et al., 2001). In particular, the elbow method determines the optimal cluster count by visualizing the point at which the sum of the squared error (SSE) between points within a cluster and the cluster centroid is minimized. The silhouette coefficient represents the average distance between a data point and other points within the same cluster relative to the mean distance to points in other clusters. Hence, higher silhouette coefficients indicate better cluster quality. The gap statistic compares the total intra-cluster variation for different cluster numbers with the expected variation under an appropriate reference null distribution (i.e., a distribution with no clustering), aiming to identify the point where the gap statistic is maximized. The number of reference null distributions for gap statistics is subjective to the user. In this study, we used 500 reference null distributions, as suggested by Iwasawa et al. (2022), where the statistics remained stable.

The optimal number of clusters was determined based on the silhouette coefficient, beginning from the intersection point at which the SSE is minimized and the gap statistics are maximized. Based on the intersection between the elbow and gap statistics, the minimum number of clusters was identified as 10, while the silhouette coefficient values suggested that the optimal number of clusters was 13. Therefore, we partitioned the 11,653 training data points into 13 clusters (Fig. 4). Northern

America was broadly divided into four regions based on distinctive features in regions such as Alaska and the Rocky Mountains. Eurasia was divided into three major regions: Europe, the Far East, and inland Central Siberia. China was divided into six regions, the largest number of clusters, where significant variability in meteorological and geographical characteristics is presented with the largest number of grid points (Carvalho et al., 2016).

230   **2.4. ML algorithm**

We utilized the RF algorithm to generate a global SWE product as introduced in the schematic of Fig. 1. The RF is an ensemble-learning method that utilizes decision trees as its foundation and has been widely used for classification and regression due to its high accuracy and speed (Breiman, 2001). The RF is based on the concept of bootstrap aggregating (i.e., bagging), which generates a bootstrapped datasets by randomly selecting samples with replacements for each decision tree.
235   This technique reduces overfitting instances by averaging over multiple decision trees, making the model less sensitive to the original training dataset, potential noise, and outliers, thus resulting in better predictive performance and greater stability (Josso et al., 2023; Isles, 2024). In RF, each decision tree is trained on an individual bootstrap sample, resulting in some data points being unused in the training process (referred to as out-of-bag samples). An advantage of this process is that these samples can be used to assess the performance of the model without the need for a separate validation set.

240   As described in Fig. 1 and Sect 2.3, we utilized the MBK algorithm to generate 13 clusters for SWE measurements. The regionalization allowed for more accurate and efficient training of the ML model owing to the inclusion of diverse in situ SWE data collected from various global sites. We employed two types of RF models to derive the global SWE. First, an individual RF regression model for each cluster is used to establish the relationship between the in situ SWE measurements and meteorological forcing and terrain data, allowing the SWE to be determined directly for each grid cell. An individual RF
245   regression model for each cluster uses data divided at a 7:3 ratio for its training and validation. Furthermore, by clustering the available SWE data, we could account for the regional variability in the SWE, leading to a more accurate estimation of the global SWE.

Second, an RF classification model is used to establish the relationship between SWE for each cluster and the corresponding meteorological forcing and terrain data. We use data divided at an 8.5:1.5 ratio for its training and validation. With a larger
250   dataset, a smaller proportion of data can be allocated to the validation set, while still having sufficient data to evaluate the

performance of the model. The grids without any in-situ SWE measurements were classified into respective clusters using the

RF classification model. After all the grids were assigned into different clusters, an individual RF regression model was

implemented for each cluster, and finally the global SWE estimation was completed.

255   The predictive performance of the models for each cluster was evaluated using three statistical metrics: the coefficient of

determination (R^2), R, and standard deviation. The obtained results were presented using a Taylor diagram, which offers a

comprehensive visual representation using multiple statistical metrics, including R and standard deviation (Taylor et al., 2001).

The standard deviation is particularly valuable for assessing the variability of the predicted data compared to in situ

measurements (Paul et al., 2023).

3 Results and discussion

260   3.1 Evaluating the performance of the RF approach

We evaluated the predictive accuracy of the RF model across the 13 clusters (Figs. 5 and 6). The R^2 values between the

SWEML and in situ SWE measurements ranged from 0.86 to 0.99 across the 13 clusters (Fig. 5), indicating strong model

reliability and good predictive accuracy. As illustrated in the Taylor diagram (Fig. 6), the normalized standard deviation (i.e.,

the ratio of the standard deviation of the predicted SWE to that of the corresponding in situ measurement) ranged from 0.89 to

265   1. This suggests that the SWEML product accurately captured the inherent variability of the observation data. Furthermore,

the R values ranged from 0.93 to 0.99, demonstrating good agreement between the predicted and observation data. Notably,

the RF model accurately represented the high variability of the SWE in snow-dominant regions such as Alaska, the Rocky

Mountains, the Alps, Scandinavia, and the Himalayas (clusters 1, 2, 5, and 10, respectively).

3.2 Comparison of SWEML with existing SWE products

270   We compared the accuracy of SWEML with those of three reference datasets (i.e., GLDAS, ESAGB, and AMSRE) during

2013 through 2017 (Fig. 7). Notably, SWEML exhibited a spatial mean RMSE of 10.80 mm, with a range from 0.15 mm to

30.02 mm, which is lower than the RMSEs of the other products (15.66 mm, 21.65 mm, and 26.06 mm for GLDAS, ESAGB,

and AMSRE, respectively) (Fig. 7a). The spatial distribution of RMSEs of different products showed that the SWEML is

substantially more accurate over snow-dominant regions, such as Alaska, the Rocky Mountains, the Scandinavian Peninsula,

275   Siberia, Northern China, and the Far East. In particular, the SWEML showed significant improvements in Alaska, the Rocky Mountains, and the Scandinavian peninsula compared to the reference datasets (Table 3). For instance, in the Rocky Mountains, the RMSE of SWEML was 16.88 mm, demonstrating reductions of 84.77%, 81.84%, and 84.26% compared to the RMSEs of GLDAS, ESAGB, and AMSRE (110.81 mm, 92.84 mm, 107.34 mm, respectively). Although the spatial mean RMSE of the GLDAS dataset was lower than that of the ESAGB dataset, it exhibited greater deviations in Alaska, the Rocky Mountains,

280   and Russia, with RMSEs that exceeded those of ESAGB by 19.89 mm, 17.97 mm, and 8.03 mm, respectively. While ESAGB exhibited a relatively low spatial mean RMSE of 23.47 mm in Russia, it is still 10 mm higher than that of SWEML. In snow-dominant regions, namely Alaska, the Rocky Mountains, and Scandinavia, ESAGB demonstrated higher errors, with RMSEs of 48.93 mm, 92.84 mm, and 183.00 mm, respectively (Table 3). Among all the four datasets considered in this study, AMSRE exhibited the highest spatial mean RMSE, particularly in Alaska, the Rocky Mountains, the Scandinavian Peninsula, Russia,

285   and China, with RMSEs of 71.28 mm, 107.34 mm, 233.58 mm, 36.61 mm and 29.14 mm, respectively. Given the abovementioned RMSE values, it was evident that SWEML most accurately captures the SWE variability.

In terms of bias (Fig. 7b), both SWEML and GLDAS tended to underestimate the SWE, with mean biases of −6.89 mm and −9.95 mm, respectively. In contrast, ESAGB and AMSRE overestimated the SWE, resulting in mean biases of 6.7 mm and 8.17 mm, respectively. Furthermore, the spatial distribution of biases showed that SWEML slightly underestimated the SWE

290   in most regions except for Alaska and the Scandinavian Peninsula, while GLDAS consistently underestimated SWE in most regions. Both ESAGB and AMSRE underestimated the SWE in North America but overestimated it across the Eurasian continent, while AMSRE also underestimated the SWE in Russia. In particular, ESAGB largely underestimated the SWE in the high-elevation regions of central Eurasia, demonstrating a high RMSE (Fig. 7a, Luojus et al., 2021). The SWEML showed a lower bias than the reference datasets in different regions with 2.22 mm, −3.58 mm, 2.57 mm, −10.97 mm, and −6.72 mm,

295   respectively (Table 3). Notably, the improvements in the SWEML were noticeable with the bias less than 53.83 mm, 67.7 mm, and 134.75 mm in Alaska, the Rocky Mountains, and the Scandinavian Peninsula, respectively. In Russia, ESAGB had the lowest bias of 4.74 mm, and high accuracy with RMSE, MAE, and R of 23.47 mm, 0.44 mm, and 0.8, respectively, outperforming the other two reference datasets. Conversely, while SWEML had a higher bias of −10.97 mm, it had a lower RMSE of 13.47 mm, MAE of 0.33 mm, and a higher R of 0.98.

300   Furthermore, the ranges of RMSE and bias of each SWE product were examined in Fig. 8. SWEML had a small interquartile

range (IQR) of RMSE and the lowest maximum RMSE of 30.02 mm, which implies its robust performance (Fig. 8a). In

contrast, the reference datasets presented a relatively wider IQR of the RMSE, indicating their weaker performance. The

maximum RMSEs of GLDAS, ESAGB, and AMSRE were 91.04 mm, 99.13 mm, and 99.94 mm, respectively. Furthermore,

the bias of SWEML varies from −29.45 mm to 28.93 mm, which is lower than that of GLDAS, with a range of bias of −78.80

305   mm to 38.30 mm (Fig. 8b). The bias of ESAGB varies from −88.39 mm to 88.61 mm. The low mean bias of ESAGB was due

to its large positive and negative biases. The bias of SWEML also had a smaller variability compared to that of ESAGB,

implying greater accuracy. AMSRE demonstrated the widest range of bias with exhibiting the highest and lowest biases of

−99.14 mm and 85.63 mm, respectively.

3.3 Elevational and seasonal characteristics of various SWE products

310   We compared the performance of SWEML with that of the three reference datasets (ESAGB, GLDAS, and AMSRE) at

different elevations (Fig. 9). Notably, SWEML demonstrated a better performance than the three reference datasets in terms

of the RMSE, normalized (N)MAE, and R across all elevations. In particular, SWEML exhibited the best performance at

elevations above 1000 m, where high levels of snow are present, with the RMSE, NMAE, and R of 8.38 mm, 0.35%, and 0.98,

respectively. In contrast, the performance of the other reference datasets at elevations above 1000 m was relatively poor.

315   Specifically, the AMSRE demonstrated the lowest accuracy at altitudes above 1000 m, with an RMSE, NMAE, and R of 36.76

mm, 1.07%, and 0.13, respectively. This dataset also demonstrated poor performance at lower elevations below 300 m (e.g.,

100 m to 200 m), with an RMSE, NMAE, R, and bias of 41.31 mm, 1.44%, 0.24, and 15.91 mm, respectively.

In Fig. 10, we compared the monthly means of SWE products with the in-situ measurements. Notably, the monthly variations

in SWEML closely follow those in the measurements, with snow accumulation beginning in September (early Fall) and

320   continuing until April (early summer) (Fig. 10a). Then, snow starts to melt in April as the weather warms up in summer. The

temporal pattern of GLDAS SWE estimates was similar to that of measurements, but it overestimated the SWE. Meanwhile,

AMSRE and ESAGB underestimated the SWE from January to March and December to March, respectively, with a sharp

increase shown in April. As illustrated in Fig. 10b, SWEML consistently demonstrated the lowest RMSEs across all months.

Furthermore, the RMSEs of SWEML were insignificant in different months, indicating its reliability throughout the year. In

325   contrast, by the start of snowmelt in February, other reference datasets exhibited a remarkable increase in the RMSE, which

makes these datasets less accurate during spring and summer.

3.4 Spatial comparison among the SWE datasets

    We analyzed the spatial distribution of the SWEML in comparison to the three reference datasets (Fig. 11). Overall, SWEML

demonstrated high SWE values over mountainous areas such as the Rocky Mountains, Alps, Himalayas, Ural, and Andes, as

330   well as in snow-dominant regions such as Alaska and the Scandinavian Peninsula (Fig. 11a). Notably, the misfit maps

comparing the predictions of SWEML with those of GLDAS and AMSRE (i.e., Figs 11b and 11d) revealed some similarities,

while the corresponding map for ESAGB exhibited slight discrepancies (Fig. 11c). The large positive values in Fig. 11b show

that GLDAS SWE was significantly lower than SWEML in most of the snow-dominant regions, such as the Rocky Mountains,

Alaska, the Scandinavian Peninsula, the Alps, the Himalayas, the Ural, and the Andes. These low SWE estimates for GLDAS

335   have also been reported by Shao et al. (2022). The SWE patterns of AMSRE and GLDAS were similar, demonstrating lower

SWE estimates compared to SWEML in most snow-dominant regions, particularly over northeastern Canada, northern Siberia,

and the Kamchatka Peninsula. Meanwhile, the ESAGB dataset, covering parts of the Northern Hemisphere above a 35˚ latitude

with numerous gaps (Fig. 11c), exhibited lower SWE values in Europe and Alaska but higher SWE values across the Eurasia

continent compared to SWEML. Previous studies reported the high accuracy of the ESAGB over Russia (Pulliainen et al.,

340   2020; Shao et al., 2022), where the SWEML also has a high accuracy as well (Fig. 11d).

4 Code and data availability

    The code used in this study is available on GitHub repository (https://github.com/krsmsuh/SWEML). The SWEML product

can be freely downloaded from Zenodo (https://doi.org/10.5281/zenodo.14195794 , Seo et al., 2024). This dataset provides

data with a daily temporal resolution and 25 km spatial resolution. Furthermore, it covers global latitudes from 90°S to 90°N

345   and longitudes from 180°W to 180°E, excluding Antarctica.

5 Conclusions

This study demonstrates a new daily gap-free global SWE product (i.e., SWEML) using a widely used ML algorithm trained using in situ SWE measurements, meteorological forcings, and terrain attributes. The performance of SWEML was consistently robust across all regions and elevations when compared to that of the reference SWE datasets (i.e., GLDAS,

350 ESAGB, and AMSRE). In particular, SWEML had the highest accuracy over snow-dominating regions as well as high-elevation regions. It also captures seasonal patterns in the SWE more effectively compared to the reference datasets, with significant improvements during the critical snowmelt period from spring to early summer.

The results of our study demonstrate the efficacy of the RF model in capturing the complexities of SWE dynamics, which is still challenging for conventional process-based models. This capability of the RF model allowed us to develop a highly

355 accurate SWE product on a global scale at a daily resolution. SWE data can be estimated readily on a global scale by using observable predictors with the trained RF approach, without explicit knowledge of the physical processes underlying the system. Moreover, this approach provides insights for developing physically based models, particularly for processes that are yet to be well understood. By ML algorithms, by learning SWE dynamics directly from input data, have the potential to uncover crucial but yet unidentified dynamics.

360 SWEML can provide important information for regions with heavy snow cover, therefore critical for the water resource management by snow-melted water. Furthermore, this dataset can contribute toward research on climate change and the cryosphere, as the high albedo values of snow-covered regions are crucial for energy partitioning. Furthermore, SWE anomaly analyses based on SWEML can assess regional variations in snow relative to climate change within the global water cycle. The SWEML dataset can also be used across various other applications, such as for improving the initialization of land surface

365 models and global climate frameworks. Furthermore, SWEML could be also useful when producing other snow-related datasets, such as snowmelt timing map with high spatial resolutions. Given the importance of snow cover in natural systems, the daily gap-free global SWE product developed in this study represents a significant resource for snow-dominant regions.

Earth System
Science
Data

Open Access
Discussions

Author contribution

JH and YK designed the study. JH collected and prepared all data, and MD developed the model code and conducted
370   simulations in collaboration with JH. JH analyzed the results under the supervision of YK and SMB. JH drafted the manuscript
with contributions from MD, JK, SMB, and YK.

Competing interests

The authors declare that they have no conflict of interest.

References

Albergel, C., Dutra, E., Munier, S., Calvet, J. C., Munoz-Sabater, J., Rosnay, P. D., and Balsamo, G.: ERA-5 and ERA-Interim
385   driven ISBA land surface model simulations: which one performs better?, Hydrol. Earth Syst. Sci., 22, 3515–3532,
https://doi.org/10.5194/hess-22-3515-2018, 2018.

Ali, S., Ghosh, N. C., and Singh, R.: Rainfall-runoff simulation using a normalized antecedent precipitation index, Hydrol. Sci.
J., 55, 266–274, https://doi.org/10.1080/02626660903546175, 2010.

Anderton, S. P., White, S. M., and Alvera, B.: Evaluation of spatial variability in snow water equivalent for a high mountain
390   catchment, Hydrol. Process., 18, 435–453, https://doi.org/10.1002/hyp.1319, 2004.

Amante, C. and Eakin, B. W.: ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis, NOAA National Geophysical Data Center [data set], https://doi.org/10.7289/V5C8276M, 2009.

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

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

Biswas, T.K., Giri, K., and Roy, S.: ECKM: An improved K-means clustering based on computational geometry, Expert Syst. Appl., 212, 118862, https://doi.org/10.1016/j.eswa.2022.118862, 2023.

Bokhorst, S., Pedersen, S. H., Brucker, L., Anisimov, O., Bjerke, J. W., Brown, R. D., Ehrich, D., Essery, R. L., Heilig, A., 400 Ingvander, S., Johansson, C., Johansson, M., Jónsdóttir, I. S., Inga, N., Luojus, K., Macelloni, G., Mariash, H., McLennan, D., Rosqvist, G. N., Sato, A., Savela, H., Schneebeli, M., Sokolov, A., Sokratov, S. A., Terzago, S., Vikhamar-Schuler, D., Williamson, S., Qiu, Y., and Callaghan, T. V.: Changing Arctic snow cover: a review of recent developments and assessment of future needs for observations, modelling and impacts, Ambio, 45, 516–537, https://doi.org/10.1007/s13280-016-0770-0, 2016.

405 Breiman, L: Random forests, Mach. Learn., 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.

Brönnimann, S., Allan, R., Atkinson, C., Buizza, R., Bulygina, O., Dahlgren, P., Dee, D., Dunn, R., Gomes, P., John, V. O., Jourdain, S., Haimberger, L., Hersbach, H., Kennedy, J., Poli, P., Pulliainen, J., Rayner, N., Saunders, R., Schulz, J., Sterin, A., Stickler, A., Titchner, H., Valente, M. A., Ventura, C., and Wilkinson, C.: Observations for Reanalyses, B. Am. Meteorol. Soc. 99, 1851–1866, https://doi.org/10.1175/BAMS-D-17-0229.1, 2018.

410 Brown, R. D., Fang, B., and Mudryk, L.: Update of Canadian historical snow survey data and analysis of snow water equivalent trends, 1967–2016, Atmos.-Ocean, 57, 149–156, https://doi.org/10.1080/07055900.2019.1598843, 2019.

Broxton, P. D., Dawson N., and Zeng, X.: Linking snowfall and snow accumulation to generate spatial maps of SWE and snow depth, Earth Space Sci., 3, 246–256, https://doi.org/10.1002.2016EA000174, 2016a.

Broxton, P. D., van Leeuwen, W. J. D., and Biederman, J. A.: Improving snow water equivalent maps with machine learning 415 of snow survey and lidar measurements, Water Resour. Res., 55, 3739–3757, https://doi.org/10.1029/2018wr024146, 2019.

Broxton, P., Zeng, X., and Dawson, N.: Why do global reanalyses and land data assimilation products underestimate snow water equivalent?, J. Hydrometeorol., 17, 2743–2761, https://doi.org/10.1175/JHM-D-16-0056.1, 2016b.

Bulygina, O. N., Groisman, P. Y., Razuvaev, V. N., and Korshunova, N. N.: Changes in snow cover characteristics over Northern Eurasia since 1966, Environ. Res. Lett. 6, 045204, https://doi.org/10.1088/1748-9326/6/4/045204, 2011.

420   Caton Harrison, T., Biri, S., Bracegirdle, T. J., King, J. C., Kent, E. C., Vignon, É., and Turner, J.: Reanalysis representation of low-level winds in the Antarctic near-coastal region, Weather Clim. Dynam., 3, 1415–1437, https://doi.org/10.5194/wcd-3-1415-2022, 2022.

Carvalho, M. J., Melo-Gonçalves, P., Teixeira, J. C., and Rocha, A.: Regionalization of Europe based on a K-means cluster analysis of the climate change of temperatures and precipitation, Phys. Chem. Earth A/B/C, 94, 22–28, 425   https://doi.org/10.1016/j.pce.2016.05.001, 2016.

Cheng, M., Fang, F., Kinouchi, T., Navon, I. M., and Pain, C. C.: Long lead-time daily and monthly streamflow forecasting using machine learning methods, J. Hydrol., 590, 125376, https://doi.org/10.1016/j.jhydrol.2020.125376, 2020.

Cho, E., Vuyovich, C. M., Kumar, S. V., Wrzesien, M. L., Kim, R. S., and Jacobs, J. M.: Precipitation biases and snow physics limitations drive the uncertainties in macroscale modeled snow water equivalent, Hydrol. Earth Syst. Sci., 26, 5721–5735, 430   https://doi.org/10.5194/hess-26-5721-2022, 2022.

Cho, K. and Kim, Y.: Improving streamflow prediction in the WRF-Hydro model with LSTM networks, J. Hydrol., 605, 127297, https://doi.org/10.1016/j.jhydrol.2021.127297, 2022.

Choi, S. and Kim, Y.: Rad-cGAN v1.0: radar-based precipitation nowcasting model with conditional generative adversarial networks for multiple dam domains, Geosci. Model Dev., 15, 5967–5985, https://doi.org/10.5194/gmd-15-5967-2022, 2022.

435   Cui, G., Anderson, M., and Bales, R.: Mapping of snow water equivalent by a deep-learning model assimilating snow observations, J. Hydrol., 616, 128835, https://doi.org/10.1016/j.jhydrol.2022.128835, 2023.

Dingman, S. L.: Physical Hydrology (third ed.), Waveland Press Inc., Long Grove, Illinois, USA, ISBN 1478611189, 2015.

Dierauer, J. R., Allen, D. M., and Whitfield, P. H.: Snow drought risk and susceptibility in the western United States and southwestern Canada, Water Resour. Res., 55, 3076–3091, https://doi.org/10.1029/2018wr023229, 2019.

440   Dodandeh, E., Choubin, B., Eidir, A. N., Nabipour, N., Panahi, M., Shamshirband, S., and Mosavi, A.: Integrated machine
       learning methods with resampling algorithms for flood susceptibility prediction, Sci. Total Environ., 705, 135983,
       https://doi.org/10.1016/j.scitotenv.2019.135983, 2020.

       Fang, K. and Shen, C: Near-real-time forecast of satellite-based soil moisture using long short-term memory with an adaptive
       data integration kernel, J. Hydrometeorol., 21, 399–413, https://doi.org/10.1175/jhm-d-19-0169.1, 2019.

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

       Frei, A., Tedesco, M., Lee, S., Foster, J., Hall, D. K., Kelly, R., and Robinson, D. A.: A review of global satellite-derived snow
       products, Adv. Space Res., 50, 1007–1029, https://doi.org/10.1016/j.asr.2011.12.021, 2012.

       Friedl, M. A., Mclver, D. K., Hodges, J. C.F., Zhang, X. Y., Muchoney, D., Strahler, A. H., Woodcook, C. E., Gospa, S.,
450   Schneider, A., Cooper, A., Baccini, A., Gao, F., and Schaaf, C.: Global land cover mapping from MODIS: algorithms and
       early results, Remote Sens. Environ., 83, 287–302, https://doi.org/10.1016/S0034-4257(02)00078-0, 2002.

       Ghiggi, G., Humphrey, V., Seneviratne, S. I., and Gudmundsson, L.: GRUN: an observation-based global gridded runoff
       dataset from 1902 to 2014, Earth Syst. Sci. Data, 11, 1655–1674, https://doi.org/10.5194/essd-11-1655-2019, 2019.

       Guo, Z., Dirmeyer, P. A., Gao, X., and Zhao, M.: Improving the quality of simulated soil moisture with a multi-model ensemble
455   approach, Q. J. R. Meteorol. Soc., 133, 731–747, https://doi.org/10.1002/qj.48, 2007.

       Harder, P. and Pomeroy, J. W.: Hydrological model uncertainty due to precipitation-phase partitioning methods. Hydrol.
       Process., 28, 4311–4327, https://doi.org/10.1002/hyp.10214, 2014.

       Henderson, G.R., Peings, Y., Furtado, J.C., and Kushner, P. J.: Snow–atmosphere coupling in the Northern Hemisphere. Nature
       Clim. Change, 8, 954–963, https://doi.org/10.1038/s41558-018-0295-6, 2018.

460   Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R.. Schepers,
       D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara,
       G. D., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger,
       L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P.

D., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J-N.: The ERA5 global reanalysis, Q. J. R. Meteorol. Soc., 146,

465   1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J-N.: ERA5 Hourly Dataset From 1940 to Present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS) [data set], https://doi.org/10.24381/cds.adbb2d47, 2023.

Hill, D. F., Burakowski, E. A., Crumley, R. L., Keon, J., Hu, J. M., Arendt, A. A., Jones, K. W., and Wolken, G. J.: Converting

470   snow depth to snow water equivalent using climatological variables, The Cryosphere, 13, 1767–1784, https://doi.org/10.5194/tc-13-1767-2019, 2019.

Hong, Y., Adler, R. F., Huffman, G. J., and Pierce, H.: Applications of TRMM-based multi-satellite precipitation estimation for global runoff prediction: prototyping a global flood modelling system, in: Satellite Rainfall Applications for Surface Hydrology, Springer, Netherlands, 245–266, https://doi.org/10.1007/978-90-481-2915-7, 2010.

475 Hu, H., Liu, J., Zhang, X., and Fang, M.: An Effective and adaptable K-means algorithm for big data cluster analysis. Pattern Recognit., 139, 109404. https://doi.org/10.1016/j.patcog.2023.109404, 2023.

Ikotun, A.M., Ezugwu, A.E., Abualigah, L., Abuhaija, B., and Heming, J.: K-means clustering algorithms: a comprehensive review, variants analysis, and advances in the era of big data, Inf. Sci. (Ny)., 622, 178–210, https://doi.org/10.1016/j.ins.2022.11.139, 2023.

480 Isles, P.D.F.: A random forest approach to improve estimates of tributary nutrient loading, Water Res., 248, 120876, https://doi.org/10.1016/j.watres.2023.120876, 2023.

Imaoka, K., Kachi, M., Fujii, H., Murakami, H., Hori., Ono, A., Igarashi, T., Nakagawa, K., Oki, T., Honda, Y., and Shimoda, H.: Global change observation mission (GCOM) for monitoring carbon, water cycles, and climate change, P IEEE, 98, 717-734, https://doi.org/10.1109/Jproc.2009.2036869, 2010.

485 Iwasawa, H., Ueno, T., Masui, T., and Tajima, S.: Unsupervised clustering for identifying spatial inhomogeneity on local electronic structures, npj Quantum Mater., 7, 24, http://doi.org/10.1038/s41535-021-00407-5, 2022.

Jennings, K. S. and Molotch, N. P.: The sensitivity of modeled snow accumulation and melt to precipitation phase methods across a climatic gradient, Hydrol. Earth Syst. Sci., 23, 3765–3786, https://doi.org/10.5194/hess-23-3765-2019, 2019.

Jennings, K. S., Winchell, T. S., Livneh, B., and Molotch, N. P.: Spatial variation of the rain–snow temperature threshold

490   across the Northern Hemisphere, Nat Commun., 9, 1148, https://doi.org/10.1038/s41467-018-03629-7, 2018.

Jiang, J., Yang, J., Zhang, C., Wu, S., Li, Z., Dai, Z., Li, X., and Qiu, Y.: Daily Snow water equivalent product with SMMR,

SSM/I   and   SSMIS   from   1980   to   2020   over   China,   Big   Earth   Data.,   6,   420-434,

https://doi.org/10.1080/20964471.2022.2032998, 2022.Josso, P., Hall, A., Williams, C., Le Bas, T., Lusty, P., and Murton,

B.: Application of random-forest machine learning algorithm for mineral predictive mapping of Fe-Mn crusts in the World

495   ocean, Ore. Geol. Rev., 162, 105671, https://doi.org/10.1016/j.oregeorev.2023.105671, 2023.

Kelly, R.: The AMSR-E snow depth algorithm: description and initial results, J. Remote Sens. Soc. Jpn., 29, 307–

317, https://doi.org/10.11440/rssj.29.307, 2009.

Kim, R. S., Kumar, S., Vuyovich, C., Houser, P., Lundquist, J., Mudryk, L., Durand, M., Barros, A., Kim, E. J., Forman, B.

A., Gutmann, E. D., Wrzesien, M. L., Garnaud, C., Sandells, M., Marshall, H.-P., Cristea, N., Pflug, J. M., Johnston, J., Cao,

500   Y., Mocko, D., and Wang, S.: Snow ensemble uncertainty Project (SEUP): quantification of snow water equivalent

uncertainty   across   North   America   via   ensemble   land   surface   modeling,   The   Cryosphere,   15,   771–791,

https://doi.org/10.5194/tc-15-771-2021, 2021.

Kinar, N. J. and Pomeroy, J. W.: Measurement of the physical properties of the snowpack, Rev. Geophys., 53, 481–544,

https://doi.org/10.1002/2015RG000481, 2015.

505   Kohler, M. A. and Linsley, R. K.: Predicting the runoff from storm rainfall, Research Paper no. 34, Weather Bureau, US Dept

of Commerce, Washington, USA, 1951.

Kraft, B., Jung, M., Körner, M., Koirala, S., and Reichstein, M.: Towards hybrid modeling of the global hydrological cycle,

Hydrol. Earth Syst. Sci., 26, 1579–1614, https://doi.org/10.5194/hess-26-1579-2022, 2022.

Lapo, K. E., Hinkelman, L. M., Raleigh, M. S., and Lundquist, J. D.: Impact of errors in the downwelling irradiances on

510   simulations of snow water equivalent, snow surface temperature, and the snow energy balance, Water Resour. Res., 51,

1649–1670, https://doi.org/10.1002/2014WR016259, 2015.

Li, X., Wei, Y., and Li, F.: Optimality of antecedent precipitation index and its application, J. Hydrol., 595, 126027,

https://doi.org/10.1016/j.jhydrol.2021.126027, 2021a.

Li, Y., Chu, X., Tian, D., Feng, J., and Mu, W.: Customer segmentation using K-means clustering and the adaptive particle

515   swarm optimization algorithm. Appl. Soft Comput., 113, 107924, https://doi.org/10.1016/j.asoc.2021.107924, 2021b.

Lundquist, J., Hughes, M., Gutmann, E., and Kpnick, S.: Our skill in modeling mountain rain and snow is bypassing the skill

of our observational networks, B. Am. Meteorol. Soc., 100, 2473–2490, https://doi.org/10.1175/bams-d-19-0001.1, 2019.

Luojus, K., Pulliainen, J., Takala, M., Lemmetyinen, J., Mortimer, C., Derksen, C., Mudryk, L., Moisander, M., Hiltunen, M.,

Smolander, T., Ikonen, J., Cohen, J., Salminen, M., Norberg, J., Veijola, K., and Venäläinen, P.: GlobSnow v3.0 Northern

520   Hemisphere snow water equivalent dataset. Sci. Data, 8, 163, https://doi.org/10.1038/s41597-021-00939-2, 2021.

Ma, T., Li, C., Lu, Z., and Wang, B.: An effective antecedent precipitation model derived from the power-law relationship

between   landslide   occurrence   and   rainfall   level,   Geomorphology,   216,   187–192,

https://doi.org/10.1016/j.geomorph.2014.03.033, 2014.

Marty, C.: GCOS SWE Data from 11 Stations in Switzerland. WSL Institute for Snow and Avalanche Research SLF [data set],

525   https://doi.org/10.16904/15, 2020.

Matsuyama, H., Saito, H., and Zemtsov, V.: Application of soil water index to landslide prediction in snowy regions: sensitivity

analysis in Japan and preliminary results from Tomsk, Russia, Prog. Earth Planet Sci., 8, 17, https://doi.org/10.1186/s40645-
021-00408-9, 2021.

Menard, C. B., Essery, R., Krinner, G., Arduini, G., Bartlett, P., Boone, A., Brutel-Vuilmet, C., Burke, E., Cuntz, M., Dai, Y.,

530   Decharme, B., Dutra, E., Fang, X., Fierz, C., Gusev, Y., Hagemann, S., Haverd, V., Kim, H., Lafaysse, M., Marke, T.,

Nasonova, O., Nitta, T., Niwano, M., Pomeroy, J., Schädler, G., Semenov, V. A., Smirnova, T., Strasser, U., Swenson, S.,

Turkov, D., Wever, N., and Yuan, H.: Scientific and human errors in a snow model intercomparison, B. Am. Meteorol. Soc.,

102, https://doi.org/10.1175/bams-d-19-0329.1, 2021.

Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E, and Houston, T. G. An overview of the global historical climatology

535   network-daily database, J. Atmospheric Ocean. Technol., 29, 897–910, https://doi.org/10.1175/JTECH-D-11-00103.1, 2012.

Merkouriadi, I., Lemmetyinen, J., Liston, G. E., and Pulliainen, J.: Solving challenges of assimilating microwave remote

sensing signatures with a physical model to estimate snow water equivalent, Water Resour. Res., 57, e2021WR030119,

https://doi.org/10.1029/2021WR030119, 2021.

Meshyk, A., Barushka, M., and Marozava, V.: Snow as a contributor to spring flooding in Belarus, Environ. Sci. Pollut. Res.,

540   28, 18826–18836, https://doi.org/10.1007/s11356-020-09638-8, 2021.

Mizukami, N., Perica, S., and Hatch, D.: Regional approach for mapping climatological snow water equivalent over the

mountainous regions of the western United States, J. Hydrol., 400, 72–82, https://doi.org/10.1016/j.jhydrol.2011.01.019,

2011.

Mortimer, C., Mudryk, L., Derksen, C., Luojus, K., Brown, R., Kelly, R., and Tedesco, M.: Evaluation of long-term Northern

545   Hemisphere snow water equivalent products, The Cryosphere, 14, 1579–1594, https://doi.org/10.5194/tc-14-1579-2020,

2020.

Mudryk, L. R., Derksen, C., Kushner, P. J., and Brown, R.: Characterization of Northern Hemisphere snow water equivalent

datasets, 1981–2010, J. Clim., 28, 8037–8051, https://doi.org/10.1175/jcli-d-15-0229.1, 2015.

Musselman, K. N., Addor, N., Vano, J. A., and Molotch, N. P.: Winter melt trends portend widespread declines in snow water

550   resources, Nat. Clim. Chang., 11, 418–424, https://doi.org/10.1038/s41558-021-01014-9, 2021.

Niedzielski, T., Szymanowski, M., Miziński, B., Spallek, W., Witek-Kasprzak, M., Ślopek, J., Kasprzak, M., Błaś, M., Sobik,

M., Jancewicz, K., Borowicz, D., Remisz, J., Modzel, P., Męcina, K., and Leszczyński, L.: Estimating snow water equivalent

using unmanned aerial vehicles for determining snow melt runoff, J. Hydrol., 578, 124046,

https://doi.org/10.1016/j.jhydrol.2019.124046, 2019.

555 NVE.: The Public Hydrological Data from NVE Sildre to Research [dataset], https://sildre.nve.no/list?paramsStationList=2003,

2023.

Osterhuber, R.: Snow Survey Procedure Manual Prepared for the California Department of Water Resources, California

Cooperative   Snow   Surveys   [data   set],   http://cdec.water.ca.gov/cgi-

progs/products/SnowSurveyProcedureManualv20141027.pdf, 2014.

560 Paul, A., Afroosa. M., Baduru, B., and Paul, B.: Showcasing model performance across space and time using single diagram,

Ocean Model., 181, 102150, http://doi.org/10.1016/j.ocemod.2022.102150, 2023.

Pulliainen, J.: Mapping of snow water equivalent and snow depth in boreal and sub-arctic zones by assimilating space-borne microwave radiometer data and ground-based observations, Remote Sens. Environ., 101, 257–269, https://doi.org/10.1016/j.rse.2006.01.002, 2006.

565   Pulliainen, J., Luojus, K., Derksen, C., Mudryk, L., Lemmetyinen, J., Salminen, M., Ikonen, J., Takala, M., Cohen, J., Smolander, T., and Norber, J.: Patterns and trends of Northern Hemisphere snow mass from 1980 to 2018, Nature, 581, 294–298, https://doi.org/10.1038/s41586-020-2258-0, 2020.

Raleigh, M. S., Lundquist, J. D., and Clark, M. P.: Exploring the impact of forcing error characteristics on physically based snow simulations within a global sensitivity analysis framework, Hydrol. Earth Syst. Sci., 19, 3153–3179,

570   https://doi.org/10.5194/hess-19-3153-2015, 2015.

Ravuri, S., Lenc, K., Wilson, M., Kangin, D., Lam, R., Mirowski, P., Fitzsimons, M., Athanassiadou, M., Kashem, S., Madge, S., Prudden, R., Mandhane, A., Clark, A., Brock, A., Simonyan, K., Hadsell, R., Robinson, N., Clancy, E., Arribas, A., and Mohamed, S.: Skilful precipitation nowcasting using deep generative models of radar, Nature, 597, 672–677, https://doi.org/10.1038/s41586-021-03854-z, 2021.

575   Rousseeuw, P. J.: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, J. Comput. Appl. Math., 20, 53–65, https://doi.org/10.1016/0377-0427(87)90125-7, 1987.

Rodell, M., Houser, P., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., and Bosilovich, M.: The global land data assimilation system, B. Am. Meteorol. Soc., 85, 381–394, https://doi.org/10.1175/BAMS-85-3-381, 2004.

580   Schirmer, M., Wirz, V., Clifton, A., and Lehning, M.: Persistence in intra-annual snow depth distribution: 1. Measurements and topographic control, Water Resour. Res., 47, W09516, https://doi.org/10.1029/2010WR009426, 2011.

Sculley, D.: Web-scale k-means clustering, in: Proceedings of the 19th International Conference on World Wide Web, Association for Computing Machinery, New York, NY, USA, 26–30 April 2010, 1177–1178, https://doi.org/10.1145/1772690.1772862, 2010.

585   Sungmin, O. and Rene O.: Global soil moisture data derived through machine learning trained with in situ measurements, Sci Data, 8, 170, https://doi.org/10.1038/s41597-021-00964-1, 2021.

Serreze, M. C., Clark, M. P., Armstrong, R. L., McGinnis, D. A., and Pulwarty, R. S.: Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data, Water Resour. Res., 35, 2145–2160, https://doi.org/10.1029/1999WR900090, 1999.

590   Seo, J. and Kim, Y.: Assessing the likelihood of drought impact occurrence with extreme gradient boosting: a case study on the public water supply in South Korea, J. Hydroinf., 25, 191–207, https://doi.org/10.2166/hydro.2023.064, 2023.

Seo, J., Panahi, M., Kim, J.H, Bateni, S. M., and Kim, Y.: Global Snow Water Equivalent Product Derived from Machine Learning Model Trained with in Situ Measurement Data, Zenodo [Dataset], https://doi.org/10.5281/zenodo.14195794 , 2024.

Shao, D., Li, H., Wang, J., Hao, X., Che, T., and Ji, W.: Reconstruction of a daily gridded snow water equivalent product for

595   the land region above 45° N based on a ridge regression machine learning approach, Earth Syst. Sci. Data, 14, 795–809, https://doi.org/10.5194/essd-14-795-2022, 2022.

Singh, V. P.: Hydrologic System Watershed Modeling, Vol. 2, Prentice-Hall, Englewood Cliffs, New Jersey, USA, ISBN 0134480287,1989.

Snauffer, A. M., Hsieh, W. W., and Cannon, A. J.: Comparison of gridded snow water equivalent products with in situ

600   measurements in British Columbia, Canada, J. Hydrol., 541, 714–726, https://doi.org/10.1016/j.jhydrol.2016.07.027, 2016.

Takala, M., Luojus, K., Pulliainen, J., Derksen, C., Lemmetyinen, J., Kärnä, J. P., Koskinen, J., and Bojkov, B.: Estimating Northern Hemisphere snow water equivalent for climate research through assimilation of space-borne radiometer data and ground-based measurements, Remote Sens. Environ., 115, 3517–3529, https://doi.org/10.1016/j.rse.2011.08.014, 2011.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res., 106, 7183–7192,

605   https://doi.org/10.1029/2000JD900719, 2001.

Tedesco, M. and Jeyaratnam, J.: AMSR-E/AMSR2 Unified L3 Global Daily 25 km EASE-Grid Snow Water Equivalent, Version 1, Boulder, Colorado USA, NASA National Snow and Ice Data Center Distributed Active Archive Center [data set], https://doi.org/10.5067/8AE2ILXB5SM6, 2019.

Tennant, C. J., Harpold, A. A., Lohse, K. A., Godsey, S. E., Crosby, B. T., Larsen, L. G., Brooks, P. D., Kirk, R. W. V., and

610   Glenn, N. F.: Regional sensitivies of seasonal snowpack to elevation, aspect, and vegetation cover in western North America, Water Resour. Res., 53, 6908–6926, https://doi.org/10.1002/2016WR019374, 2017.

Thorndike, R. L.: Who belongs in the family?, Psychometrika, 18, 267–276, https://doi.org/10.1007/BF02289263, 1953.

Tibshirani, R., Walther, G., and Hastie, T.: Estimating the number of clusters in a data set via the gap statistic, J. R. Statist. Soc. B., 63, 2, 411–423, https://doi.org/10.1111/1467-9868.00293, 2001.

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

Tsang, L., Durand, M., Derksen, C., Barros, A. P., Kang, D.-H., Lievens, H., Marshall, H.-P., Zhu, J., Johnson, J., King, J., Lemmetyinen, J., Sandells, M., Rutter, N., Siqueira, P., Nolin, A., Osmanoglu, B., Vuyovich, C., Kim, E., Taylor, D., Merkouriadi, I., Brucker, L., Navari, M., Dumont, M., Kelly, R., Kim, R. S., Liao, T.-H., Borah, F., and Xu, X.: Review 620 article: global monitoring of snow water equivalent using high-frequency radar remote sensing, The Cryosphere, 16, 3531–3573, https://doi.org/10.5194/tc-16-3531-2022, 2022.

USDA NRCS.: National Water and Climate Center Interactive Map—Cooperator Snow Sensors [data set], https://data.nal.usda.gov/dataset/national-water-and-climate-center-interactive-map, 2022a.

USDA NRCS.: SNOwpack TELemetry Network (SNOTEL) [data set], https://data.nal.usda.gov/dataset/snowpack-telemetry-625 network-snotel, 2022b.

USDA NRCS.: Soil Climate Analysis Network (SCAN) [data set], https://data.nal.usda.gov/dataset/soil-climate-analysis-network-scan, 2022c.

Varhola, A., Coops, N. C., Weiler, M., and Moore, R. D.: Forest canopy effects on snow accumulation and ablation: an integrative review of empirical results, J. Hydrol., 392, 219–233, https://doi.org/10.1016/j.jhydrol.2010.08.009, 2010.

630 Viviroli, D., Archer, D. R., Buytaert, W., Fowler, H. J., Greenwood, G. B., Hamlet, A. F., Huang, Y., Koboltschnig, G., Litaor, M. I., López-Moreno, J. I., Lorentz, S., Schädler, B., Schreier, H., Schwaiger, K., Vuille, M., and Woods, R.: Climate change and mountain water resources: overview and recommendations for research, management and policy, Hydrol. Earth Syst. Sci., 15, 471–504, https://doi.org/10.5194/hess-15-471-2011, 2011.

Wang, Y. H., Broxton, P., Fang, Y., Behrangi, A., Barlage, M., Zeng, X., and Niu, G. Y.: A wet-bulb temperature-based rain-635 snow partitioning scheme improves snowpack prediction over the drier western United States, Geophys. Res. Lett., 46, 13825–13835, https://doi.org/10.1029/2019GL085722, 2019.

Walker, A., Brasnett, B., and Brown, R.: Canadian Meteorological Centre (CMC) Daily Gridded Snow Depth Analysis for Northern Hemisphere, 1998–2008 [data set], https://doi.org/10.5443/10916, 2011.

Wrzesien, M. L., Pavelsky, T. M., Durand, M. T., Dozier, J., and Lundquist, J. D.: Characterizing biases in mountain snow accumulation from global data sets, Water Resour. Res., 55, 9873–9891, https://doi.org/10.5194/hess-15-471-2011, 2019.

Xie, J., Jonas, T., Rixen, C., Jong, R. D., Garonna, I., Notarnicola, C., Asam, S., Schaepman, M. E., and Kneubühler, M.: Land surface phenology and greenness in Alpine grasslands driven by seasonal snow and meteorological factors, Sci. Total Environ., 725, 138380, https://doi.org/10.1016/j.scitotenv.2020.138380, 2020.

Ye, H., Cohen, J., and Rawlins, M.: Discrimination of solid from liquid precipitation over Northern Eurasia using surface atmospheric conditions. J. Hydrometeorol. 14, 1345–1355, https://doi.org/10.1175/JHM-D-12-0164.1, 2013.

Yueh, S. H., Shah, R., Xu, X., Stiles, B., and Bosch-Lluis, X.: A satellite synthetic aperture radar concept using P-band signals of opportunity, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 14, 2796–2816, https://doi.org/10.1109/jstars.2021.3059242, 2021.

Zêzere, J. L., Trigo, R. M., and Trigo, I. F.: Shallow and deep landslides induced by rainfall in the Lisbon region (Portugal): assessment of relationships with the North Atlantic oscillation, Nat. Hazard Earth Syst. Sco., 5, 3, 331–344, https://doi.org/10.5194/nhess-5-331-2005, 2005.

Zhang, Y. and Ma, N.: Spatiotemporal variability of snow cover and snow water equivalent in the last three decades over Eurasia, J. Hydrol. 559, 238–251, https://doi.org/10.1016/j.jhydrol.2018.02.031, 2018.
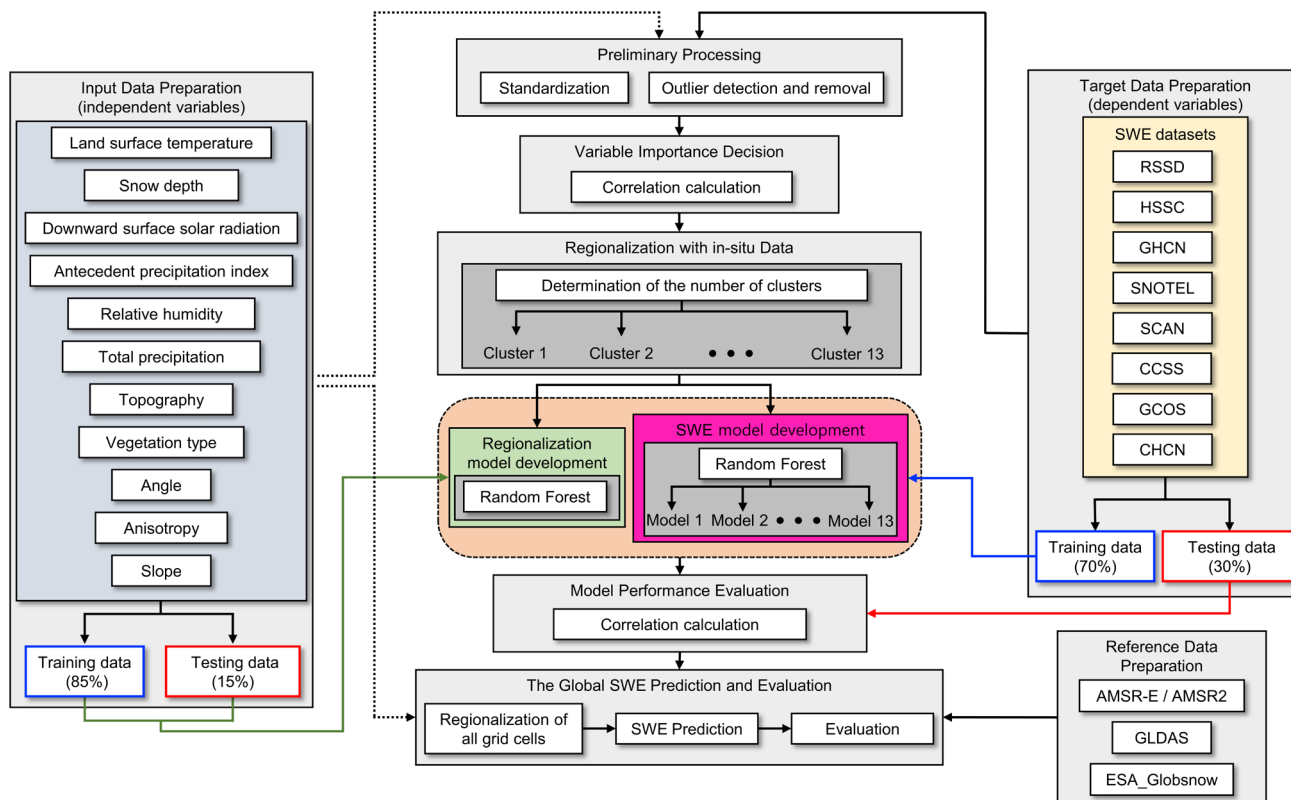
655

Figure 1: Schematic overview of the process of the RF-based global-scale gridded SWE (SWEML) product. First, point-scale in situ SWE measurements are allocated to a 0.25°-degree using the long-term mean and standard deviation of SWE data from ERA5-land. Subsequently, grid cells containing in situ SWE measurements are regionalized into 13 clusters (Random Forest classification). A Random Forest regression model is then implemented on each cluster to predict SWE data. The resulting global SWE predictions

660 **are validated against in situ SWE measurements and compared with reference datasets, including GLDAS, AMSRE, and ESAGB.**
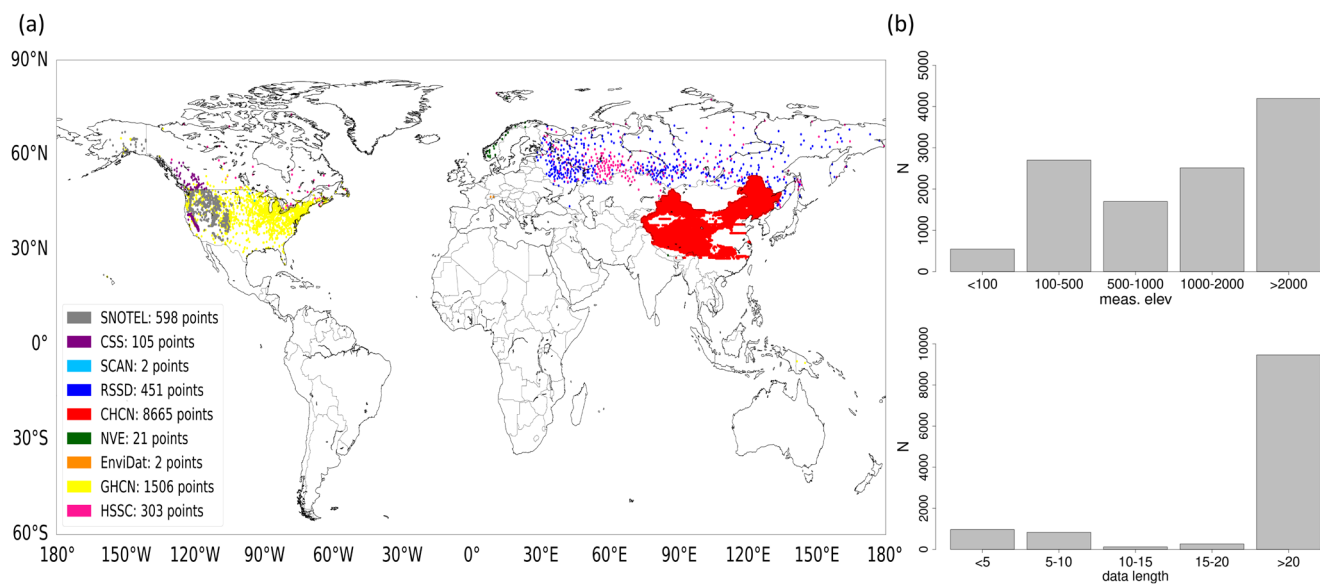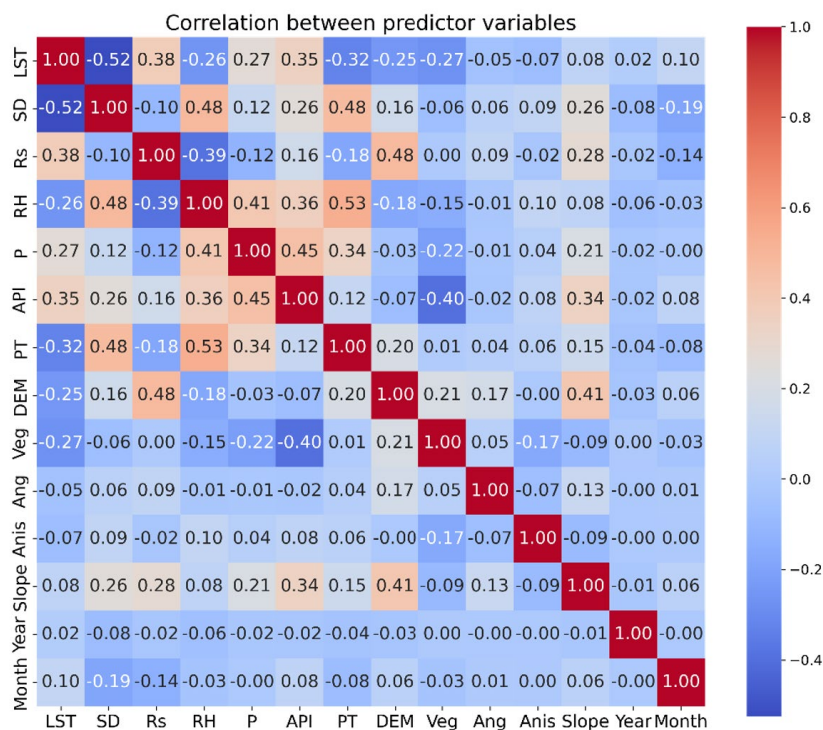
Figure 2: (a) Spatial distribution of the in situ SWE measurements utilized in this study (11,653 grid points). (b) Distribution of grid

665 **cells based on elevation and data period over the study period from 1980 through 2020.**

Figure 3: Correlation matrix demonstrating the relationships between several input predictors: LST: land surface temperature, SD: snow depth, Rs: downward solar radiation, RH: relative humidity, P: precipitation, API: antecedent precipitation index, PT: precipitation type, DEM: digital elevation map, Veg: vegetation type, Anis: anisotropy of orography, and Slope: slope of orography.
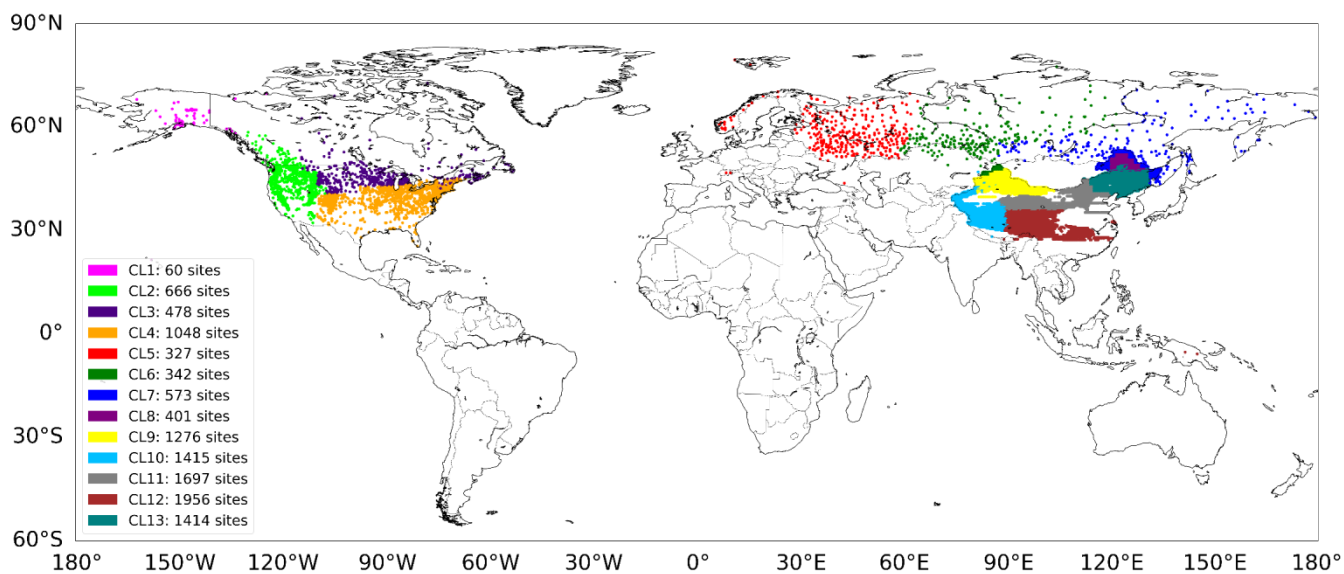
Figure 4: Regionalization map generated by the MiniBatch K-means (MBK) clustering, presenting the number of in situ measurement sites in each cluster (CL).
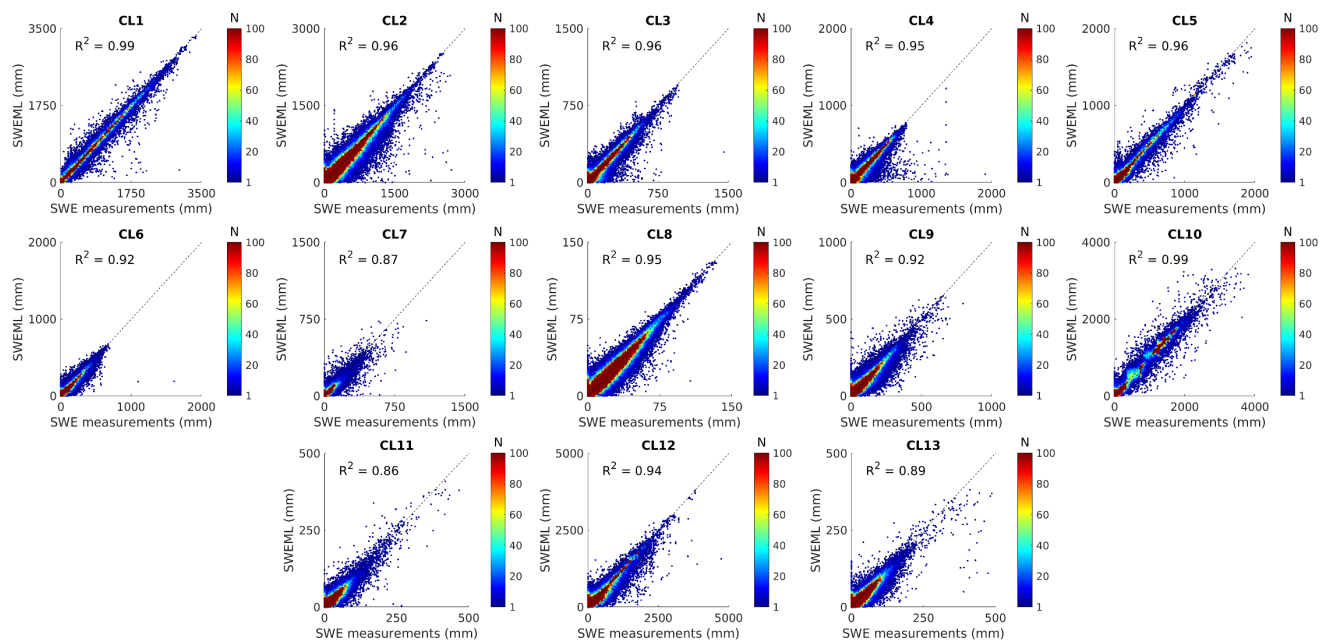
680

Figure 5: Scatterplots comparing the estimations of SWEML with corresponding in situ measurements over all grid cells within each cluster.
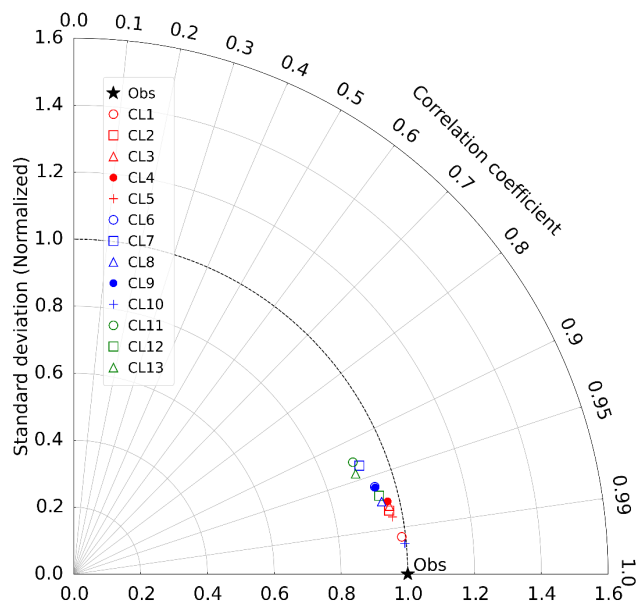
685

Figure 6: Taylor diagram for comparison between the SWEML and in situ measurements over all grid cells in each cluster.
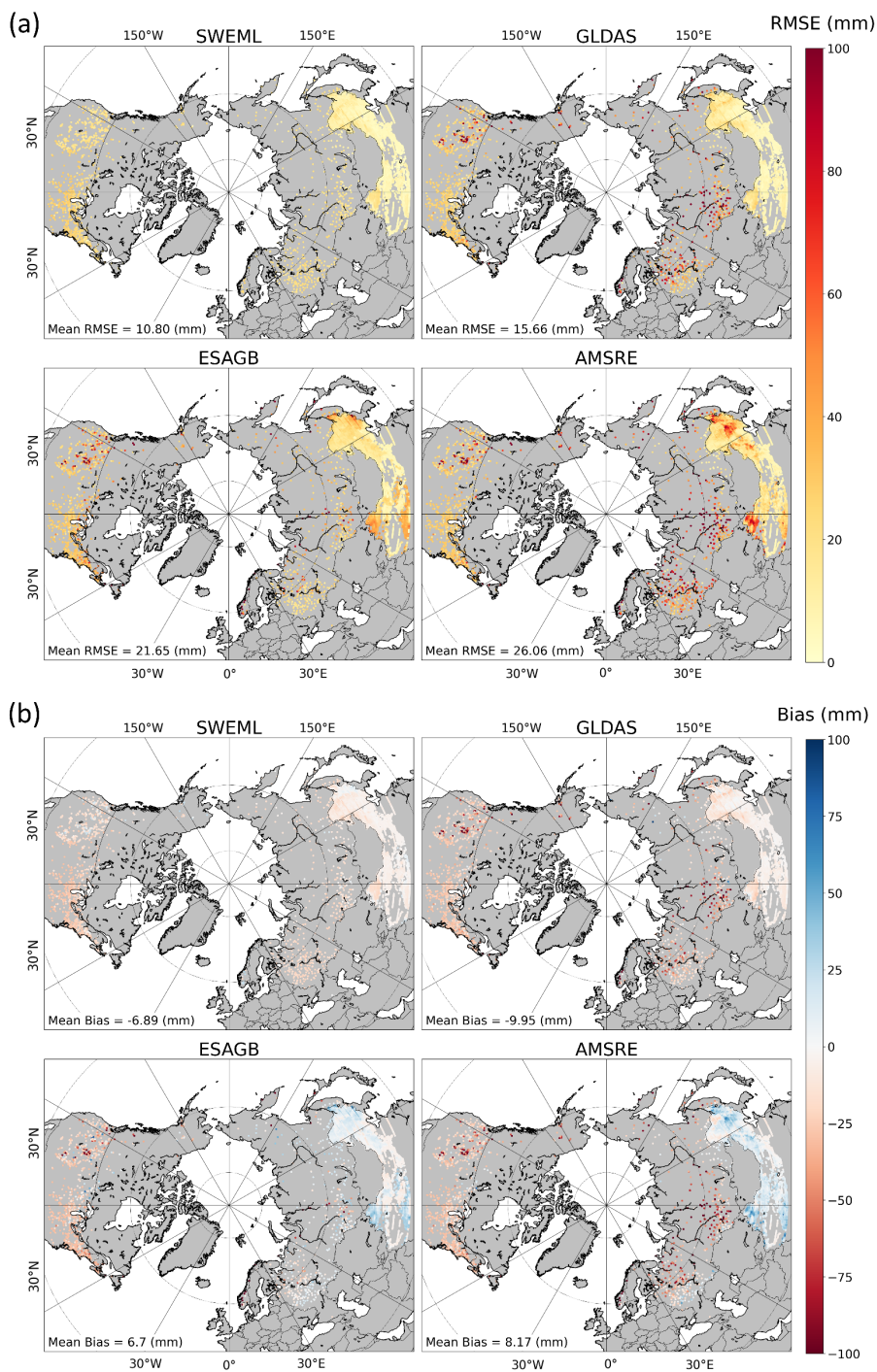
Figure 7: Comparison of SWEML with reference SWE products (GLDAS, ESAGB, and AMSRE): (a) RMSE (mm) and (b) bias
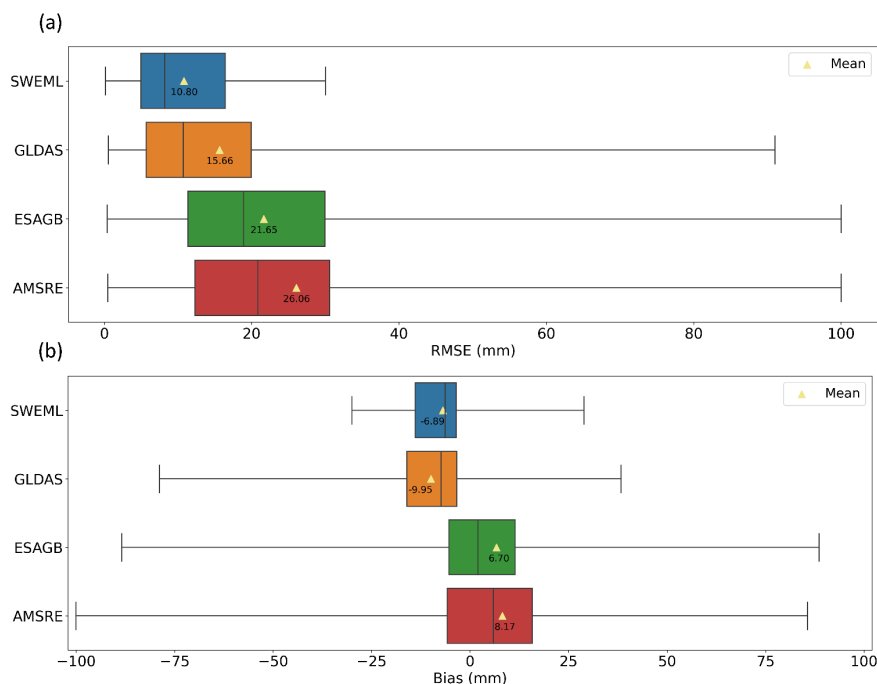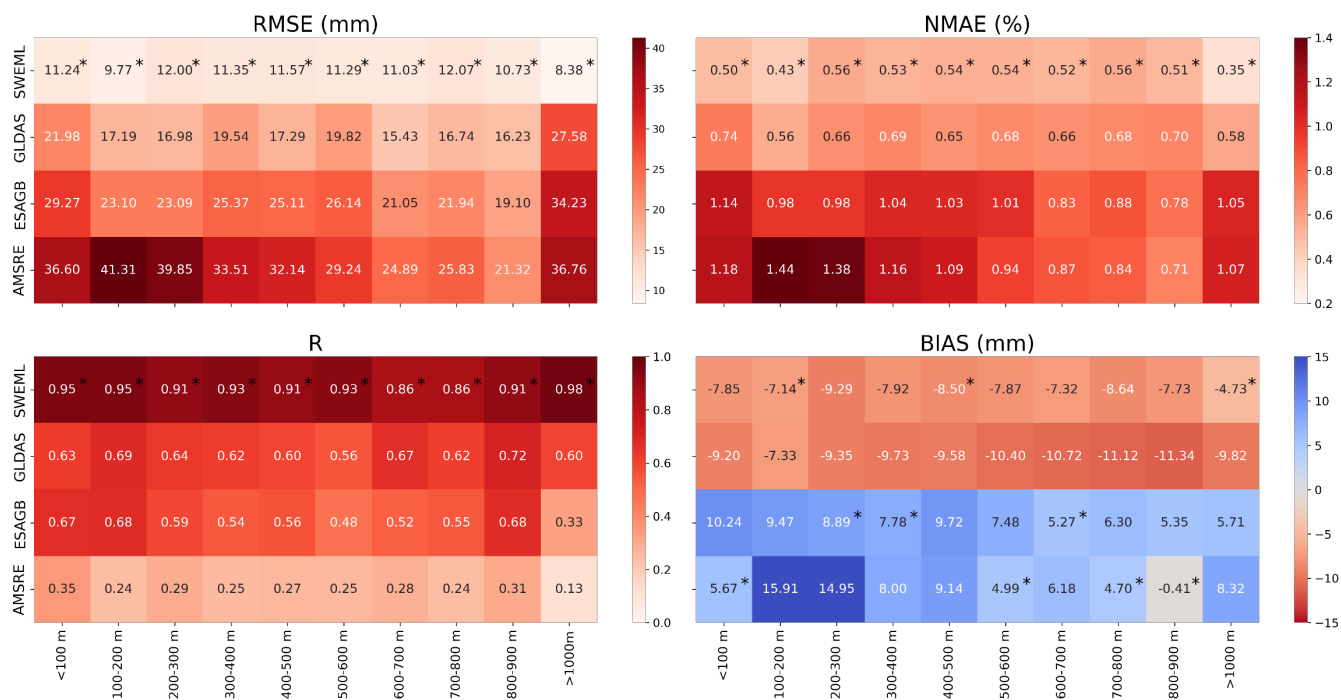
690   **(mm).**

Figure 8: Spear boxplot of the (a) RMSE and (b) bias of each SWE product at all in situ measurement sites. The left and right sides of each box represent the 25th and 75th percentiles, respectively. The solid vertical line within each box indicates the median value, while error bars denote the minimum and maximum values.

Figure 9: Comparison of SWEML with the reference products (GLDAS, ESAGB, and AMSRE) across different elevations. Numbers with an asterisk (*) correspond to the best-performing dataset at the given elevation.
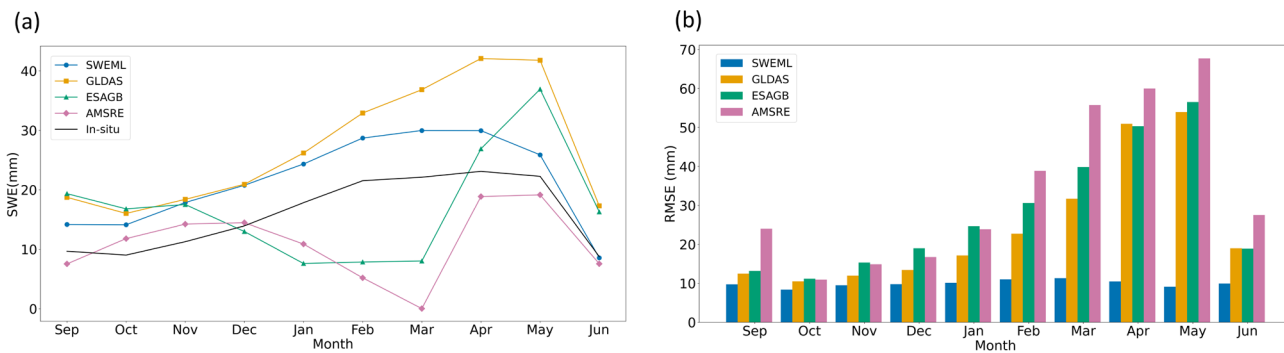
700   **Figure 10: (a) Spatial mean values of the SWE datasets and in situ measurements for each month. (b) RMSE of the SWE datasets**

compared to the corresponding in situ measurements for each month.

Earth System
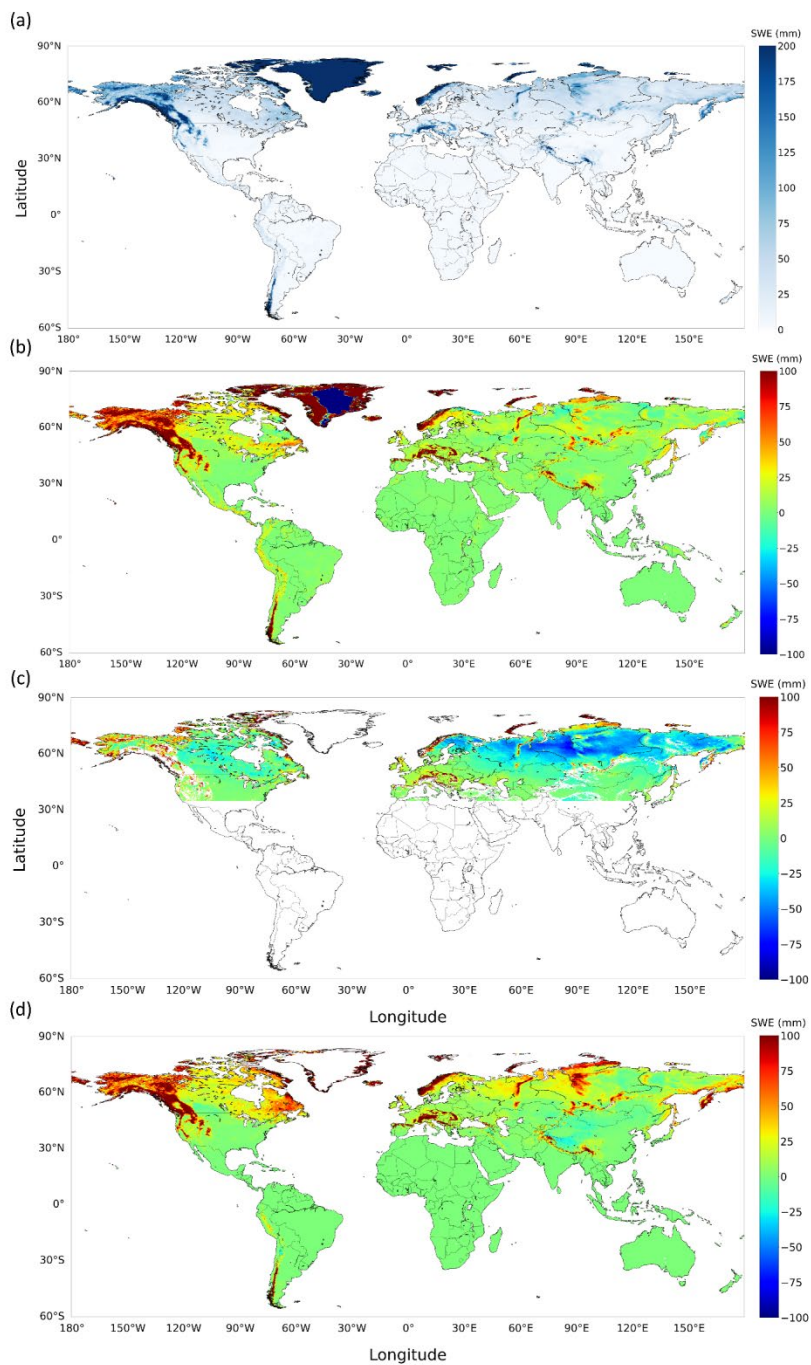Science
Data

Open Access
Discussions

Figure 11: (a) Annual mean SWE values derived from SWEML. Differences between SWEML and (b) GLDAS, (c) ESAGB, and (d) AMSRE SWE datasets.

705   **Table 1. List of SWE data sources used in this study.**

Data type	Data	Data periods	Temporal resolution	Spatial resolution	Spatial coverage
Remote sensing	AMSR-E (https://nsidc.org/data/ae_dysno/versions/2) / AMSR2 (https://nsidc.org/data/ae_dysno/versions/1) (Aqua/GCOM-W1)	2012-07 through 2020-12	Daily	25 km	Global (except Greenland)
In situ	HSSC (Finnish Meteorological Institute; FMI)	1980-01 through 2020-12	5–15 daily	3064 data points	Canada, Finland, and former Soviet Union/Russia (FSU)
	RSSD (Russia Research Institute of Hydrometeorological Information; RIHMI-WDC)		5–15 daily	517 sites	Russia
	SNOTEL (USDA NRCS)		Daily	846 sites	Western U.S.
	SCAN (USDA NRCS)			2 sites	
	CCSS (USDA NRCS)			144 sites	
	GCOS (EnviDat)		15 daily	11 sites	Switzerland
	NVE (Norwegian Water Resources and Energy Directorate)		Daily	29 sites	Norway
	GHCN (National Oceanic and Atmospheric Administration; NOAA)		Daily	14,690 data points	Eastern U.S.
Data assimilation dataset	GLDAS (https://disc.gsfc.nasa.gov)	1980-01 through 2020-12	Daily	25 km	Global
Reanalysis	ESA GlobSnow Reanalysis (https://www.globsnow.info/swe)	1980-01 through 2020-12	Daily	25 km	North Hemisphere above 35º latitude (except Greenland)
	Daily SWE Dataset (Jian, L et al., 2022) [included China Meteorology Administration (CMA)]	1980-01 through 2020-12	1–5 daily	25 km	China
	ERA5-Land SWE	1980-01 through 2020-12	Hourly	10 km	Global

Table 2. Input data used in the RF model.

	Variable	Source	Description
Dynamic	Precipitation (P)	ERA5	Daily meteorological forcing data obtained from the ECMWF reanalysis
	Precipitation type (PT)		
	Relative humidity (RH)		
	Snow depth (SD)		
	Land surface temperature (LST)		
	Downward solar radiation (Rs)		
	Antecedent precipitation index (API)		Derived from precipitation measurements conducted on the given day and previous days
	Year		Year of in situ SWE measurements
	Month		Month of in situ SWE measurements
Invariant	Vegetation type (Veg)	GLDAS (https://ldas.gsfc.nasa.gov/gldas/vegetation-class-mask)	Predominant vegetation type (MODIS-derived) within each grid cell
	Digital elevation map (DEM)	ETOPO1	Mean and standard deviation of sub-grid scale elevation values within each grid cell
	Angle of orography (Ang)	ERA5	Mean and standard deviation of sub-grid scale angle values within each grid cell
	Anisotropy of orography (Anis)		Mean and standard deviation of sub-grid scale anisotropy values within each grid cell
	Slope of orography (Slope)		Mean and standard deviation of sub-grid scale slope values within each grid cell

710   **Table 3. Statistical metrics of each SWE product based on region.**

Region	Statistical metric	SWE Product			
		SWEML	GLDAS	ESAGB	AMSRE
Alaska	RMSE (mm)	16.33	68.82	48.93	71.28
	MAE (mm)	0.16	0.61	0.43	0.67
	R	0.95	0.26	0.57	0.47
	Bias (mm)	2.22	−43.28	−10.28	−56.05
Rocky Mountains	RMSE (mm)	16.88	110.81	92.84	107.34
	MAE (mm)	0.17	0.85	0.71	0.85
	R	0.99	0.44	0.58	0.54
	Bias (mm)	−3.58	−71.28	−54.83	−70.82
Scandinavia	RMSE (mm)	17.53	98.80	183.00	233.58
	MAE (mm)	0.09	0.38	0.69	0.88
	R	0.99	0.94	0.86	0.62
	Bias (mm)	2.57	−57.78	−109.49	−137.32
Russia	RMSE (mm)	13.47	31.50	23.47	36.61
	MAE (mm)	0.33	0.51	0.44	0.52
	R	0.98	0.67	0.80	0.34
	Bias (mm)	−10.97	−16.36	4.74	−11.97
China	RMSE (mm)	9.93	13.08	23.71	29.14
	MAE (mm)	0.52	0.63	1.11	1.22
	R	0.84	0.64	0.51	0.37
	Bias	−6.72	−8.10	8.46	11.95

-