

## Response to Referee #3 Comments

Jing's paper produced a daily cloud-free Normalized Difference Snow Index (NDSI) product with 500 m spatial resolution based on MODIS C6 snow cover datasets in China. So far as we know, the NDSI threshold is the crucial parameter for snow detection by use of optical remote sensing data. The paper in its current version needs major revision and resubmission to meet the level expected of ESSD, for the following reasons:

Firstly, the importance of NDSI needs to be clarified in introduction and using data. NDSI is different from NDVI. The readers are more concerned about binary snow cover or fractional snow cover than NDSI itself. Therefore, it is difficult for me to evaluate whether this dataset is uniqueness or usefulness. Secondly, the current validation scheme is insufficient to support the Spatio-Temporal Adaptive fusion method. The two issues must be addressed for this dataset to be published on ESSD.

### Response:

Thank you very much for the critical comments and suggestions regarding our article. Considering all the constructive comments, we carefully revised the Introduction and significantly improved the evaluation experiments in the Results and Discussion. Note that the NDSI\_Snow\_Cover (hereafter referred to as NDSI) scientific data set with a range of 0, 10 to 100 was used in this study.

**In summary, the main points of the response include:** (1) The importance of NDSI was analyzed from three aspects (Section 1). (2) The MODIS C6 NDSI data were compared with the atmospheric corrected NDSI calculated from the surface reflectance bands of MOD09GA products. (3) The overall classification accuracy and monthly average classification accuracy of STAR NDSI collection were compared with those of TAC NDSI and NIEER AVHRR SCE datasets based on in-situ snow depth measurements (Section 3.1). The accuracy of STAR NDSI collection during the snow period was emphasized. (4) The effectiveness of STAR method on cloud removal under different simulated cloud conditions was analyzed based on Landsat NDSI maps (Section 4). Details are presented below.

### General Comments:

1. The importance of the NDSI research is insufficiently described, and why NDSI is more important than binary and FSC products should be further described in the introduction.

**Response:** Thank you for the critical comment. The importance of NDSI research mainly includes three aspects: (1) NDSI is a more accurate description of the snow detection as compared to SCA and FSC (Riggs and Hall, 2015; Riggs et al., 2017); (2) As a basic data, NDSI has the significant advantage of allowing users to more

accurately determine SCA or FSC for their particular study areas and application requirements (Hall et al., 2019); (2) Cloud-free SCA and FSC datasets produced by composite algorithms are frequently released, while high-quality cloud-free NDSI datasets are still scarce. In addition, we also consider all binary SCA, FSC and NDSI products to be important. For decades, they have facilitated a variety of snow-related researches. Finally, we added a description of the importance of NDSI research in the Introduction as follows.

Increasing studies have moved to the MODIS C6 suite since its release in 2016. In C6 data, snow cover is reported as NDSI rather than SCA and FSC. NDSI is an index that is related to the snow presence in a pixel and is a more accurate description of snow fraction as compared to SCA and FSC (Riggs and Hall, 2015; Riggs et al., 2017). The clear-sky accuracy of C6 NDSI datasets is robust compared to higher resolution remote-sensed images (such as Landsat and Sentinel series) and in-situ measurements (Crawford, 2015; Zhang et al., 2019; Aalstad et al., 2020). As a basic data, it has the significant advantage of allowing users to more accurately determine SCA or FSC for their particular study areas and application requirements (Hall et al., 2019). For example, several optimal classification thresholds for SCA (Huang et al., 2018; Malmros et al., 2018; Tong et al., 2020) and specially tuned mapping methods for FSC (Kuter et al., 2018; Hou et al., 2020; Zhang et al., 2021) were designed to generate regional SCA and FSC datasets from NDSI snow cover datasets, which were superior to the globally harmonized algorithms in C5 data. However, severe cloud contamination also limits the application of NDSI datasets, resulting in many studies only considering cloud-free areas (Kuter et al., 2018; Malmros et al., 2018; Tong et al., 2020; Hou et al., 2020; Zhang et al., 2021). Since the aforementioned cloud removal methods were generally designed for binary SCA, their applicability to NDSI with more complicated spatio-temporal characteristics should be improved. Thus, several gap-filling methods with an associated concern of spatial and temporal correlations of snow presence were proposed to remove clouds from NDSI (Jing et al., 2019; Chen et al., 2020; Li et al., 2020). Among these methods, the spatio-temporal feature-based methods with relatively high robustness are more effective for improving NDSI datasets (Jing et al., 2019).

Many studies on snow monitoring in China are available, and most of these studies focus on binary SCA mapping. On the regional scale, QTP, which is known as the world's third pole, plays a key role in the global climate system. Nevertheless, snow cover mapping is particularly challenging over QTP due to the frequent cloud cover resembling fragmented snow. A large number of studies have demonstrated that the snow cover variability over QTP is extremely complex, with significant spatio-temporal heterogeneity (Gao et al., 2012; Tang et al., 2013; Yu et al., 2016; Liang et al., 2017; Zhang et al., 2012). NX (Wang et al., 2008) and NC (Che et al., 2016) located in mid-latitude areas are dominated by seasonal snow cover. Che et al. (2019) presented an integrated snow cover dataset from a distributed hydrometeorological observation network in the Heihe River Basin, which achieved a

prominent demonstration of data synthesis at a watershed scale. In addition, the large-scale transient snow cover areas increase the level of challenge for generating high-quality snow cover datasets. On the national scale, Huang et al. (2016) obtained a long-term cloud-removed SCA product using a multi-source data fusion method. Despite many relevant studies, only a few cloud-free snow cover datasets have been released publicly.

Several typical long-term cloud-free snow cover products available online are listed in Table 1 (datasets are referenced via DOI), which cover most snow-dominated regions in China. Huang (2020) provided MODIS daily cloudless SCA products with relatively accurate snow detection capabilities in Northern Hemisphere based on multi-source data. Muhammad and Thapa (2020, 2021) obtained eight-day/daily MODIS SCA and glacier composite datasets for High Mountain Asia by aggregating seasonal, temporal, and spatial filters, which can serve as a valuable input for hydrological and glaciological investigations. Hao et al. (2021; 2022) yielded two long-term daily SCA datasets over China through a series of processes such as quality control, cloud detection, snow discrimination, and gap-filling (including hidden Markov random field and snow-depth interpolation techniques). Their releases and updates promoted the research of snow cover characteristics in China. Qiu et al. (2017) yielded a daily FSC dataset with detailed snow cover information over High Mountain Asia with MDC and spatial filtering. Additionally, the global cloud-gap-filled MODIS NDSI dataset (MOD10A1F) is available online since 2020, where cloud-covered grids in the MODIS Terra NDSI product are filled by retaining clear-sky observations from previous days (Hall and Riggs, 2020). However, this dataset performs poorly in China, where periodic and transient snow is dominant. In general, cloud-free SCA datasets produced by composite algorithms are frequently released, while high-quality cloud-free NDSI datasets are still scarce.

**Table 1. Typical long-term cloud-free snow cover products covering most snow-dominated regions in China.**

References	Type	Spatial coverage	Temporal coverage	Temporal resolution	Spatial resolution	DOI
Hao et al. (2021)	SCA	China	1981–2019	Daily	~5 km	10.11888/Snow.tpd.271381
Hao et al. (2022)	SCA	China	2000–2020	Daily	~500 m	10.12072/ncdc.I-SNOW.db0001.2020
Huang (2020)	SCA	Northern hemisphere	2000–2015	Daily	~1 km	10.12072/ncdc.CCI.db0044.2020
Muhammad and Thapa (2021)	SCA	High Mountain Asia	2002–2019	Daily	~500 m	10.1594/PANGAEA.918198
Qiu et al. (2017)*	FSC	High Mountain Asia	2002–2018	Daily	~500 m	10.11922/sciencedb.457
Hall and Riggs (2020)	NDSI	Global coverage	2000–present	Daily	~500 m	10.5067/MODIS/MOD10A1F.061

\*Cloud coverage is less than 10%.

To this end, this study generates a spatiotemporally continuous Terra–Aqua MODIS NDSI product with satisfactory accuracy for China, fully considering the spatio-temporal characteristics of regional snow cover

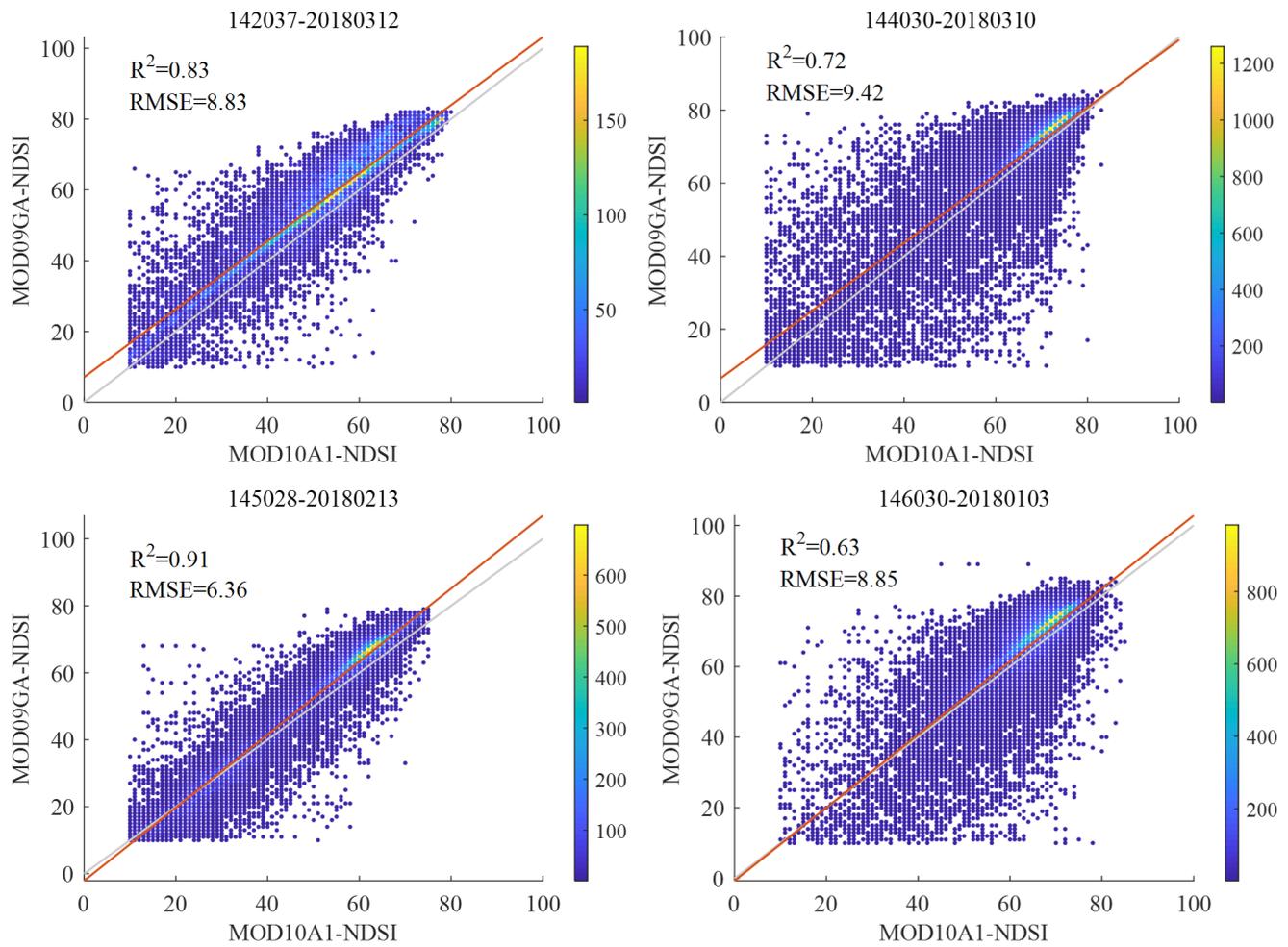
variability. A Spatio-Temporal Adaptive fusion method with error correction (STAR) improved from our previous work (Jing et al., 2019) is utilized to eliminate cloud obscuration. The long-term detailed snow cover extent dataset will facilitate snow-related scientific studies and practical applications in China. The rest of this paper is arranged as follows. Firstly, Section 2 describes the input data and the proposed cloud removal method. Then, Section 3 presents the verification accuracy of STAR NDSI collection, with a subsequent analytical application. The cloud removal effectiveness under different cloud coverages is discussed in Section 4. Finally, the data availability and the conclusions are provided in Section 5 and Section 6, respectively.

2. The NDSI value in either MODIS C5 or C6 is the NDSI without atmospheric correction. How this NDSI differs from NDSI by the atmospheric corrected from MOD09GA/MYD09GA? Has the author compared it, and which NDSI value is more useful to readers?

**Response:** Thank you for the critical comment. The objective of this study is to improve the temporal and spatial continuity of MODIS NDSI products. The accuracy of the original data in clear-sky areas is very important but seems to be beyond the scope of this study. However, we still added a comparative experiment of MOD09GA NDSI and MOD10A1 NDSI datasets (C6). The results reveal that the accuracy of MOD10A1 NDSI dataset is slightly higher than that of MOD09GA NDSI dataset compared to Landsat NDSI maps. The reason may be that the NDSI algorithm specially designed for MOD10 products, is more applicable to them. In addition, Riggs et al. (2017) discovered that “compared to true color (bands 1, 4, 3) image of MOD09GA, all the snow-cover extent is detected in C6 NDSI\_Snow\_Cover by the revised algorithm; however, significant snow-cover extent was missed in C5 FSC”. Therefore, MOD10A1 C6 NDSI products are relatively reliable, while MOD10A1 C5 SCA/FSC products are underestimated. Due to space limitations, this comparative experiment was not added to the manuscript. The details of the experiment are as follows.

The atmospheric corrected NDSI data were calculated from the surface reflectance bands of MOD09GA products based on the algorithm of MOD10A1 NDSI dataset (C6). Supplementary Figure 1 shows four scatter diagrams of MOD10A1 NDSI and MOD09GA NDSI datasets in snow-cover areas. The two NDSI datasets are partially inconsistent, with  $R^2$  ranging from 0.63 to 0.91 and RMSEs ranging from 6.36 to 9.42. For an in-depth evaluation, the performance statistics for MOD09GA NDSI and MOD10A1 NDSI datasets against Landsat NDSI maps are shown in Supplementary Table 1. Compared with Landsat NDSI maps, their average CCs are almost equal. In addition, in terms of RMSEs, AEs and SRDs, MOD10A1 NDSI data are slightly superior to MOD09GA NDSI data. Despite the limited spatio-temporal scope of the samples, this comparative experiment can reflect the reliability of MODIS C6 NDSI products. However, more comprehensive comparisons are needed for the specific

researches on the original accuracy in clear-sky areas.



**Supplementary Figure 1. Scatter diagram of MOD10A1 NDSI and MOD09GA NDSI datasets in snow-cover areas.** Note that *142037\_20180312* denotes region (Worldwide Reference System of Landsat) and date.

**Supplementary Table 1. Performance statistics for MOD09GA NDSI and MOD10A1 NDSI datasets against Landsat NDSI maps.**

Region_Date	CC		RMSE		AE		SRD (%)	
	MOD09GA NDSI	MOD10A1 NDSI						
119027_20180223	0.98	0.99	6.29	5.77	1.65	1.57	-1.67	-1.47
119028_20180311	0.97	0.97	11.37	8.95	6.93	4.69	-0.96	-3.22
119029_20180311	0.77	0.81	2.05	1.80	-0.01	-0.02	0.03	0.04
119030_20180311	0.58	0.66	1.98	1.82	-0.16	-0.12	-0.95	-0.60
122027_20180316	0.91	0.94	12.34	9.31	0.18	0.10	-18.36	-12.64
123037_20180203	0.81	0.82	1.94	3.03	-0.07	0.71	-0.83	3.94
138039_20180401	0.80	0.83	15.19	14.60	-2.29	-5.80	-6.19	-6.62
139030_20180102	0.98	0.97	10.91	8.50	5.17	4.17	1.84	5.29
139035_20180307	0.91	0.93	10.65	7.61	3.85	1.96	-4.07	2.69
139036_20180307	0.79	0.83	9.77	7.73	0.59	-0.93	-3.95	-1.82
140035_20180226	0.92	0.94	8.49	7.75	-0.20	-3.13	-2.73	-2.36

140039_20180125	0.95	0.96	8.53	7.04	2.59	2.16	-0.87	0.60
141034_20180305	0.87	0.84	10.57	11.04	-0.37	-4.05	-5.30	-7.05
141035_20180217	0.93	0.95	5.28	3.97	0.48	-0.53	-1.66	-2.50
141035_20180305	0.84	0.87	11.13	10.55	2.78	-6.12	-3.64	-6.00
142035_20180107	0.67	0.72	4.53	4.04	-0.25	-0.20	-2.24	-0.70
142036_20180107	0.91	0.95	6.63	4.62	0.15	0.32	-5.75	-0.77
142037_20180312	0.97	0.97	9.55	7.18	3.69	1.30	-0.76	0.43
144028_20180105	0.89	0.89	24.83	20.75	24.65	20.58	0.00	0.00
144029_20180105	0.50	0.35	24.42	20.54	23.94	19.91	-0.10	0.00
144030_20180105	0.65	0.61	24.59	21.13	23.10	19.86	-0.61	-0.16
144030_20180310	0.91	0.92	17.23	15.43	10.00	8.89	3.68	5.94
145028_20180213	0.94	0.94	20.01	17.72	16.54	15.52	4.98	11.33
145035_20180128	0.72	0.74	8.83	9.37	1.52	2.66	1.84	7.04
146029_20180103	0.79	0.80	24.77	21.30	23.89	20.43	0.14	0.27
146029_20180220	0.92	0.91	16.46	14.36	12.34	10.91	3.45	7.78
146030_20180103	0.66	0.64	26.96	23.61	25.47	22.32	0.07	0.69
146031_20180220	0.97	0.97	8.03	6.94	2.37	1.79	-2.44	-1.92
146035_20180103	0.82	0.83	11.37	12.27	2.66	5.12	-1.10	9.70
<b>Average</b>	<b>0.84</b>	<b>0.85</b>	<b>12.23</b>	<b>10.65</b>	<b>6.59</b>	<b>4.97</b>	<b>-1.66</b>	<b>0.27</b>

3. [The current validation plan \(in-situ snow depth observations and Landsat NDSI maps\) is insufficient to support the Spatio-Temporal Adaptive fusion method. Please add the improved validation plan.](#)

**Response:** Thank you for the critical comment. We carefully redesigned the evaluation experiments and substantially revised the evaluation part in the manuscript. **In the Results**, the improved evaluation experiments included: **(1)** The overall classification accuracy and monthly average classification accuracy of STAR NDSI collection were compared with those of TAC NDSI and NIEER AVHRR SCE datasets based on in-situ snow depth measurements (Section 3.1). **(2)** The numerical accuracy of STAR NDSI collection during snow period was compared with those of TAC NDSI and MODIS CGF NDSI datasets based on Landsat NDSI maps (Section 3.2). **(3)** The evaluation in clear-sky areas and cloud-cover areas was performed based on Landsat NDSI maps, to highlight the accuracy of the recovered pixels in STAR NDSI collection during the snow period (Section 3.2). Besides, **in the Discussion**, the cloud removal effectiveness of STAR method under different simulated cloud conditions was analyzed based on Landsat NDSI maps. The new evaluation experiments were presented in the response to minor comments 7 and 8.

### Minor comments:

1. L 95. “The daily snow cover datasets of C6 were used in this study.” There are NDSI\_Snow\_Cover and NDSI scientific data sets in the C6 by MODIS C6 User Guide (Riggs, 2015). The NDSI\_Snow\_Cover and NDSI is

different, the author need to describe the data used in the study. This is related to the subsequent results.

**Response:** Thank you for the suggestion. The NDSI\_Snow\_Cover data were used in this study. We added a description of the scientific data set in the Data and Methods as follows.

...The NDSI\_Snow\_Cover (hereafter referred to as NDSI) scientific data set with a range of 0, 10 to 100 was used in this study...

2. Fig.1. It is recommended to remove the NC snow area cover. This is only an administrative division rather than a snow region (<https://essd.copernicus.org/articles/13/4711/2021/>). The in-situ observations of this area were not used in this study. In addition, TP suggests replacing by QTP?

**Response:** Thank you for the suggestion. Figure 1 was revised. The NC snow cover area was removed, XJ was replaced by NX, and TP was replaced by QTP (The names in the text were also revised).

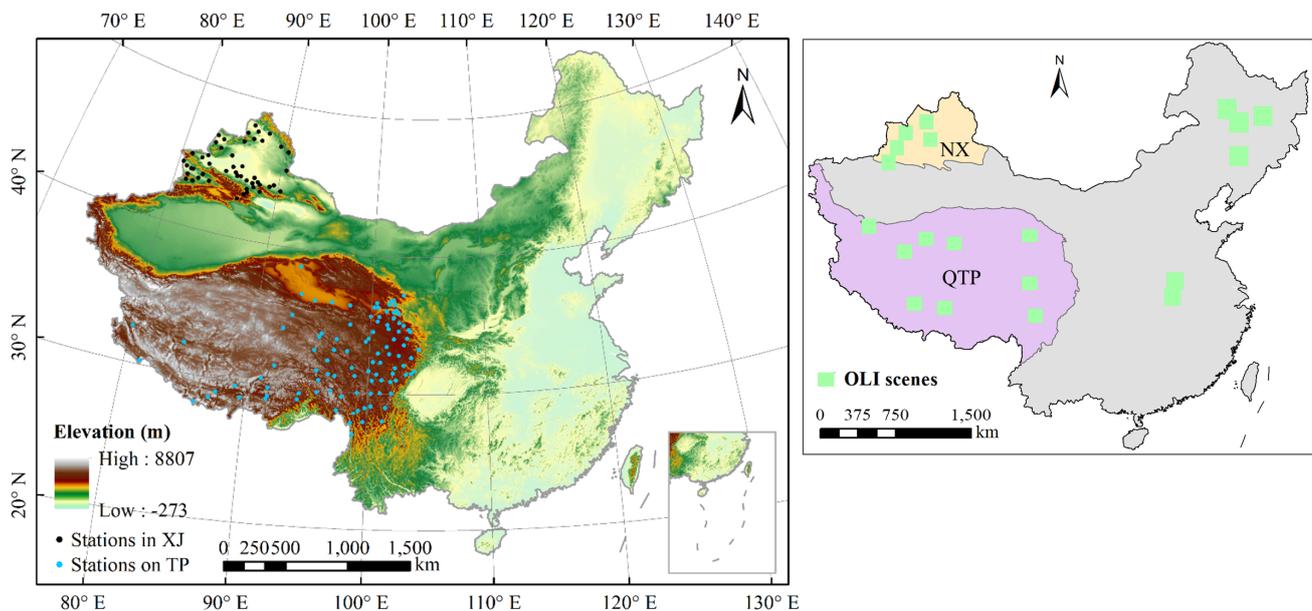


Figure 1. Topographic relief of China, meteorological stations in NX and QTP, and Landsat OLI scenes used for validation.

3. What does the dashed half-frame line in Fig.2 mean?

**Response:** Thank you for the comment. The steps surrounded by the dashed half-frame line constitute our Spatio-Temporal Adaptive fusion with error correction (STAR) method, including spatio-temporal adaptive fusion (STAF) and error correction (EC). This is an iterative cloud removal process until no cloud remains.

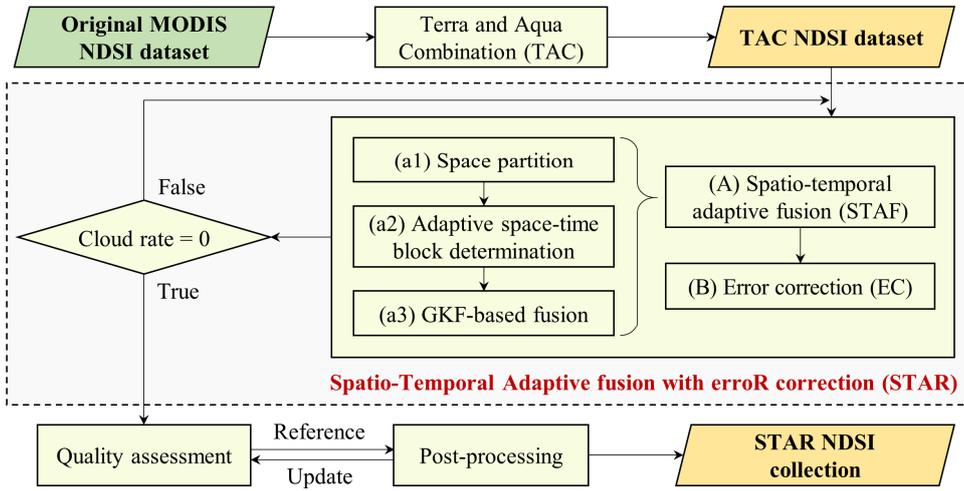


Figure 2. Schematic of the generation procedure of STAR NDSI collection.

4. L 118. The description of fusion method and rules is not clear, only the priority is determined at L 123. It's better to describe the fusion method first, and then introduced the interpolation used by Aqua.

**Response:** Thank you for the suggestion. The description of TAC was revised as follows.

TAC blends the same-day snow maps deriving from MODIS sensors onboard Terra and Aqua satellites. Its cornerstone is the unlikely significant changes of the snow pattern within the data-acquired time interval (approximately 3 h). Since TAC can efficiently decrease the cloud fraction by 5%–20% with negligible precision sacrifice (Li et al., 2019), it is introduced as a pre-processing to reduce cloud coverage preliminarily. Its priority scheme is determined as high value > low value > cloud.

$$\begin{aligned} NDSI^P &= NDSI^{Terra} \text{ IF } (NDSI^{Terra} > NDSI^{Aqua} \text{ OR } NDSI^{Aqua} \text{ is cloud}), \\ NDSI^P &= NDSI^{Aqua} \text{ IF } (NDSI^{Aqua} > NDSI^{Terra} \text{ OR } NDSI^{Terra} \text{ is cloud}), \end{aligned} \quad (1)$$

where  $NDSI^{Terra}$  and  $NDSI^{Aqua}$  are MODIS NDSI datasets from Terra and Aqua satellites, respectively.  $NDSI^P$  represents the pre-processed NDSI maps after TAC (referred to as TAC NDSI dataset in subsequent sections). The snow in low altitude and low latitude areas during summer is reversed to no snow to alleviate commission errors inherited from the original data. In addition, since the Aqua dataset is available since July 2002, the key-process STAR is directly used to remove clouds from Terra MODIS NDSI dataset between August 2000 and May 2002. Particularly, the improved Aqua MODIS C6 NDSI dataset significantly enhances the effectiveness of TAC due to the successful restoration of the absent Aqua MODIS band 6 data by the quantitative image restoration method (Gladkova et al., 2012).

5. L158. What does NDSIP mean?

**Response:** Thank you for the comment.  $NDSI^P$  represents the pre-processed NDSI map after the pre-

process TAC. We added a description to the manuscript as follows.

...Specifically, the residual errors of the intersecting cloud-free areas of the pre-processed and fused NDSI maps ( $NDSI^P$  after TAC and  $NDSI^F$  after STAF) are diffused to other cloud-free areas of the fused NDSI maps using the triangulation-based natural neighbor interpolation...

6. L 216. What does “snow-clad pixels” mean? Are there any reference?

**Response:** Thank you for the comment. To clarify the text, “snow-clad pixels” was revised to “snow-covered pixels”.

7. L 214. Section 3.1 The validation method need to be improved.

The in situ snow depth derived from 49 and 92 CMA station from BJ and QTP. However, the validation date need to be clear. Due to snow-free period is long, many stations record no snow in one year. In fact, the most useful and most concerned should be the NDSI recovery during the snow cover period. The author should focus on the NDSI recovery during the snow cover period and a detailed confusion matrix needs to be given. In addition, the authors need to focus on the accuracy comparison of the product itself (TAC, L 218, the cloud-covered areas in the TAC NDSI dataset are considered to be snow-free. Here the cloud-covered areas should be eliminated without comparison) and the final spatial continuous product (STAR). The reader is concerned with the loss of NDSI accuracy after STAR interpolation.

**Response:** Thank you for the critical comment. The comment pointed to two issues, which we responded to separately.

# **Issue 1.** For the validation dates, we carefully revised the validation against in-situ snow depth measurements, including: **(1) the detailed confusion matrices were added to Table 4 and Table 5 so that the numbers of snow-cover and snow-free samples in NX and QTP during their entire validation dates are clear.** In addition, OEs ( $OE = SN/(SN + SS)$ ) also reflected the detection accuracy of snow pixels; **(2) the monthly classification accuracies in a hydrological year were added to Fig.4, and their temporal characteristics were analyzed in detail.** The corresponding revisions in the manuscript are as follows.

Table 4 demonstrates that NIEER AVHRR SCE and STAR NDSI datasets preeminently capture the snow dynamics in NX referring to the in-situ measurements, with OAs more than 90%. However, TAC NDSI dataset is insufficient to accurately describe the snow cover variability. Although CEs perform well regardless of the snow depth threshold, OEs of TAC NDSI collection are extremely high, indicating that many cloud-covered areas are dominated by snow. NIEER AVHRR SCE dataset partially retrieves snow pixel under cloud obstruction with an

OE decreased by ~43%. STAR NDSI collection completely removes clouds and accurately presents snow distribution, with an OE further decreased from ~17% to ~7%. The generation procedure in NX has two strengths. Firstly, the satellite-borne sensors can accurately capture the snow events on the ground due to the generally thick snow averaging approximately 20 cm. Secondly, the gap-filling approach with comprehensive consideration of spatial and temporal correlation has outstanding reliability due to the significant periodicity of snow variation. It can be inferred that the NDSI datasets in NC have high accuracy because of the similar snow conditions, despite the lack of in-situ data in this region.

By contrast, despite the satisfactory performance of OAs and CEs, the OEs of three snow cover datasets over QTP are as remarkably high as 72%, 40%, and 39% even at the snow depth threshold of 1 cm (Table 5). This finding indicates the omission of a large number of snow-covered pixels. The specific reasons are as follows. Firstly, the original MODIS NDSI maps frequently underestimate the snow presence throughout the snow period because discriminating the shallow snow pixels with an averaged snow depth of approximately 4 cm over QTP is challenging. Secondly, the credibility of the spatio-temporal contextual information is relatively low because the patchy snow rapidly and irregularly varies due to the extremely complex topographic and climatic conditions, leading to a further decrease in the accuracy of the gap-filled results. Lastly, the meteorological stations over QTP are unevenly distributed and are mostly located in low- and medium-altitude/latitude areas dominated by transient snow. Consequently, the evaluation results slightly exaggerate the real OEs.

For the in-depth analysis of the temporal characteristics, the monthly classification accuracies of TAC NDSI, NIEER AVHRR SCE and STAR NDSI products in NX and QTP are shown in Fig. 4 (the horizontal axis is the month in a hydrological year). In NX (group a), the monthly snow fraction in the in-situ samples is greater than 85% from December to next February. Therefore, the clouds in TAC NDSI dataset seriously affect the snow cover estimation, while both cloud-free products exhibit superior OAs. Compared to NIEER AVHRR SCE product, STAR NDSI collection has slightly higher CEs but relatively lower OEs. The OEs of STAR NDSI collection typically occur during snow accumulation and ablation periods, and almost disappear during stable snow-cover and snow-free periods. In QTP, the snow period is generally from October to next May, with the monthly snow fraction of less than 10% in the in-situ samples. Consequently, the underestimation of snow coverage caused by the clouds in TAC NDSI dataset is slight. All three datasets perform well in OA and CE but perform significantly worse in OE. All three products achieve outstanding OAs and CEs but exhibit relatively poor OEs. In NIEER AVHRR SCE and STAR NDSI datasets, these OEs are generally observed outside the snow period. As mentioned above, there are three reasons for this phenomenon. Nonetheless, STAR NDSI collection presents superior classification accuracy to TAC NDSI and NIEER AVHRR SCE datasets.

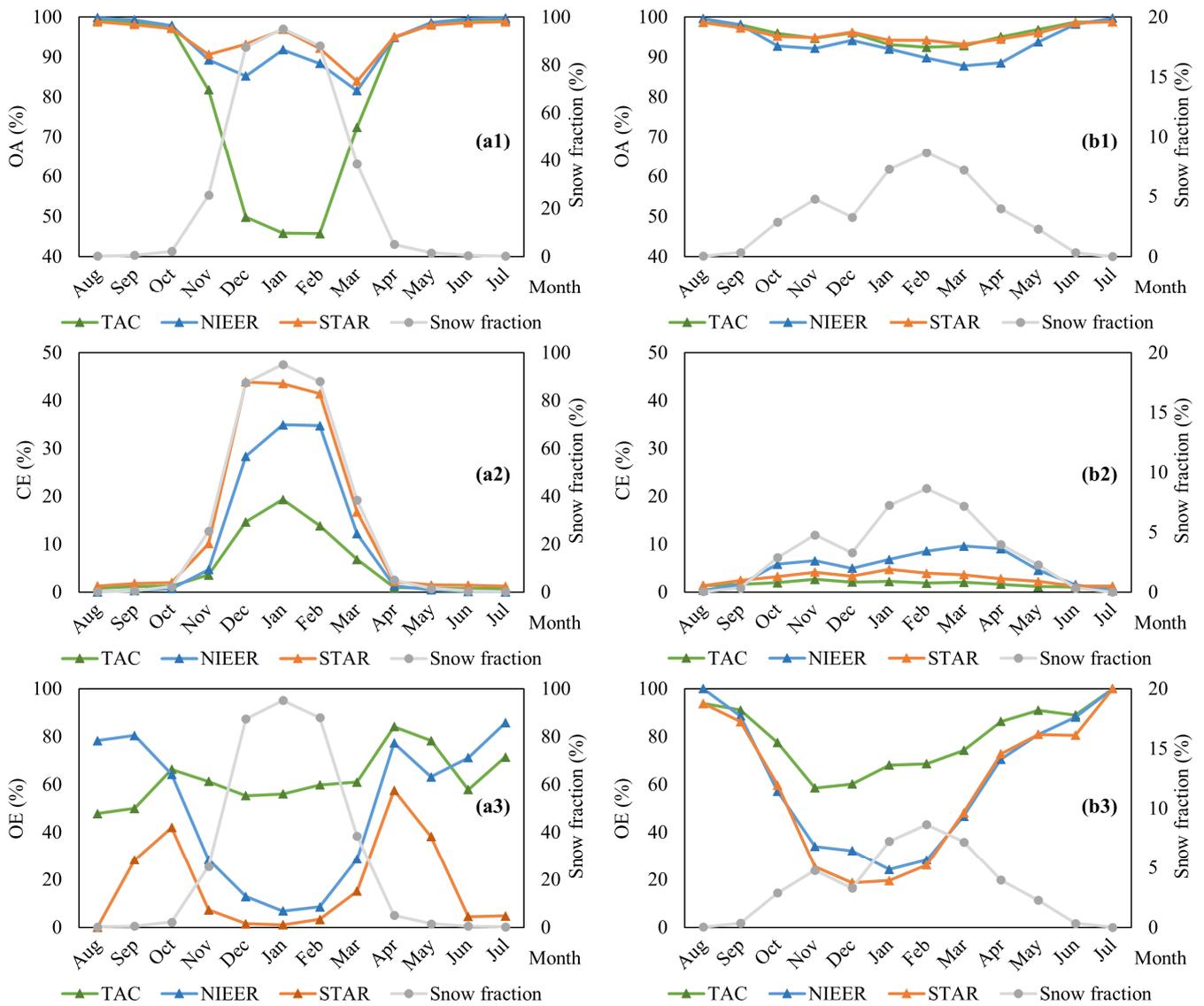
**Table 4. Confusion matrices between TAC NDSI, NIEER AVHRR SCE, STAR NDSI datasets and in-situ snow-depth (SD) data in NX from 1 January 2001 to 31 August 2007.**

Station	TAC			NIEER			STAR		
<b>Indicator</b>	Snow	No snow	Total	Snow	No snow	Total	Snow	No snow	Total
Snow (SD > 0 cm)	13466	19836	33302	27656	5646	33302	30955	2347	33302
<b>OE</b>	40%	<b>60%</b>		83%	<b>17%</b>		93%	<b>7%</b>	
No snow	1269	76909	78178	1384	76794	78178	2741	75437	78178
<b>CE</b>	<b>2%</b>	98%		<b>2%</b>	98%		<b>4%</b>	96%	
Total	14735	96745	111480	29040	82440	111480	33696	77784	111480
<b>OA</b>			<b>81%</b>			<b>94%</b>			<b>95%</b>
Snow (SD > 1 cm)	13136	18390	31526	26899	4627	31526	29909	1617	31526
<b>OE</b>	42%	<b>58%</b>		85%	<b>15%</b>		95%	<b>5%</b>	
No snow	1599	78355	79954	2141	77813	79954	3787	76167	79954
<b>CE</b>	<b>2%</b>	98%		<b>3%</b>	97%		<b>5%</b>	95%	
Total	14735	96745	111480	29040	82440	111480	33696	77784	111480
<b>OA</b>			<b>82%</b>			<b>94%</b>			<b>95%</b>

**Table 5. Confusion matrices between TAC NDSI, NIEER AVHRR SCE, STAR NDSI datasets and in-situ snow-depth (SD) data in QTP from 1 August 2000 to 31 December 2013.**

Station	TAC			NIEER			STAR		
<b>Indicator</b>	Snow	No snow	Total	Snow	No snow	Total	Snow	No snow	Total
Snow (SD > 0 cm)	5189	18095	23284	10900	12384	23284	11219	12065	23284
<b>OE</b>	22%	<b>78%</b>		47%	<b>53%</b>		48%	<b>52%</b>	
No snow	6145	404624	410769	17663	393106	410769	9338	401431	410769
<b>CE</b>	<b>1%</b>	99%		<b>4%</b>	96%		<b>2%</b>	98%	
Total	11334	422719	434053	28563	405490	434053	20557	413496	434053
<b>OA</b>			<b>94%</b>			<b>93%</b>			<b>95%</b>
Snow (SD > 1 cm)	4126	10357	14483	8391	6092	14483	8813	5670	14483
<b>OE</b>	28%	<b>72%</b>		58%	<b>42%</b>		61%	<b>39%</b>	
No snow	7208	412362	419570	20172	399398	419570	11744	407826	419570
<b>CE</b>	<b>2%</b>	98%		<b>5%</b>	95%		<b>3%</b>	97%	
Total	11334	422719	434053	28563	405490	434053	20557	413496	434053
<b>OA</b>			<b>96%</b>			<b>94%</b>			<b>96%</b>

Overall, STAR NDSI collection is capable of snow status estimation, eliminating cloud contamination in TAC NDSI dataset, and capturing more snow events than NIEER AVHRR SCE dataset. However, the accuracy of STAR NDSI collection has regional and temporal heterogeneity. Firstly, the accuracy over QTP is lower than that of NX, which is consistent with the characteristic of the original MODIS NDSI maps. Then, the permanent and periodic snow regime regions reconstructed by STAR have prominently high accuracy, while the transient snow-covered regions are easily omitted. Fortunately, the monitoring of permanent and periodic snow plays a key role in most snow-related investigations. Finally, the accuracy of stable snow-cover and snow-free periods is slightly better than that of snow accumulation and ablation periods.



**Figure 4. Monthly classification accuracy of TAC NDSI, NIEER AVHRR SCE, and STAR NDSI products on NX (group a) and QTP (group b). Note that the optimal values for OA, CE and OE are 100%, 0% and 0%, respectively.**

**# Issue 2.** For the validation areas, all the cloud-covered areas were eliminated without comparison in the previous version of the manuscript. But the evaluation experiments were redesigned according to the comment from Chief editor Kirsten Elger. To fairly evaluate these products in the same areas, the cloud-covered areas in TAC NDSI dataset were considered to be snow-free. **On the one hand, we briefly introduced the evaluation results for cloud-free areas in the previous version**, which demonstrated that STAR NDSI collection can completely remove clouds without a significant loss of accuracy. **On the other hand, an internal comparison of STAR NDSI collection in clear-sky areas and cloud-cover areas was performed based on Landsat NDSI maps in the current version**, which highlighted the accuracy of the recovered pixels in STAR NDSI collection. The results reveal that the accuracy of recovered areas is inevitably slightly lower than that of clear-sky areas. Although the average CC decreases from 0.85 to 0.73 and the average RMSE increases from 13.48 to 16.30 compared with

clear-sky areas, the accuracy of recovered areas is satisfactory. Since many recovered areas inherit errors from clear-sky areas because the cloud removal procedure completely relies on the original dataset, a slight decrease in accuracy is reasonable. This finding highlights that STAR NDSI collection can completely remove clouds with satisfactory accuracy.

**The evaluation results for cloud-free areas in the previous version are as follows.**

### 3.1 Validation against in situ snow depth measurements

...Table 4 demonstrates that both NDSI datasets preeminently capture the snow dynamics in NX referring to the in-situ measurements, with the OAs reaching 0.97 and 0.95, respectively. CEs and OEs perform well regardless of the snow depth threshold, highlighting that remote-sensed NDSI datasets are capable of snow status estimation in NX. The generation procedure in NX has two strengths. First, the satellite-borne sensors can accurately capture the snow events on the ground due to the generally thick snow averaging approximately 20 cm. Second, the gap-filling approach with comprehensive consideration of spatial and temporal correlation has outstanding reliability due to the significant periodicity of snow variation. It can be inferred that the NDSI datasets in NC have high accuracy because of the similar snow conditions, despite the lack of in-situ data in this region.

By contrast, despite the satisfactory performance of OAs and CEs, the OEs of two NDSI datasets over QTP are as remarkably high as 28% and 39% even at the snow depth threshold of 1 cm (Table 5). This finding indicates the omission of a large number of snow-covered pixels. The specific reasons are as follows. First, the original MODIS NDSI maps frequently underestimate the snow presence throughout the snow period because discriminating the shallow snow pixels with an averaged snow depth of approximately 4 cm over QTP is challenging. Second, the credibility of the spatio-temporal contextual information is relatively low because the snow rapidly and irregularly varies due to the extremely complex topographic and climatic conditions, leading to a further decrease in the accuracy of the gap-filled results. Last, the meteorological stations over QTP are unevenly distributed and are mostly located in low- and medium-altitude/latitude areas dominated by transient snow. Consequently, the evaluation results slightly exaggerate the real OEs...

**Supplementary Table 4. Classification statistics in NX.**

Indicators	Snow depth > 0 cm		Snow depth > 1 cm	
	TAC	STAR	TAC	STAR
Snow fraction	21%	30%	21%	28%
OA	0.97	0.95	0.97	0.95
CE	0.02	0.04	0.03	0.05
OE	0.06	0.06	0.05	0.05

**Supplementary Table 5. Classification statistics over QTP.**

Indicators	Snow depth > 0 cm		Snow depth > 1 cm	
	TAC	STAR	TAC	STAR
Snow fraction	3%	5%	2%	3%
OA	0.96	0.95	0.97	0.96
CE	0.02	0.03	0.03	0.03
OE	0.45	0.52	0.28	0.39

### 3.2 Validation based on Landsat NDSI maps

...The snow dynamics presented by TAC NDSI and STAR NDSI datasets are highly consistent with Landsat NDSI maps, with an average CC of approximately 0.84. This finding highlights that STAR NDSI collection can completely remove clouds without sacrificing accuracy. The average RMSEs of TAC NDSI and STAR NDSI datasets are 13.48 and 14.64, respectively, which are mainly due to systematic overestimation (Landsat NDSI values are generally low). In terms of snow coverage, TAC NDSI and STAR NDSI datasets are respectively slightly overestimated and underestimated, with corresponding average SRDs of 0.76% and -1.48% (SRD indicates the difference of snow rate compared with the Landsat NDSI map)...

Supplementary Table 6. Performance statistics for two MODIS NDSI datasets against Landsat NDSI maps.

Region_Date	CC		RMSE		AE		SRD (%)		NCR		SRD × NCR (%)	
	STAR	TAC	STAR	TAC	STAR	TAC	STAR	TAC	STAR	TAC	STAR	TAC
NC1_20180225	0.87	0.95	17.10	16.89	15.07	15.23	1.02	2.91	0.73	0.28	0.75	0.82
NC2_20180311	0.83	0.89	13.87	12.27	12.14	11.78	-1.68	-0.10	0.72	0.41	-1.22	-0.04
NC3_20180311	<b>0.86</b>	0.92	8.79	2.31	0.15	-0.09	-2.77	-1.02	0.65	0.43	-1.80	-0.43
NC4_20180318	0.98	0.98	11.33	10.29	6.08	4.88	-0.97	-1.23	0.70	0.59	-0.68	-0.73
CCR1_20180203	0.93	0.83	5.33	3.37	1.33	0.68	2.46	3.48	0.62	0.53	1.53	1.84
CCR2_20180203	<b>0.73</b>	0.55	8.43	10.93	0.10	6.80	-5.39	32.63	0.36	0.02	-1.93	0.58
QTP1_20180322	0.83	0.75	10.70	10.22	0.77	0.44	1.31	0.11	0.72	0.46	0.94	0.05
QTP2_20180225	<b>0.82</b>	0.86	15.27	13.50	-0.30	0.89	-9.12	-7.49	0.85	0.66	-7.75	-4.95
QTP3_20180320	0.74	0.73	7.91	3.74	-1.49	-0.40	-3.48	-1.07	0.74	0.63	-2.59	-0.67
QTP4_20180401	0.79	0.79	16.73	16.56	-3.71	-4.35	-8.15	-8.14	0.45	0.32	-3.71	-2.61
QTP5_20180307	0.92	0.94	13.87	13.82	7.64	8.01	1.07	1.68	0.59	0.34	0.64	0.57
QTP6_20180305	0.78	0.79	14.54	15.03	4.66	1.75	-2.93	-4.79	0.65	0.23	-1.91	-1.10
QTP7_20180107	0.75	0.98	18.12	15.17	-0.53	10.28	-19.25	-1.00	0.65	0.25	-12.47	-0.25
QTP8_20180128	0.82	0.75	10.79	11.13	2.74	3.27	4.39	7.62	0.64	0.42	2.81	3.19
NX1_20180105	0.85	0.89	24.87	24.47	24.67	24.29	0.07	0.00	0.93	0.44	0.07	0.00
NX2_20180213	0.92	0.95	20.81	20.47	18.29	18.27	7.25	9.00	0.70	0.53	5.05	4.80
NX3_20180220	0.86	0.93	18.78	17.45	16.06	15.37	1.89	4.02	0.72	0.32	1.37	1.28
NX4_20180103	<b>0.74</b>	0.64	28.86	29.54	26.66	28.09	-0.23	0.61	0.49	0.37	-0.11	0.23
NX5_20180220	0.92	0.97	11.99	8.92	2.28	3.26	-7.96	-1.31	0.69	0.37	-5.51	-0.49
Average	0.84	0.85	14.64	13.48	6.98	7.81	-2.24	1.89	0.66	0.40	-1.48	0.76

Note that NCR is the intersecting non-cloud rate. SRD is the difference in snow rate. Red and blue bold values respectively indicate that STAR NDSI is an improvement and degradation compared with TAC NDSI.

#### 8. L 248. Section 3.2. The validation method need to be improved.

The focus of validation in the study should be whether the STAR method is reliable. Therefore, a reasonable verification scheme is to select actually cloud-free the Landsat NDSI maps as a reference maps, then

artificially set a random 20%, 50% or 80% cloud cover (only my suggestion) on the corresponding MODIS data. The different cloud ratio maps were recovery after STAR interpolation and validated by reference Landsat NDSI maps, and the conclusions is convincing.

**Response:** Thank you for the critical comment. Since the validation scheme of this manuscript was designed from a product perspective, we did not include a simulated experiment in the Results. However, we also believe that the effectiveness of STAR method under different cloud conditions is important. Therefore, we added the simulated experiment to the Discussion as follows.

To elaborate the cloud removal effectiveness of the proposed STAR method, the performance statistics under different simulated cloud conditions are shown in Table 9. Four TAC NDSI maps with little cloud cover during the snow period were used in the simulated experiment. Cloud masks from other dates were added to the target maps, with different fractions of about 20%, 50% and 80%. Subsequently, the artificially cloud-covered maps were recovered by STAR and validated by Landsat NDSI maps. The quantitative results indicate that the recovery effectiveness of STAR typically declines significantly when cloud coverage is greater than 80%. As a result, STAR can completely remove clouds with little loss of accuracy. Only in the *NC4\_20180318* scene, high overestimation occurs when cloud coverage reaches 55%. The phenomenon is caused by high cloud coverage and rapid snow variation in space and time. Therefore, users are recommended to refer to the QA maps of STAR NDSI collection during snow accumulation and ablation periods, in which Bit 7 reflects the cloud coverage of the space-time block.

**Table 9. The cloud removal effectiveness of STAR compared to Landsat NDSI maps under different simulated cloud conditions.** Note that blue bold values indicate a significant degradation of the accuracy under the current cloud cover compared to the previous one.

Region_Date	Snow fraction (%)	Added cloud (%)	CC		RMSE		AE		SRD (%)	
			TAC	STAR	TAC	STAR	TAC	STAR	TAC	STAR
NC4_20180318	44%	11%		0.98		10.44		4.62		-1.78
		<b>55%</b>	0.98	<b>0.89</b>	10.29	<b>19.11</b>	4.88	11.17	-1.23	<b>12.96</b>
		80%		0.86		21.24		13.03		16.31
QTP2_20180225	83%	16%		0.85		13.77		0.28		-7.90
		44%	0.86	0.84	13.50	14.21	0.89	-0.31	-7.49	-8.86
		<b>81%</b>		<b>0.78</b>		<b>17.64</b>		<b>-3.72</b>		<b>-17.17</b>
QTP9_20180125	42%	19%		0.90		10.51		0.89		-3.14
		49%	0.89	0.87	10.56	11.34	0.93	0.65	-3.16	-3.64
		<b>80%</b>		<b>0.54</b>		<b>19.35</b>		<b>-9.59</b>		<b>-25.03</b>
NX2_20180213	83%	18%		0.94		20.84		18.64		9.71
		48%	0.95	0.94	20.47	21.26	18.27	18.86	9.00	9.15
		75%		0.93		22.24		19.69		8.18

## References

- Aalstad, K., Westermann, S., and Bertino, L.: Evaluating satellite retrieved fractional snow-covered area at a high-Arctic site using terrestrial photography, *Remote Sens. Environ.*, 239, 111618, <https://doi.org/10.1016/j.rse.2019.111618>, 2020.
- Che, T., Dai, L. Y., Zheng, X. M., Li, X. F., and Zhao, K.: Estimation of snow depth from passive microwave brightness temperature data in forest regions of northeast China, *Remote Sens. Environ.*, 183, 334-349, <https://doi.org/10.1016/j.rse.2016.06.005>, 2016.
- Che, T., Li, X., Liu, S., Li, H., Xu, Z., Tan, J., Zhang, Y., Ren, Z., Xiao, L., Deng, J., Jin, R., Ma, M., Wang, J., and Yang, X.: Integrated hydrometeorological, snow and frozen-ground observations in the alpine region of the Heihe River Basin, China, *Earth Syst. Sci. Data*, 11, 1483-1499, <https://doi.org/10.5194/essd-11-1483-2019>, 2019.
- Chen, S., Wang, X., Guo, H., Xie, P., and Sirelkhatim, A. M.: Spatial and temporal adaptive gap-filling method producing daily cloud-free NDSI time series, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 13, 2251-2263, <https://doi.org/10.1109/JSTARS.2020.2993037>, 2020.
- Crawford, C. J.: MODIS Terra Collection 6 fractional snow cover validation in mountainous terrain during spring snowmelt using Landsat TM and ETM, *Hydrol. Processes*, 29, 128-138, <https://doi.org/10.1002/hyp.10134>, 2015.
- Gao, J., Williams, M. W., Fu, X. D., Wang, G. Q., and Gong, T. L.: Spatiotemporal distribution of snow in eastern Tibet and the response to climate change, *Remote Sens. Environ.*, 121, 1-9, <https://doi.org/10.1016/j.rse.2012.01.006>, 2012.
- Gladkova, I., Grossberg, M., Bonev, G., Romanov, P., and Shahriar, F.: Increasing the accuracy of MODIS/Aqua snow product using quantitative image restoration technique, *IEEE Geosci. Remote Sens. Lett.*, 9, 740-743, <https://doi.org/10.1109/LGRS.2011.2180505>, 2012.
- Hall, D. K., Hall, D. K., Riggs, G. A., Digirolamo, N. E., and Román, M. O.: Evaluation of MODIS and VIIRS cloud-gap-filled snow-cover products for production of an Earth science data record, *Hydrol. Earth Syst. Sci.*, 23, 5227-5241, <https://doi.org/10.5194/hess-23-5227-2019>, 2019.
- Hao, X., Huang, G., Che, T., Ji, W., Sun, X., Zhao, Q., Zhao, H., Wang, J., Li, H., and Yang, Q.: The NIEER AVHRR snow cover extent product over China - A long-term daily snow record for regional climate research, *Earth Syst. Sci. Data*, 13, 4711-4726, <https://doi.org/10.5194/essd-13-4711-2021>, 2021.
- Hao, X., Huang, G., Zheng, Z., Sun, X., Ji, W., Zhao, H., Wang, J., Li, H., and Wang, X.: Development and validation of a new MODIS snow-cover-extent product over China, *Hydrol. Earth Syst. Sci.*, 26, 1937-1952, <https://doi.org/10.5194/hess-26-1937-2022>, 2022.
- Hou, J., Huang, C., Zhang, Y., and Guo, J.: On the Value of Available MODIS and Landsat8 OLI Image Pairs for MODIS Fractional Snow Cover Mapping Based on an Artificial Neural Network, *IEEE Trans. Geosci. Remote Sens.*, 58, 4319-4334, <https://doi.org/10.1109/TGRS.2019.2963075>, 2020.
- Huang, X. D., Deng, J., Ma, X. F., Wang, Y. L., Feng, Q. S., Hao, X. H., and Liang, T. G.: Spatiotemporal dynamics of snow cover based on multi-source remote sensing data in China, *The Cryosphere*, 10, 2453-2463, <https://doi.org/10.5194/tc-10-2453-2016>, 2016.
- Huang, Y., Liu, H. X., Yu, B. L., We, J. P., Kang, E. L., Xu, M., Wang, S. J., Klein, A., and Chen, Y. N.: Improving MODIS snow products with a HMRF-based spatio-temporal modeling technique in the Upper Rio Grande Basin, *Remote Sens. Environ.*, 204, 568-582, <https://doi.org/10.1016/j.rse.2017.10.001>, 2018.
- Jing, Y., Shen, H., Li, X., and Guan, X.: A Two-Stage Fusion Framework to Generate a Spatio-Temporally Continuous MODIS NDSI Product over the Tibetan Plateau, *Remote Sensing*, 11, 2261, <https://doi.org/10.3390/rs11192261>, 2019.
- Kuter, S., Akyurek, Z., and Weber, G. W.: Retrieval of fractional snow covered area from MODIS data by multivariate adaptive regression splines, *Remote Sens. Environ.*, 205, 236-252, <https://doi.org/10.1016/j.rse.2017.11.021>, 2018.
- Li, M., Zhu, X., Li, N., and Pan, Y.: Gap-Filling of a MODIS normalized difference snow index product based on the similar pixel selecting algorithm: A case study on the Qinghai-Tibetan Plateau, *Remote Sensing*, 12, 1077, <https://doi.org/10.3390/rs12071077>, 2020.
- Li, X. H., Jing, Y. H., Shen, H. F., and Zhang, L. P.: The recent developments in cloud removal approaches of MODIS snow cover product, *Hydrol. Earth Syst. Sci.*, 23, 2401-2416, <https://doi.org/10.5194/hess-23-2401-2019>, 2019.

- Liang, H., Huang, X. D., Sun, Y. H., Wang, Y. L., and Liang, T. G.: Fractional Snow-Cover Mapping Based on MODIS and UAV Data over the Tibetan Plateau, *Remote Sensing*, 9, 1332, <https://doi.org/10.3390/rs9121332>, 2017.
- Malmros, J. K., Mernild, S. H., Wilson, R., Tagesson, T., and Fensholt, R.: Snow cover and snow-albedo changes in the central Andes of Chile and Argentina from daily MODIS observations (2000-2016), *Remote Sens. Environ.*, 209, 240-252, <https://doi.org/10.1016/j.rse.2018.02.072>, 2018.
- Muhammad, S., and Thapa, A.: An improved Terra-Aqua MODIS snow cover and Randolph Glacier Inventory 6.0 combined product (MOYDGL06\*) for high-mountain Asia between 2002 and 2018, *Earth Syst. Sci. Data*, 12, 345-356, <https://doi.org/10.5194/essd-12-345-2020>, 2020.
- Muhammad, S., and Thapa, A.: Daily Terra-Aqua MODIS cloud-free snow and Randolph Glacier Inventory 6.0 combined product (M\*D10A1GL06) for high-mountain Asia between 2002 and 2019, *Earth Syst. Sci. Data*, 13, 767-776, <https://doi.org/10.5194/essd-13-767-2021>, 2021.
- Riggs, G. A., and Hall, D. K.: MODIS Snow Products Collection 6 User Guide, 2015.
- Riggs, G. A., Hall, D. K., and Roman, M. O.: Overview of NASA's MODIS and Visible Infrared Imaging Radiometer Suite (VIIRS) snow-cover Earth System Data Records, *Earth Syst. Sci. Data*, 9, 765-777, <https://doi.org/10.5194/essd-9-765-2017>, 2017.
- Tang, Z. G., Wang, J., Li, H. Y., and Yan, L. L.: Spatiotemporal changes of snow cover over the Tibetan plateau based on cloud-removed moderate resolution imaging spectroradiometer fractional snow cover product from 2001 to 2011, *J. Appl. Remote Sens.*, 7, 073582, <https://doi.org/10.1117/1.jrs.7.073582>, 2013.
- Tong, R., Parajka, J., Komma, J., and Blöschl, G.: Mapping snow cover from daily Collection 6 MODIS products over Austria, *J. Hydrol.*, 590, 125548, <https://doi.org/10.1016/j.jhydrol.2020.125548>, 2020.
- Wang, X. W., Xie, H. J., and Liang, T. G.: Evaluation of MODIS snow cover and cloud mask and its application in Northern Xinjiang, China, *Remote Sens. Environ.*, 112, 1497-1513, <https://doi.org/10.1016/j.rse.2007.05.016>, 2008.
- Yu, J. Y., Zhang, G. Q., Yao, T. D., Xie, H. J., Zhang, H. B., Ke, C. Q., and Yao, R. Z.: 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, <https://doi.org/10.1109/tgrs.2015.2496950>, 2016.
- Zhang, G., Xie, H., Yao, T., Liang, T., and Kang, S.: Snow cover dynamics of four lake basins over Tibetan Plateau using time series MODIS data (2001-2010), *Water Resour. Res.*, 48, W10529, <https://doi.org/10.1029/2012WR011971>, 2012.
- Zhang, H., Zhang, F., Zhang, G., Che, T., Yan, W., Ye, M., and Ma, N.: Ground-based evaluation of MODIS snow cover product V6 across China: Implications for the selection of NDSI threshold, *Sci. Total Environ.*, 651, 2712-2726, <https://doi.org/10.1016/j.scitotenv.2018.10.128>, 2019.
- Zhang, H., Zhang, F., Zhang, G., Yan, W., and Li, S.: Enhanced scaling effects significantly lower the ability of MODIS normalized difference snow index to estimate fractional and binary snow cover on the Tibetan Plateau, *J. Hydrol.*, 592, 125795, <https://doi.org/10.1016/j.jhydrol.2020.125795>, 2021.