Science of the Total Environment 724 (2020) 138156

Contents lists available at ScienceDirect

Science of the Total Environment

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

Comparative evaluation of VIIRS daily snow cover product with MODIS for snow detection in China based on ground observations

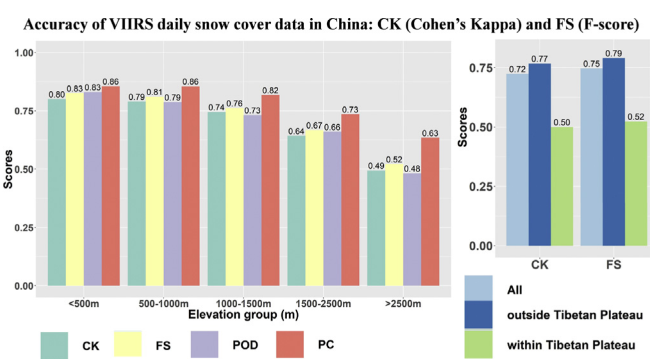
Hongbo Zhang [a,b,*], Fan Zhang [c,d,e], Tao Che [d,f], Shijin Wang [b]

[a] College of Water Resources & Civil Engineering, China Agricultural University, Beijing, China
[b] State Key Laboratory of Cryospheric Science, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou, China
[c] Key Laboratory of Tibetan Environment Changes and Land Surface Processes, Institute of the Tibetan Plateau Research, Chinese Academy of Sciences (CAS), Beijing, China
[d] CAS Center for Excellence in the Tibetan Plateau Earth Sciences, Beijing, China
[e] University of Chinese Academy of Sciences, Beijing, China
[f] Key Laboratory of Remote Sensing of Gansu Province, Heihe Remote Sensing Experimental Research Station, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China

HIGHLIGHTS

- 0.14 is the optimal NDSI threshold for applying VIIRS snow cover data in China.
- VIIRS shows much better accuracy than MODIS AQUA in China.
- VIIRS shows lower accuracy than MODIS TERRA in stations within the Tibetan Plateau.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 24 December 2019
Received in revised form 4 March 2020
Accepted 22 March 2020
Available online 23 March 2020

Editor: Ouyang Wei

Keywords:
VIIRS
MODIS
Daily snow cover
Tibetan Plateau
China

ABSTRACT

Accurate spatiotemporal information of snow cover not only is important for investigating the mechanisms of climate change but also greatly contributes to hydrological modelling in mountainous regions. The Suomi-National Polar-orbiting Partnership (S-NPP) Visible Infrared Imaging Radiometer Suite (VIIRS) (referred to as VNP) daily snow cover product is recently released and expected to take place of Moderate Resolution Imaging Spectroradiometer (MODIS) snow cover products in near future. As an important addition to the widely used MODIS products, there is also an urgent need for a reliable accuracy evaluation and comparison of VNP for future large-scale daily snow cover mapping. This study for the first time evaluates the accuracy of VNP daily snow cover data in China using daily snow depth observations from 330 stations. The accuracy of VNP data is generally good with the averaged CK (Cohen's Kappa) and FS (F-Score) as high as 0.72 and 0.75, respectively, but considerably decreases to 0.50 and 0.52 for the Tibetan Plateau. VNP shows slightly better accuracy than MODIS TERRA for stations outside the Tibetan Plateau owing to its higher spatial resolution, but its accuracy is lower than TERRA for those within the Tibetan Plateau possibly due to its longer time interval between ground observation and satellite overpass time. By contrast, VNP shows much better accuracy than MODIS AQUA in China including both outside

* Corresponding author at: China Agricultural University, College of Water Resources & Civil Engineering, 100083 Beijing, China.
E-mail address: zhanghongbo@cau.edu.cn (H. Zhang).

https://doi.org/10.1016/j.scitotenv.2020.138156

and within the Tibetan Plateau. This study provides important implications for optimal use of VNP and MODIS daily snow cover products in China, which may further contribute to more accurate snow variation information for climate analysis and cryospheric hydrological modelling.

1. Introduction

Accurate spatial and temporal snow information is crucial for investigating Earth's system changes because snow is considered to have substantial impacts on climate change due to its high albedo. Many studies have identified the snow-albedo feedback as an important mechanism for explaining the phenomenon of elevation-dependent-warming in high-elevation areas (Guo et al., 2016; Pepin et al., 2015). As an important component of water resources in mountainous basins, snow cover also provides crucial calibration and validation information for cryospheric hydrological modelling (Immerzeel et al., 2009; Zhang et al., 2015). However, the limited number of stations in such mountainous or remote areas poses challenges for fully investigating the snow variation on regional or global scales.

Remote sensing snow cover (Grody and Basist, 1996; Hall et al., 2002), snow water equivalent (i.e., the amount of water that can be released if the snowpack totally melts (Singh et al., 2011)) or snow depth products (Bo and Feng, 2000; Che et al., 2008; Kelly et al., 2003) are thus developed and widely used for snow related researches including hydrological simulation (calibration and validation) (Thirel et al., 2013; Zhang et al., 2015) and climate analysis (Liu and Chen, 2011; Yeo et al., 2017) in recent decades. Among them, the MODIS daily snow cover product may be the most popularly used due to its relatively high spatial (~500 m) and temporal (1 day) resolutions. Numerous studies have evaluated their accuracy around the world (Arsenault et al., 2014; Huang et al., 2011; Parajka and Blöschl, 2006; Parajka and Blöschl, 2012) and their new version's (i.e. version 6) accuracy (F-score) are reported to be 0.81 and 0.63 for TERRA and AQUA in China, respectively (Zhang et al., 2019). Recently, a new daily snow cover product — the NASA Suomi-National Polar-orbiting Partnership (S-NPP) Visible Infrared Imaging Radiometer Suite (VIIRS) snow cover data product (referred to as "VNP" here) — is released, and is expected to take place of MODIS data series (already beyond the design lifetime) for continually providing daily snow observations in future, though MODIS is still working well (Riggs et al., 2017).

Careful validation is very meaningful for optimal use of remote sensing snow cover data (Hall and Riggs, 2007), there is thus an urgent need to investigate the accuracy information about the new VNP product. To our knowledge, very few studies have evaluated the accuracy of the new snow cover product (Key et al., 2013; Thapa et al., 2019), especially for China. The three daily snow cover products (i.e. VNP, MODIS TERRA and AQUA) coexisting after 2012 poses an important question of how to optimally use them for future snow mapping in China. It may thus be very useful and interesting to find out whether there are significant differences in the accuracy between MODIS and VNP: is VNP better than MODIS TERRA and AQUA because of its higher spatial resolution (~375 m versus ~500 m)? Since many efforts have been made to combine snow cover data from various satellite products such as MODIS TERRA, MODIS AQUA and IMS (Interactive Multisensor Snow and Ice Mapping System) for overcoming the serious problem of cloud blockage (Huang et al., 2017; Yu et al., 2016), the emergence of VNP may provide a new option. In addition, Zhang et al. (2019) find that MODIS TERRA and AQUA snow cover products show relatively poor accuracy in high-elevation areas, especially for the Tibetan Plateau, thus it is also very interesting to investigate whether VNP can help improve the snow mapping accuracy in those areas.

Similar to MODIS products, the direct binary data of snow or no-snow classification are no longer provided and only the NDSI is available in the new VNP product. It means that the "end-users" themselves have to determine the NDSI threshold above which the corresponding pixel should be identified as snow covered (Riggs et al., 2017). It should be noted that MODIS and VIIRS have different band locations, and they use very different wavelength bands for calculating NDSI (Justice et al., 2013). Thus, VNP may have different NDSI thresholds though our previous study shows that 0.1 could be the optimal NDSI threshold for applying MODIS snow cover products in China (Zhang et al., 2019).

As an important extension of our previous study of assessing the performance of an updated version of MODIS snow cover products in China (Zhang et al., 2019), this study uses more recent snow depth observations to evaluate the new VNP snow cover product and investigate their accuracy differences in China. The main focuses of this study include: (1) evaluating the accuracy of VNP daily snow cover product in China using ground observations, especially for the Tibetan Plateau; (2) finding out whether VNP is better than MODIS (including TERRA and AQUA) snow cover data. To achieve these goals, daily snow depth data of 2012–2018 from 330 stations across China are used for providing ground truth and the optimal NDSI threshold is selected for VNP data based on a sensitivity test. The accuracy of VNP data is also compared with MODIS TERRA and AQUA daily snow cover products.

2. Data

2.1. Ground measurements

Daily snow depth data during 2012.1–2018.12 from >800 China Meteorology Administration (CMA) stations across China mainland were collected at the initial stage. The snow depth records are manually measured using a ruler at 8 AM and rounded to the nearest whole centimeter (CMA, 2017). Those with snow depth <0.5 cm are recorded as thin or trace snow. Since thin snow may bring much uncertainty (Ke et al., 2016), such observations are not considered here. A filter process was then implemented to remove the "invalid" stations with greater-than-or-equal-to-1 cm snow depth observations (while the corresponding pixel is not cloud-covered) <20 days according to suggestions from Metsämäki (2016). The remaining 330 stations are actually used for this study (Fig. 1). Daily precipitation, air temperature, relative humidity and wind speed observations are also obtained from the same stations for analyzing factors affecting the optimal NDSI thresholds or validation accuracy. It should be noted that all the meteorological variables are aggregated from daily data to multiyear average values during the same period for each station (i.e. one station corresponds to one value). Elevation, longitude and latitude information are also acquired from CMA.

2.2. VIIRS snow cover data

The VIIRS/NPP Snow Cover Daily L3 Global 375 m SIN Grid V001 dataset (here referred to as "VNP") is evaluated in this study. VIIRS is onboard the Suomi National Polar-orbiting Partnership satellite which was launched in Oct. 28, 2011. In the snow cover product, the NDSI is used for snow detection and calculated following Eq. (1).

$$NDSI = (B1 - B3)/(B1 + B3) \quad (1)$$

where, B1 is the VIIRS red light band I1 (wavelength: 0.640 μm), B3 is the VIIRS shortwave near-infrared band I3 (wavelength: 1.61 μm) (Riggs et al., 2016b). The valid range of NDSI is 0–1, but it is actually

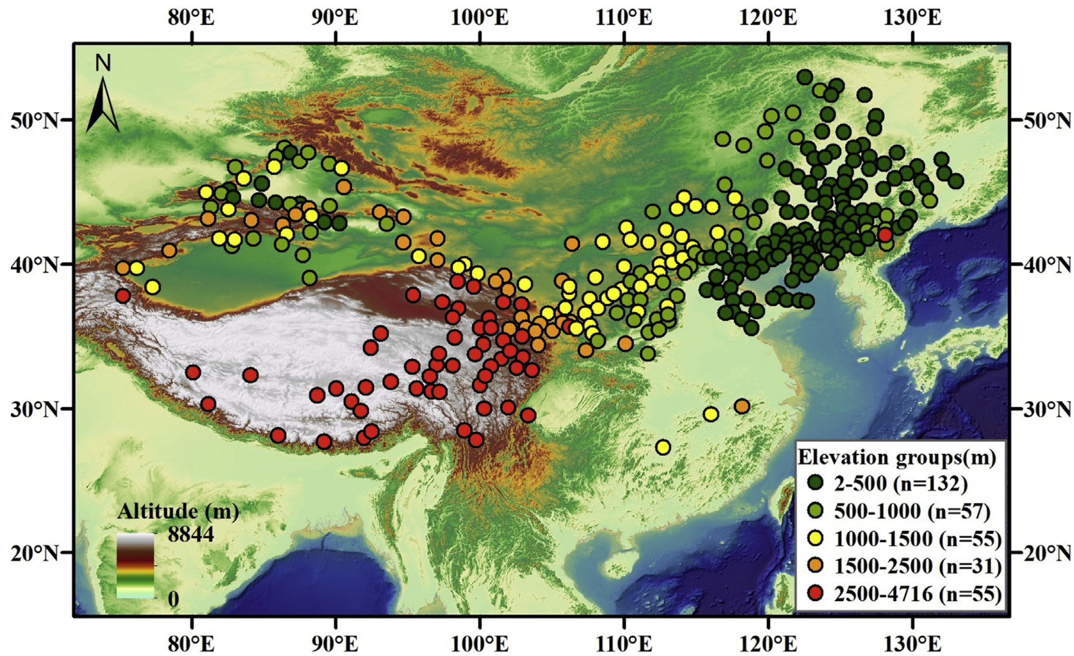


Fig. 1. Distributions and elevations of the 330 stations across China mainland. The layer of elevations is derived from Shuttle Radar Topography Mission (SRTM) dataset. The number of stations corresponding to each elevation group is also shown in the legend.

0.1–1 for VNP because the NDSI values <0.1 are set to 0 for reducing the uncertainty about snow detection (Riggs et al., 2016b). In addition, some other screening methods such as reversing snow detection results for low-elevation (< 1300 m) pixels with large brightness temperature (≥ 281 K) and for low visible reflectance conditions are also implemented. The detailed descriptions about how VNP snow cover data are developed can be found in Riggs et al. (2016b). A threshold of NDSI is then needed for distinguishing snow from snow-free land and the NDSI threshold of 0.4 was commonly used worldwide before. However, as mentioned in Introduction, the "0.4" is no longer used in the new snow cover product and the NDSI threshold should be determined according to the study area (Riggs et al., 2017). Since VNP dataset is only available from 1/19/2012, the overlapping period of 1/19/2012–12/31/2018 is selected as the study period.

2.3. MODIS data

Both MODIS TERRA (MOD10A1) and MODIS AQUA (MYD10A1) daily snow cover products version 6 at about 500 m are used here for comparison with VNP data. MODIS snow cover data employs a very similar algorithm to VNP for snow detection which is mainly based on Eq. (2).

$$NDSI = (B4 - B6)/(B4 + B6) \quad (2)$$

where, B4 is MODIS visible light band 4 (wavelength: 0.555 μm), B6 is MODIS shortwave infrared band 6 (wavelength: 1.640 μm) (Riggs et al., 2016a). Due to the non-functional band 6 of AQUA, B6 has been restored using the Quantitative Image Restoration (QIR) algorithm (Gladkova et al., 2012) for AQUA NDSI calculation. In the daytime, the overpass times of MODIS TERRA and AQUA are 10:30 AM and 1:30 PM, respectively. It should be noted that the overpass time of VIIRS is almost the same as MODIS AQUA (Justice et al., 2013). Similar to VNP data, the NDSI threshold of MODIS snow cover data has to be determined by users. In addition to the snow cover data, the "Terra and Aqua combined Moderate Resolution Imaging Spectroradiometer (MODIS) Land Cover Type (MCD12Q1) Version 6" dataset (Sulla-Menashe and Friedl, 2018) is also used for extracting the land cover types: for each station, the annual land cover types (following the International Geosphere-Biosphere Programme classification schemes) during 2012–2018 are extracted and the most frequent type of the seven years is selected.

3. Methods

The method for determining the optimal NDSI threshold of VNP snow cover data and how its accuracy is evaluated in China is briefly described in Fig. 2. Based on the daily snow depth observations from 330 stations, the optimal NDSI threshold is firstly determined for VNP snow cover data through a detailed sensitivity test. The accuracy of VNP data is then evaluated based on the optimal NDSI threshold and by comparison with MODIS TERRA and AQUA snow cover data which are also with their optimal NDSI threshold determined from similar sensitivity tests. The geographic and meteorological variables as well as land cover types are finally used for analyzing the factors and uncertainties of the accuracy of VNP snow cover data. The detailed processes are described in following sections.

3.1. Accuracy metrics

Station observed daily snow depth data are used here as the ground truth for evaluating VNP and MODIS snow cover data. Station based snow depth observations are generally considered as reliable validation data for evaluating remotely sensed snow cover products including MODIS TERRA and AQUA (Huang et al., 2011; Parajka and Blöschl, 2006; Zhang et al., 2019), though there are also some shortcomings including mismatched observation time and scale issue which will be discussed later. The snow depth observations from stations are directly compared with the snow cover data of the VNP or MODIS pixels where the stations are located.

A commonly used confusion matrix is adopted here for comparing VNP/MODIS snow cover data with ground observations: 1) snow depth observation ≥1 cm and the VNP/MODIS pixel is snow covered, defined as "*TP*" (true positive); 2) snow depth observation ≥1 cm while the VNP/MODIS pixel is snow-free, defined as "*FN*" (false negative); 3) snow depth observation = 0 cm while the VNP/MODIS pixel is

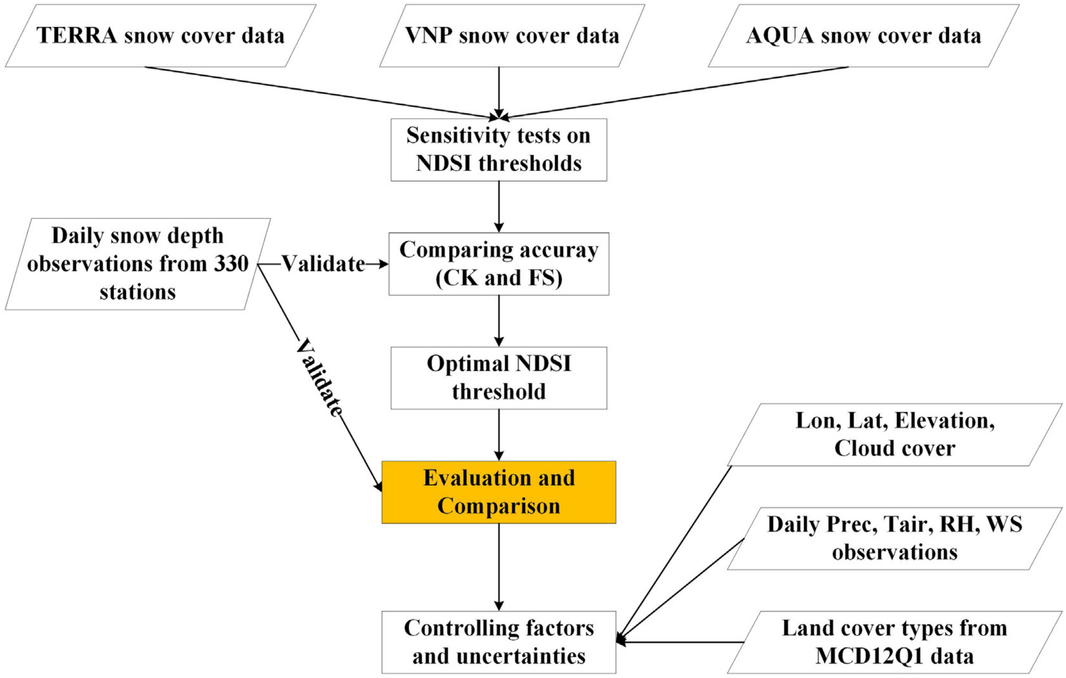


Fig. 2. Flow chart describing how the comparative evaluation of VNP snow cover data is conducted. Prec: precipitation; Tair: air temperature; RH: relative humidity; WS: wind speed.

snow covered, defined as "*FP*" (false positive); 4) snow depth observation = 0 cm and the VNP/MODIS pixel is snow-free, defined as "*TN*" (true negative). Four kinds of accuracy metrics including POD (probability of detection), PC (Precision), FS (F-score) and CK (Cohen's Kappa) are then calculated based on the confusion matrix with their detailed definitions listed in Table 1. A comprehensive comparison of them shows that CK and FS are more reliable and useful than some other metrics in evaluating MODIS snow cover data such as OA (overall accuracy) and FAR (false alarming rate) (Zhang et al., 2019). In addition, either POD or PC only reflects the snow omission (the product predicts no snow whereas snow cover is actually observed) or commission (the product predicts snow-covered whereas actually non-snow) errors, respectively (Zhou et al., 2013). In fact, POD plus omission errors is equal to 1, and PC plus commission errors is equal to 1. Thus, CK and FS are used as the main evaluating metrics, and POD and PC are also calculated for better analyzing the error sources.

Table 1
Definitions of a confusion matrix and 4 kinds of accuracy metrics for VNP/MODIS snow cover estimates versus ground snow depth observations.

		VNP/MODIS	
		Snow	Non-snow
Station-observed	Snow depth ≥ 1 cm	*TP*	*FN*
	Snow depth = 0 cm	*FP*	*TN*

Metric	Definition
POD	$\frac{TP}{TP+FN}$
PC	$\frac{TP}{TP+FP}$
FS	$\frac{2\times POD\times PC}{POD+PC}$
CK	$\frac{OA-Pr(e)}{1-Pr(e)}$ Where, $Pr(e)=(\frac{TP+FP}{Total}\times\frac{TP+FN}{Total})+(\frac{TN+FP}{Total}\times\frac{TN+FN}{Total})$ $Total = TP + FN + FP + TN$

Note: *TP*, *FP*, *FN* and *TN* are all numbers. *TP* is "true positive", *FP* is "false positive", *FN* is "false negative" and *TN* is "true negative".

3.2. Finding the optimal NDSI threshold for VNP and MODIS snow cover data

Though 0.4 has been widely used as the NDSI threshold in various remotely sensed snow cover products for the recent 20 years, some have found that 0.4 may not be the optimal one in all regions (Dong et al., 2014; Riggs et al., 2017), especially in China where obviously lower thresholds should be used (Dong et al., 2014; Hao et al., 2008; Wang et al., 2012; Zhang et al., 2019). Based on daily snow depth observations from 330 stations, the optimal NDSI threshold for VNP data is determined based on a sensitivity test following Zhang et al. (2019): varying the NDSI threshold from 0.1 to 0.9 with a step of 0.01 resulting to 81 iterations; for each iteration, the six accuracy metrics including CK and FS are calculated for all the 330 stations and their averaged scores are recorded; the optimal NDSI threshold is finally determined as the one with the highest averaged CK and FS. The same procedures are applied to MODIS TERRA and MODIS AQUA snow cover products.

3.3. Evaluating VNP snow cover data and its comparison with MODIS snow cover products

After the optimal NDSI threshold is determined for VNP, MODIS TERRA and MODIS AQUA, their averaged CKs based on the optimal NDSI thresholds are considered as the final performance measurements. Considering the large topographic gradients of China, the accuracy of VNP snow cover data are further revealed by dividing all the 330 stations into 5 elevation groups including elevations <500, 500–1000, 1000–1500, 1500–2500 and >2500 m. The number of stations corresponding to each elevation group and their locations are shown in Fig. 1. The spatial distributions of accuracy for VNP snow cover data are thus analyzed by investigating the effects of multiple geographic and meteorological variables including latitude, longitude, elevation, snow depth, air temperature and cloud cover. The cloud cover is calculated by counting the number of cloudy dates according to the VNP snow cover product and taking the percentage of cloudy days as proxy. To investigate the influences of different variables on the accuracy of VNP snow cover data, both the correlation and linear regression analysis are conducted. The linear regression is a popular method of predicting a dependent variable by incorporating one or multiple independent variables (or factors) (Luo et al., 2018b; Stahl et al., 2006; Zhang et al., 2016) and the calculated R-squared (defined as the proportions of variances explained by the independent variables in a regression model) can be used for measuring the relationship between the dependent and the independent variables (Cristobal et al., 2008; Luo et al., 2018a; Lyu et al., 2019), though some assumptions of the linear regression (e.g. the multivariate normality) may not always be met and could thus introduce some uncertainties. Uncertainties about land cover types and snow depth thresholds are also investigated.

It should be noted that the numbers of valid stations after initial filtering are different for the three snow cover products due to different cloudy dates, and they are 330, 263 and 252 for VNP, TERRA and AQUA, respectively. For a fair comparison, only the overlapping 245 stations are used for comparing the accuracy between them. Then, a multiple comparison based on the 245 stations is conducted using paired *t*-test with Holm correction (Holm, 1979) using CK as the decisive index according to Zhang et al. (2019) to detect whether there are significant accuracy differences between the three snow cover products.

4. Results and discussions

4.1. Selection of NDSI threshold for VNP snow cover product and its spatial uncertainty

Fig. 3 shows how the performances of different snow cover products react to different thresholds of NDSI. It is clear that the optimal NDSI thresholds for both MODIS TERRA and MODIS AQUA are 0.1 (i.e. the minimum NDSI threshold). However, it should be noted that this finding is not new (Zhang et al., 2019), and it is reconfirmed here based on more recent daily snow depth observations. The optimal NDSI threshold for VNP is 0.14, which is a little higher than that for MODIS. Compared with the globally used 0.4, using 0.14 as the NDSI threshold greatly increases CK and FS from 0.54 and 0.57 to 0.72 and 0.75, respectively, strongly suggesting that the NDSI threshold of 0.4 could not be reasonable for VNP snow cover data in China.

This is further verified by the station-based optimal NDSI thresholds at 330 stations as plotted in Fig. 4a with the averaged threshold of ~0.16, which is close to 0.14. About 61% of the optimal NDSI thresholds are <0.15, those >0.2 account for only about 21% and only ~2% of them are ≥ 0.4. Though not all the optimal NDSI thresholds are exactly 0.14, the overall performance of using 0.14 as the NDSI threshold is demonstrated to be the best in China as shown in Fig. 3. There seems to be no obvious relationship between the station-based optimal NDSI thresholds and their corresponding CKs as shown in the supporting information (Fig. S1a), and most points are clustered at areas with relatively low NDSI thresholds and high CKs. A multiple comparison based on the non-parametric Kruskal-Wallis test (Siege and Castellan, 1988) is also conducted between five elevation groups whereas showing no significant differences (Fig. S1b). In addition, the correlation analysis shows that very weak or non-significant correlation is found between the optimal NDSI thresholds and the geographic and meteorological variables including elevation, longitude, latitude, snow depth, precipitation, air temperature, wind speed and relative humidity with the highest correlation coefficient only as 0.28 indicating that the optimal NDSI threshold for VNP in a specific area is very hard to predict. Thus, the NDSI threshold of 0.14 is suggested when applying VNP snow cover data for snow detection in China.

4.2. The accuracy of VNP snow cover data in China including the Tibetan Plateau and the possible factors

Generally, VNP shows a good accuracy in China with the averaged CK and FS as high as 0.72 and 0.75 (Fig. 5), respectively, considering that the CK values of between 0.61 and 0.8 and between 0.81 and 1 can be interpreted as "substantial" and "almost perfect" by Landis and Koch (1977), respectively. Fig. 4b plots the distribution of CK across China and many of the stations with low accuracy are located within the Tibetan Plateau. It is further verified in Fig. 5a that the averaged CK and FS of stations within the Tibetan Plateau are only 0.50 and 0.52, respectively. The results are basically consistent with other studies evaluating MODIS snow cover data in the Tibetan Plateau that relatively low accuracy is reported (Pu et al., 2007) (e.g. a low CK of 0.61 (0.35) for MODIS TERRA (AQUA) snow cover data by Zhang et al. (2019)).

One of the major factors affecting the accuracy of VNP snow cover data could be elevation as shown by Fig. 5b. The accuracy comparison between five elevation groups shows that the average CK (FS) decreases significantly from 0.8 (0.83) to 0.49 (0.52) as elevation increases. The

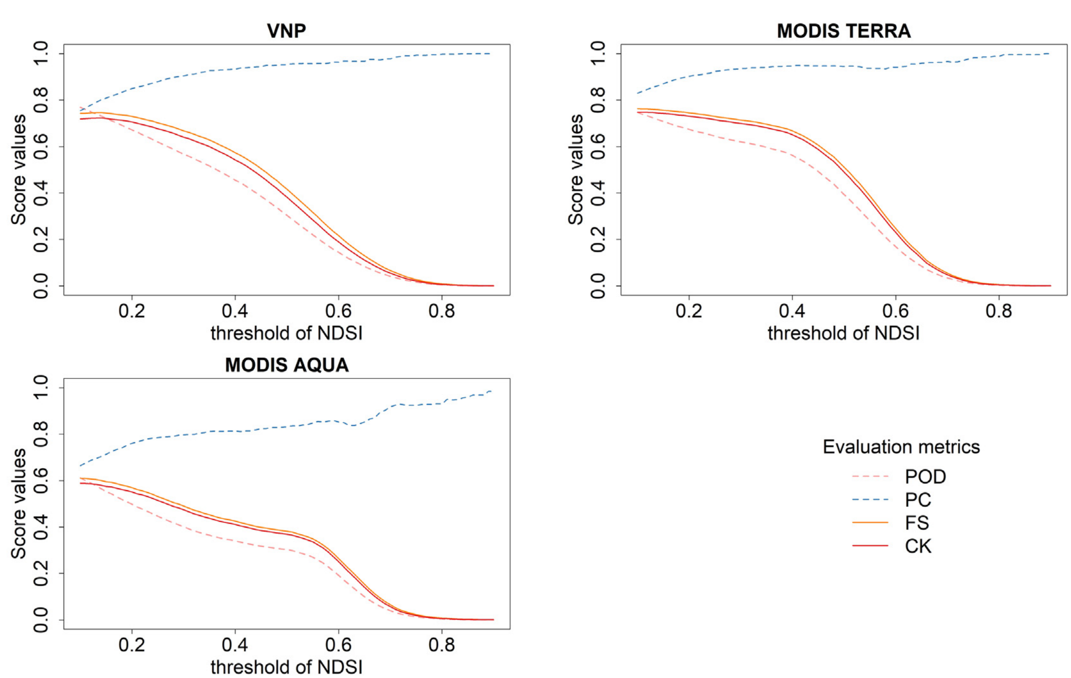


Fig. 3. Sensitivity tests of NDSI threshold for VIIRS, MODIS TERRA and MODIS AQUA snow cover data. The values shown here are the average of a given metric from stations.

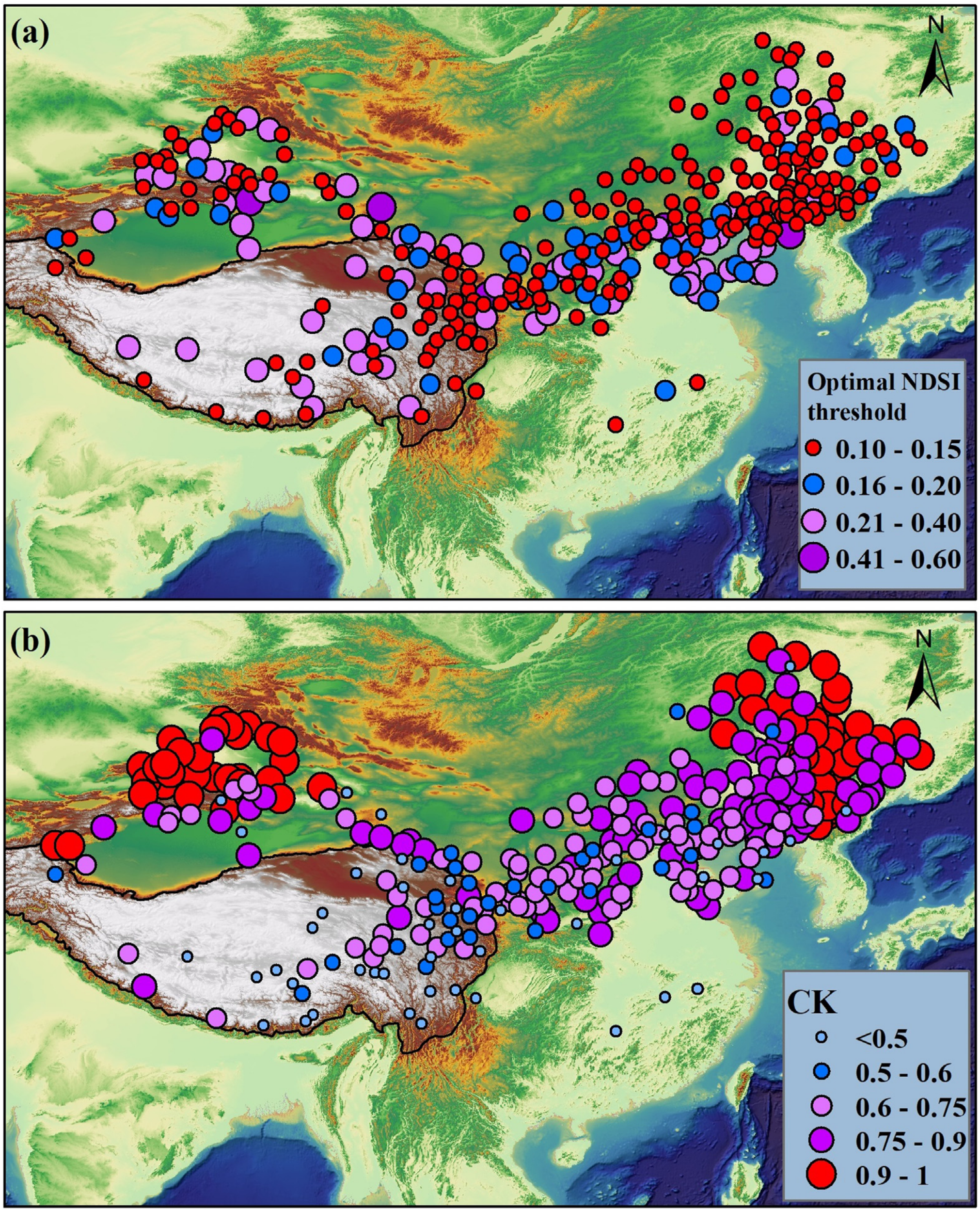


Fig. 4. Distribution of station-based optimal NDSI thresholds (a) and validation accuracy (CK) (b) for VNP snow cover data. The black lines outline the extent of the Tibetan Plateau. The background layer of elevations is derived from Shuttle Radar Topography Mission (SRTM) dataset.

problem of pixel heterogeneity can be amplificated in high-elevation areas usually with complex terrains (e.g. the Tibetan Plateau) and leads to lower accuracy of snow cover estimates. Previous studies show that air temperature (Dong et al., 2014), clouds (Arsenault et al., 2014) and snow depth (Yang et al., 2015) may also affect the snow cover detection errors. There is also a clear pattern that the accuracy is higher in the north whereas lower in the south (Fig. 4b). Thus, in order to further explore the factors affecting the spatial distribution of VNP's accuracy, the Pearson correlation analysis is conducted between the station-based CK and 6 kinds of geographic and meteorological variables mentioned in Section 3.3. The results show that in addition to elevation (correlation coefficient: −0.56), snow depth seems to be another major factor with a strong positive (0.54) correlation with the CK of VNP snow cover data (Fig. 6a). Small snow depth may not be able to cover the whole pixel or produce a big enough NDSI value for detection which is also apt to snow omission errors. Though latitude shows the highest correlation as 0.67, we consider it reflects the comprehensive effects of multiple variables including elevation, snow depth, temperature and cloud cover, because latitude and CK show very similar correlations with these variables (Fig. 6a). This hypothesis is further verified by the linear regression analysis showing that the model based on the four variables (i.e. Model B in Fig. 6b) can explain 49% of the total variances of CK which is very close to that of the one based on only latitude (45%). In addition, elevation and snow depth (Model C) can explain 48% of total variances which is only 6% less than Model E that considers all the six variables. Model E is also demonstrated to be the best model with the highest R-squared based on the comparison between the 63 models considering all the possible combinations of the six independent variables (i.e., 2^6 - 1). By contrast, temperature and cloud cover (Model D) can only explain 5%. In conclusion, elevation and snow depth are considered as two major factors controlling the accuracy of VNP snow cover data in China.

The reason why VNP shows obviously lower accuracy in the Tibetan Plateau can thus be attributed to its very high elevation (mean:

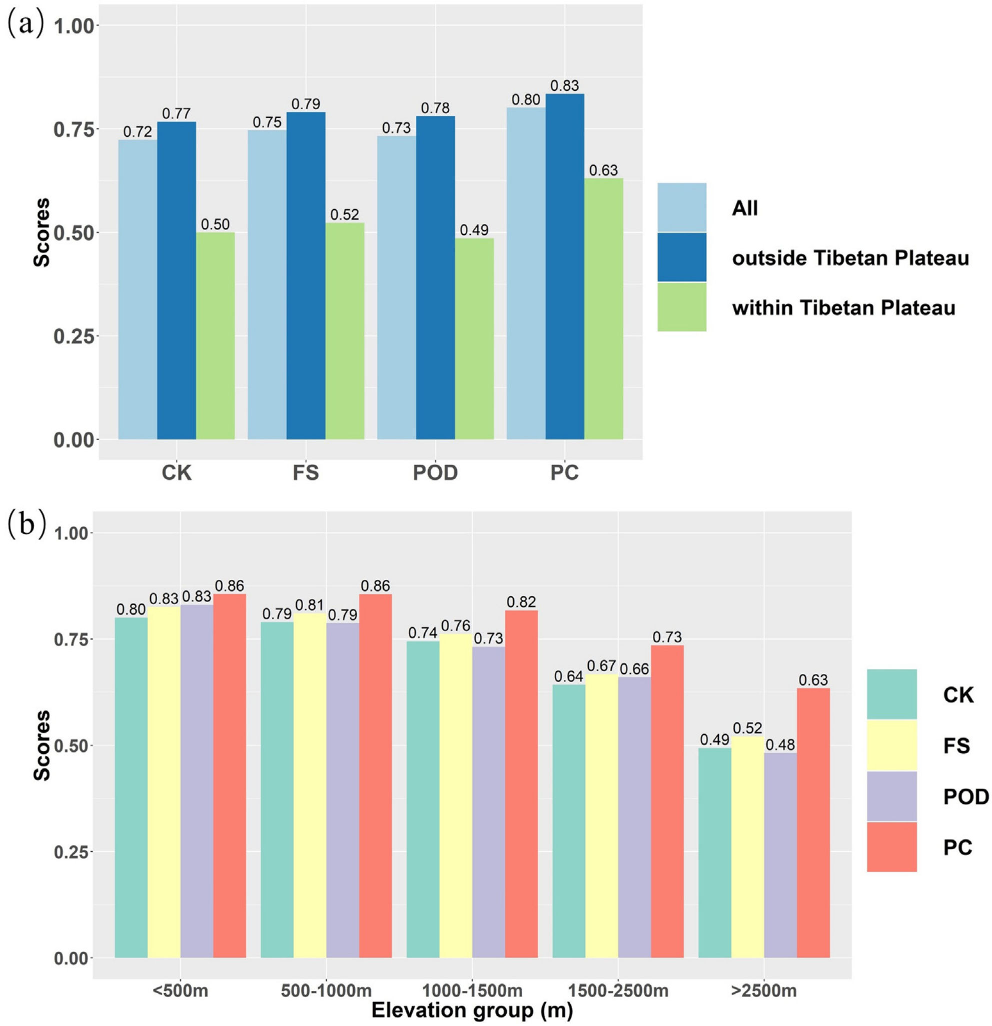


Fig. 5. Overall accuracy of VNP snow cover data in China (including outside and inside the Tibetan Plateau) (a) and the accuracy comparison between five elevation groups (b).

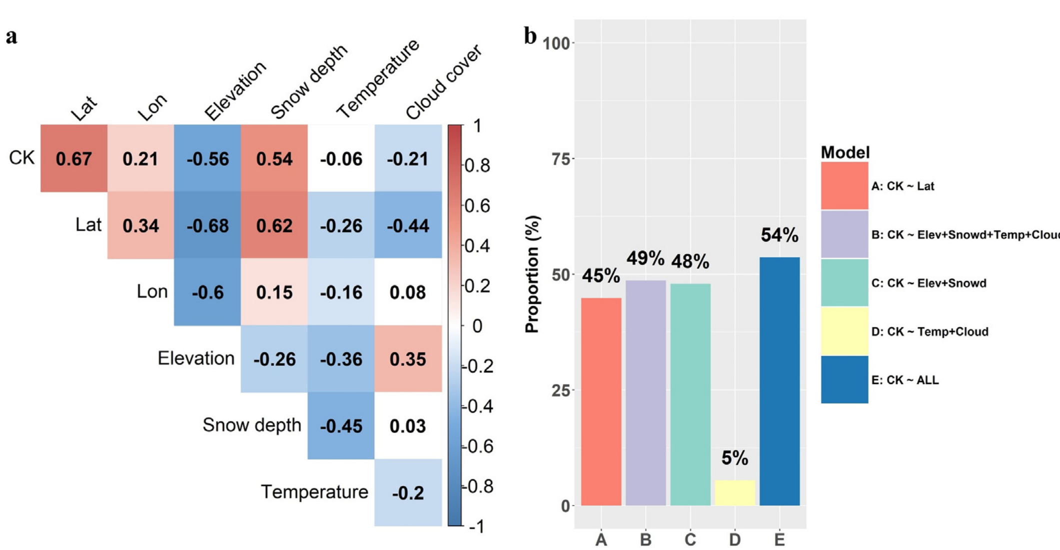


Fig. 6. The possible factors affecting the accuracy of VNP snow cover product. The left panel shows the correlation matrix of validation accuracy (CK) and six geographic and meteorological variables (a). The right panel displays the proportions of variances of validation accuracy explained by models considering (1) only latitude (Model A), (2) elevation, snow depth and two other variables (Model B), (3) only elevation and snow depth (Model C), (4) temperature and cloud cover (Model D), and (5) all the six variables (Model E). Lat: latitude; Lon: longitude; Elev: elevation; Snowd: snow depth; Temp: temperature; Cloud: cloud cover. For the correlation matrix, significantly (at 0.01 significance level) positive correlation values are shown in red; significantly (at 0.01 significance level) negative correlation values are shown in blue; insignificant correlation values are filled as blank.

~3653 m) and much smaller snow depth (mean: ~0.3 cm), considering that the averaged elevation and snow depth of stations outside the Tibetan Plateau are ~718 m and ~1.9 cm, respectively. It is clear that the much larger omission errors with the very low POD of 0.49 are the primary causes of the low CK and FS in the Tibetan Plateau, given that PC is relatively high (0.63) (Fig. 5). This indicates that many of actual snow-covered events in the Tibetan Plateau are not detected by the VNP daily snow cover product though a very low NDSI threshold (i.e. 0.14) has already been used. The scale issue enhanced by the high elevation and the problem about the time interval between ground observation and the satellite overpass time worsened by relatively short snow depth (snow melts or sublimates) may jointly lead to such large omission errors.

4.3. Comparison of the performance between VNP and MODIS (TERRA and AQUA) in China

In general, the accuracy of VNP (mean CK: 0.74) is as good as that of MODIS TERRA (mean CK: 0.75) snow cover product without significant difference, whereas is obviously much better than that of MODIS AQUA (mean CK: 0.59) product based on the multiple comparison results (Fig. 7a). However, it may be surprising that the accuracy of VNP (mean CK: 0.51) is significantly lower than that of MODIS TERRA (mean CK: 0.60) in the Tibetan Plateau (Fig. 7b), though VNP has a higher spatial resolution. The comparison results based on stations outside the Tibetan Plateau (Fig. 7c) further indicate that VNP is in trouble mainly in the Tibetan Plateau. More detailed information about the comparison are shown in Table 2. It should be noted that the accuracy of VNP snow cover data shown here is a little different from that reported in Section 4.2 because 85 stations, which are only valid for VNP but not valid for TERRA or AQUA, are removed for a fair comparison. For stations outside the Tibetan Plateau, VNP actually shows slightly higher accuracy (CK) than MODIS TERRA and both its snow omission and commission errors are lower, which is owing to its higher resolution. However, when it goes to the Tibetan Plateau, both snow omission and commission errors from VNP are much larger than TERRA. Considering that both VNP and TERRA are using a similar snow detection algorithm and the similar NDSI thresholds, we thus attribute this "abnormal" phenomenon in the Tibetan Plateau to the bigger mismatches of observation time and satellite overpass time of VNP. TERRA observes in the morning which is close to the ground observation time (8 AM), whereas VNP observes in the afternoon. As mentioned before, the Tibetan Plateau is featured with small snow depth and snow may be observed in the morning and disappear in the afternoon due to snow melt or sublimation. Since VNP and AQUA have almost the same overpass time, the much lower accuracy of MODIS AQUA for areas both inside and outside the Tibetan Plateau could be attributed to the disfunction of band 6 used in AQUA snow cover product though the band 6 is claimed to be restored properly.

The results above imply that when using remotely sensed daily snow cover products in China, there is a priority order: VNP or MODIS TERRA is first, then MODIS AQUA. This is very important for future studies aimed at producing more accurate high-resolution snow cover data by combining multiple data sources, since VNP shows generally good accuracy and is much more reliable than MODIS AQUA which is frequently used as an important candidate in previous studies (Gao et al., 2011; Huang et al., 2014; Yu et al., 2016). It should be noted that either VNP or MODIS snow cover products are used for detecting whether snow is present within a pixel, not for extracting other important characters of snow such as snow depth.

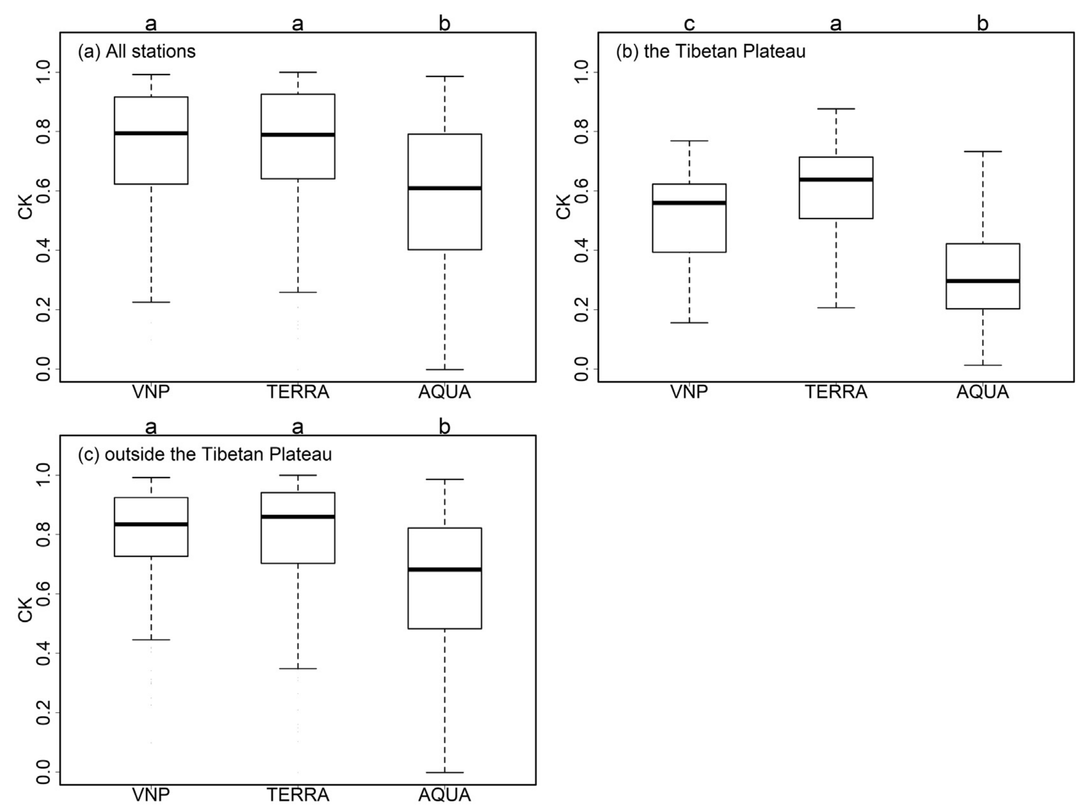


Fig. 7. Comparison of the accuracy (CK) between VNP, MODIS TERRA and MODIS AQUA snow cover products. The multiple comparisons are conducted based on the station-based CKs from all stations (a), those within the Tibetan Plateau (b) and those outside the Tibetan Plateau (c). Letters at the top indicate the significance of the differences: the datasets with a same letter at the top indicate insignificant difference, otherwise indicate significant differences. E.g., in panel (a) the letters at the top of VNP and TERRA are the same indicating insignificant difference in their station-based CKs, but those of VNP and AQUA are different indicating significant differences. The box and whiskers show the distributions of station-based CK values: the lower and upper edges of the box are the first and third quartiles, respectively; the horizonal line in the box indicate the median; the lower and upper whiskers indicate the minimum and maximum values but are drawn no >1.5 times the interquartile range from the first and third quartiles, respectively.

Table 2
Summary of comparison results between VNP, MODIS TERRA and MODIS AQUA snow cover products based on the overlapping 245 stations.

Datasets		CK	FS	POD	PC
ALL	VNP	0.74	0.77	0.74	0.84
	MODIS/TERRA	0.75	0.77	0.74	0.84
	MODIS/AQUA	0.59	0.62	0.62	0.67
Tibetan Plateau	VNP	0.51	0.53	0.48	0.67
	MODIS/TERRA	0.60	0.61	0.52	0.79
	MODIS/AQUA	0.33	0.36	0.35	0.42
Outside Tibetan Plateau	VNP	0.80	0.82	0.80	0.88
	MODIS/TERRA	0.79	0.80	0.79	0.86
	MODIS/AQUA	0.65	0.67	0.68	0.73

Note: all the accuracy metrics are spatially averaged based on stations.

4.4. Uncertainties about clouds, data missing, snow depth thresholds and land cover

Similar to MODIS snow cover products, clouds are also a major problem for VNP and the averaged cloud cover is 42.9 ± 8.6% across 330 stations during the study period. The total data missing caused by clouds, "no decision", "fill" and other problems defined in Riggs et al. (2019) account for 43.3 ± 8.5%. The high cloud cover may have moderate effects on the results with a negative correlation coefficient of −0.21 with the accuracy of VNP snow cover data (Fig. 6a).

The constant snow depth threshold of 1 cm is used throughout this study; however, its value may affect the accuracy assessment of snow cover products (Hao et al., 2018; Wang et al., 2018). It is inspiring that the results of sensitivity tests on NDSI thresholds are very consistent for four kinds of snow depth thresholds including 1, 2, 3 and 4 cm which all indicate that a small NDSI threshold (ranging from 0.14 to 0.18) correspond to the highest accuracy (Fig. 8a). It is also reasonable that the accuracy increases with snow depth thresholds (Fig. 8b) as snow depth is found to have significantly positive effects on the accuracy as mentioned before.

Land cover may also affect the accuracy of VNP snow cover data, especially the forests where sometimes only low NDSI values can be detected for snow cover due to the forest canopy (Riggs et al., 2016b). Since most stations are located in urban or grassland areas, some land cover types featured with trees including forests and savannas are combined as "Trees" here. Only land cover types with number of stations >5 are considered. The comparison results between different land cover types are shown in Fig. 8c. The averaged CK of Croplands is the highest, which is consistent with some other studies finding the highest accuracy of MODIS snow cover products in croplands or agricultural areas (Hall and Riggs, 2007; Yang et al., 2015). However, the comparison results may be biased because of the very limited number of stations for some land cover types such as the trees-covered and barren areas. A multiple comparison based on the Kruskal-Wallis test is further conducted to detect whether there are significant differences in the CKs between the five land cover types. The non-parametric Kruskal-Wallis test is selected here because it makes fewer assumptions and also well applies to situations with unequal sample sizes (Lowry, 2014). It is clear that the accuracy of croplands is significantly higher than grasslands and trees, however, the differences between croplands, urban and barren are not statistically significant possibly due to the very few stations of barren. Thus, ground observations covering more types of land covers should be collected in future to alleviate this problem.

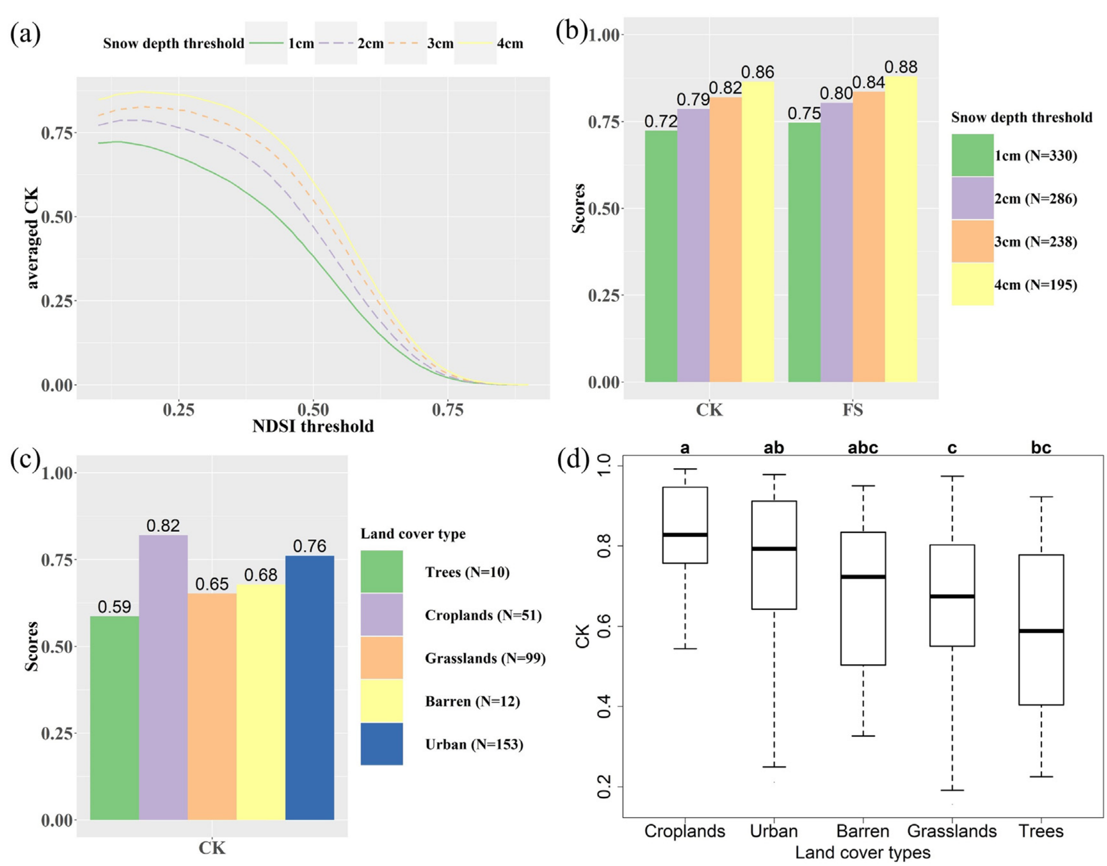


Fig. 8. The effects of snow depth thresholds on sensitivity tests of NDSI thresholds (a) and on the evaluation accuracy (b), the effects of land cover types on the accuracy (c), and the multiple comparison of station-based accuracy between different land cover types. "N" is number of stations. In panel (b), letters at the top indicate the significance of the differences: the land cover types with a same letter at the top indicate insignificant difference, otherwise indicate significant differences. The box and whiskers show the distributions of station-based CK values: the lower and upper edges of the box are the first and third quartiles, respectively; the horizontal line in the box indicate the median; the lower and upper whiskers indicate the minimum and maximum values but are drawn no >1.5 times the interquartile range from the first and third quartiles, respectively.

5. Conclusion

This study investigates for the first time which NDSI threshold should be used for VNP snow cover data based on ground snow depth observations from 330 stations during 1/19/2012–12/31/2018 across China mainland. It is also the first study of evaluating the accuracy of VNP snow cover data based on ground observations in China. The accuracy of VNP data is also compared with those of MODIS TERRA and AUQA snow cover products.

It is concluded that 0.14 is generally the optimal NDSI threshold for VNP snow cover data in China based on the sensitivity tests of NDSI thresholds on the validation accuracy and the results are basically consistent among four kinds of snow depth conditions (i.e. snow depth thresholds of 1, 2, 3 and 4 cm). The averaged Cohen's Kappa (CK) and F-Score (FS) of VNP snow cover data are 0.72 and 0.75, respectively, indicating good accuracy in China. However, its accuracy on the Tibetan Plateau including CK and FS is largely decreased to 0.50 and 0.52, respectively, which is considered to be caused by the much lower snow depth and significantly higher elevation of the Tibetan Plateau. Our study also provides important implications about combined use of VNP and MODIS snow cover products that VNP snow cover data show obviously higher accuracy than AQUA in China including both outside and inside the Tibetan Plateau, however, VNP shows only slightly higher accuracy than TERRA for outside the Tibetan Plateau and significantly lower accuracy than TERRA for stations inside the Tibetan Plateau possibly due to the longer time interval between station observation and satellite overpass time. Future studies may need to collect more detailed samples at the overpass time of satellites to find out the real accuracy difference between VNP and TERRA products and more reliable accuracy information on different land covers.

CRediT authorship contribution statement

Hongbo Zhang: Writing - original draft, Conceptualization, Methodology, Software. **Fan Zhang:** Data curation, Writing - review & editing. **Tao Che:** Data curation, Writing - review & editing. **Shijin Wang:** Data curation, Writing - review & editing.

Declaration of competing interest

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

Acknowledgements

This study is supported by National Natural Science Foundation of China (Grant Number: 91547210), the Class A Strategic Pilot Science and Technology Special Subtopics of the Chinese Academy of Sciences (Grant Number: XDA19070503), State Key Laboratory of Cryospheric Science, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences (Grant Number: SKLCS-OP-2020-13) and National Natural Science Foundation of China (Grant Number: 41701079). The authors are grateful to the Chinese Meteorology Administration (http://data.cma.cn/) for providing daily meteorological observations including snow depth, precipitation, air temperature, wind speed and relative humidity. The VNP and MODIS snow cover products can be downloaded from the Earthdata website (http://earthdata.nasa.gov/).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2020.138156.

References

Arsenault, K.R., Houser, P.R., De Lannoy, G.J., 2014. Evaluation of the MODIS snow cover fraction product. Hydrol. Process. 28, 980–998.
Bo, Y.C., Feng, X.Z., 2000. The Retrieval of Snow-depth Using SSM/I Remotely Sensed Data and its Results Assessment in Qinghai-Xizang (Tibet) Plateau. IEEE, New York.
Che, T., Li, X., Jin, R., Armstrong, R., Zhang, T.J., 2008. Snow depth derived from passive microwave remote-sensing data in China. In: Schneebeli, M. (Ed.), Annals of Glaciology. 49. Int Glaciological Soc, Cambridge, pp. 145–154 2008. 49.
CMA, 2017. Specifications for Surface Meteorological Observation: Snow Depth and Snow Pressure. China Meteorological Press, Beijing.
Cristobal, J., Ninyerola, M., Pons, X., 2008. Modeling air temperature through a combination of remote sensing and GIS data. J. Geophys. Res.-Atmos. 113.
Dong, J., Ek, M., Hall, D., Peters-Lidard, C., Cosgrove, B., Miller, J., et al., 2014. Using air temperature to quantitatively predict the MODIS fractional snow cover retrieval errors over the Continental United States. J. Hydrometeorol. 15, 551–562.
Gao, Y., Lu, N., Yao, T., 2011. Evaluation of a cloud-gap-filled MODIS daily snow cover product over the Pacific Northwest USA. J. Hydrol. 404, 157–165.
Gladkova, I., Grossberg, M., Bonev, G., Romanov, P., Shahriar, F., 2012. Increasing the accuracy of MODIS/Aqua snow product using quantitative image restoration technique. IEEE Geosci. Remote Sens. Lett. 9, 740–743.
Grody, N.C., Basist, A.N., 1996. Global identification of snowcover using SSM/I measurements. IEEE Trans. Geosci. Remote Sens. 34, 237–249.
Guo, D., Yu, E., Wang, H., 2016. Will the Tibetan Plateau warming depend on elevation in the future? Journal of Geophysical Research: Atmospheres 121, 3969–3978.
Hall, D.K., Riggs, G.A., 2007. Accuracy assessment of the MODIS snow products. Hydrol. Process. 21, 1534–1547.
Hall, D.K., Riggs, G.A., Salomonson, V.V., DiGirolamo, N.E., Bayr, K.J., 2002. MODIS snow-cover products. Remote Sens. Environ. 83, 181–194.
Hao, X.H., Wang, J., Hong-Yi, L.I., 2008. Evaluation of the NDSI threshold value in mapping snow cover of MODIS——a case study of snow in the middle Qilian Mountains. Journal of Glaciology & Geocryology 30, 132–138.
Hao, X., Luo, S., Che, T., Wang, J., Li, H., Dai, L., et al., 2018. Accuracy assessment of four cloud-free snow cover products over the Qinghai-Tibetan plateau. International Journal of Digital Earth 1–19.
Holm, S., 1979. A simple sequentially rejective multiple test procedure. Scand. J. Stat. 65–70.
Huang, X., Liang, T., Zhang, X., Guo, Z., 2011. Validation of MODIS snow cover products using Landsat and ground measurements during the 2001–2005 snow seasons over northern Xinjiang, China. Int. J. Remote Sens. 32, 133–152.
Huang, X., Hao, X., Feng, Q., Wang, W., Liang, T., 2014. A new MODIS daily cloud free snow cover mapping algorithm on the Tibetan Plateau. Sci. Cold Arid Reg 6, 0116–0123.
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. Remote Sens. Environ. 190, 274–288.
Immerzeel, W.W., Droogers, P., de Jong, S.M., Bierkens, M.F.P., 2009. Large-scale monitoring of snow cover and runoff simulation in Himalayan river basins using remote sensing. Remote Sens. Environ. 113, 40–49.
Justice, C.O., Román, M.O., Csiszar, I., Vermote, E.F., Wolfe, R.E., Hook, S.J., et al., 2013. Land and cryosphere products from Suomi NPP VIIRS: overview and status. Journal of Geophysical Research: Atmospheres 118, 9753–9765.
Ke, C.-Q., Li, X.-C., Xie, H., Dong-Hui, M., Liu, X., Cheng, K., 2016. Variability in snow cover phenology in China from 1952 to 2010. Hydrol. Earth Syst. Sci. 20, 755.
Kelly, R.E., Chang, A.T., Tsang, L., Foster, J.L., 2003. A prototype AMSR-E global snow area and snow depth algorithm. IEEE Trans. Geosci. Remote Sens. 41, 230–242.
Key, J.R., Mahoney, R., Liu, Y., Romanov, P., Tschudi, M., Appel, I., et al., 2013. Snow and ice products from Suomi NPP VIIRS. Journal of Geophysical Research: Atmospheres 118, 12,816–12,830.
Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics 159–174.
Liu, J., Chen, R., 2011. Studying the spatiotemporal variation of snow-covered days over China based on combined use of MODIS snow-covered days and in situ observations. Theor. Appl. Climatol. 106, 355–363.
Lowry, R., 2014. Concepts and Applications of Inferential Statistics.
Luo, P., Mu, D., Xue, H., Ngo-Duc, T., Dang-Dinh, K., Takara, K., et al., 2018a. Flood inundation assessment for the Hanoi Central Area, Vietnam under historical and extreme rainfall conditions. Sci. Rep. 8, 1–11.
Luo, P., Zhou, M., Deng, H., Lyu, J., Cao, W., Takara, K., et al., 2018b. Impact of forest maintenance on water shortages: hydrologic modeling and effects of climate change. Sci. Total Environ. 615, 1355–1363.
Lyu, J., Mo, S., Luo, P., Zhou, M., Shen, B., Nover, D., 2019. A quantitative assessment of hydrological responses to climate change and human activities at spatiotemporal within a typical catchment on the Loess Plateau, China. Quat. Int. 527, 1–11.
Metsämäki, S., 2016. Report on Validation of VIIRS-FSC Products Against In-situ Observations. Finnish Environment Institute.
Parajka, J., Blöschl, G., 2006. Validation of MODIS snow cover images over Austria. Hydrol. Earth Syst. Sci. Discuss. 3, 1569–1601.
Parajka, J., Blöschl, G., 2012. MODIS-based snow cover products, validation, and hydrologic applications. In: Chang, N.-B., Hong, Y. (Eds.), Multiscale Hydrologic Remote Sensing: Perspectives and Applications.
Pepin, N., Bradley, R., Diaz, H., Baraër, M., Caceres, E., Forsythe, N., et al., 2015. Elevation-dependent warming in mountain regions of the world. Nat. Clim. Chang. 5, 424–430.
Pu, Z., Xu, L., Salomonson, V.V., 2007. MODIS/Terra observed seasonal variations of snow cover over the Tibetan Plateau. Geophys. Res. Lett. 34.
Riggs, G.A., Hall, D.K., Román, M.O., 2016a. MODIS Snow Products User Guide for Collection 6.

Riggs, G.A., Hall, D.K., Román, M.O., 2016b. VIIRS Snow Products Algorithm Theoretical Basis Document (ATBD).
Riggs, G.A., Hall, D.K., Román, M.O., 2017. Overview of NASA's MODIS and Visible Infrared Imaging Radiometer Suite (VIIRS) snow-cover Earth System Data Records. Earth System Science Data 9, 765.
Riggs, G.A., Hall, D.K., Román, M.O., 2019. NASA S-NPP VIIRS Snow Cover Products Collection 1 (C1) User Guide.
Siege, S., Castellan, N., 1988. Non-parametric Statistics for the Behavioural Sciences. MacGraw Hill Int., New York.
Singh, V.P., Haritashya, U.K., Shroder, J.F., 2011. Encyclopedia of Snow, Ice and Glaciers. Springer.
Stahl, K., Moore, R.D., Floyer, J.A., Asplin, M.G., McKendry, I.G., 2006. Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density. Agric. For. Meteorol. 139, 224–236.
Sulla-Menashe, D., Friedl, M.A., 2018. User Guide to Collection 6 MODIS Land Cover (MCD12Q1 and MCD12C1) Product. USGS, Reston, VA, USA, pp. 1–18.
Thapa, S., Chhetri, P.K., Klein, A.G., 2019. Cross-comparison between MODIS and VIIRS snow cover products for the 2016 hydrological year. Climate 7, 57.
Thirel, G., Salamon, P., Burek, P., Kalas, M., 2013. Assimilation of MODIS snow cover area data in a distributed hydrological model using the particle filter. Remote Sens. 5, 5825–5850.
Wang, X.L., Wang, W., Feng, Q.S., Zhi-Bang, L.V., Liang, T.G., 2012. A snow cover mapping algorithm based on MODIS data in Qinghai Province. Acta Pratacul. Sin. 24, 293–299.
Wang, X., Jian, W., Tao, C., Huang, X., Li, H., 2018. Snow cover mapping for complex mountainous forested environments based on a multi-index technique. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing 1–9.
Yang, J., Jiang, L., Ménard, C.B., Luojus, K., Lemmetyinen, J., Pulliainen, J., 2015. Evaluation of snow products over the Tibetan Plateau. Hydrol. Process. 29, 3247–3260.
Yeo, S.-R., Kim, W., Kim, K.-Y., 2017. Eurasian snow cover variability in relation to warming trend and Arctic Oscillation. Clim. Dyn. 48, 499–511.
Yu, J., Zhang, G., Yao, T., Xie, H., Zhang, H., Ke, C., et al., 2016. Developing daily cloud-free snow composite products from MODIS Terra-Aqua and IMS for the Tibetan Plateau. IEEE Trans. Geosci. Remote Sens. 54, 2171–2180.
Zhang, F., Zhang, H., Hagen, S.C., Ye, M., Wang, D., Gui, D., et al., 2015. Snow cover and runoff modelling in a high mountain catchment with scarce data: effects of temperature and precipitation parameters. Hydrol. Process. 29, 52–65.
Zhang, H., Zhang, F., Ye, M., Che, T., Zhang, G., 2016. Estimating daily air temperatures over the Tibetan plateau by dynamically integrating MODIS LST data. Journal of Geophysical Research: Atmospheres 121, 11,425–11,441.
Zhang, H., Zhang, F., Zhang, G., Che, T., Yan, W., Ye, M., et al., 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.
Zhou, H., Aizen, E., Aizen, V., 2013. Deriving long term snow cover extent dataset from AVHRR and MODIS data: Central Asia case study. Remote Sens. Environ. 136, 146–162.