Response to Referee #2 Comments

General Comment:

The manuscript titled "STAR NDSI collection: A cloud-free MODIS NDSI dataset (2001–2020) for China" estimates cloud-free snow data for China. The authors use Spatio-Temporal Adaptive fusion method with erroR correction (STAR) to derive snow cover. Cloud cover is the main obstacle in passive remote sensing snow monitoring and is important to overcome. The study is important but there are few major issues in the present form which needs to be addressed.

General Response:

Thank you very much for the critical comments and suggestions regarding our article. Considering all the constructive comments, we carefully revised the manuscript and supplemented a comparative experiment. Based on the experimental results, the priority scheme of the pre-process TAC and the accuracy of STAR NDSI collection were discussed in detail.

In summary, the main points of the response include: (1) Combined with the comments of referee #1, we searched various existing cloud-free snow cover products covering China and added descriptions in the Introduction. **(2)** A comparative experiment between the modified Terra and Aqua combination (TAC) and the original TAC was performed. **(3)** The uncertainty caused by the pre-process TAC and the accuracy of STAR NDSI collection were analyzed. **(4)** The proposed cloud removal method was re-emphasized. **(5)** The code was uploaded to Zenode.

Major comments:

Q1. The authors use combined Terra and Aqua MODIS data in this manuscript. They combine the data first and then use STAR method consisting of spatio-temporal adaptive fusion (STAF) and error correction (EC). Combining Terra and Aqua this way potentially overestimates snow (Muhammad and Thapa, 2020, 2021). The authors are suggested to either revise the TAC or explain the potential uncertainty.

Response: Thank you for the critical comment. Firstly, we carefully learned the cloud removal method proposed by Muhammad and Thapa (2020, 2021). Then, we analyzed the necessity of TAC using the maximum strategy. Finally, we performed a comparative experiment between the modified TAC and the original TAC to demonstrate the analysis. The priority schemes of the modified TAC and the original TAC are determined as *low* *value* > *high value* > *cloud* and *high value* > *low value* > *cloud*, respectively.

Muhammad and Thapa (2020, 2021) suggested merging Terra and Aqua 8-day binary snow cover products in a way that considered snow only where pixels in both the products are classified as snow (i.e., *no snow* > *snow* > *cloud*). A multi-step combination method was used to remove clouds from MODIS snow cover area (SCA) products, consisting of seasonal filtering, temporal filtering, spatial filtering, and TAC. This priority scheme of TAC is an inter-verification of Terra and Aqua 8-day snow cover products which generally overestimate the snow cover extent. It also helps to avoid uncertainty produced using spatial filtering. Therefore, it is superior to the priority scheme of *snow* > *no snow* > *cloud*. However, they also suggested that "**We do not recommend this merging criteria for daily snow products** in mountainous areas due to the error of omission which may be further increased because of the off-nadir-view acquisition and edge pixels".

For STAR NDSI collection, the pre-process TAC is used to combine the Terra and Aqua MODIS daily NDSI datasets. Since there is no significant overestimation in the daily original datasets and no uncertainty caused by spatial filtering, the priority scheme is set as *high value* > *low value* > *cloud*. The four quantitative evaluations in the manuscript are shown below (Revised results based on the comments of Referee #1). Compared with NIEER AVHRR SCE (Hao et al., 2021) and MODIS CGF NDSI (Hall and Riggs, 2020) products, our STAR NDSI collection presents optimal classification accuracy and numerical accuracy. All the results reveal that our STAR NDSI collection tends to slightly underestimate rather than overestimate snow cover area. The omission errors (OEs) in Table 4 and Table 5 are relatively significant and the differences in snow rate (SRDs, MODIS − Landsat) in Table 6 and Table 7 are generally negative. These findings are in line with their assessment of daily snow products (Muhammad and Thapa, 2020). Particularly, the absolute errors (AEs, MODIS − Landsat) in the evaluations based on Landsat NDSI maps are generally positive. This is mainly due to the different spectral response curves of the bands on different sensors.

3.1 Validation against in-situ snow depth measurements

Table 4. Classification statistics based on in XJ.

Indicators	Snow depth > 0 cm (Snow fraction = 30%)			Snow depth > 1 cm (Snow fraction = 28%)		
	TAC	NIEER	STAR	TAC	NIEER	STAR
OA	0.81	0.94	**0.95**	0.82	0.94	**0.95**
CE	0.02	0.02	0.04	0.02	0.03	0.05
OE	0.60	0.17	**0.07**	0.58	0.15	**0.05**

Table 5. Classification statistics over TP.

Indicators	Snow depth > 0 cm (Snow fraction = 5%)			Snow depth > 1 cm (Snow fraction = 3%)		
	TAC	NIEER	STAR	TAC	NIEER	STAR
OA	0.94	0.93	**0.95**	0.96	0.94	**0.96**
CE	0.01	0.04	0.02	0.02	0.05	0.03
OE	0.78	0.53	**0.52**	0.72	0.42	**0.39**

Note: TAC, NIEER, and STAR represent TAC NDSI, NIEER AVHRR SCE, and STAR NDSI datasets, respectively.

3.2 Validation based on Landsat NDSI maps

Table 6. Performance statistics for TAC NDSI, MODIS CGF NDSI, and STAR NDSI datasets against Landsat NDSI maps.

Region_Date	Cloud cover (%)	CC			RMSE			AE			SRD (%)		
		TAC	CGF	STAR	TAC	CGF	STAR	TAC	CGF	STAR	TAC	CGF	STAR
NC1_20180225	61.4	0.49	0.90	0.87	25.64	20.20	17.10	-11.59	18.35	15.07	-60.1	1.0	1.0
NC2_20180311	43.6	0.69	0.72	0.83	26.75	19.77	13.87	-9.36	15.72	12.14	-43.1	-5.7	-1.7
NC3_20180311	34.6	0.19	0.86	0.86	18.61	20.97	8.79	-10.26	10.78	0.15	-34.2	0.0	-2.8
NC4_20180318	16.3	0.73	0.95	0.98	21.50	15.03	10.25	-3.33	8.26	5.42	-16.6	0.5	-1.1
CCR1_20180203	14.9	0.20	0.93	0.95	11.53	14.91	5.04	-3.56	6.55	1.54	-11.4	4.2	2.8
CCR2_20180203	95.0	-0.07	0.52	0.73	14.26	38.77	8.43	-9.07	32.76	0.10	-47.8	29.8	-5.4
TP1_20180322	36.3	0.39	0.75	0.83	18.03	14.26	10.70	-5.30	2.02	0.77	-13.8	1.2	1.3
TP2_20180225	22.4	0.54	0.71	0.82	25.50	17.92	15.27	-9.66	-2.62	-0.30	-27.8	-11.5	-9.1
TP3_20180320	15.3	0.29	0.31	0.74	11.40	11.64	7.91	-2.89	-2.37	-1.49	-6.7	-5.9	-3.5
TP4_20180401	29.5	0.47	0.51	0.79	30.94	29.60	16.64	-16.86	-16.36	-3.71	-31.6	-29.2	-8.3
TP5_20180307	42.5	0.47	0.92	0.92	30.00	13.80	13.80	-11.76	7.67	7.67	-36.0	1.0	1.0
TP6_20180305	64.9	0.17	0.76	0.78	42.26	17.26	14.53	-32.30	-4.49	4.67	-66.1	-10.9	-2.9
TP7_20180107	60.8	0.44	0.75	0.80	28.13	18.09	18.08	-15.95	-0.48	6.04	-60.0	-19.2	-12.6
TP8_20180128	34.6	0.49	0.38	0.82	11.98	28.32	10.65	0.67	-4.91	2.68	1.1	-47.7	4.3
XJ1_20180105	52.2	0.79	0.89	0.86	27.53	22.78	22.81	-4.00	20.21	22.18	-52.2	8.4	0.1
XJ2_20180213	23.4	0.64	0.82	0.92	23.55	10.65	20.81	7.13	2.68	18.29	-14.8	4.3	7.2
XJ3_20180220	56.0	0.56	0.73	0.86	26.35	24.83	18.78	-9.87	21.24	16.06	-51.2	2.8	1.9
XJ4_20180103	23.5	0.70	0.65	0.74	28.94	25.58	28.86	15.70	22.93	26.66	-21.8	-1.1	-0.2
XJ5_20180220	46.1	0.55	0.87	0.92	23.55	15.95	11.99	-10.44	5.59	2.28	-32.7	-4.1	-8.0
Average	40.7	**0.46**	**0.73**	**0.84**	**23.50**	**20.02**	**14.44**	**-7.51**	**7.55**	**7.17**	**-33.0**	**-4.3**	**-1.9**

Note: TAC, CGF, and STAR represent TAC NDSI, MODIS CGF NDSI, and STAR NDSI datasets, respectively.

Table 7. Performance statistics for STAR NDSI collection against Landsat NDSI maps in clear-sky and cloud-cover areas according to the TAC dataset. Note: Red and blue bold values respectively indicate an improvement and degradation of cloud-cover areas compared with clear-sky areas (corresponding to four groups in Fig. 4).

Region_Date	CC		RMSE		AE		SRD (%)	
	Clear-sky	Cloud-cover	Clear-sky	Cloud-cover	Clear-sky	Cloud-cover	Clear-sky	Cloud-cover
NC1_20180225	0.95	0.76	16.89	17.23	15.23	14.97	2.91	-0.2
NC2_20180311	0.89	0.71	12.27	15.71	11.78	12.62	-0.10	-3.7
NC3_20180311	**0.92**	**0.42**	2.31	14.60	-0.09	0.61	-1.02	-6.1
NC4_20180318	0.98	0.86	10.29	14.75	4.88	11.30	-1.23	-0.2
CCR1_20180203	0.83	0.68	3.37	10.29	0.68	6.50	3.48	-0.9
CCR2_20180203	**0.55**	**0.74**	10.93	8.28	6.80	-0.25	32.63	-7.4
TP1_20180322	0.75	0.87	10.22	11.49	0.44	1.35	0.11	3.4
TP2_20180225	**0.86**	**0.64**	13.50	20.24	0.89	-4.40	-7.49	-14.8
TP3_20180320	0.73	0.69	3.74	18.19	-0.40	-7.52	-1.07	-16.9
TP4_20180401	0.79	0.78	16.56	17.08	-4.35	-2.17	-8.14	-8.3
TP5_20180307	0.94	0.88	13.82	13.86	8.01	7.16	1.68	0.1

TP6_20180305	0.79	0.76	15.03	14.26	1.75	6.24	-4.79	-1.9
TP7_20180107	0.98	0.63	15.17	19.74	10.28	3.27	-1.00	-20.0
TP8_20180128	0.75	0.89	11.13	9.92	3.27	1.67	7.62	-1.7
XJ1_20180105	0.89	0.62	24.47	21.17	24.29	20.25	0.00	0.1
XJ2_20180213	0.95	0.74	20.47	21.88	18.27	18.37	9.00	1.5
XJ3_20180220	0.93	0.75	17.45	19.77	15.37	16.60	4.02	0.2
XJ4_20180103	**0.64**	**0.58**	29.54	26.53	28.09	22.00	0.61	-2.9
XJ5_20180220	0.97	0.86	8.92	14.81	3.26	1.13	-1.31	-15.7
Average	0.85	0.73	13.48	16.30	7.81	6.83	1.89	-5.0

A comparative experiment was performed to further demonstrate the effectiveness of the original TAC, as shown in Supplementary Table 1 (not added to the manuscript). In this experiment, the priority schemes of the modified TAC and the original TAC are determined as *low value* > *high value* > *cloud* and *high value* > *low value* > *cloud*, respectively. Combining Terra and Aqua daily NDSI products with a minimum strategy, the average AE and RMSE are slightly decreased by 1.8 and 0.21, respectively. However, the SRDs are worsened compared with Landsat NDSI maps, especially on Tibetan Plateau (TP). The OEs are also expected to be worsened compared with in-situ snow depth measurements. In addition, this underestimation of snow cover area will be further amplified after the key-process STAR due to less snow cover information. As a result, we also do not recommend the pre-process TAC with a minimum strategy in the cloud removal of daily MODIS NDSI products.

Supplementary Table 1. Performance statistics for the modified TAC (TAC_MIN) and the original TAC (TAC_MAX) in clear-sky areas according to the TAC dataset. Note: Red bold values indicate a significant degradation of the modified TAC compared with the original TAC used for STAR NDSI collection. Purple and bold Regions/Dates are the newly added Regions/Dates on TP.

Region_Date	CC		RMSE		AE		SRD (%)	
	TAC_MAX	TAC_MIN	TAC_MAX	TAC_MIN	TAC_MAX	TAC_MIN	TAC_MAX	TAC_MIN
NC1_20180225	0.95	0.95	16.76	16.44	15.17	14.27	2.91	0.17
NC2_20180311	0.90	0.90	12.21	12.20	11.75	11.74	-0.08	-0.08
NC3_20180311	0.92	0.92	2.34	2.34	-0.09	-0.12	-1.03	-1.24
NC4_20180318	0.98	0.98	10.34	9.52	4.90	4.11	-1.20	-1.99
CCR1_20180203	0.83	0.88	3.35	2.37	0.71	0.09	3.67	-0.25
CCR2_20180203	0.54	0.56	11.00	10.40	6.90	6.34	33.43	31.59
TP1_20180322	0.75	0.76	10.26	9.43	0.51	-1.01	**0.51**	**-4.07**
TP2_20180225	0.86	0.82	13.50	15.97	0.86	-4.84	**-7.51**	**-18.59**
TP3_20180320	0.75	0.75	3.69	3.77	-0.40	-0.57	-1.10	-1.93
TP4_20180401	0.80	0.69	16.31	22.89	-4.40	-11.44	**-8.42**	**-20.90**
TP5_20180307	0.94	0.94	13.57	12.94	8.21	6.26	1.83	-1.97
TP6_20180305	0.80	0.76	14.87	16.97	1.79	-0.03	**-4.75**	**-8.12**
TP7_20180107	0.98	0.98	15.13	14.51	10.27	9.16	-0.93	-3.81
TP8_20180128	0.76	0.82	10.58	7.96	3.12	1.48	7.51	2.36
XJ1_20180105	0.89	0.85	24.47	24.07	24.28	23.88	0.00	0.00
XJ2_20180213	0.95	0.95	20.53	18.46	18.35	16.35	9.07	8.05

XJ3_20180220	0.93	0.91	17.49	16.21	15.50	13.90	4.15	3.74
XJ4_20180103	0.67	0.65	29.53	23.54	28.18	22.23	0.62	0.56
XJ5_20180220	0.98	0.97	8.63	7.75	3.15	2.46	-1.24	-1.79
TP9_20180125	0.90	0.84	10.35	11.96	0.93	-3.29	**-3.16**	**-12.33**
TP10_20180305	0.83	0.80	12.97	14.15	-3.44	-5.08	**-5.73**	**-10.40**
TP11_20180107	0.93	0.94	12.14	11.70	4.74	3.93	**-0.79**	**-5.08**
TP12_20180312	0.97	0.97	10.24	8.86	5.11	2.85	1.13	-0.41
TP13_20180307	0.94	0.94	12.12	12.12	5.30	4.61	**-1.76**	**-4.63**
TP14_20180226	0.82	0.81	12.11	12.76	-4.72	-5.41	**-5.73**	**-8.67**
Average	**0.86**	**0.85**	**12.98**	**12.77**	**6.27**	**4.47**	**0.86**	**-2.39**

Q2. The authors missed to share the code to generate STAR NDSI dataset. It is incomplete without sharing the code. The code is also required to evaluate the methodology as well.

Response: Thank you for the critical comment. We uploaded the code to Zenode and added a description in the Data availability as follows. We promise to update Closed Access to Open Access when the manuscript is accepted.

The improved cloud-free MODIS NDSI collection (STAR NDSI collection) for China from 1 August 2000 to 31 July 2020, including STAR NDSI and STAR QA data, is available for download at https://doi.org/10.5281/zenodo.5644386 (Jing et al., 2021). The dataset is provided using a WGS 84 / UTM zone 48N projection, with a tag image file format (TIFF). Users can discuss and respond to issues that arise during the use of this dataset. New versions can be released in consideration of user comments. **In addition, a source code for this collection is available at https://doi.org/10.5281/zenodo.6396149** (Jing, 2022).

Q3. The C6 snow is in NDSI ranging between 0 and 100. It is not explained how the authors reconstructed the snow data. It is a challenge to improve the data on how to replace the cloudy pixel, so it is significant to understand the way the value is replaced.

Response: Thank you for the comment. 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. This cloud removal method is detailed in the Algorithm description (Section 2.2).

For the methodology, the clouds are generally removed by a multi-step combination method in existing snow cover products. An accumulated cloud-free result can be obtained by combining the spatial methods and temporal methods alternately. However, the independent and successive utilization cannot take full consideration of the spatio-temporal information. Therefore, STAR NDSI collection is generated by a novel Spatio-Temporal Adaptive fusion method with erroR correction (STAR), which consists of space partition, adaptive space-time block determination, Gaussian kernel function-based fusion, and error correction. This method comprehensively considers spatial and temporal contextual information and thus is promising for the cloud removal of NDSI products. The cloud removal procedure is described as follows.

2.2 Algorithm description

MODIS NDSI datasets are unable to represent the daily conditions of snow accumulation and ablation accurately because the optical remote-sensed images are subject to severe cloud pollution. Therefore, a Spatio-Temporal Adaptive fusion method with erroR correction (STAR), which is derived from our two-stage spatio-temporal fusion method (Jing et al., 2019), is presented to produce a spatio-temporal continuous snow collection. As shown in Fig. 2, the generation procedure comprises the pre-process TAC and the key-process STAR. Then, a quality assessment (QA) approach is presented to provide a data reliability profile for users. On this basis, post-processing is used to further improve the data quality in individual abnormal areas.

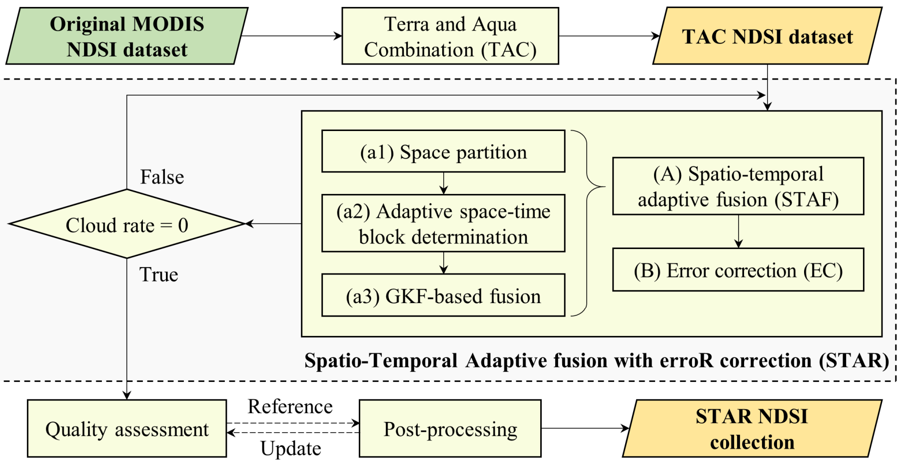

Figure 2. Schematic of the generation procedure of STAR NDSI collection.

2.2.1 Terra and Aqua combination (TAC)

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). 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). TAC can efficiently decrease the cloud fraction by 5%–20% with negligible precision sacrifice (Li et al., 2019). Thus, this method is introduced as a pre-processing to reduce cloud coverage preliminarily. Its priority scheme is determined as high value > low value > cloud. Particularly, 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.

2.2.2 Spatio-Temporal Adaptive fusion with erroR correction (STAR)

Many regions with persistent clouds are out of the scope of TAC. To this end, an advanced STAR method, which comprehensively utilizes spatio-temporal contextual information, is proposed to remove the clouds thoroughly. As shown in Fig. 3, the method performs in two passes: spatio-temporal adaptive fusion (STAF) and error correction (EC).

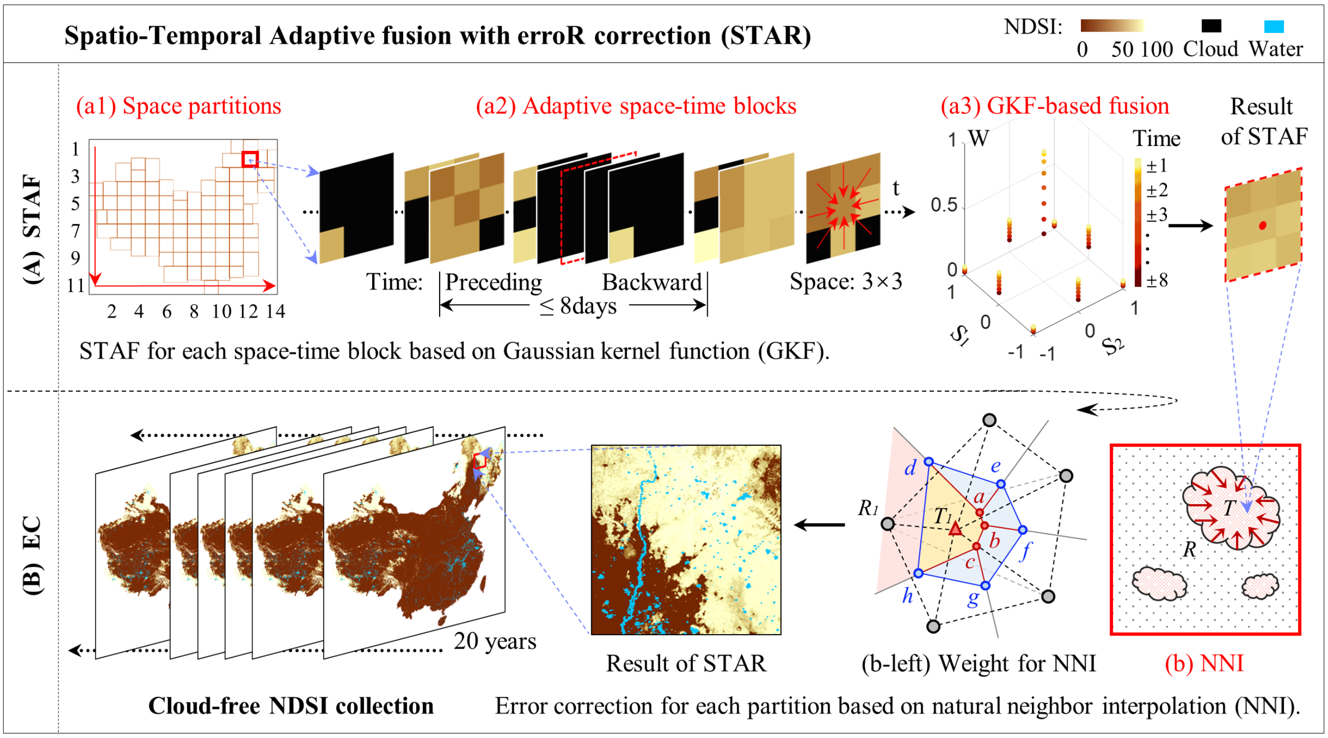

Figure 3. Detailed flowchart of the Spatio-Temporal Adaptive fusion with erroR correction (STAR).

The first pass involves the generation of new NDSI maps by adaptively merging the spatio-temporal contextual information, including space partition, adaptive space-time block determination, and Gaussian kernel function (GKF)-based fusion. The research area is first segmented into dozens of partitions considering the spatial heterogeneity of snow patterns. Thus, the subsequent processes can be performed on a partition basis. Moreover, the optimal query partitions (*Q*) to each target partition (*T*) are determined by a comprehensive consideration of temporal distance (*t*), regional correlation (*r*), and cloud-free fraction (*f*) concerning the temporal complexity of snow variations. The following optimal parameters are derived from the extensive experiments.

$$\begin{cases} Scheme\ 1:\ r > 0.7,\ if\ f^{C\&T} > 0.3 \\ Scheme\ 2: \max(t^{-1} + f^{C}),\ \text{others} \end{cases}, \quad (1)$$

where the regional correlation between the candidate partition (*C*) within an eight-day window and the target partition is considered representative if the fraction of the intersecting cloud-free areas ($f^{C\&T}$) is higher than 0.3.

The candidate partition is then determined as a query partition according to Scheme 1 when the regional correlation is larger than 0.7. Otherwise, Scheme 2 is activated. Two query partitions with short distance and high cloud-free fraction are identified within the preceding eight days and the backward eight days, respectively. Subsequently, the 3 × 3 neighborhoods for each pixel of the target partition in all the associated query partitions are determined as the space-time reference block. Last, each pixel is reassigned a fused value from the related space-time block, as expressed in Eq. (2):

$$NDSI_i^F = \sum_{t=1}^{M}\sum_{s=1}^{N} w_{(i,st)} \times NDSI_{(i,st)}^Q, \quad where\ \ \mathrm{W}_{(i,st)} = r_t^2 \times \exp\left(\frac{-((\varepsilon \times \Delta s_{(i,s)})^2 + \Delta t_{(i,t)}^2))}{2 \times \sigma^2}\right), \tag{2}$$

where $NDSI_i^F$ denotes the fused NDSI of Pixel i in the target partition. $NDSI_{(i,st)}^Q$ is the pre-processed NDSI in associated query partitions. M is the number of query partitions, each of which contains N reference pixels. In addition, the weight $\mathrm{W}_{(i,st)}$ is assigned by a two-dimensional GKF involving the spatial distance ($\Delta s_{(i,s)}$) and the temporal distance ($\Delta t_{(i,t)}$), which is then normalized to $w_{(i,st)}$. σ is the standard deviation of GKF. ε characterizes the dimensional difference, which is equal to σ_t/σ_s with an expression of each single-dimensional GKF. r_t represents the regional correlation between the query and target partitions if Scheme 1 works; otherwise, it is ignored (i.e., $r_t = 1$). The constant term ($\varepsilon/(2\pi\sigma^2)$) of GKF is ignored due to the normalization process. The important parameters in STAF are listed in Table 2.

Table 2. Description and default values of STAF parameters.

Parameter	Description	Value
W_T	Temporal window for query partition	±8 day
W_N	Neighboring window for reference pixel	3 × 3
r	Minimum regional correlation for query partition	0.7
σ	Standard deviation in the GKF	0.5
ε	Dimensional difference coefficient (σ_t/σ_s)in the GKF	25/9

The second pass corrects the fused NDSI maps considering the spatial correlation within a partition. Specifically, the residual errors of the intersecting cloud-free areas of the pre-processed and fused NDSI maps (refer to as $NDSI^P$ and $NDSI^F$) are diffused to other cloud-free areas of the fused NDSI maps using the triangulation-based natural neighbor interpolation (Sibson, 1981). Then, the high-quality NDSI maps ($NDSI^H$) can be generated by removing all errors from the fused NDSI maps. The process is formulated as follows:

$$\begin{cases} E_R = NDSI_R^F - NDSI_R^P \\ E_{T(i)} = \sum_{n=1}^{N'} \phi_{(i,n)} E_{R(i,n)}, \\ NDSI_T^H = NDSI_T^F - E_T \end{cases} \tag{3}$$

where R indicates the reference area which is the boundary of the intersecting cloud-free areas. T indicates the target area. The dynamic weights in the error diffusion from E_R to E_T are based on the Voronoi diagrams. As expressed in Fig. 3 (b-left), the original Voronoi cells (bounded by red and gray solid lines) of the reference pixels (gray dots) intersect with the new Voronoi cells (bounded by blue and gray solid lines) of the reference and target pixels. Taking the target pixel T_1 with the reference pixel R_1 as an example, the weight is assigned as the ratio of the area of the intersecting Voronoi cell (A_{dabch}) to that of the new Voronoi cell (A_{defgh}).

$$\phi_{(1,1)} = \frac{A_{dabch}}{A_{defgh}}. \tag{4}$$

After all the partitions are processed in sequence, the next iteration of STAR begins until the clouds are completely removed.

Q4. The authors indicate they have derived data between 2000 and 2020. The Aqua data is available from July 2002, the authors should clearly mention the observed period. As the data is combined Terra and Aqua, therefore, it should be between 2002 and 2020 not starting from the year 2000.

Response: Thank you for the comment. Since the Aqua data is available from July 2002, cloud-free NDSI data before July 2002 cannot be produced if only TAC is used. In this study, TAC is introduced as a pre-processing to reduce cloud coverage preliminarily, since it can efficiently decrease the cloud fraction by 5%–20% with negligible precision sacrifice (Li et al., 2019). However, this pre-processing is not essential. We used the key-process STAR to completely remove clouds from Terra MODIS NDSI dataset between 1 August 2000 and 3 July 2002. Then, both the pre-process TAC and the key-process STAR are used to produce cloud-free NDSI data with Terra and Aqua MODIS NDSI datasets. Consequently, an improved cloud-free NDSI collection (STAR NDSI collection) for China from 1 August 2000 to 31 July 2020 can be generated. We added a clear description to the manuscript as follows.

(Lines 145-146) 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.

Q5. One of the major issues is the remaining overestimation. The authors have to consider the existence of overestimation mainly due to the larger solar zenith angle. It is, therefore, necessary to estimate the overestimation in the combined Terra and Aqua as in the combined product the uncertainty increases.

Response: Thank you for the critical comment. More rigorous evaluation and more detailed analysis were added to the Results (Section 3) and shown in the response to Q1. The related discussion is as follows:

A comparative experiment was performed to further demonstrate the effectiveness of the original TAC, as shown in Supplementary Table 1 (not added to the manuscript). In this experiment, the priority schemes of the modified TAC and the original TAC are determined as *low value* > *high value* > *cloud* and *high value* > *low value* > *cloud*, respectively. Combining Terra and Aqua daily NDSI products with a minimum strategy, the average AE and RMSE are slightly decreased by 1.8 and 0.21, respectively. However, the SRDs are worsen compared with Landsat NDSI maps, especially on Tibetan Plateau (TP). The OEs are also expected to be worsen compared with in-situ snow depth measurements. In addition, this underestimation of snow cover area will be further amplified after the key-process STAR due to less snow cover information. As a result, we also do not recommend the pre-process TAC with a minimum strategy in the cloud removal of daily MODIS NDSI products.

Supplementary Table 1. Performance statistics for the modified TAC (TAC_MIN) and the original TAC (TAC_MAX) in clear-sky areas according to the TAC dataset. Note: Red bold values indicate a significant degradation of the modified TAC compared with the original TAC used for STAR NDSI collection. Purple and bold Regions/Dates are the newly added Regions/Dates on TP.

Region_Date	CC		RMSE		AE		SRD	
	TAC_MAX	TAC_MIN	TAC_MAX	TAC_MIN	TAC_MAX	TAC_MIN	TAC_MAX	TAC_MIN
NC1_20180225	0.95	0.95	16.76	16.44	15.17	14.27	2.91	0.17
NC2_20180311	0.90	0.90	12.21	12.20	11.75	11.74	-0.08	-0.08
NC3_20180311	0.92	0.92	2.34	2.34	-0.09	-0.12	-1.03	-1.24
NC4_20180318	0.98	0.98	10.34	9.52	4.90	4.11	-1.20	-1.99
CCR1_20180203	0.83	0.88	3.35	2.37	0.71	0.09	3.67	-0.25
CCR2_20180203	0.54	0.56	11.00	10.40	6.90	6.34	33.43	31.59
TP1_20180322	0.75	0.76	10.26	9.43	0.51	-1.01	**0.51**	**-4.07**
TP2_20180225	0.86	0.82	13.50	15.97	0.86	-4.84	**-7.51**	**-18.59**
TP3_20180320	0.75	0.75	3.69	3.77	-0.40	-0.57	-1.10	-1.93
TP4_20180401	0.80	0.69	16.31	22.89	-4.40	-11.44	**-8.42**	**-20.90**
TP5_20180307	0.94	0.94	13.57	12.94	8.21	6.26	1.83	-1.97
TP6_20180305	0.80	0.76	14.87	16.97	1.79	-0.03	**-4.75**	**-8.12**
TP7_20180107	0.98	0.98	15.13	14.51	10.27	9.16	-0.93	-3.81
TP8_20180128	0.76	0.82	10.58	7.96	3.12	1.48	7.51	2.36
XJ1_20180105	0.89	0.85	24.47	24.07	24.28	23.88	0.00	0.00
XJ2_20180213	0.95	0.95	20.53	18.46	18.35	16.35	9.07	8.05
XJ3_20180220	0.93	0.91	17.49	16.21	15.50	13.90	4.15	3.74
XJ4_20180103	0.67	0.65	29.53	23.54	28.18	22.23	0.62	0.56
XJ5_20180220	0.98	0.97	8.63	7.75	3.15	2.46	-1.24	-1.79
TP9_20180125	0.90	0.84	10.35	11.96	0.93	-3.29	**-3.16**	**-12.33**
TP10_20180305	0.83	0.80	12.97	14.15	-3.44	-5.08	**-5.73**	**-10.40**
TP11_20180107	0.93	0.94	12.14	11.70	4.74	3.93	**-0.79**	**-5.08**
TP12_20180312	0.97	0.97	10.24	8.86	5.11	2.85	1.13	-0.41
TP13_20180307	0.94	0.94	12.12	12.12	5.30	4.61	**-1.76**	**-4.63**
TP14_20180226	0.82	0.81	12.11	12.76	-4.72	-5.41	**-5.73**	**-8.67**
Average	**0.86**	**0.85**	**12.98**	**12.77**	**6.27**	**4.47**	**0.86**	**-2.39**

Overall, it can be concluded that the uncertainty caused by the pre-process TAC is slight in this study. Undoubtedly, other approaches are likely to be valuable for different research areas and applications, such as Terra NDSI data being used alone and TAC with a minimum strategy.

Minor comments:

Q1. Line 45-65: The authors missed to point out one of the most recent cloud-free 8-day (Muhammad and Thapa, 2020 - https://doi.org/10.5194/essd-12-345-2020) and daily (Muhammad and Thapa 2021 - https://doi.org/10.5194/essd-13-767-2021) snow data, combining Terra and Aqua satellites data, reducing up to 50% of uncertainty. These datasets uniquely combine Terra and Aqua, to avoid overestimation after temporal and spatial filters are applied to individual products for clouds removal. The authors are advised to add these important papers.

Response: Thank you very much for the suggestion. Combined with the suggestion of Referee #1, we added and listed the existing important snow cover products in the Introduction as follows.

…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 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) 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. 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.

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…

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. (2020)	SCA	China	2000–2020	Daily	~500 m	10.12072/ncdc.I-SNOW.db0001.2020
Hao et al. (2021)	SCA	China	1981–2019	Daily	~5 km	10.11888/Snow.tpdc.271381
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%.

Q2. MODIS is onboard on Terra and Aqua satellites, the authors are advised to clearly mention which constellation they use in e.g. to clearly mention in A daily spatio-temporal continuous MODIS C6 NDSI dataset with a spatial resolution of 500 m for China (Fig. 1) from 2001 to 2020 is generated for the first time.

Response: Thank you very much for the suggestion. We added multiple descriptions of the constellations to the manuscript as follows.

(Lines 14-16) In this study, a recent 20-year stretch seamless **Terra–Aqua** MODIS NDSI collection in China is generated using a Spatio-Temporal Adaptive fusion method with erroR correction (STAR), which comprehensively considers spatial and temporal contextual information.

(Lines 101-102) 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.

(Lines 145-146) 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.

(Lines 379-380) The improved cloud-free **Terra–Aqua** MODIS NDSI collection (STAR NDSI collection) for China from 1 August 2000 to 31 July 2020.

(Lines 385-386) STAR NDSI collection is derived from **Terra–Aqua** MODIS NDSI datasets using an optimized STAR from our last research.

References

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.

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.

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.

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.

Jing, Y. H., Li, X. H., and Shen, H. F.: STAR NDSI collection: A cloud-free MODIS NDSI dataset (2001–2020) for China (Version 01) [Data set], Zenodo, https://doi.org/10.5281/zenodo.5644386, 2021.

Jing, Y. H.: A spatio-temporal adaptive fusion method with error correction for cloud-free MODIS NDSI estimation (Version 01) [Software], Zenodo, https://doi.org/10.5281/zenodo.6396149, 2022.

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.

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.

Sibson, R.: A brief description of natural neighbor interpolation (Chapter 2), in: Interpolating Multivariate Data, edited by: Barnett, V., John Wiley, Chichester, 21-36, 1981.