Response to Reviewer #2

Comments to the Author:

Snow cover plays an essential role in climate change and the hydrological cycle of the Tibetan Plateau. Currently optical sensors are severely affected by clouds, resulting in a gap in snow products. Using MODIS snow cover product and HMRF algorithm, this work produced daily cloud-free snow cover dataset from 2002 to 2021 over the Tibetan Plateau. In order to validate the accuracy of the dataset, the authors used snow depth data from ground meteorological stations and Landsat-8 images as reference data to systematically evaluate the accuracy of snow products produced from different altitudes and slopes. And this work improved the elevation representing environmental information of the original HMRF model with solar radiation based on the experience of actual field experiments, and the validation results showed its great effect on the accuracy improvement. However, there are still some issues needed to be justified clearly.

Response: We are very grateful for the constructive suggestions and comments from the reviewer. We have significantly improved the manuscript in this revision. In this revision, we have refitted the empirical relationship between snow fraction and NDSI, and reprocessed the input data for HMRF modeling. We have regenerated a more rigorous daily gap-free snow cover dataset. In addition, we have added longer time series and terrain-corrected Landsat images for validation, including Landsat-5 TM, Landsat-7 ETM+, and Landsat-8 OLI images. The new accuracy assessment demonstrates the effect and potential applications of our new daily snow cover dataset. Please see our responses below.

The threshold of NDSI used in this work is set as 0.4, while in the work of Zhang et al., (2020), they used the value of NDSI as 0.1 to determine snow or not in the Tibetan Plateau. So, I'd like suggest the authors have to compare these two threshold on the determination of snow cover in Tibetan Plateau.

Reference: Zhang, H., Zhang, F., Che, T., & Wang, S. (2020). Comparative evaluation of VIIRS daily snow cover product with MODIS for snow detection in China based on ground observations. Science of The Total Environment, 724, 138156.

Response: We appreciate the reference and suggestion from the reviewer. In this revision, we first compared the extracted results by using the threshold of NDSI as 0.1 and 0.4 with *in situ* observation (Table R1). The overall accuracy of NDSI with the threshold of 0.4 (97.39%) was higher than that of NDSI with the threshold of 0.1 (95.24%). However, we also found that the accuracy of snow category with NDSI threshold of 0.4 (68.37%) was lower than that with NDSI threshold of 0.1 (83.11%). The reason for the high overall accuracy with NDSI threshold of 0.4 is that the accuracy of non-snow category with this threshold is higher, and the large non-snow samples (89 times the snow category) enhanced the overall accuracy. Due to the small number of snow category samples using *in situ* observation (only 882), we further selected Landsat series data with similar amount of snow and non-snow samples for further verification (Table R2). In this case, the overall accuracy with NDSI threshold of 0.4 (83.50%) was still higher than that with NDSI threshold of 0.1 (77.28%). When the threshold is set as 0.4, although the accuracy of snow category (82.64%) is lower than that of 0.1 threshold (94.93%), the accuracy for non-snow category (84.36%) is much higher than that of 0.1 threshold (59.63%). In our experiments, using 0.1 as the NDSI threshold in the Tibetan Plateau may cause too many non-snow pixels to be misclassified as snow pixels.

We further explored why our finding is different from Zhang et al. (2020)' results. The study area of Zhang et al. (2020) is the entire China, while ours is the Tibetan Plateau. First, the thresholds of the measured *in situ* snow depth data used for validation are different. The *in situ* snow depth used for verification in Zhang et al., (2020) is divided by a 1 cm threshold. Ke et al. (2016) demonstrates that thin snow depth reduces the reliability of snow-related studies in China. In our study, a 3 cm threshold was utilized to classify the *in situ* snow depth (Huang et al., 2022). Different snow depth thresholds lead to various snow classifications, which lead to different results. In addition, the numbers of snow and non-snow samples used for validation are also different between our two studies. We selected the Landsat series data with similar amount of snow and non-snow samples for verification, while Zhang et al., (2020) used the *in situ* snow depth with more non-snow samples.

We also compared our results with other snow studies over the Tibetan Plateau. Gao et al. (2019) explored the optimal NDSI threshold for snow cover identification on the Tibetan Plateau under different land cover types, and verified the accuracy with Landsat-5 TM and Landsat-8 OLI data. Their results show that the optimal NDSI thresholds are 0.33, 0.40, and 0.47 under grassland, sparse vegetation surface types, and other underlying surface types, respectively. Since our study did not divide the Tibetan Plateau into different land cover types, a threshold of 0.4 was selected based on our experimental results and as referenced to existing literature. The suggestion of the reviewer has given us a good inspiration. In our future research, we will explore other optimal threshold of NDSI for snow identification in the Tibetan Plateau.

Table R1 Confusion matrices between MODIS snow products with different threshold of NDSI, and *in situ* observation during 2002–2021.

In situ observation	NDSI with a threshold of 0.4			NDSI with a threshold of 0.1		
	Snow	Non-snow	Total	Snow	Non-snow	Total
Snow	603 **(68.37%)**	279 (31.63%)	882	733 **(83.11%)**	149 (16.89%)	882
Non-snow	1789 (2.28%)	76696 (97.72%)	78485	3629 (4.62%)	74856 (95.38%)	78485
Total	2392	76975	79367	4362	75005	79367
Overall accuracy	**97.39%**			**95.24%**		

Table R2 Confusion matrices between MODIS snow products with different threshold of NDSI, and snow cover mapped from Landsat series observation during 2002–2021.

Landsat series	NDSI with a threshold of 0.4			NDSI with a threshold of 0.1		
	Snow	Non-snow	Total	Snow	Non-snow	Total
Snow	239056 **(82.64%)**	50227 (17.36%)	289283	274602 **(94.93%)**	14681 (5.07%)	289283

Non-snow	45235 (15.64%)	244048 (84.36%)	289283	116790 (40.37%)	172493 (59.63%)	289283
Total	284291	294275	578566	391392	187174	578566
Overall accuracy	**83.50%**			**77.28%**		

References:

Gao, Y., Hao, X. H., He, D. C.,Huang, G. H., Wang, J., Zhao, H. Y., Wei., Y. R., Shao., D. H., Wang., W. G.: Snow cover mapping algorithm in the Tibetan Plateau based on NDSI threshold optimization of different land cover types, Journal of Glaciology and Geocryology, 41(5), 1162-1172, doi: 10.7522/j.issn.1000-0240.2019.1155, 2019. (in Chinese)

Huang, Y., Song, Z. C., Yang, H. X., Yu, B. L., Liu, H. X., Che, T., Chen, J., Wu, J. P., Shu, S., Peng, X. B., Zheng, Z. J., and Xu, J. H.: Snow cover detection in mid-latitude mountainous and polar regions using nighttime light data, Remote Sensing of Environment, 268, doi: 10.1016/j.rse.2021.112766, 2022.

Ke, C. Q., Li, X. C., Xie, H. J., Ma, D. H., Liu, X., Cheng, K.,: Variability in snow cover phenology in China from 1952 to 2010, Hydrology and Earth System Sciences, 20, 755, doi: 10.5194/hess-20-755-2016, 2016.

The Tibetan Plateau has high altitude and complex terrain, and Landsat-8 data used for reference data is 30 m, which will be affected by terrain and mountain shadow. Have you considered the terrain effect on Landsat-8 snow cover? And how’s it affect validation results?

Response: We agree that Landsat data may be influenced by topography and mountain shadow to some extent, which can result in underestimation of snow cover in the Tibetan Plateau. In this revision, we applied a classic topographic correction model, C correction model (Teilet et al., 1982), to correct for the terrain effect on all Landsat series data used in this study (now in Line 151-152).

Reference:

Teilet, P. M., Guindon, B., Goodenough, D. G.: On the slope-aspect correction of multispectral scanner data, Canadian Journal of Remote Sensing, 8(2), 1537-1540, doi: 10.1080/07038992.1982.10855028, 1982.

Why use solar radiation not net radiation to represent environmental effect? Net radiation might be more related with snow surface than solar radiation here.

Response: Net radiation is defined as the difference between incoming and outgoing radiation flux. Solar radiation we used in this work applies latitude, slope, aspect, date, and interval time as inputs, and estimates direct, diffuse, reflected solar energy received by the ground. The complex topography of the Tibetan Plateau determines the availability of radiation at specific locations. Compared with the net radiation, solar radiation takes into account the effect of terrain (i.e., latitude, slope, aspect) and seasons (i.e., date) more comprehensively, which is very necessary for the Tibetan Plateau with complex terrain conditions. So, we used solar radiation to represent environmental effect.

The snow fraction estimated method used in the equation (2) was derived through other regions, and many studies have shown that the linear relationship has limited accuracy in the Tibetan Plateau region. If possible, I'd like suggest the authors re-fit that empirical relationship between snow fraction and NDSI in the Tibetan Plateau region. In addition, the fitting relations of Terra and Aqua satellites are different. If the same equation was used for Terra and Aqua, it might cause error on snow cover determination.

Response: According to the reviewer's comments, we have re-fitted the empirical relationship between snow fraction and NDSI of Terra and Aqua in the Tibetan Plateau using Landsat series data over 20 years (Eq.1 and Eq.2):

$$P(x_i|\beta_1)_{Terra} = (1.222 \times NDSI + 0.038)/100 \quad (1)$$

$$P(x_i|\beta_1)_{Aqua} = (1.164 \times NDSI + 0.058)/100 \quad (2)$$

The sample points used for re-fit Terra and Aqua were 972884, and 952221, respectively, and the correlation coefficients of the empirical relationship of Terra and Aqua satellites were 0.86 and 0.89, respectively (now in Line 192-199). Due to the re-fit of the FSC based on Eq.1 and Eq.2, we also recalculated the optimal parameters and reproduced the dataset. The new calculated optimal parameters of the HMRF*solar* model for spectral, spatial-temporal, and environmental information of the TP were 0.117, 1.294, and 0.532, respectively (now in Line 397-403). The overall accuracy of the reproduced dataset was 91.36%, which increased by 2.06% compared with the overall accuracy of original MODIS products (Table 3, in Sections 4.1 and 4.2).

Table 3. Confusion matrices between HMRF$_{solar}$-based snow products, HMRF$_{ele}$-based snow products, original MODIS snow products, and snow cover mapped from Landsat series data products for gap-free pixels during 2002–2021.

Landsat series data	HMRF$_{solar}$-based snow products			HMRF$_{ele}$-based snow products			Original MODIS snow products		
	Snow	Non-snow	Total	Snow	Non-snow	Total	Snow	Non-snow	Total
Snow	916593 (85.06%)	160936 (14.94%)	1077529	901202 (83.64%)	176327 (16.36%)	1077529	881646 (81.82%)	195883 (18.18%)	1077529
Non-snow	108214 (5.31%)	1931065 (94.69%)	2039279	120590 (5.91%)	1918689 (94.09%)	2039279	137391 (6.74%)	1901888 (93.26%)	2039279
Total	1024807	2092001	3116808	1021792	2095016	3116808	1019037	2097771	3116808
Overall accuracy	91.36%			90.47%			89.31%		

Landsat-8 images was not enough to demonstrate the current results. If possible, please add more validation Landsat images, such as Landsat-5/7 images.

Response: In this revision, we have added more Landsat images for validation, including Landsat-5 TM, Landsat-7 ETM+, and Landsat-8 OLI images. The detailed information of the Landsat images is shown in Table A1.

Table A1. Landsat series images used for assessment of the HMRF-based snow cover products in this study.

Image pair No.	Sensor	Tile path/row	Date of acquisition	Cloud cover (%)
1	ETM+	131/38	2002-11-22	1%
2	ETM+	136/38	2003-1-28	0%
3	ETM+	132/41	2003-2-17	1%
4	TM	136/33	2003-8-16	0%
5	TM	141/35	2003-9-20	0%
6	TM	135/33	2004-8-27	0%
7	TM	137/39	2004-12-15	1%
8	TM	132/38	2005-1-13	1%
9	TM	136/39	2005-3-14	1%
10	TM	132/34	2005-4-3	1%
11	TM	142/34	2005-6-28	0%
12	TM	136/36	2005-10-24	1%
13	TM	133/38	2005-11-4	1%
14	TM	135/39	2006-2-6	1%
15	TM	135/33	2006-8-1	0%
16	TM	141/39	2006-9-28	1%
17	TM	132/34	2006-10-31	2%
18	TM	134/37	2006-11-30	1%
19	TM	136/39	2006-12-14	1%
20	TM	134/33	2007-3-6	1%
21	TM	141/34	2007-7-13	1%
22	TM	139/35	2007-9-17	1%
23	TM	132/37	2008-2-23	1%
24	TM	132/42	2008-3-10	1%
25	TM	134/36	2008-5-11	1%
26	TM	145/35	2008-6-25	1%
27	TM	142/34	2008-8-7	1%
28	TM	150/33	2008-10-2	1%
29	TM	130/37	2008-11-7	1%
30	TM	133/37	2008-12-14	1%
31	TM	132/38	2009-3-13	0%
32	TM	132/37	2009-4-14	1%
33	TM	147/35	2009-8-13	1%
34	TM	151/33	2009-9-10	1%
35	TM	138/37	2009-10-17	1%
36	TM	134/36	2009-11-22	1%
37	TM	133/38	2010-2-19	0%
38	TM	135/39	2010-3-21	1%
39	TM	150/32	2010-11-9	1%
40	TM	134/39	2011-3-1	1%
41	TM	141/35	2011-8-25	0%
42	TM	132/37	2011-10-29	0%
43	OLI	147/37	2013-4-18	2%
44	OLI	149/34	2013-5-18	2%
45	OLI	146/36	2013-8-1	1%

46	OLI	145/36	2013-9-27	2%
47	OLI	141/35	2013-11-18	1%
48	OLI	133/40	2014-1-13	1%
49	OLI	136/38	2014-2-19	1%
50	OLI	136/33	2014-7-13	0%
51	OLI	144/35	2014-8-22	1%
52	OLI	139/38	2015-1-10	1%
53	OLI	143/39	2015-3-11	1%
54	OLI	151/33	2015-10-13	1%
55	OLI	136/38	2015-12-23	1%
56	OLI	132/34	2016-5-3	2%
57	OLI	151/33	2016-6-25	2%
58	OLI	143/34	2016-9-21	1%
59	OLI	133/37	2016-11-18	2%
60	OLI	146/38	2017-2-1	1%
61	OLI	151/33	2017-4-9	0%
62	OLI	144/35	2017-7-29	1%
63	OLI	133/37	2017-11-5	1%
64	OLI	131/36	2018-4-16	1%
65	OLI	133/38	2018-11-8	0%
66	OLI	137/40	2018-12-22	0%
67	OLI	131/38	2019-3-18	1%
68	OLI	136/33	2019-8-28	0%
69	OLI	151/33	2019-9-22	0%
70	OLI	134/36	2019-11-2	1%
71	OLI	132/38	2019-12-6	1%
72	OLI	135/39	2020-1-12	1%
73	OLI	136/38	2020-2-4	1%
74	OLI	151/33	2020-8-23	1%
75	OLI	149/35	2020-12-31	1%
76	OLI	132/34	2021-1-25	1%
77	OLI	143/34	2021-7-17	1%
78	OLI	151/33	2021-9-27	0%
79	OLI	135/36	2021-10-29	1%
80	OLI	143/36	2021-11-22	0%
81	OLI	151/33	2021-12-16	1%

Why the validation accuracy of HMRF*solar* or HMRF*dem* is higher than MODIS? In my opinion, HMRF just filled the data gap, why the accuracy is also improved a lot. Please justify it.

Response: Our HMRF-based framework can exploit spatial and temporal contextual information and environmental association information, in addition to the MODIS spectral information that was used in the standard NASA algorithm to produce the original MODIS snow products. The category of all pixels (including data-gap pixels and gap-free pixels) on the entire initial MODIS snow cover products were determined by employing the optimal parameters and HMRF algorithm. As demonstrated in our previous study (Huang et al., 2018), our HMRF framework not only fills the data gaps, but also improves the snow cover estimate accuracy of original MODIS snow cover products.

Reference:

Huang, Y., Liu, H., Yu, B., Wu, J., Kang, E. L., Xu, M., Wang, S., Klein, A., and Chen, Y.: Improving MODIS snow products with a HMRF-based spatio-temporal modeling technique in the Upper Rio Grande Basin, Remote Sensing of Environment, 204, 568-582, doi: 10.1016/j.rse.2017.10.001, 2018.

I have concerned that the weight used in Equation (1), such as U_{xi}, U_{st}, U_{ev} are negative defined in Equation (3), (4) and (13).

Response: In Equation (1), U_{xi}, U_{st}, and U_{ev} are the spectral, spatiotemporal, and environmental energy functions, respectively. Because the probabilities modeled by HMRF are equivalent to the energies characterized by a Gibbs random field (Geman and Geman, 1984), the maximization of the probability can be realized by minimizing total energy function (Huang et al., 2018, Equation (S1)):

$$\underset{C_2}{\mathrm{Max}}\{P(\beta_n \,|x_i, N_{xi}, N_{st}, I_{ev})\} = \underset{C_2}{\mathrm{Min}}\{\frac{1}{Z} e^{-[U_T(\beta_n, x_i, N_{xi}, N_{st}, I_{ev})]}\} \tag{S1}$$

where Z is a constant; U_T is the total energy function, detailed derivation procedure and description can be found in Huang et al. (2018).

Thus, in previous Equation (3), (4), and (13) (now in Equation (4), (5), and (14)), U_{xi}, U_{st}, U_{ev} are negative defined by using spectral probability, spatiotemporal probability, and environmental probability.

Reference:

Geman, S., Geman, D.: Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-6(6), 721-741, 1984.

Huang, Y., Liu, H., Yu, B., Wu, J., Kang, E. L., Xu, M., Wang, S., Klein, A., and Chen, Y.: Improving MODIS snow products with a HMRF-based spatio-temporal modeling technique in the Upper Rio Grande Basin, Remote Sensing of Environment, 204, 568-582, doi: 10.1016/j.rse.2017.10.001, 2018.

Specific comments/suggestions:

Please provide the definition and equation of accuracy evaluation index (OA, OE et al.)

Response: The definition and equation of accuracy evaluation index have been provided in Line 270-278.

Figure 8. Please add the latitude and longitude information.

Response: The latitude and longitude have been added in Figure 8.

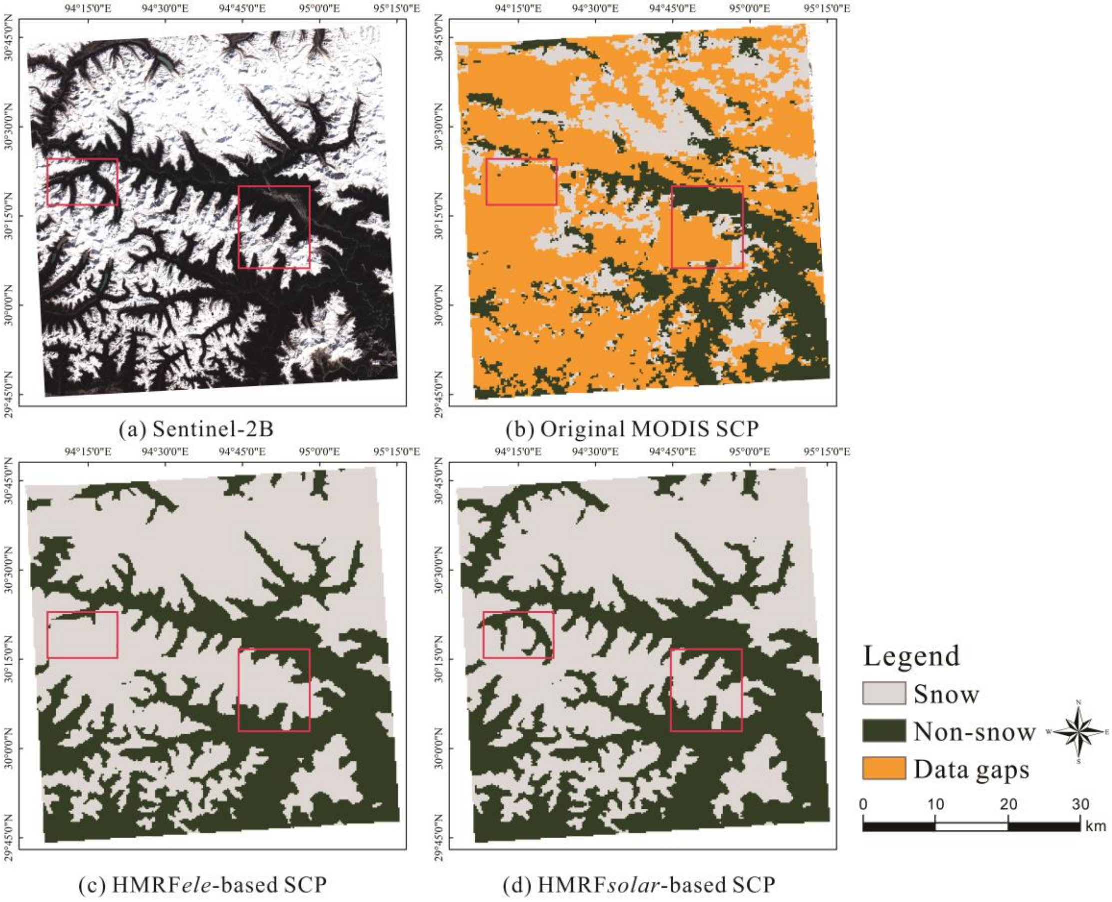
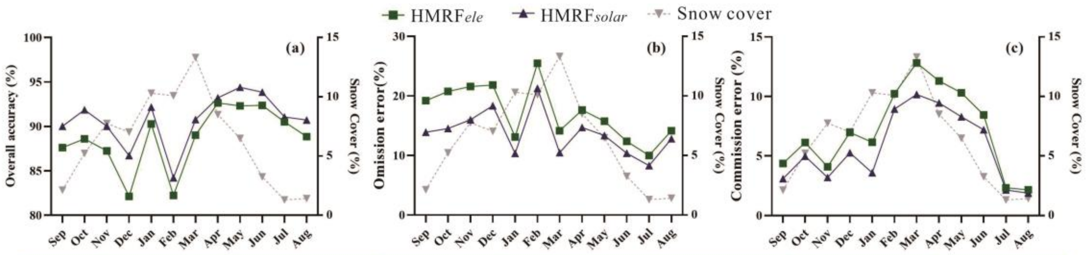


Figure 8: Comparison between true-color Sentinel-2B imagery and(a), original MODIS snow products (b), HMRF$_{ele}$-based snow products (c) HMRF$_{solar}$-based snow products.

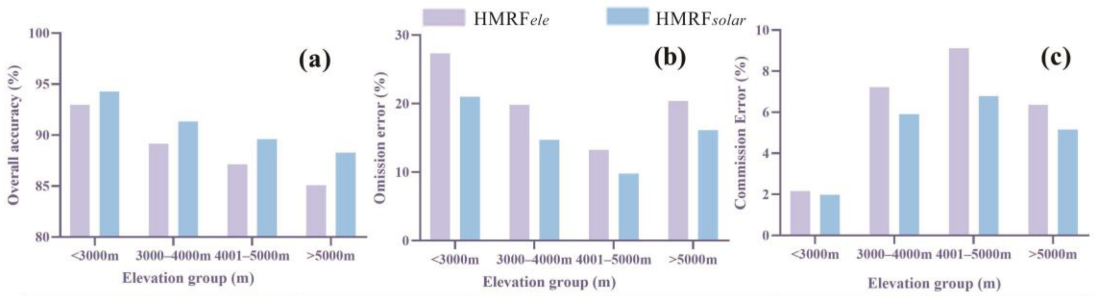
The Figure 4,5,7,8 resolution is too low, please check whether the Figure format meets the requirements of the journal.

Response: The resolution of Figure 4,5,7,8 have been improved



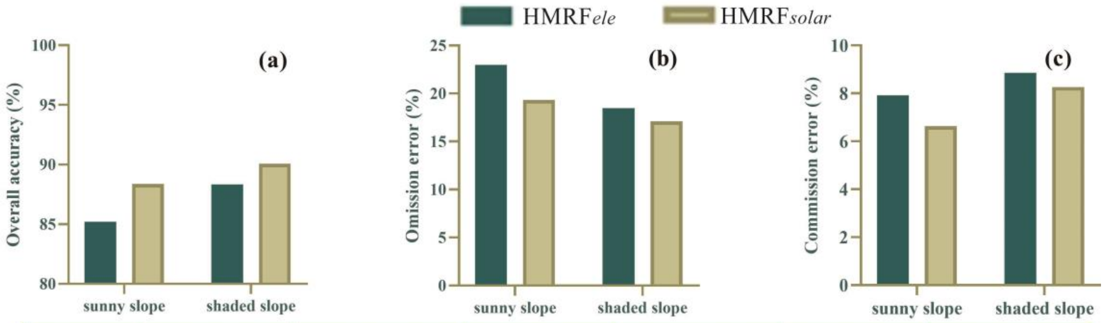
Month	OA-HMRF*ele*	OA-HMRF*solar*	OE-HMRF*ele*	OE-HMRF*solar*	CE-HMRF*ele*	CE-HMRF*solar*	Snow cover
Sep	87.62%	90.02% (+2.40%)	19.21%	13.92% (-5.29%)	4.37%	3.09% (-1.28%)	5.23%
Oct	88.61%	91.87% (+3.26%)	20.76%	14.53% (-6.23%)	6.13%	4.98% (-1.15%)	7.75%
Nov	87.26%	90.03% (+1.91%)	21.59%	16.01% (-5.58%)	4.09%	3.19% (-0.90%)	7.03%
Dec	82.14%	86.72% (+4.58%)	21.84%	18.37% (-3.47%)	6.98%	5.27% (-1.71%)	10.30%
Jan	90.29%	92.17% (+1.88%)	13.11%	10.36% (-2.75%)	6.15%	3.59% (-2.56%)	10.10%
Feb	82.24%	84.22% (+1.98%)	25.49%	21.29% (-4.20%)	10.23%	8.94% (-1.29%)	13.32%
Mar	89.03%	90.78% (+1.75%)	14.17%	10.49% (-3.68%)	12.82%	10.17% (-2.65%)	8.51%
Apr	92.64%	93.22% (+0.70%)	17.64%	14.72% (-2.92%)	11.31%	9.46% (-1.85%)	6.49%
May	92.32%	94.40% (+2.08%)	15.76%	13.36% (-2.40%)	10.31%	8.29% (-2.02%)	3.25%
Jun	92.37%	93.86% (+1.22%)	12.38%	10.39% (-1.99%)	8.46%	7.21% (-1.25%)	1.30%
Jul	90.53%	91.07% (+2.69%)	10.01%	8.31% (-1.70%)	2.33%	2.14% (-0.19%)	1.42%
Aug	88.86%	90.73% (+1.87%)	14.15%	12.82% (-1.33%)	2.16%	1.88% (-0.28%)	2.13%

Figure 4: Temporal variations in OA (a), OE (b), and CE (c) of HMRF$_{ele}$- and HMRF$_{solar}$-based snow products from 2002–2021.



Elevation (m)	OA-HMRF*ele*	OA-HMRF*solar*	OE-HMRF*ele*	OE-HMRF*solar*	CE-HMRF*ele*	CE-HMRF*solar*
<3000m	92.96%	94.26% (+1.30%)	27.31%	20.98% (-6.33%)	2.16%	1.98% (-0.18%)
3000–4000m	89.15%	91.33% (+2.18%)	19.84%	14.69% (-5.15%)	7.23%	5.91% (-1.32%)
4001–5000m	87.12%	89.60% (+2.48%)	13.24%	9.77% (-3.47%)	9.12%	6.79% (-2.33%)
>5000m	85.09%	88.25% (+3.16%)	20.37%	16.12% (-4.25%)	6.36%	5.16% (-1.20%)

Figure 5: Effect of elevation on OA (a), OE (b), and CE (c) of HMRF_{ele}- and HMRF_{solar}-based snow products from 2002–2021.



Aspect	OA-HMRF*ele*	OA-HMRF*solar*	OE-HMRF*ele*	OE-HMRF*solar*	CE-HMRF*ele*	CE-HMRF*solar*
Sunny slope	85.21%	88.37% (+3.16%)	22.96%	19.33% (-3.63%)	7.91%	6.63% (-1.28%)
Shaded slope	88.32%	90.07% (+1.75%)	18.47%	17.10% (-1.37%)	8.85%	8.26% (-0.59%)

Figure 7: Effect of aspect on OA (a), OE (b), and CE (c) of HMRF_{ele}- and HMRF_{solar}-based snow products from 2002–2021.

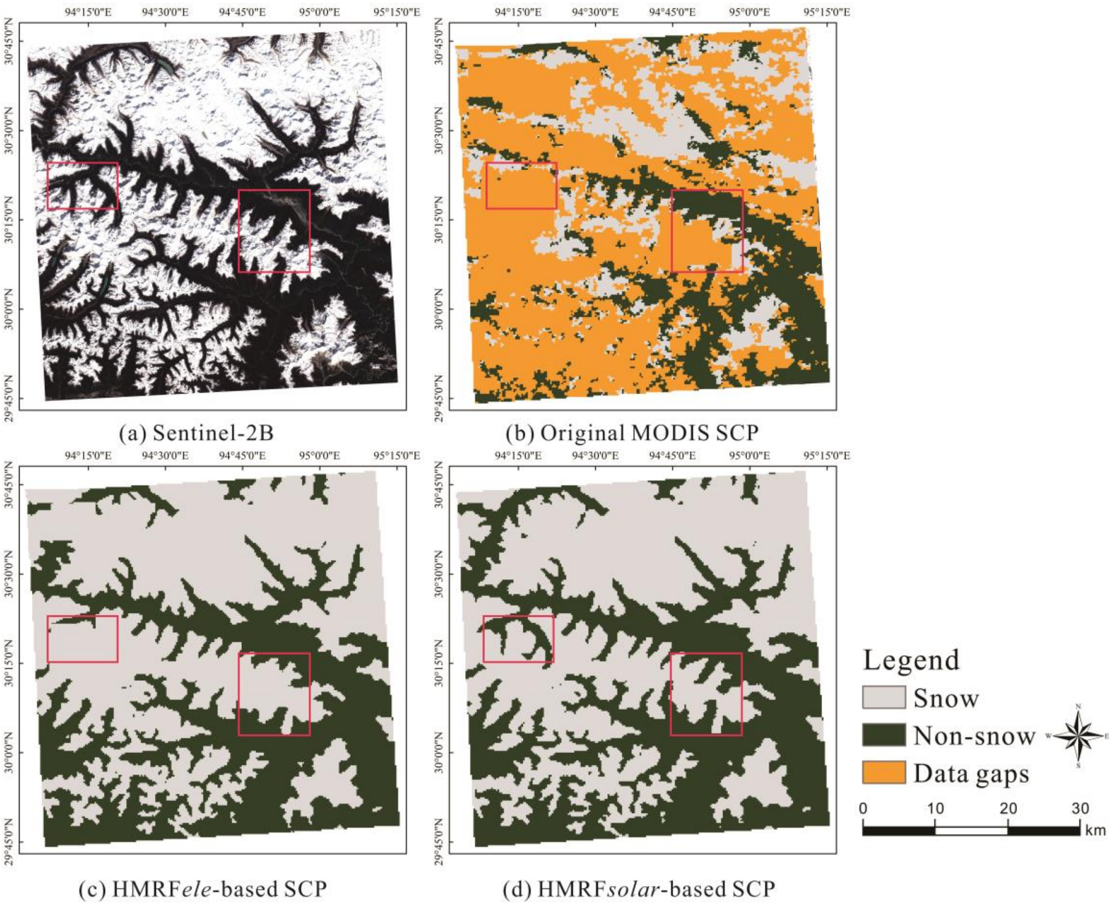


Figure 8: Comparison between true-color Sentinel-2B imagery and(a), original MODIS snow products (b), HMRF_{ele}-based snow products (c) HMRF_{solar}-based snow products.

Line115-116, “the values of 211, 237, and 239 in the NDSI_Snow_Cover_Class band were reclassified as non-snow”. From #Line111, “the values of 211, 237 and 239” indicate “night time”, “inland water”, “ocean”. So it is not reasonable that the pixels with three values are determined as non-snow.

Response: Thank you for your careful reading. In our data processing, the value of 211(night time) was determined as data-gap. As for the values of 237 (inland water) and 239 (ocean), we determined them as non-snow, as referenced as Huang et al. (2022). We have revised the classification scheme in Line 121.

Reference:

Huang, Y., Song, Z. C., Yang, H. X., Yu, B. L., Liu, H. X., Che, T., Chen, J., Wu, J. P., Shu, S., Peng, X. B., Zheng, Z. J., and Xu, J. H.: Snow cover detection in mid-latitude mountainous and polar regions using nighttime light data, Remote Sensing of Environment, 268, doi: 10.1016/j.rse.2021.112766, 2022.