Science of Remote Sensing 10 (2024) 100163

Contents lists available at ScienceDirect

Science of Remote Sensing

journal homepage: www.sciencedirect.com/journal/science-of-remote-sensing

Improved estimation of daily blue-sky snow shortwave albedo from MODIS data and reanalysis information

Anxin Ding [a], Shunlin Liang [b,*], Han Ma [b], Tao He [c], Aolin Jia [d], Qian Wang [e]

[a] School of Resources and Environmental Engineering, Hefei University of Technology, Hefei 230009, China
[b] Jockey Club STEM Laboratory of Quantitative Remote Sensing, Department of Geography, University of Hong Kong, Hong Kong, 999077, China
[c] Hubei Key Laboratory of Quantitative Remote Sensing of Land and Atmosphere, School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, 430079, China
[d] Department of Environment Research and Innovation (ERIN), Luxembourg Institute of Science and Technology (LIST), Belvaux, L-4362, Luxembourg
[e] State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing, 100875, China

ARTICLE INFO

Keywords:
MODIS
GLASS
VIIRS
ERA5-Land
Snow albedo
Prior knowledge
Direct estimation algorithm
XGBoost algorithm

ABSTRACT

Snow albedo is a key geophysical parameter that controls the energy exchanges between the atmosphere and Earth's surfaces and has been widely utilized in climatic and environmental change studies. However, recent studies have demonstrated that current albedo satellite products still have large uncertainties in snow-covered areas. In this study, we estimated the blue-sky shortwave albedo of snow surfaces using the eXtreme Gradient Boosting (XGBoost) algorithm with Moderate Resolution Imaging Spectroradiometer (MODIS) top-of-atmosphere (TOA) reflectance values, ERA-5 land reanalysis snow parameters (e.g., snow cover, snow density and snow depth water equivalent) and in situ measurements. In the XGBoost model, the MODIS MCD43 albedo values were input as prior knowledge, and the random sample validation results showed that the R^2 and root mean square error (RMSE) values of this model were approximately 0.953 and 0.044, respectively. The typical sites for independent validation were subjected to in situ measurements at the UPE_L, AWS5, and CA_ARB sites. Finally, the retrieved XGBoost albedo values were compared with the official NASA MODIS (MCD43, collection 6), the Global Land Surface Satellite (GLASS), and the National Oceanic and Atmospheric Administration (NOAA) Visible Infrared Imaging Radiometer Suite (VIIRS) SURFALB albedo products. The validation results indicated that the proposed approach achieved much greater accuracy (RMSE = 0.052, bias = 0.002) than did the corresponding official MODIS (RMSE = 0.087, bias = −0.033), GLASS (RMSE = 0.089, bias = −0.031) and VIIRS SURFALB albedo (RMSE = 0.100, bias = −0.032) products. The improved shortwave albedo captured the rapid temporal changes in surface snow conditions.

1. Introduction

Snow albedo is indispensable to the intricate dynamics of the global energy budget, climate, and environmental shifts due to its pivotal role in regulating both regional and global energy distributions (Qu et al., 2013, 2016; Wang et al., 2012, 2014). With its extensive spatial coverage and rapid spatial and temporal evolution, snow exhibits pronounced interannual and seasonal variations. Notably, snow typically exhibits a significantly greater albedo than other terrestrial surfaces, such as soil and vegetation (Ding et al., 2022a, 2022b), with freshly fallen snow often reflecting more than 80% of incoming solar radiation (Ding et al., 2019a, 2019b). Consequently, temporal and spatial fluctuations in snow albedo are intricately linked with global climate variations and the dynamics of regional meteorological systems (Burakowski et al., 2015; Serreze and Barry, 2011; Wang et al., 2012). Snow albedo, constituting a fundamental component of surface energy equilibrium, exerts a potent positive feedback on surface temperatures and the phenomenon of global warming, which is particularly evident at higher latitudes (Stroeve et al., 2005;2013; Liang, 2005; Jafariserajehlou et al., 2020). Therefore, achieving high precision in the retrieval of snow albedo holds significant utility in meeting diverse user requirements, encompassing the global energy equilibrium, climate variability, and hydrological processes (Qu et al., 2015; Zhang et al., 2022).

Currently, a multitude of surface albedo products with diverse temporal and spatial resolutions have been disseminated globally and extensively utilized within the realm of remote sensing (Zhang et al.,

* Corresponding author.
E-mail address: shunlin@hku.hk (S. Liang).

https://doi.org/10.1016/j.srs.2024.100163
Received 8 May 2024; Received in revised form 9 July 2024; Accepted 8 September 2024
Available online 16 September 2024

2022). Researchers have increasingly focused on enhancing and validating land surface albedo (LSA) products, especially for snow albedo (Jia et al., 2022, 2023; Lin et al., 2018; Lu et al., 2021; Wu et al., 2016). The validation accuracy of surface albedo products in both snow-free and snow-covered regions is summarized in Table 1. These products exhibit consistent systematic underestimations in snow-covered areas, necessitating further enhancement in snow albedo accuracy compared to that for snow-free regions. Notable products such as the MODerate Resolution Imaging Spectroradiometer (MODIS) (Corbea-Pérez et al., 2021; Liu et al., 2017; Wang et al., 2018), Global LAnd Surface Satellite (GLASS) (Liu et al., 2013), and Visible Infrared Imaging Radiometer Suite (VIIRS) albedo products (Liu et al., 2017; Wang et al., 2017) are available. Several factors contribute to the relatively poorer performance in snow regions, primarily stemming from the challenges associated with the bidirectional reflectance distribution function (BRDF) of snow and albedo estimation. First, snow predominantly occurs in mid-to high-latitude regions, where higher incidence angles pose significant challenges for the construction of an analytical snow BRDF model and albedo estimation, despite increased acquisition frequency at these latitudes (Schaaf et al., 2011b). For instance, the operational MODIS BRDF/albedo product is not recommended for use under large solar geometries (i.e., solar zenith angles, SZAs, greater than or equal to 70° and view zenith angles, VZAs, greater than or equal to 70°), with highly angular scenarios typically flagged with low quality indicators (Ding et al., 2019b). To solve this problem, Jiao et al. (2019) improved the MODIS algorithm based on the ART model, and Ding et al. (2023) fully verified the feasibility of the improved algorithm in retrieving snow BRDF/Albedo/NBAR. Second, rapid weather fluctuations at high latitudes hinder the comprehensive capture of rainfall and snowfall processes, necessitating relatively high temporal satellite resolutions. In the current albedo estimation algorithm, it is usually necessary to accumulate multi-angle reflectance observations, and fit the snow BRDF model to estimate the albedo. For example, MODIS accumulates observations under clear sky conditions within 16 days, and POLDER accumulates multi-angle reflectance observations within one month. This makes it difficult for many albedo products to capture snowfall and snowmelt processes. Third, the issue of mixed pixels is exacerbated in snow-covered areas, posing substantial challenges in discriminating between clouds and snow (Hall et al., 1995). In addition, the physical mechanism of snow mixing with other surface types is not clear, and how other surface types affect the snow reflection characteristics in the case of mixed pixels, such as thin snow and snow-covered forest. Consequently, these factors frequently result in lower-quality albedo products in snow-covered regions than in snow-free regions, despite high-quality flagged albedo demonstrating good agreement with field data (Stroeve et al., 2005, 2013; Wang et al., 2014; Wright et al., 2014).

Compared with the above methods, the direct estimation algorithm directly correlates top-of-atmosphere (TOA) reflectance with the LSA through comprehensive simulations of radiative transfer processes to estimate snow albedo (Liang, 2001; 2003). For example, Liang (2005) used this algorithm to retrieve the daily snow albedo from MODIS data, indicating the feasibility of this method to estimate snow albedo. Subsequent developments included the production of GLASS products from Advanced Very High Resolution Radiometer (AVHRR) and MODIS data (Liang et al., 2013, 2021). Further enhancements and applications have been extended to estimating albedo from VIIRS and Landsat data (He et al., 2018; Qu et al., 2014; Wang et al., 2017). Recently, Lin et al. (2022) utilized a direct estimation algorithm to estimate 10 m surface albedo from Sentinel-2 satellite observations, revealing the inadequacy of MODIS BRDF model parameters in snow-covered areas. Typically, the direct estimation algorithm leverages high-quality MODIS BRDF model parameters as prior knowledge to establish the TOA reflectance and surface broadband albedo relationship (He et al., 2018). Consequently, researchers often restrict the use of the direct estimation algorithm to estimate albedo solely in snow-free areas due to limitations inherent in MODIS BRDF model parameters (Ma et al., 2022; Zhang et al., 2020). Moreover, the current direct estimation algorithm establishes the relationship between the simulation database and TOA reflectance without accounting for errors in the simulated albedo. The uncertainty associated with simulation data may thus propagate into the inversion albedo, significantly constraining the enhancement of the snow BRDF/albedo product accuracy. Consequently, there is burgeoning interest in developing novel direct estimation algorithms focused on refining albedo accuracy in snow-covered regions.

Machine learning techniques provide practical methods for remote sensing inversion, particularly given the complexities and challenges associated with model intricacies and the existence of multiple optimal solutions (Chen et al., 2021; Chen and Guestrin, 2016). The initial direct estimation algorithm was primarily used to establish the relationship between a simulation database and TOA reflectance. Although uncertainties in simulation data could introduce ambiguity into inversion outcomes, this concern has been alleviated to a large extent by the abundance of available data. In the contemporary era, satellite observations provide vast quantities of remote sensing data, while globally distributed ground-based radiation flux observation networks provide continuous radiation measurements at various sites. Consequently, machine learning algorithms are now directly employed to connect satellite observations with global site albedo measurements, showing significant potential for snow albedo estimation. In this study, a machine learning algorithm (i.e, eXtreme Gradient Boosting, XGBoost) designed to directly estimate blue-sky snow shortwave albedo utilizing MODIS data alongside ancillary information is introduced. The methodology incorporates reanalysis datasets to ensure sufficient snow information, establishing a direct linkage between satellite observations and global site albedo measurements to circumvent uncertainties associated with simulation data. Moreover, it leverages the NASA operational MODIS MCD43 albedo (MCD43) product as prior knowledge to enhance algorithm robustness. Consequently, this approach mitigates the propagation of uncertainty across multiple stages, such as atmospheric correction, BRDF angular modelling, simulation database usage, and narrow-to-broadband conversion.

The structure of this paper is delineated as follows. Section 2 details

Table 1
Summary of the verification accuracy of surface albedo products in snow-covered and snow-free areas.

References	Products	Snow-free albedo	Snow albedo
Chen et al. (2015)	MISR	R^2 = 0.747, RSE = 0.051	R^2 = 0.172, RSE = 0.265
Stroeve et al. (2005, 2013)	MODIS	/	RMSE = 0.069, bias = −0.05
Liu et al. (2013)	GLASS	RMSE = 0.030, bias = −0.002	RMSE = 0.126, bias = 0.005
Liu et al. (2017)	MODIS	RMSE = 0.020, bias = −0.004	RMSE = 0.063, bias = −0.028
	VIIRS	RMSE = 0.020, bias = −0.004	RMSE = 0.049, bias = −0.016
Wang et al. (2017)	VIIRS	RMSE = 0.023, bias = 0.002	RMSE = 0.050, bias = 0.032
Wu et al. (2018)	MODIS	RMSE = 0.032, bias = −0.016	RMSE = 0.054, bias = −0.028
	MuSyQ	RMSE = 0.033, bias = −0.014	RMSE = 0.054, bias = −0.013
He et al. (2018)	Landsat	RMSE = 0.027, bias = −0.012	RMSE = 0.071, bias = −0.011
Corbea-Pérez et al. (2021)	MODIS	/	RMSE = 0.070, bias = −0.010

Note: Coefficient of determination (R^2), residual standard error (RSE), root mean square error (RMSE).

the diverse snow data sources, encompassing site albedo measurements, satellite data, and ERA5-Land data. Section 3 outlines the overarching framework of the algorithm and model training process. In Section 4, we conduct a comprehensive analysis of the results obtained through the proposed method, employing a range of data sources. Finally, we address potential limitations and draw conclusions based on the primary findings in Section 5 and Section 6.

2. Data

To estimate shortwave snow albedo using the XGBoost algorithm with MODIS observations, we used a variety of data sources, including collecting in situ shortwave albedo measurements on a global scale for model training and validation, satellite data as model inputs and product comparisons (e.g., surface variables and geolocation information), and a reanalysis dataset (i.e., ERA5-Land) as additional model inputs. Each type of data (i.e., in situ measurements, satellite data, and ERA5-Land data) and the corresponding processing steps are described in the following sections.

2.1. Ground measurements

In situ shortwave radiation measurements were collected at eight observation networks from 2002 to 2019 to obtain sufficient shortwave snow albedo for model training and validation. The eight observation networks are the AmeriFlux network, the FLUXNET network, the AsiaFlux network, the European Fluxes Database Cluster (EFDC), the Coordinated Energy and Water Cycle Observations Project (CEOP) network, the Baseline Surface Radiation Network (BSRN) network, the Programme for Monitoring of the Greenland Ice Sheet (PROMICE), and the Marine and Atmospheric Research (IMAU). There were some overlapping or spatially close sites between networks. We selected one set of data for each site to avoid data repetition and confusion. Strict data quality control was carefully conducted before the aggregation of in situ shortwave albedo data. First, we removed the raw data records labeled with a bad quality flag. Then, we checked the temporal continuity and removed the individual sites with few continuity records. Finally, we manually inspected and removed any unreasonable radiation values. In total, we selected 101 sites for model training and validation. There were 4 sites from the EFDC, 12 sites from FLUXNET, 40 sites from AmeriFlux, 8 sites from the BSRN, 2 sites from the CEOP, 24 sites from the PROMICE, 9 sites from the IMAU, and 2 sites from AsiaFlux. Fig. 1 shows the spatial distribution of the 101 in situ observation sites used in this study. The white areas labeled IGBP = 15 (i.e., Snow and Ice) represent perennial snow albedo sites, primarily located in Greenland and Antarctica. The stations in other regions are predominantly seasonal snow albedo sites, mainly distributed across North America and Europe. We estimated the daily mean albedo from three hourly averages of shortwave downwelling and upwelling irradiance centred at noon (11:00, 12:00, and 13:00) (Jin et al., 2003; Wright et al., 2014). To smooth the albedo variation caused by blowing wind, the addition of fresh snow, and other disturbances, we utilized a 3-h average irradiance around noon to calculate the daily snow albedo.

2.2. Satellite data

The satellite data used in this study are summarized in Table 2. We collected MODIS 1 km TOA Level-1B calibrated radiance observations (i.

Table 2
Summary of the satellite data used in this study.

Products	Variables	Resolution (spatial/ temporal)
MOD021KM, MYD021KM	TOA reflectance of B1-B7	1 km/instant
MOD03, MYD03	VZA, SZA, RAA	1 km/instant
MOD10A1, MYD10A1	NDSI	500 m/daily
MCD43A2, MCD43A3	MODIS MCD43 albedo	500 m/daily
GLASS02A06	GLASS albedo	500 m/4-days
SURFALB	VIIRS SURFALB albedo	1 km/daily
GTOPO30, MERIT	Elevation, slope and aspect	1 km/1997, 90 m/2003

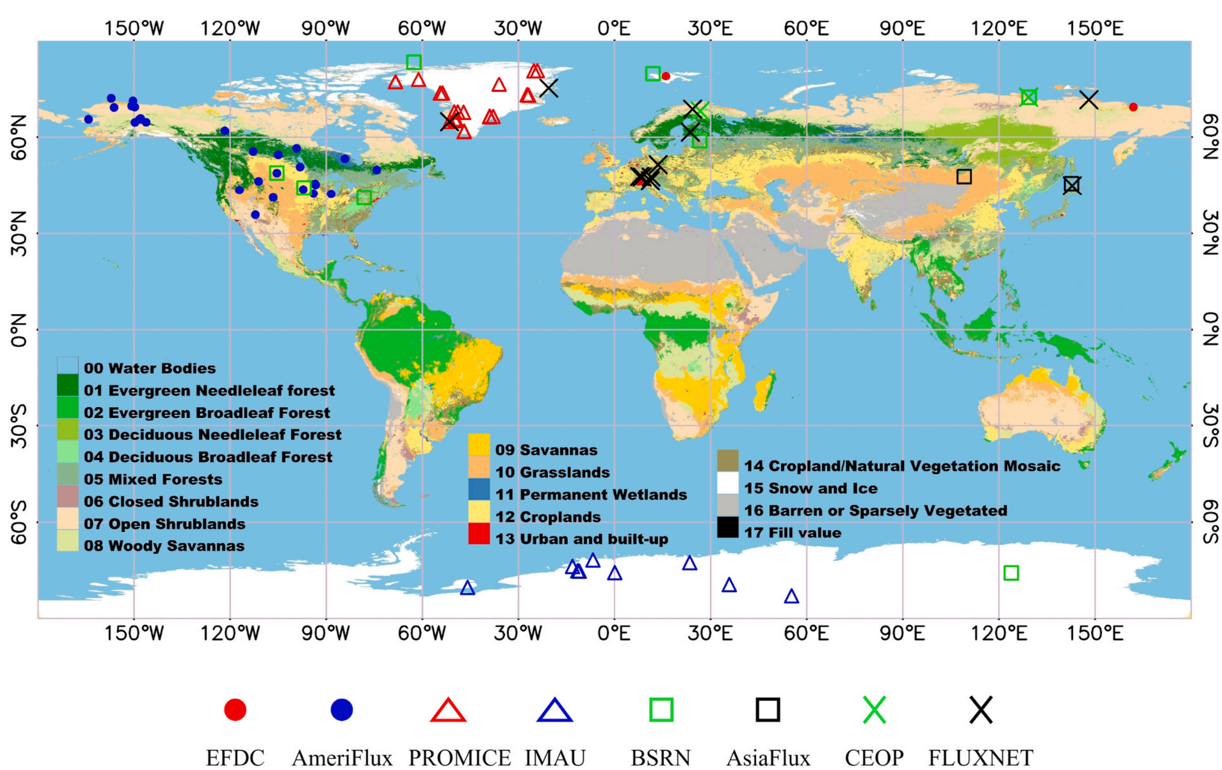

Fig. 1. Spatial distribution of 101 in situ sites from eight radiation observation networks.

e., MOD021KM and MYD021KM) and corresponding geolocation data (i.e., MOD03 and MYD03) from 2002 to 2019, which are available at https://search.earthdata.nasa.gov. The MOD021KM and MYD021KM products with a 1 km resolution provide the MODIS TOA reflectance from visible to near-infrared wavelengths. The MOD03 and MYD03 products offer latitude, longitude, surface elevation and angular geometry information, and the latitude and longitude information can be used to match the in situ measurements with satellite observations. In this study, the TOA reflectance, angle information of the seven shortwave bands (i.e., B1-B7), and surface elevation were utilized to estimate the blue-sky shortwave albedo in snow-covered areas. The daily MODIS snow cover products (i.e., MOD10A1 and MYD10A1) from Collection 6 were utilized to identify snow conditions.

The MODIS, GLASS and VIIRS SURFALB albedo products have been used to improve the global energy budget and climate change studies and track ephemeral snowfall and snowmelt processes. The NASA operational MODIS Collection V006 daily BRDF/Albedo products (MCD43) at a 500 m gridded resolution provide an improved daily temporal resolution compared with the previous 8-day MODIS Collection V005 products at a 500 m resolution, which can effectively improve the temporal monitoring of vegetation phenology and snowmelt (Schaaf et al., 2002; Wang et al., 2012, 2014). Therefore, the MODIS MCD43 albedo was used as prior knowledge for model training to improve the accuracy of snow albedo estimation. The GLASS albedo products suite includes albedo from AVHRR at 0.05° and from MODIS at 250 and 500 m (Liang et al., 2021). The latest version (V40) of the GLASS albedo product was used because it has a spatial resolution of 500 m and a temporal resolution of 4 days. The National Oceanic and Atmospheric Administration (NOAA) VIIRS surface albedo (SURFALB) is a novel approach for directly estimating daily blue-sky albedo from a single directional observation and provides daily values at a 1 km spatial resolution (Wang et al., 2013, 2017), while the direct estimation algorithm retrieves daily albedo from VIIRS observations by determining the relationship between LSA and TOA reflectance. The NOAA VIIRS SURFALB albedo has been validated with ground measurements (Zhou et al., 2016), and the current retrieval errors are relatively small and are regularly updated. This VIIRS SURFALB product is available from https://www.star.nesdis.noaa.gov/jpss/index.php.

The GTOPO30 data were from a global digital elevation model (DEM), which was released in 1996 with a grid spacing of 30 arc seconds (~1 km). The elevation of the GTOPO30 data ranged from −407 to 8752 m, and the ocean area was designated as a fill value. The resolution and accuracy of the GTOPO30 data were not as high as those of the multi-error-removed improved-terrain (MERIT) DEM data, but they had global coverage, which was better than that of the MERIT DEM data. The accuracy of the GTOPO30 data does not have a unified standard, as it depends on the accuracy of the source data, which is generally not greater than ±30 m. In this study, we used the multi-error-removed improved-terrain (MERIT) DEM to fill in the missing data of the GTOPO30 data to eliminate major error components from existing spaceborne DEMs.

2.3. ERA5-land data

ERA5-Land is a reanalysis dataset that can offer a consistent view of the evolution of surface variables over several decades. These reanalysis data combine observations from all over the world with model data to produce a global complete and consistent dataset. The ERA5-Land data offer many variable products at the global scale, including soil, vegetation, snow, temperature, radiation and heat. The ERA5-Land data cover the period from 1981 to the present at a 10 km resolution and offer a temporal resolution of either hourly or monthly data. The temporal and spatial resolutions of the ERA5-Land data make it very useful for all kinds of land surface applications, such as rainfall, snowfall, flood and drought forecasting. In this study, we used the hourly snow parameters of ERA5-Land to capture the rapid variation in snow albedo, including snow albedo (snw_alb), snow cover (snw_cov), snow density (snw_den), snow depth (snw_dep), snow depth water equivalent (snw_equ), temperature of the snow layer (tem_lay), snowfall and snowmelt.

3. Methods

The overall framework of the process developed in this study is shown in Fig. 2. The procedure included three components. First, 90% of the samples were compiled for model training and validation, and the remaining samples were used as an independent dataset for model evaluation. Second, the mean decrease impurity (MDI) method was used to remove redundant variables. Finally, grid research was combined with a random search to determine the parameters of the final model when the variables were determined.

3.1. XGBoost algorithm

The XGBoost algorithm is a machine learning system for tree boosting proposed by Chen and Guestrin in 2016 (Chen and Guestrin, 2016). The algorithm is an efficient implementation of a gradient boosting framework, including efficient linear model solvers and tree learning algorithms, supporting multiple objective functions such as regression, classification, and ranking. The XGBoost algorithm is based on boosting integration technology and develops a strong learner by combining a group of weak learners with additive strategies, which has the advantages of high speed and efficiency, outstanding performance, support for multiple input types, and customizable functions. The Python package of the XGBoost algorithm was used to implement this method. The main parameters that determined the structure of the model included the number of gradient-boosted trees and the maximum tree depth, which were adjusted using a grid search method based on 10-fold cross-validation. The training accuracy of the XGBoost algorithm is generally the same as that of the RF algorithm, but it usually has a more efficient running speed.

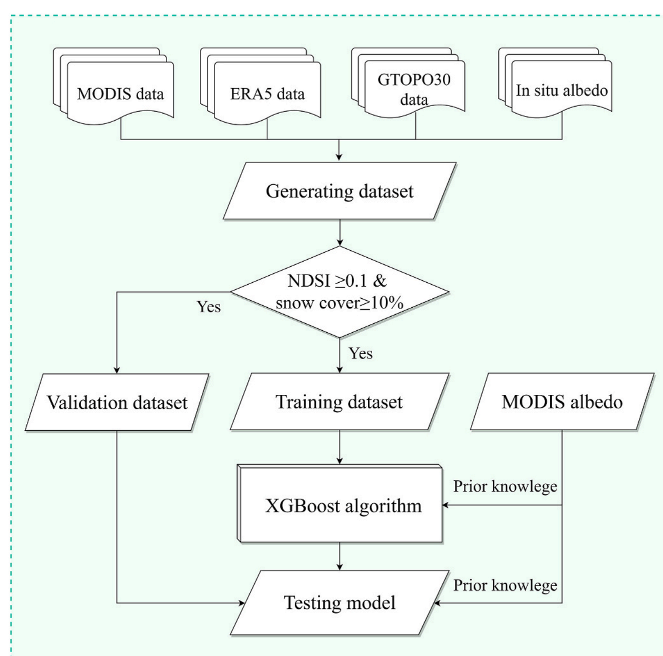

Fig. 2. Flow chart of the XGBoost algorithm used to estimate the shortwave albedo of snow.

3.2. Determination of model parameters

All the samples from 2002 to 2019 were compiled, and the data from 2013 to 2014 were reserved for independent validation. The samples were randomly divided into two groups: the training dataset was used to obtain an XGBoost model (80%), and the validation dataset was used to select the optimal model (20%). The following variables were selected: TOA reflectance of b1-b7, angle information (i.e., SZA, VZA, RAA), NDSI, terrain data (i.e., elevation, slope and aspect), and snow parameters of the ERA5-Land data (i.e., snw_alb, snw_cov, snw_den, snw_dep, snw_equ, tem_lay, snowfall, and snowmelt). Note that the albedo is the integral of the BRDF over the entire hemispherical geometry. Therefore, the TOA reflectance and angle information are essential in this process. Numerous studies have shown that the distribution, type, and albedo variations of snow are highly dependent on elevation (Grünewald et al., 2014; Huang et al., 2017; Jain et al., 2009; Trujillo et al., 2012). Additionally, for high-reflectance surfaces like snow, it is necessary to consider topographic influences (Shi and Xiao, 2022). Consequently, terrain data is also incorporated into the estimation of snow albedo. The snow parameters of the ERA5-Land data can be used to reflect the state of the snow surface, such as new snow, old snow, and melting snow. The MODIS NDSI (NDSI $\geq$0.1) and snow cover (snw_cov $\geq$10%) of the ERA5-Land data were utilized to identify snow-covered conditions and prevent nonsnow pixels from introducing errors (Kouki et al., 2023; Zhang et al., 2019).

The mean decrease in impurity (MDI) index was used to select the optimal parameters of the model to prevent the model from being too complex, which can effectively reflect the contribution of parameters to the model and eliminate some variables with low contribution rates. Fig. 3 shows that the temperature of the snow layer was dominant, followed by snow cover and DOY, and, finally, TOA reflectance-related information. The MDI values of TOA reflectance and angle information were low, which may be because snow recognition and conditions have a greater essential influence on snow albedo. The variables with low MDI values did not indicate an insufficient correlation with snow albedo but indicated a greater correlation with the variables at the higher rankings. There were no obvious changes in the model accuracy after feature selection, which demonstrated that the eliminated variables were redundant for the model construction. After feature selection, the snow albedo can be estimated as follows:

$$Snow\_albedo = F(DOY, SZA, NDSI, elevation, slope, aspect, snw\_alb, snw\_den, snw\_dep, snw\_equ, tem\_lay) \tag{1}$$

In addition, we used the MODIS albedo (i.e., BSA and WSA) as prior knowledge to improve the accuracy of the model training.

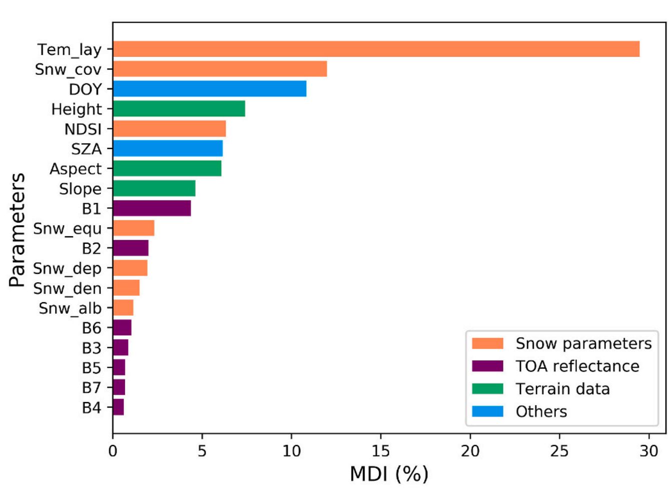

Fig. 3. The mean decrease in impurity (MDI) results for the determination of model parameters.

3.3. Evaluation approaches

In this study, we proposed a machine learning algorithm for improving the MODIS shortwave albedo product in snow-covered areas. First, 10-fold cross-validation (CV) and independent validation samples from two years (i.e., 2013 and 2014) were used to validate the performance of the model. Then, several official albedo products (i.e., MODIS, GLASS and VIIRS SURFALB) were applied for comparison with our proposed method in terms of independent sites and spatial patterns. Finally, we selected the root mean square error (RMSE), mean relative error (MRE) and bias values as the quality assessment indices, which are expressed as shown in Eqs. (2)–(4):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_2 - y_1)^2}{n}} \tag{2}$$

$$MRE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_2 - y_1}{y_1}\right| \times 100\% \tag{3}$$

$$bias = \frac{\sum_{i=1}^{n}(y_2 - y_1)}{n} \tag{4}$$

where y_1 represents the site albedo, y_2 represents the snow albedo retrieved using the proposed method, and n represents the amount of site albedo.

4. Results and analysis

4.1. XGBoost model training and validation

In this section, we first explore the training accuracy of the use of the NASA MODIS MCD43 albedo as prior knowledge for estimating the new snow shortwave albedo. Fig. 4 shows a comparison of model validation for the XGBoost algorithm without the MODIS MCD43 albedo as prior knowledge and for the XGBoost algorithm with the MODIS MCD43 albedo as prior knowledge. The use of the XGBoost algorithm to directly model site albedo and satellite observation data has proven feasible, as evidenced by the algorithm's exceptionally high training accuracy. The predicted albedo of the XGBoost algorithm without the MODIS albedo as prior knowledge was shown to be quite accurate, presenting a high correlation coefficient (R^2 = 0.932; RMSE = 0.054) and a negligible bias. However, the predicted albedo of the XGBoost algorithm with the MODIS albedo as prior knowledge showed a slightly greater accuracy than the site albedo (R^2 = 0.953 and RMSE = 0.044). Meanwhile, the MRE value derived using the XGBoost algorithm with the MODIS MCD43 albedo as prior knowledge (MRE = 5.490%) was smaller than that of the result of the XGBoost algorithm without the MODIS MCD43 albedo as prior knowledge (MRE = 7.253%). In addition, the result of Fig. 4a strayed from the 1:1 line since there was a lack of constraint on the MODIS MCD43 albedo, especially at low snow albedo. The result of Fig. 4b was more concentrated on the 1:1 line due to the constraint of the MODIS MCD43 albedo. The model validation results indicate that at perennial snow sites, the accuracy is relatively high even without incorporating MODIS albedo as prior knowledge. Thus, adding MODIS albedo does not significantly enhance model accuracy at these sites. However, in non-pure snow regions, where surface cover types are complex and mixed pixel issues are prominent, incorporating MODIS albedo as prior knowledge during model training provides a strong constraint. Consequently, the validation results show that the improvement in model accuracy is more pronounced in non-pure snow regions when MODIS albedo is included as prior knowledge compared to perennial snow sites. These results demonstrate that the MODIS MCD43 albedo performs as well as prior knowledge and can improve the accuracy of model training, especially at low snow albedo values.

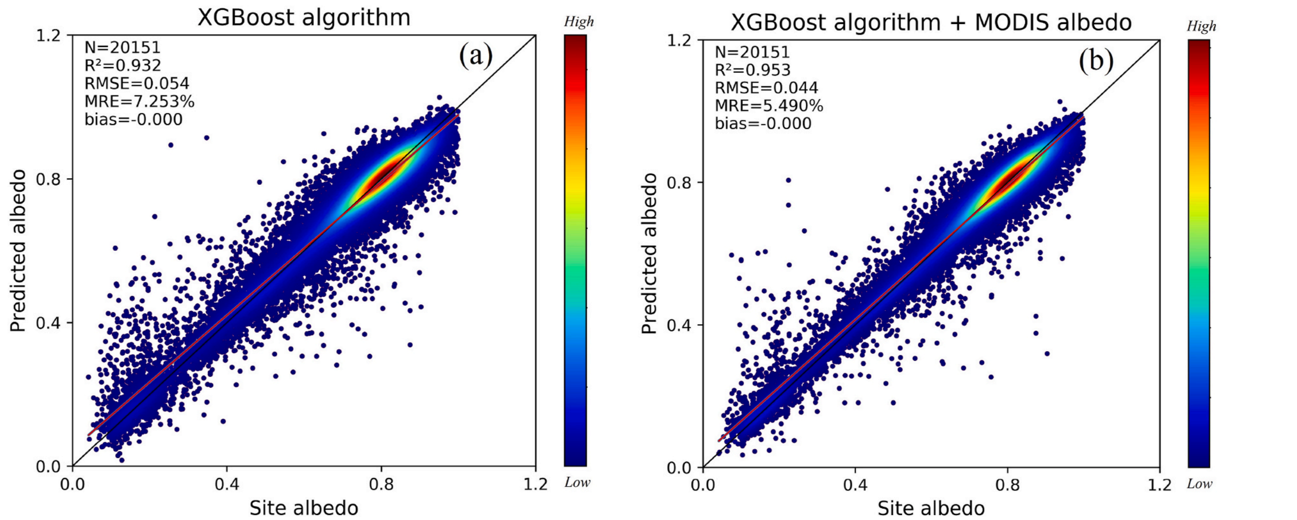

Fig. 4. Density scatterplots of model validation for the XGBoost algorithm without the MODIS MCD43 albedo as prior knowledge (a) and for the XGBoost algorithm with the MODIS MCD43 albedo as prior knowledge (b).

4.2. Analysis of the influence of the main parameters

We analyzed the impact of main parameters (i.e., snow layer temperature, snow cover, DOY, height, NDSI, SZA) on the training results through density scatter plots, as shown in Fig. 5. The error source for the temperature of the snow surface may occur at approximately 273.15 K. At this temperature, snow begins to melt, exposing more of the nonsnow surface, which is an unusual challenge in estimating snow albedo. The overall bias range was within 0.2 when the temperature was less than 273.15 K. For the snow cover parameters, the overall difference was relatively large when the snow cover was approximately 20%. Additionally, there were some large deviations, particularly when the snow cover was approximately 100%, which may be due to the uncertainty of the ERA5-Land product. Due to variations in the DOY, there was a large uncertainty in the spring and autumn due to snow detection and conditions during snowfall and snowmelt periods, which led to greater deviations. The variation in snow albedo was relatively complicated, and the overall uncertainty was greater when the surface height was relatively low. The variation in the snow albedo was relatively small, and the overall deviation was relatively small when the snow height was relatively high. There was no significant difference in the range of the NDSI, but there was a greater deviation at low NDSI. The possible reason is that there may be greater uncertainty in snow recognition at low NDSI values. Surprisingly, there was no obvious difference in the overall deviation within the entire SZA variation range, especially when the SZA was greater than 70°, although the MODIS MCD43 albedo products are usually of low quality at these SZAs (Schaaf et al., 2011a). In general, the error source of model training appeared to mainly come from snow detection and the processes of snowfall and snowmelt. Moreover, the model training accuracy was greater in purely snow-covered areas.

4.3. Analysis of model performance at individual sites

In this section, we assessed the performance of the proposed approach in retrieving snow albedo compared with current albedo products (i.e., MODIS MCD43, GLASS, and VIIRS SURFALB albedo) and with site albedo measurements. In the validation process, we ignored the MODIS quality control information because the MODIS MCD43 albedo is usually low quality at high latitudes. Fig. 6 shows comparisons of the trend and histogram results of the MODIS MCD43, GLASS, VIIRS SURFALB, and albedo retrieved by the proposed XGBoost approach with the site albedo measurements. For the sites in the Greenland region in Fig. 6a, the current albedo products and predicted albedo had very good agreement with the site albedo at the UPE_L site in 2014 and could very effectively capture the process of melting at DOY = 140–250. However, the current albedo products appeared to have more obvious underestimations at DOY = 168–234 during snowfall. The XGBoost-predicted albedo using the proposed approach appeared to solve this underestimation problem, resulting in a higher consistency with the site albedo than with the current albedo products. For the sites in the Antarctic region shown in Fig. 6b, snow cover existed year-round, and the variation trends of snow albedo were generally similar. These albedo products showed high consistency compared with the site albedo at the AWS5 site in 2013, and the albedo predicted by the proposed approach also performed well. However, there was an obvious underestimation of these albedo products when approaching the Southern Hemisphere in winter at DOY = 97–107. The possible reason is that the snow stops melting during this period, the albedo increases, and the current albedo products do not seem to capture this phenomenon. In contrast, the albedo predicted by the proposed approach captured this variation well. In addition, our proposed approach can effectively improve the unreasonable values displayed by the MODIS MCD43 albedo (i.e., DOY = 62–72) at this particular site. For the variation in seasonal snow shown in Fig. 6c, the snow melting process occurred after snowfall at the CA_ARB site in 2013, which caused the snow albedo to increase suddenly and then decrease. The albedo predicted by the proposed approach model captured this variation well. Moreover, our proposed approach effectively improved the underestimation of the current albedo products at DOY = 1–100, especially for the VIIRS SURFALB albedo. In general, the snow albedo retrieved by our proposed approach matched quite well with the site albedo, as it also captured the rapidity of the snowfall and snowmelt process.

Fig. 7 summarizes the results for all the sites. The MODIS MCD43, GLASS, and VIIRS SURFALB albedo generally matched well with all of the site albedos. The validation accuracies of the MODIS MCD43 and GLASS albedo were comparable, and their validation results were slightly better than that of the VIIRS SURFALB albedo, which may be because the VIIRS SURFALB albedo product used only one satellite's observations and provided the daily mean values. The mean differences between these albedo products and the site albedo were approximately 0.032, with RMSE values varying from 0.087 to 0.100 and MRE values ranging from 10.634% to 13.203%. These results indicate that these albedo products are significantly underestimated in snow-covered areas compared with site albedo, especially in nonpure snow areas. This is most likely due to the field data capturing significant understorey snow that is missed by coarser resolution satellite-derived products. Since the

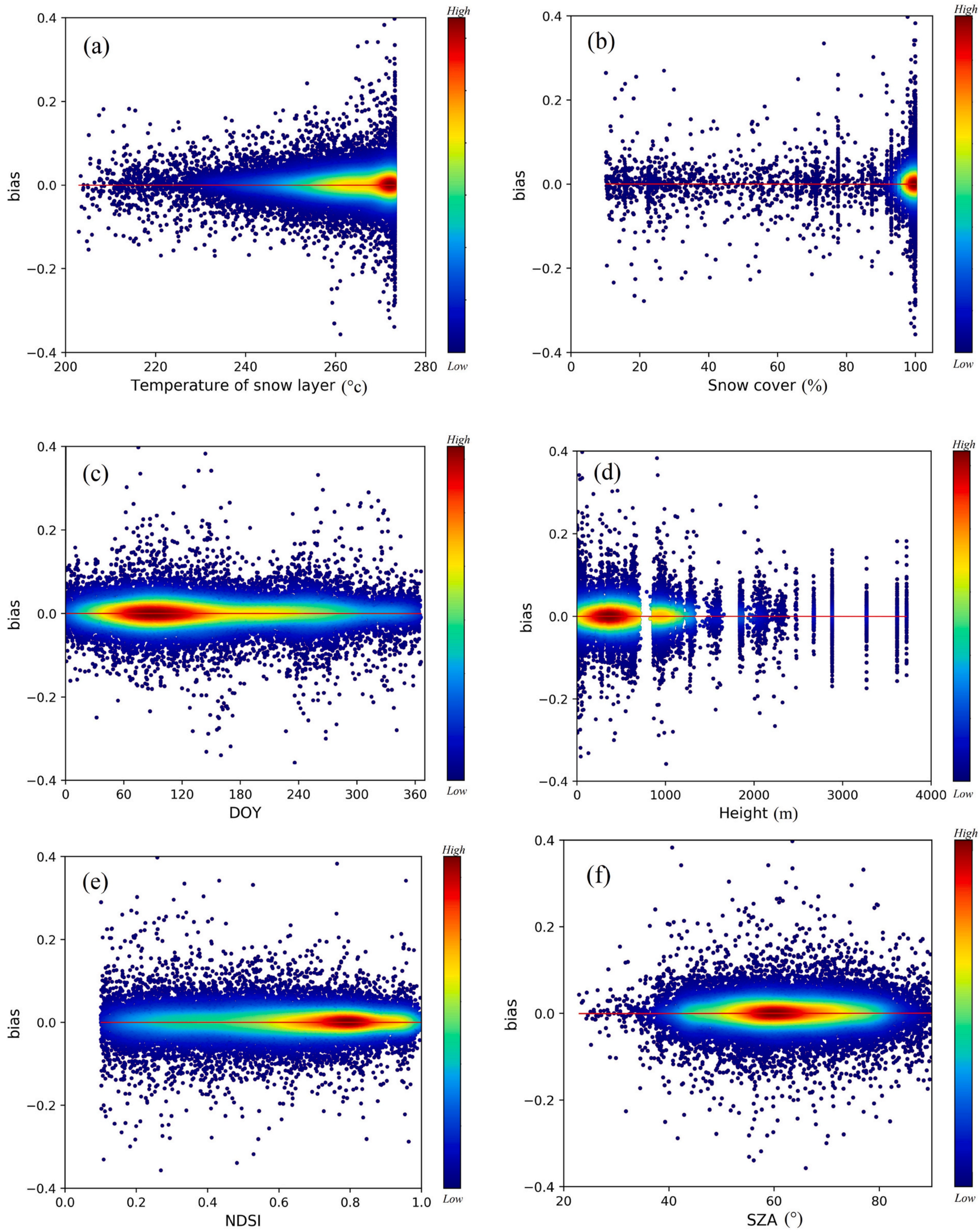

Fig. 5. The density scatter plot showing the impact of main parameters on the training results: the temperature of the snow layer (a), snow cover (b), DOY (c), height (d), NDSI (e), and SZA (f).

GLASS albedo directly establishes the relationship between TOA reflectance and simulated albedo by utilizing high-quality MODIS MCD43 BRDF/Albedo products, this product may have similar problems. Compared with those of the current satellite-derived albedo products, the mean difference between the albedo retrieved by the proposed XGBoost approach and the site albedo was low, with an RMSE of approximately 0.052 and an MRE of approximately 7.527%. The albedo predicted by the proposed XGBoost approach seemed to address the underestimation of the current albedo products very well and showed a greater consistency with the site albedo. In general, however, these albedo products showed higher accuracies in areas with permanent snow cover but still indicated a slight underestimation compared with the site albedo. The current albedo products need to be further improved in seasonal snow regions, where it is difficult to capture timely

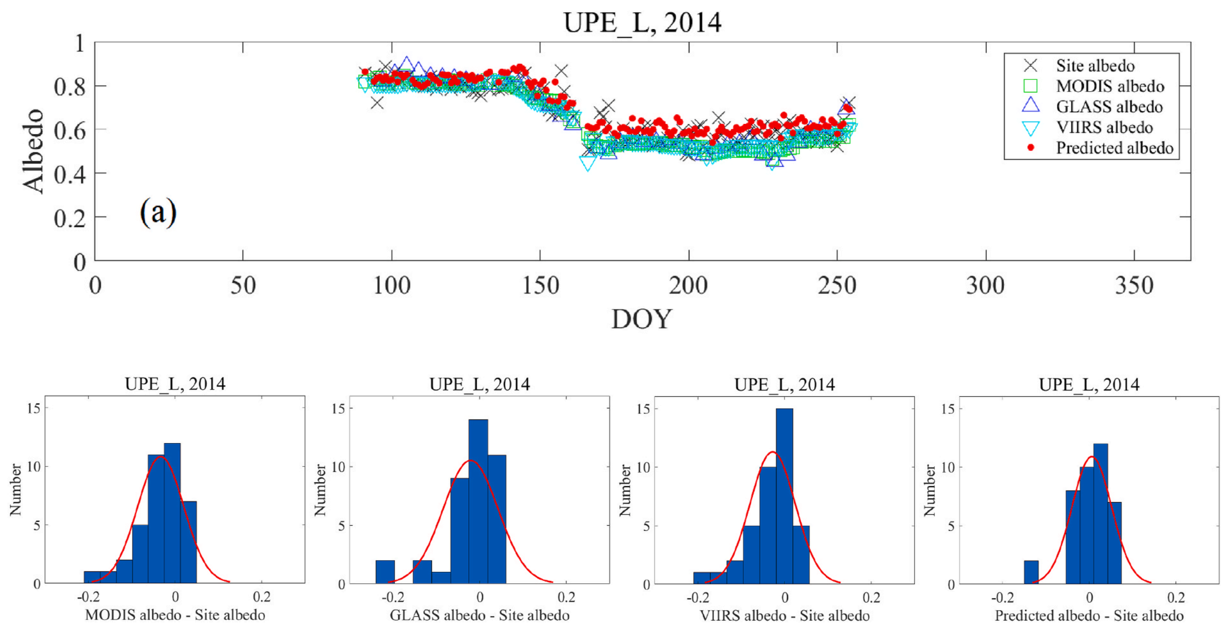

Fig. 6A. Comparison of the trend and histogram results of the MODIS MCD43, GLASS, and VIIRS SURFALB products and the albedo retrieved with the proposed approach for the site albedo at the UPE_L site in 2014.

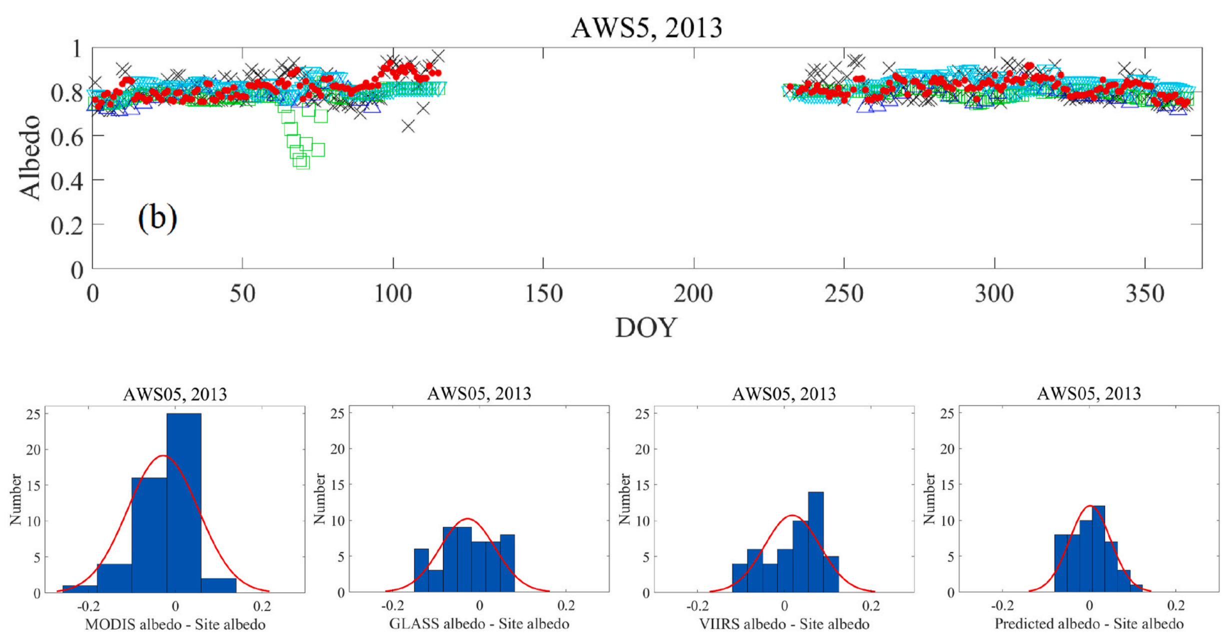

Fig. 6b. Comparison of the trend and histogram results of the MODIS MCD43, GLASS, and VIIRS SURFALB products and the albedo retrieved with the proposed approach for the site albedo at the AWS5 site in 2013.

estimates of snowfall and snowmelt processes and therefore may represent a larger underestimation. Another possible reason is that the MODIS MCD43 albedo is obtained using multidate clear-sky surface reflectance values over 16 days to fit the RTLSR BRDF model, although the date of interest is emphasized to capture the conditions on that particular day. The proposed XGBoost method directly links TOA reflectance to instantaneous LSA and does not require atmospheric correction or accumulation of observations over a certain period. Therefore, it can characterize the temporal variation in LSA better, especially when the LSA changes rapidly. In general, these current satellite albedo products were more consistent with the retrieved albedo in Antarctica but still slightly underestimated compared with the site albedo, which may be due to the inability to capture surface roughness and variability.

4.4. Analysis of spatial patterns

In this section, we compared the spatial patterns of the NASA operational MODIS MCD43 albedo, GLASS albedo, and NOAA VIIRS SURFALB albedo with the retrieved albedo using the proposed XGBoost approach in 2013 for a single day at DOY = 184 for the h17v01 tiles and DOY = 328 for the h18v16 tiles. These two typical tiles cover the different regions of Greenland and Antarctica. Fig. 8 shows a comparison of the spatial distribution between the current satellite-derived albedo

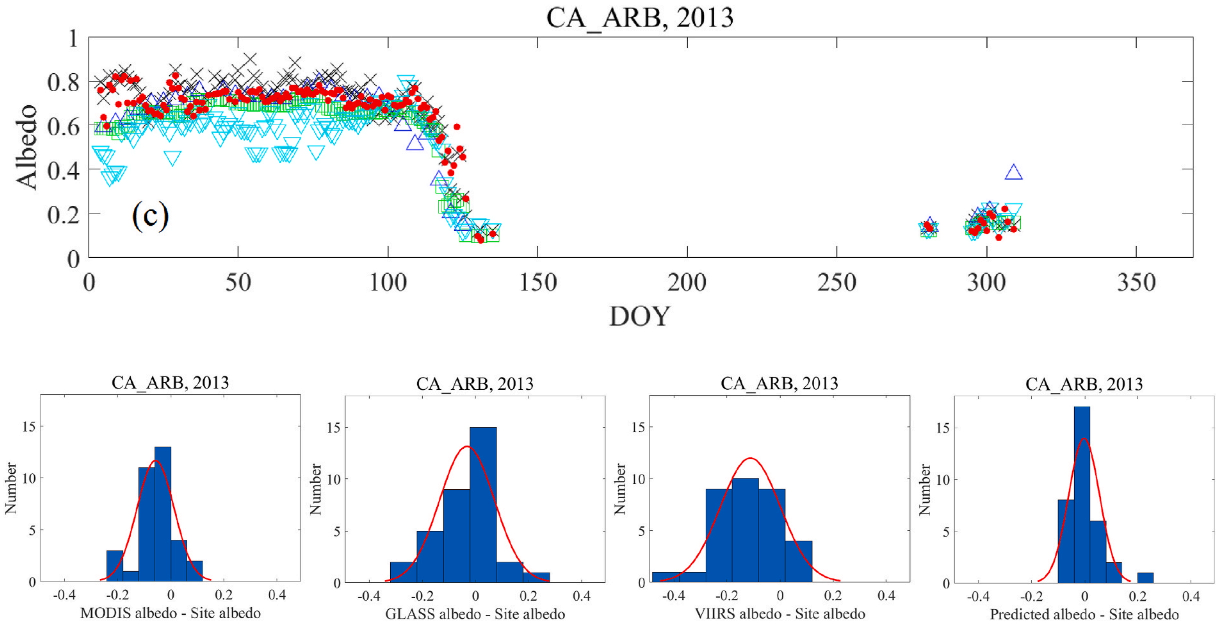

Fig. 6c. Comparison of the trend and histogram results of the MODIS MCD43, GLASS, and VIIRS SURFALB products and the albedo retrieved with the proposed approach for the site albedo at the CA_ARB site in 2013.

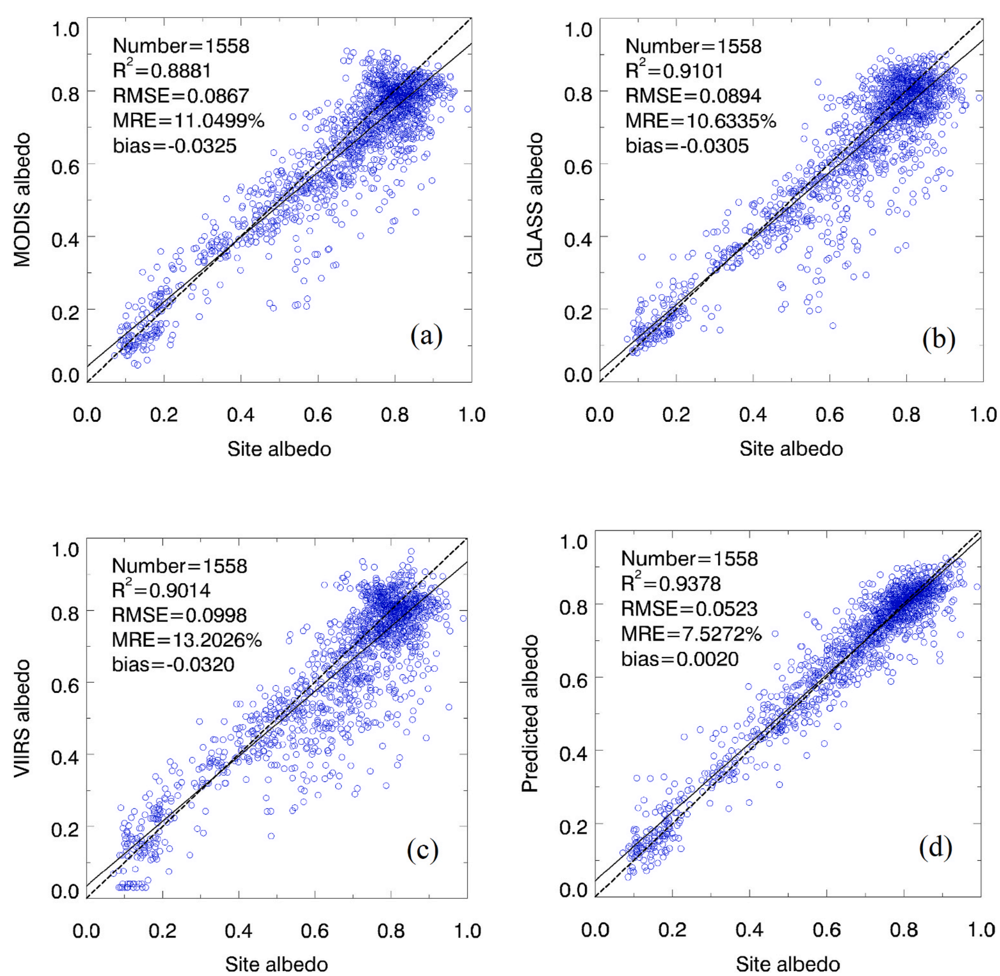

Fig. 7. Comparison of all site albedo based on the MODIS MCD43 albedo (a), GLASS albedo (b), VIIRS SURFALB albedo (c) products and the albedo (d) predicted with the proposed approach.

products and the retrieved albedo using the proposed XGBoost approach in 2013 at DOY = 184 for tile h17v01, Greenland. It is clear that the retrieved XGBoost albedo was in high agreement with these other albedo products, but there was a slight difference, especially for the VIIRS SURFALB albedo. The VIIRS SURFALB albedo exhibited higher values at the top of the h17v01 tile and lower values at the bottom of the h17v01 tile compared with the other albedo results. In addition, the retrieved albedo using the proposed XGBoost approach was in better agreement with the MODIS MCD43 albedo, which may be caused by the use of the MODIS MCD43 albedo as prior knowledge and direct links of satellite observations to site albedo using the XGBoost algorithm. However, the retrieved XGBoost albedo values were slightly greater than the MODIS MCD43 albedo values, especially at the top of tile h17v01. By comparing Figs. 8 and 9, we can see that the spatial pattern of the h18v16 tile was more consistent than the pattern of the h17v01 tile, which may be caused by snow cover year round in the Antarctic region, with a smaller overall variation in snow albedo. However, there was a larger difference between these currently derived satellite albedo products and the retrieved XGBoost albedo in the bottom right of the h18v16 tile, and the VIIRS SURFALB albedo exhibited higher albedo values than any other albedo at high latitudes in the h18v16 tile. In general, the albedo retrieved by the proposed method was in very good agreement with the current albedo products in this study.

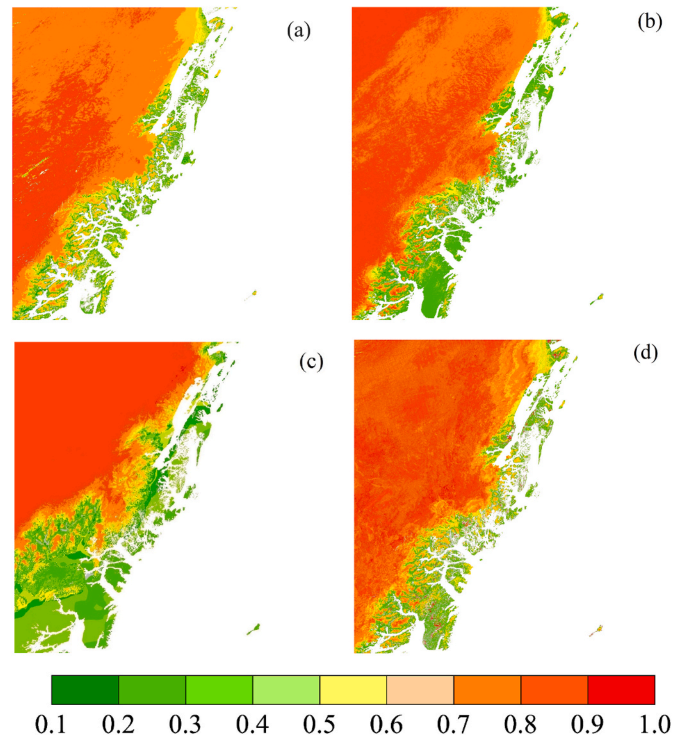

Fig. 8. Comparison of the spatial distributions of the MODIS MCD43 albedo (a), GLASS albedo (b), VIIRS SURFALB albedo (c) and XGBoost albedo (d) retrieved in this study using the proposed approach on DOY = 184 in the h17v01 tile.

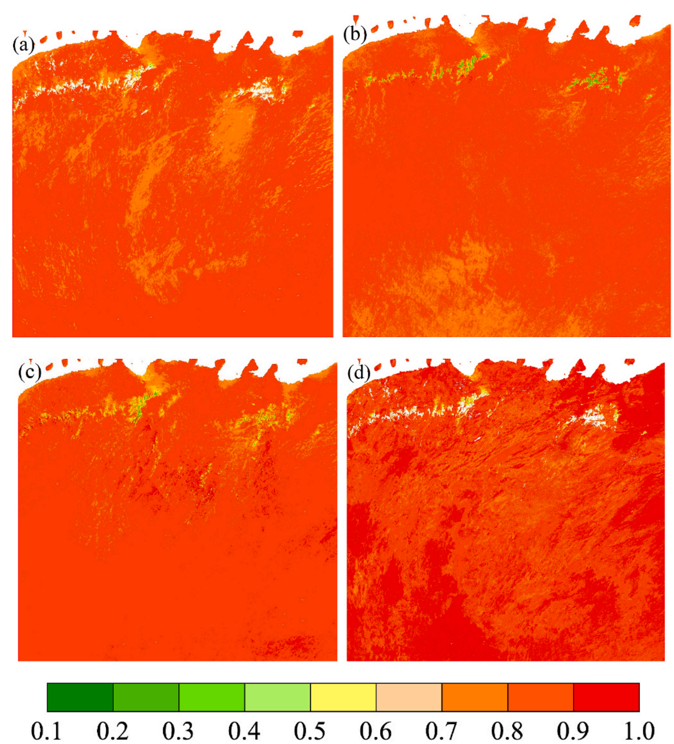

Fig. 9. Comparison of spatial distribution between the MODIS MCD43 albedo (a), GLASS albedo (b), VIIRS SURFALB albedo (c) and albedo (d) retrieved in this study using the proposed approach on DOY = 328 in the h18v16 tile.

5. Discussion

5.1. Model training with Terra and Aqua data

In this section, we explored the training accuracy for the Terra and Aqua data separately for estimating snow shortwave albedo. Fig. 10 shows density scatterplots of model validation for the Terra and Aqua data using the MODIS MCD43 albedo as prior knowledge (acknowledging that the MCD43 product uses both Terra and Aqua data). The R^2 value of the Terra data was approximately 0.933, and the RMSE value was approximately 0.053 and had a negligible bias, which indicated that the training accuracy for the Terra data showed high accuracy with site albedo. The R^2 value of the Aqua data was approximately 0.928, and the RMSE value was approximately 0.055 and had a negligible bias, indicating that the training accuracy for the Aqua data also showed high accuracy with site albedo. Both Terra and Aqua data performed well for model training for estimating snow shortwave albedo. The MRE value derived from the Terra data was approximately 6.772%, which was reduced by 0.162% compared to the result derived from the Aqua data. The accuracy of the model training was greater than that of using the Terra and Aqua satellites alone (i.e., Fig. 4b) when the data from the Terra and Aqua satellites were combined. The R^2 value of the Terra and Aqua data was approximately 0.953, and the RMSE value was approximately 0.044, which was much better than the results of the Terra or Aqua data alone. There was more information when the data from the morning Terra and afternoon Aqua satellites were combined. Moreover, the MRE value was approximately 5.490% when the Terra and Aqua satellite data were combined, which was reduced by 1.280% and 1.444%, respectively, compared to the results derived from the Terra and Aqua data alone. These results demonstrate that both the Terra data performed slightly better than the Aqua data for estimating snow shortwave albedo. The accuracy of model training was greater when both the Terra and Aqua satellites were used than when the Terra and Aqua satellites were used alone. Additionally, the MODIS albedo produced by full inversion and magnitude inversion algorithms was not distinguished in the model training. This is because, in high-latitude regions, the MODIS albedo product mainly comes from the magnitude inversion algorithm, accounting for 74.60% (Ding et al., 2023). Therefore, the MODIS albedo product is divided into sets associated with the full inversion and magnitude inversion algorithms, which have a negligible impact on the overall training results.

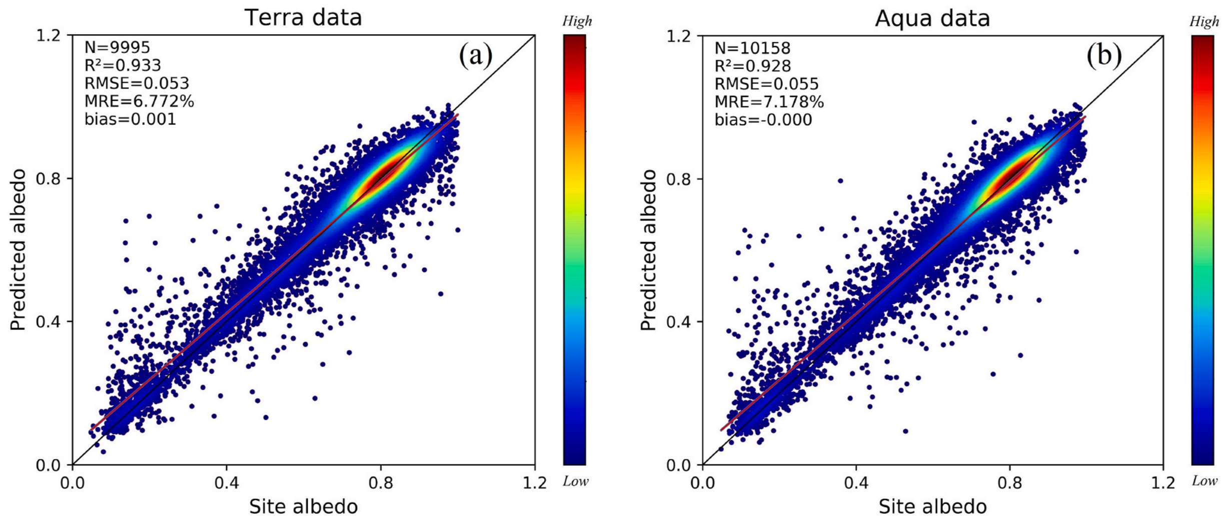

Fig. 10. Density scatterplots of model validation for the Terra (a) and Aqua (b) data using the MODIS MCD43 albedo product (which uses both Terra and Aqua) as prior knowledge.

5.2. *Uncertainty of snow data*

The in situ measurements of LSA were calculated as the ratio of total upwards to total downwards radiation, which was measured by an automatic weather station (AWS). For the field measurements, the instrument has a spectral range of 300–2500 nm in the PROMICE network (Fausto et al., 2021), which has a measurement uncertainty of 5% for daily totals. The IMAU network is equipped with Kipp and Zonen CNR1 or CNR4 radiometers in Antarctica. The CNR1 radiometer has a spectral range of 305–2800 nm, and the CNR4 radiometer has a spectral range of 300–2800 nm. The measurement uncertainty of these two radiometers is approximately 10% (Jakobs et al., 2020). The AmeriFlux and FLUXNET networks have a spectral range of 280–2800 nm, and data are recorded every 30 min. The BSRN network has a spectral range of 280–3000 nm, which is released at 3- and 1-min intervals (Zhang et al., 2019). The measurement uncertainty of these networks is approximately 10%. The AWS uses a single-frequency global positioning system (GPS) receiver to measure the position and elevation of each station. Some sites on the ice surface have been repositioned during maintenance visits over distances larger than several tens of metres. The location change affects the radiation measured since the stations have moved away from the opening crevasses. In addition, the footprints of the tower measurements were calculated based on the instrument height and effective field of view (Román et al., 2013), and the tower height was generally used to approximate the instrument height. The tower heights of different sites were usually different, which implied that they had different spatial representations. We carried out strict quality control for various data sources, and the measurement uncertainty of the instrument, the spectral range, and the spatial representations of the site data were also acknowledged as sources of uncertainty in the model inversion.

5.3. *Analysis of uncertainties in snow albedo estimation*

Quantifying and understanding uncertainties in the estimation of snow shortwave albedo is crucial. This paper discusses data uncertainty and model uncertainty and analyzes their impact on retrieval results. The inherent uncertainties in the training data are one of the main sources of uncertainty in the model's retrieval results. Therefore, cross-validation and uncertainty analysis methods are employed to analyze data uncertainties. By using these methods, we can clearly understand the range of uncertainties in the training data and consider this factor in subsequent model evaluations. Model uncertainty primarily reflects how uncertainties in the training data affect retrieval results. On this basis, we further quantify model uncertainty. Using uncertainty propagation formulas, we transfer the uncertainties from the training data to the model's retrieval results, calculating the uncertainty for each retrieval result. This analysis allows us to quantify and correct these impacts, thereby improving the model's reliability (Wu et al., 2019, 2023).

In the estimation of snow shortwave albedo, the uncertainties at different temporal and spatial scales are crucial for snow albedo. We discussed the uncertainties analysis in different periods of data, such as seasonal and inter-annual variations of uncertainty, and identified the uncertainty patterns in long-term monitoring (Du et al., 2023; Wen et al., 2023). For different geographic regions, we analyzed spatial uncertainties, identified the characteristics of uncertainties in different areas, and explored possible causes. The above uncertainty analysis can further enhance the robustness of the method. Specific improvement measures include: (1) Strengthening data processing methods: by introducing more uncertainty quantification methods, improve the reliability and stability of training data. (2) Optimizing model design: considering uncertainty propagation in model design enhances the model's adaptability to data uncertainty and reduces the impact of uncertainty on retrieval results. (3) Continuous evaluation: establishing a long-term evaluation mechanism, regularly assessing and correcting uncertainties, and improving the long-term application value of the model. We believe these improvements will make the method more reliable and assist in more accurately interpreting and applying retrieval data. These uncertainty analyses provide important directions for future research. In future studies, we will further study and quantify uncertainties to improve the accuracy and applicability of the method.

6. Conclusion

In this study, we proposed a machine learning algorithm for retrieving snow shortwave albedo from MODIS data and other ancillary information. This approach used NASA operational MODIS MCD43 albedo data as prior knowledge and then directly linked satellite observations to global site albedo measurements via the XGBoost algorithm. Validation analyses showed that our proposed approach improved the estimation of shortwave albedo and was quite accurate, with a high correlation coefficient ($R^2 = 0.938$) and a negligible bias between the predicted XGBoost albedo and the tower site albedo. The MODIS MCD43 albedo used as prior knowledge performed quite well in terms of improving the accuracy of model validation, especially at low snow albedo amounts. Therefore, the proposed XGBoost approach is effective at estimating snow shortwave albedo. In addition, both the Terra and Aqua data performed well in model training for estimating snow shortwave albedo, with the Terra data performing slightly better than the Aqua data for estimating snow shortwave albedo. However, the accuracy of the model validation was greater when the Terra and Aqua data were combined than when the Terra and Aqua data were used alone.

In summary, this work presents a promising approach using the XGBoost algorithm to estimate snow shortwave albedo, and we demonstrate the ability of this approach to retrieve the intrinsic albedo. Therefore, this novel direct estimation algorithm, which utilizes both site data and prior knowledge from the current MODIS MCD43 albedo products, has the potential to estimate improved snow shortwave albedo for many applications, e.g., energy budget and climate change studies in snow-covered regions.

CRediT authorship contribution statement

Anxin Ding: Writing – original draft, Validation, Resources, Methodology, Investigation, Data curation. **Shunlin Liang:** Writing – review & editing, Resources, Methodology, Data curation, Conceptualization. **Han Ma:** Writing – review & editing, Validation, Resources, Methodology, Data curation. **Tao He:** Writing – review & editing, Validation, Resources, Funding acquisition. **Aolin Jia:** Writing – review & editing, Validation, Formal analysis, Data curation. **Qian Wang:** Writing – review & editing, Methodology, Data curation.

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.

Data availability

Data will be made available on request.

Acknowledgements

This study was partially supported by the National Natural Science Foundation of China (No. 42301363), Anhui Provincial Natural Science Foundation (No. 2308085QD118), National Key Research and Development Program of China (No. 2016YFA0600103), National Natural Science Foundation of China (No. 42090011), National Natural Science Foundation of China (No. 41971288), and Fundamental Research Funds for the Central Universities (Nos. JZ2024HGTB0254, JZ2023HGQA0148). We are thankful for the data support from the National Earth System Science Data Center, National Science and Technology Infrastructure of China. (http://www.geodata.cn). We would like to thank Prof. Crystal B. Schaaf for her valuable comments and suggestions.

References

Burakowski, E.A., Ollinger, S.V., Lepine, L., Schaaf, C.B., Wang, Z., Dibb, J.E., Hollinger, D.Y., Kim, J.H., Erb, A., Martin, M., 2015. Spatial scaling of reflectance and land surface albedo over a mixed-use, temperate forest landscape during snow-covered periods. Rem. Sens. Environ. 158, 465–477.

Chen, Y., Liang, S., Ma, H., Li, B., He, T., Wang, Q., 2021. An all-sky 1 km daily land surface air temperature product over mainland China for 2003-2019 from MODIS and ancillary data. Earth Syst. Sci. Data 13, 4241–4261.

Chen, T., Guestrin, C., 2016. Xgboost: a scalable tree boosting system. Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining 785–794.

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., 2015. Xgboost: extreme gradient boosting. R package version 0 4–2, 1–4.

Corbea-Pérez, A., Calleja, J.F., Recondo, C., Fernández, S., 2021. Evaluation of the MODIS (C6) daily albedo products for livingston island, antarctic. Rem. Sens. 13, 2357.

Ding, A., Jiao, Z., Dong, Y., Qu, Y., Zhang, X., Xiong, C., He, D., Yin, S., Cui, L., Chang, Y., 2019a. An assessment of the performance of two snow kernels in characterizing snow scattering properties. Int. J. Rem. Sens. 40, 6315–6335.

Ding, A., Jiao, Z., Dong, Y., Zhang, X., Peltoniemi, J.I., Mei, L., Guo, J., Yin, S., Cui, L., Chang, Y., Xie, R., 2019b. Evaluation of the snow albedo retrieved from the snow kernel improved the Ross-Roujean BRDF model. Rem. Sens. 11 (45), 85–96.

Ding, A., Jiao, Z., Zhang, X., Dong, Y., Kokhanovsky, A., Guo, J., Jiang, H., 2023. A practical approach to improve the MODIS MCD43A products in snow-covered areas. Journal of Remote Sensing 3, 57. https://doi.org/10.34133/remotesensing.0057.

Ding, A., Liang, S., Jiao, Z., Ma, H., Kokhanovsky, A.A., Peltoniemi, J., 2022a. Improving the asymptotic radiative transfer model to better characterize the pure snow hyperspectral bidirectional reflectance. IEEE Trans. Geosci. Rem. Sens. 60, 1–16. https://doi.org/10.1109/TGRS.2022.3144831.

Ding, A., Ma, H., Liang, S., He, T., 2022b. Extension of the Hapke model to the spectral domain to characterize soil physical properties. Rem. Sens. Environ. 269, 112843.

Du, X., Wu, X., Tang, R., Zeng, Q., Li, Z., Wang, J., Xiao, Q., 2023. An improved upscaling method of in-situ measurements with consideration of their uncertainty for the spatial scale match between satellite and in-situ measurements. IEEE Trans. Geosci. Rem. Sens. 61, 1–17.

Fausto, R.S., Dirk, V.A., Kenneth, D.M., Baptiste, V., Michele, C., Andreas, Ahlstrøm, P., Signe, B.A., William, C., Nanna, B.K., Kristian, K.K., Niels, J., Korsgaard, S.H.L., Søren, N., Allan, Ø.P., Christopher, L.S., Anne, Solgaard, M., And, J.E.B., 2021. PROMICE automatic weather station data. Earth Syst. Sci. Data 80, 1–41.

Grünewald, T., Bühler, Y., Lehning, M., 2014. Elevation dependency of mountain snow depth. Cryosphere 8 (6), 2381–2394.

Hall, D.K., Riggs, G.A., Salomonson, V.V., 1995. Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data. Rem. Sens. Environ. 54, 127–140.

He, T., Liang, S., Wang, D., Cao, Y., Gao, F., Yu, Y., Feng, M., 2018. Evaluating land surface albedo estimation from Landsat MSS, TM, ETM+, and OLI data based on the unified direct estimation approach. Rem. Sens. Environ. 204, 181–196.

Huang, X., Deng, J., Wang, W., Feng, Q., Liang, T., 2017. Impact of climate and elevation on snow cover using integrated remote sensing snow products in Tibetan Plateau. Rem. Sens. Environ. 190, 274–288.

Jafariserajehlou, S., Rozanov, V.V., Vountas, M., Gatebe, C.K., Burrows, J.P., 2020. On the retrieval of snow grain morphology, the accuracy of simulated reflectance over snow using airborne measurements in the Arctic. Atmos. Meas. Tech. 32, 11–23. https://doi.org/10.5194/amt-2020-58.

Jain, S.K., Goswami, A., Saraf, A.K., 2009. Role of elevation and aspect in snow distribution in Western Himalaya. Water Resour. Manag. 23, 71–83.

Jakobs, S., Reijmer, C.H., Smeets, P., Trusel, L.D., Wessem, J.M.V., 2020. A benchmark dataset of in situ antarctic surface melt rates and energy balance. J. Glaciol. 66, 256–281.

Jia, A., Wang, D., Liang, S., Peng, J., Yu, Y., 2022. Global daily actual and snow-free blue-sky land surface albedo climatology from 20-year MODIS products. J. Geophys. Res. Atmos. 127 (8). https://doi.org/10.1029/2021jd035987.

Jia, A., Wang, D., Liang, S., Peng, J., Yu, Y., 2023. Improved cloudy-sky snow albedo estimates using passive microwave and VIIRS data. ISPRS J. Photogrammetry Remote Sens. 196, 340–355. https://doi.org/10.1016/j.isprsjprs.2023.01.004.

Jiao, Z., Ding, A., Kokhanovsky, A., Schaaf, C., Bréon, F., Dong, Y., Wang, Z., Liu, Y., Zhang, X., Yin, S., Cui, L., Chang, Y., 2019. Development of a snow kernel to model the anisotropic reflectance of snow in a kernel-driven BRDF model framework. Rem. Sens. Environ. 221, 198–209.

Jin, Y., Schaaf, C.B., Gao, F., Li, X., Strahler, A.H., Lucht, W., Liang, S., 2003. Consistency of MODIS surface bidirectional reflectance distribution function and albedo retrievals: 1. Algorithm performance. J. Geophys. Res. 108 (D5). https://doi.org/10.1029/2002jd002803.

Kouki, K., Luojus, K., Riihelä, A., 2023. Evaluation of snow cover properties in ERA5 and ERA5-Land with several satellite-based datasets in the Northern Hemisphere in spring 1982–2018. Cryosphere Discuss. 2023, 1–33.

Liang, S., 2001. Narrowband to broadband conversions of land surface albedo I Algorithms. Rem. Sens. Environ. 76 (2), 213–238. https://doi.org/10.1016/S0034-4257(00)00205-4.

Liang, S., 2003. A direct algorithm for estimating land surface broadband albedo from MODIS imagery. IEEE Trans. Geosci. Rem. Sens. 41 (1), 136–145.

Liang, S., 2005. Mapping daily snow/ice shortwave broadband albedo from Moderate Resolution Imaging Spectroradiometer (MODIS): the improved direct retrieval algorithm and validation with Greenland in situ measurement. J. Geophys. Res. 110.

Liang, S., Cheng, J., Jia, K., Jiang, B., Liu, Q., Xiao, Z., Yao, Y., Yuan, W., Zhang, X., Zhao, X., Zhou, J., 2021. The global LAnd surface satellite (GLASS) product suite. Bull. Am. Meteorol. Soc. 102, E323–E337.

Liang, S., Zhao, X., Liu, S., Yuan, W., Cheng, X., Xiao, Z., Zhang, X., Liu, Q., Cheng, J., et al., 2013. A long-term Global LAnd Surface Satellite (GLASS) data-set for environmental studies. International Journal of Digital Earth 6, 5–33.

Lin, X., Wen, J., Liu, Q., Xiao, Q., You, D., Wu, S., Hao, D., Wu, X., 2018. A multi-scale validation strategy for albedo products over rugged terrain and preliminary application in Heihe river basin, China. Rem. Sens. 10, 156. https://doi.org/10.3390/rs10020156.

Lin, X., Wu, S., Chen, B., Lin, Z., Yan, Z., Chen, X., Yin, G., You, D., Wen, J., Liu, Q., Xiao, Q., Liu, Q., Lafortezza, R., 2022. Estimating 10-m land surface albedo from Sentinel-2 satellite observations using a direct estimation approach with Google Earth Engine. ISPRS J. Photogrammetry Remote Sens. 194, 1–20. https://doi.org/10.1016/j.isprsjprs.2022.09.016.

Liu, Q., Liang, S., Qu, Y., Liu, N., Liu, S., Tang, H., Liang, S., 2013. Preliminary evaluation of the long-term GLASS albedo product. International Journal of Digital Earth 6 (1), 69–95.

Liu, Y., Wang, Z., Sun, Q., Erb, A.M., Li, Z., Schaaf, C.B., Zhang, X., Román, M.O., Scott, R.L., Zhang, Q., Novick, K.A., Syndonia Bret-Harte, M., Petroy, S., SanClements, M., 2017. Evaluation of the VIIRS BRDF, Albedo and NBAR products suite and an assessment of continuity with the long term MODIS record. Rem. Sens. Environ. 201, 256–274.

Lu, Y., Wang, L., Hu, B., Zhang, M., Qin, W., Zhou, J., Tao, M., 2021. Evaluation of satellite land surface albedo products over China using ground-measurements. International Journal of Digital Earth 14 (11), 1493–1513.

Ma, Y., He, T., Liang, S., Wen, J., Gastellu-Etchegorry, J.P., Chen, J., Ding, A., Feng, S., 2022. Landsat snow-free land surface albedo estimation over sloping terrain: algorithm development and evaluation. IEEE Trans. Geosci. Rem. Sens. 60, 1–14. https://doi.org/10.1109/TGRS.2022.3149762.

Qu, Y., Liang, S., Liu, Q., He, T., Liu, S., Li, X., 2015. Mapping surface broadband albedo from satellite observations: a review of literatures on algorithms and products. Rem. Sens. 7 (1), 990–1020. https://doi.org/10.3390/rs70100990.

Qu, Y., Liang, S., Liu, Q., Li, X., Feng, Y., Liu, S., 2016. Estimating Arctic sea-ice shortwave albedo from MODIS data. Rem. Sens. Environ. 186, 32–46.

Qu, Y., Liu, Q., Liang, S., Wang, L., Liu, N., Liu, S., 2014. Direct-estimation algorithm for mapping daily land-surface broadband albedo from MODIS data. IEEE Trans. Geosci. Rem. Sens. 52, 907–919.

Román, M.O., Gatebe, C.K., Shuai, Y., Wang, Z., Gao, F., Masek, J.G., He, T., Liang, S., Schaaf, C.B., 2013. Use of in situ and airborne multiangle data to assess MODIS and landsat-based estimates of directional reflectance and albedo. IEEE Trans. Geosci. Rem. Sens. 51, 1393–1404.

Schaaf, C.B., Gao, F., Strahler, A.H., Lucht, W., Li, X., Tsang, T., et al., 2002. First operational BRDF, albedo nadir reflectance products from MODIS. Rem. Sens. Environ. 83, 135–148.

Schaaf, C.B., Liu, J., Gao, F., Strahler, A.H., 2011a. MODIS MCD43 albedo and reflectance anisotropy products from Aqua and Terra, in land remote sensing and global environmental change: NASA's Earth observing system and the science of ASTER and MODIS. Remote Sensing and Digital Image Processing Series 11, 873–894.

Schaaf, C.B., Wang, Z., Strahler, A.H., 2011b. Commentary on Wang and Zender-MODIS snow albedo bias at high solar zenith angles relative to theory and to in situ observations in Greenland. Rem. Sens. Environ. 115, 1296–1300.

Serreze, M.C., Barry, R.G., 2011. Processes and impacts of Arctic amplification: a research synthesis. Global Planet. Change 77, 85–96.

Shi, H., Xiao, Z., 2022. Exploring topographic effects on surface parameters over rugged terrains at various spatial scales. IEEE Trans. Geosci. Rem. Sens. 60, 1–16. https://doi.org/10.1109/TGRS.2021.3098607.

Stroeve, J., Box, J.E., Gao, F., Liang, S., Nolin, A., Schaaf, C., 2005. Accuracy assessment of the MODIS 16-day albedo product for snow: comparisons with Greenland in situ measurements. Rem. Sens. Environ. 94, 46–60.

Stroeve, J., Box, J.E., Wang, Z., Schaaf, C., Barrett, A., 2013. re-evaluation of MODIS MCD43 Greenland albedo accuracy and trends. Rem. Sens. Environ. 138, 199–214.

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.

Wang, D., Liang, S., He, T., Yu, Y., 2013. Direct estimation of land surface albedo from VIIRS data: algorithm improvement and preliminary validation. J. Geophys. Res. 118, 12577–12586.

Wang, D., Liang, S., Zhou, Y., He, T., Yu, Y., 2017. A new method for retrieving daily land surface albedo from VIIRS data. IEEE Trans. Geosci. Rem. Sens. 55 (3), 1–11.

Wang, Z., Schaaf, C.B., Chopping, M.J., Strahler, A.H., Wang, J., Román, M.O., Rocha, A. V., Woodcock, C.E., Shuai, Y., 2012. Evaluation of Moderate-Resolution Imaging Spectroradiometer (MODIS) snow albedo product (MCD43A) over tundra. Rem. Sens. Environ. 117, 264–280.

Wang, Z., Schaaf, C.B., Strahler, A.H., Chopping, M.J., Román, M.O., Shuai, Y., Woodcock, C.E., Hollinger, D.Y., Fitzjarrald, D.R., 2014. Evaluation of MODIS MCD43 albedo product (MCD43A) over grassland, agriculture and forest surface types during dormant and snow-covered periods. Rem. Sens. Environ. 140, 60–77.

Wang, Z., Schaaf, C.B., Sun, Q., Shuai, Y., Román, M.O., 2018. Capturing rapid land surface dynamics with Collection V006 MODIS BRDF/NBAR/Albedo (MCD43) products. Rem. Sens. Environ. 207, 50–64.

Wen, J., Wu, X., You, D., Ma, X., Ma, D., Wang, J., Xiao, Q., 2023. The main inherent uncertainty sources in trend estimation based on satellite remote sensing data. Theor. Appl. Climatol. 151 (1), 915–934.

Wright, P., Bergin, M., Dibb, J., Lefer, B., Domine, F., Carman, T., Carmagnola, C., Dumont, M., Courville, Z., Schaaf, C., Wang, Z., 2014. Comparing MODIS daily snow albedo to spectral albedo field measurements in central Greenland. Rem. Sens. Environ. 16, 125–142.

Wu, X., Wen, J., Tang, R., Wang, J., Zeng, Q., Li, Z., Xiao, Q., 2023. Quantification of the uncertainty in multiscale validation of coarse-resolution satellite albedo products: a study based on airborne CASI data. Rem. Sens. Environ. 287, 113465.

Wu, X., Wen, J., Xiao, Q., Liu, Q., Peng, J., Dou, B., Liu, Q., 2016. Coarse scale in situ albedo observations over heterogeneous snow-free land surfaces and validation strategy: a case of MODIS albedo products preliminary validation over northern China. Rem. Sens. Environ. 184, 25–39.

Wu, X., Wen, J., Xiao, Q., You, D., Dou, B., Lin, X., Hueni, A., 2018. Accuracy assessment on MODIS (V006), GLASS and MuSyQ land-surface albedo products: a case study in the Heihe river basin, China. Rem. Sens. 10, 2045. https://doi.org/10.3390/rs10122045.

Wu, X., Wen, J., Xiao, Q., You, D., Lin, X., Wu, S., Zhong, S., 2019. Impacts and contributors of representativeness errors of in situ albedo measurements for the validation of remote sensing products. IEEE Trans. Geosci. Rem. Sens. 57 (12), 9740–9755.

Zhang, H., Zhang, F., Zhang, G., Che, T., Yan, W., Ye, M., Ma, N., 2019. Ground-based evaluation of MODIS snow cover product V6 across China: implications for the selection of NDSI threshold. Sci. Total Environ. 651, 2712–2726.

Zhang, X., Jiao, Z., Dong, Y., He, T., Ding, A., Yin, S., Zhang, H., Cui, L., Chang, Y., Guo, J., Xie, R., 2020. Development of the direct-estimation albedo algorithm for snow-free Landsat TM albedo retrievals using field flux measurements. IEEE Trans. Geosci. Rem. Sens. 58, 1550–1567. https://doi.org/10.1109/TGRS.2019.2946598.

Zhang, X., Jiao, Z., Zhao, C., Qu, Y., Liu, Q., Zhang, H., Tong, Y., Wang, C., Li, S., Guo, J., Zhu, Z., Yin, S., Cui, L., 2022. Review of land surface albedo: variance characteristics, climate effect and management strategy. Rem. Sens. 14 (6), 1382. https://doi.org/10.3390/rs14061382.

Zhou, Y., Wang, D., Liang, S., Yu, Y., He, T., 2016. Assessment of the Suomi NPP VIIRS land surface albedo data using station measurements and high-resolution albedo maps. Rem. Sens. 8, 137. https://doi.org/10.3390/rs8020137.