To Editor:

We greatly appreciate the reviewers taking the time to provide constructive feedback and valuable suggestions. The implementation of these suggestions has greatly improved the quality of the manuscript, enabling us to make significant improvements. Every revision suggestion and opinion proposed by the reviewer has been carefully considered. Of course, there are also some parts that we do not agree with, and we will annotate them in segments.

We have also prepared and supplemented the documents one by one according to your six requirements. Original comments and suggestions quoted from this decision letter highlighted in black italics, and our response is in blue. The revised content is marked in red text in the manuscript.

Response to the comments

General comments:

1 *This paper proposed a high spatial resolution (500 meters) snow depth estimation method based on AMSR-2 brightness temperature data and automated machine learning (AutoML), significantly improving the accuracy of snow depth monitoring in the complex terrain areas of the Tibetan Plateau. Through the Pearson correlation coefficient, 19 key influencing factors were selected, including AMSR-2 brightness temperature, slope, and surface roughness. The method comprehensively considers the impact of various geographic and topographic factors on snow depth, enhancing the model's robustness and accuracy. But the authors ignore influence of the snow properties. How to asses the influence?However, the paper still needs to improve the writing style and enhance the readability and standardization of the charts, and the details need to check carefully. In conclusion, we suggest a major revision.*

Reply: Thank you for your comprehensive evaluation and critical questions—your concern about snow properties and manuscript standardization is crucial for improving the study's rigor.

Undoubtedly, snow characteristics such as snow density, snow particle size, and pore water content have significant impacts on snow depth inversion. However, this study is limited by the basic dataset of the research area and time scale. The existing datasets are mostly accumulated data from station observations, which are sparsely distributed in the Qinghai Tibet Plateau region at 2012 to 2021 and are not sufficient for automatic machine learning inversion. Therefore, this study did not include it as an influencing factor. We have supplemented this explanation in the Discussion section:

(Lines: 515-521)

(1) It is acknowledged that there are numerous factors influencing SD retrieval, apart from the topographical conditions considered in this study. The accuracy of SD retrieval can be significantly affected by snow characteristics such as snow density, grain size, liquid water content, as well as vegetation canopy (Ni et al., 2024; Zhang et al., 2021). Nevertheless, the study is constrained by the fundamental dataset of the research domain and time span. The extant datasets are predominantly composed of accumulated data from station observations, which are sparsely distributed in the QTP from 2012 to 2021 and are inadequate for AutoML inversion. Consequently, this study did not incorporate it as a potential influencing factor.

Specific comments:

1 *Line 20 'Compared with Landsat-8, the estimated SD spatial distribution is consistent with the* *snow cover extent on optical images, which can provide reliable data for monitoring snow cover changes in mountainous regions'. Sd consistent with SCE, and the logic is not proper.*

Reply: Thank you for pointing out this logical ambiguity—this was due to incomplete expression. We have revised the sentence to clarify that the comparison is between "SD spatial distribution" and "SCE derived from Landsat-8", ensuring logical consistency:

(Lines: 21-23)

Compared with the snow cover extent (SCE) derived from Landsat-8 optical images, the estimated SD spatial distribution is consistent with the SCE, which can provide reliable data for monitoring snow cover changes in mountainous regions.

2 Line 27 'Roof of the World,' Please remove ','. It would be better use '' instead of ", which is the Chinese usage. Please check the whole manuscript.

Reply: This was indeed caused by our negligence, so we have checked the entire manuscript and made revisions to ensure that there are no more similar errors. We have deleted the comma and corrected it to 'Roof of the World' (line 27). We also replaced ' with " throughout the manuscript.

3 Line 62 Support Vector Machine (SVR), SVM would be better than SVR..

Reply: Thank you for bringing this issue to our attention. It was caused by a typing error. Having checked the entire manuscript, we found two similar errors and replaced 'SVR' with 'SVM' throughout.

4 In the last part of Introduction, the author need the add the objectives of your work and the contribution or potential usage to this field.

Reply: Thank you very much for your constructive suggestion. It is indeed due to our unclear expression in the introduction section. We have supplemented the last paragraph of the Introduction with research objectives, innovations, and contributions:

(Lines: 94-108)

The objective of this study is to propose a PMW SD estimation method based on the AutoML (Pycaret model) for complex mountainous areas characterised by sparse and unevenly distributed observational data. Firstly, the Che-AMSR2 downscaled SD data and ground-based SD observations are used as input data (dependent variables) for the Pycaret model, whilst the AMSR-2 BT data and 28 factors, such as slope and surface roughness, are used as independent variables. Then, a total of 19 key factors were screened through the utilisation of the Pearson correlation coefficient method. For the four distinct snow underlying surface types (forests, grasslands, water bodies, and unused land) training was conducted on the sample data using the selected input variables. Finally, the optimal AutoML model is obtained for each snow subsurface coverage type is subsequently selected to estimate SD on the QTP. The study employs snow cover products to identify the presence or absence of snow in 500 m pixels. For snow-free pixels, the SD value is set to 0, while for snow-covered pixels, the proposed SD estimation method is utilized to obtain the SD values anew. The innovation of this study lies in (1) the adoption of a multi-source sample combination strategy, the integration of downscaled and meteorological station SD data as target variables, (2) in addition to the first application of the PyCaret AutoML framework for complex mountainous SD inversion. The findings of this study provide a reference for SD inversion in other cold regions, such as the European Alps. The remainder of this study is organized as follows: Sections 2 and 3 introduces the study region, data and methods; Section 4 describes the results; Sections 5 and 6 present the discussion and conclusions.

5 The art-to state progress of SD estimation is not enough, and the author need to improve it.

Reply: Thank you for pointing out this deficiency. I am very sorry for the inconvenience caused to you. This is due to our unclear description of the snow depth inversion algorithm. For this reason, we have provided a clearer description of the methods described in the article.

(Lines: 95-103)

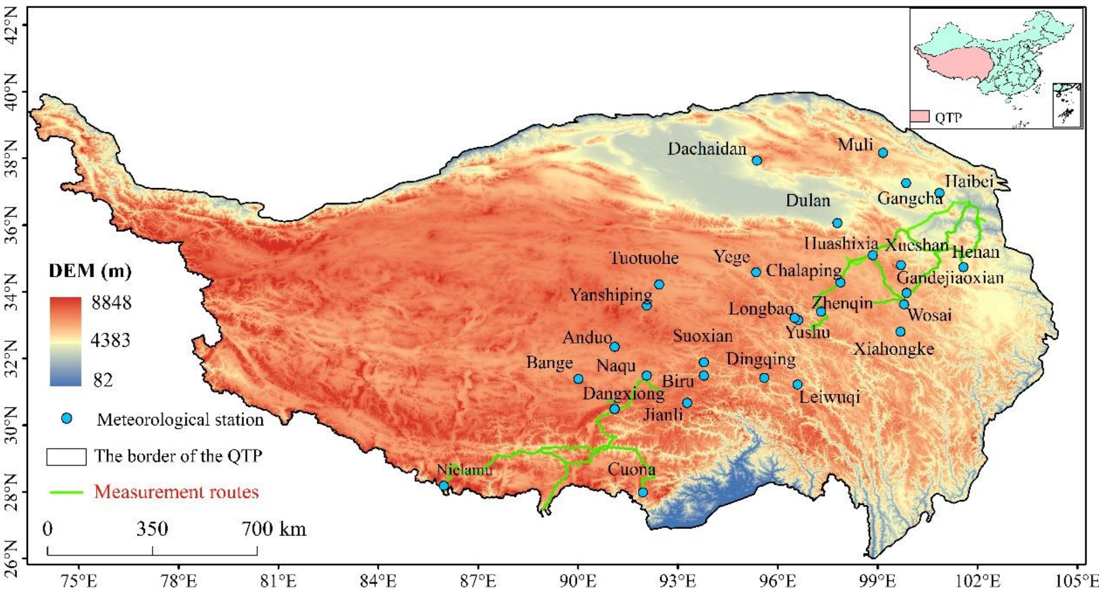
Firstly, the Che-AMSR2 downscaled SD data and ground-based SD observations are used as input data (dependent variables) for the Pycaret model, whilst the AMSR-2 BT data and 28 factors, such as slope and surface roughness, are used as independent variables. Then, a total of 19 key factors were screened through the utilisation of the Pearson correlation coefficient method. For the four distinct snow underlying surface types (forests, grasslands, water bodies, and unused land) training was conducted on the sample data using the selected input variables. Finally, the optimal AutoML model is obtained for each snow subsurface coverage type is subsequently selected to estimate SD on the QTP. The study employs snow cover products to identify the presence or absence of snow in 500 m pixels. For snow-free pixels, the SD value is set to 0, while for snow-covered pixels, the proposed SD estimation method is utilized to obtain the SD values anew.

6 What is the amount of SD measurement from climate stations and line measurement? Line measurement is not a proper name. The site is mainly distributed in the middle and the east part and rarely distributed in the west part. What is the influence of limited materials on the remote sensing inversion.

Reply: Thank you for your questions about data quantity and distribution—these are key to evaluating inversion reliability.

It is undeniable that meteorological stations in the Tibet Plateau are mainly distributed in the central and eastern parts of the site, with few in the western part. This practical problem has indeed brought great difficulties to remote sensing inversion and downscaling. In response to this issue, we adopt a strategy of multi-source data fusion, using the downscaled product Che_SMSR2-NSD data as supplementary reference data, which is used as the true value along with meteorological station data. To some extent, it solves the problem of limited observational data in reality.

We have changed ‘Line measurement’ to ‘Measurement routes’ (Figure 1) and checked the entire manuscript for similar expression errors. The modification is as follows:



7 Lin 109 Table1 application: usage. Please sperate the usage of auxiliary data.

Reply: Thank you very much for your improvement suggestion. We have changed 'application' to 'usage'(Table1). We found that the purpose of the auxiliary data was indeed unclear, so we made modifications to the parts of the table that were not described clearly. The revised part is now attached:

Datasets		Spatial Resolution	Data period	Data sources	Usage
AMSR-2 BT		10 km	2012.10～2021.03	https://gportal.jaxa.jp/	Establish model
Che_AMSR2_NSD		500 m	2012.10～2018.03	-	Input data
Daily cloud-free snow cover dataset		500 m	2012.10～2021.03	https://poles.tpdc.ac.cn/zh-hans/	Snow Identification
SD observations	Meteorological station	-	2015～2019	https://data.tpdc.ac.cn/home/	Input data and verification
	Measurement routes	-	2018～2019	https://www.csdata.org https://www.ncdc.ac.cn/	
Auxiliary Data	SRTM DEM	90 m	-	https://earthexplorer.usgs.gov/	Extract terrain parameters
	CNLUCC	1 km	2020	https://www.resdc.cn/	Obtain land cover types
	ERA5-Land	1 km	2012.10～2021.03	https://climate.copernicus.eu/	The monthly average temperature
	Landsat-8	30 m	2012.10～2018.03	https://www.usgs.gov	Obtain snow cover extent

8 Line 114-116 Please sperate the long sentence.

Reply: Thank you for correcting this readability issue. We have split the long sentence into two concise ones and checked the entire manuscript for similar problems:

(Lines: 128-130)

AMSR-2 is a microwave sensor mounted on the GCOM-W1 satellite, which was launched by the Japan Aerospace Exploration Agency (JAXA) (Imaoka et al., 2012). It conducts observations at seven frequencies, and each frequency has horizontal and vertical polarization modes (Imaoka et al., 2012).

9 How to deal with the different spatial resolution between different data source.

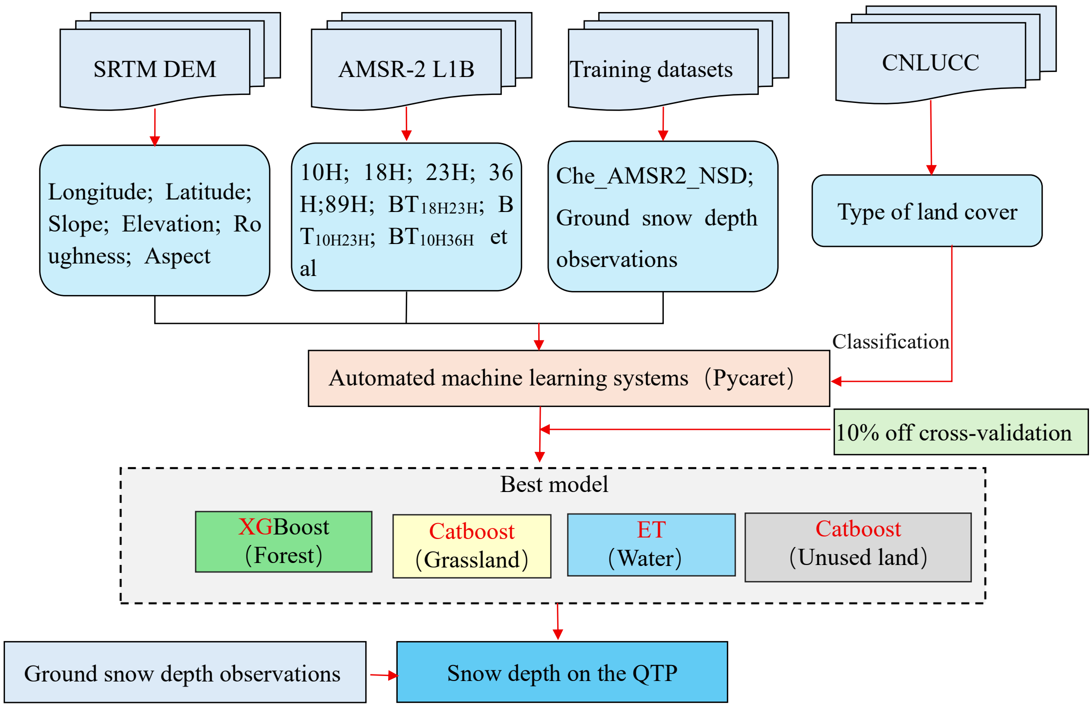
Reply: Thank you for taking note of this issue. In this article, we reprogram all data to the UTM 45 Zone coordinate system and use resampling to unify the data resolution to 500m. We have added corresponding descriptions in the data section of the article:

(Lines: 121-124)

As shown in Table 1, the dataset used for this research experiment comprises five main categories: AMSR-2 (Advanced Microwave Scanning Radiometer 2) BT; downscaled SD data (Che_AMSR2_NSD); daily cloud-free snow cover products; ground-based SD observations; and other auxiliary data. All data in this study was projected onto the Universal Transverse Mercator (UTM) zone 45, and resampled to achieve a uniform resolution of 500m.

10 The width of flow line is not consistent.

Reply: hank you for pointing out this formatting flaw. we have fixed Figure 3.



11 This study marks the introduction of the PyCaret automated machine learning framework into snow depth estimation. It automates data processing and model selection, reducing human intervention while enhancing the efficiency of model selection and parameter optimization. Figure 3 is too simple and needs to improve. How to solve the overfitting?

Reply: Thank you very much for your valuable comments and suggestions on our manuscript. Your feedback has helped us further clarify the limitations of existing methods and improve the rigor of our study.

In response to your comment that "Figure 3 is too simplistic and requires further improvement", we have optimized Figure 3 as follows:

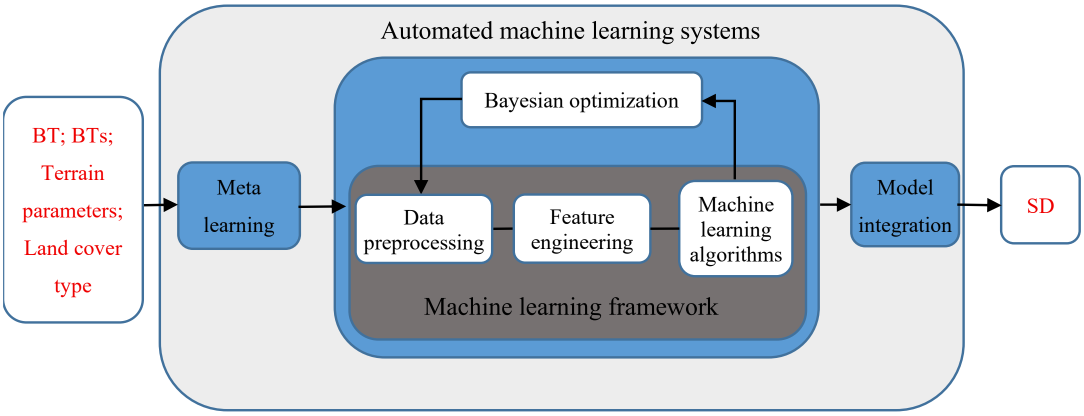


Figure 3: Flowchart of this algorithm.

To address the overfitting issue you raised, our study introduces the AutoML framework to tackle it from three key aspects:

1.Automated Data Preprocessing: AutoML automatically performs data cleaning and standardization avoiding subjective biases in manual processing, ensuring the quality of input data, and reducing "over-learning" caused by data noise.

2.Algorithmic Feature Selection: It leverages algorithms to automatically screen key features, eliminating redundant variables and simplifying input dimensions to prevent the model from learning non-essential correlations.

3.Systematic Model Optimization: AutoML automatically explores various types of models and optimizes hyperparameters— for instance, matching models with pruning or regularization strategies. This balances model complexity and generalization ability, and avoids the tendency of overemphasizing training set performance during manual model selection.This content has been supplemented in the manuscript:

(Lines: 76-85)

Generally, single or several ML models are used to train data for specific regions, and there are many challenges in data processing, feature selection, and the selection of the best model, which are accomplished through intuition or trial and error (Du et al., 2020). This usually leads to overfitting and low model robustness when training ML models, as many researchers tend to prioritize achieving better model performance (Du et al., 2020; Feurer et al., 2015; Hernandez et al., 2025). Without human intervention, Automated Machine Learning (AutoML) can autonomously execute a series of processes, including data processing and model performance evaluation, and ultimately identify the optimal ML model (Ribeiro et al., 2024; Hernandez et al., 2025). Indeed, the fundamental design philosophy of AutoML is predicated on the objective of reducing human bias, thus enhancing the generalisability and stability of models (Benghzial et al., 2023; Ribeiro et al., 2024; Hernandez et al., 2025).

12 ***Line 235-240:*** *The introduction of SD estimation in the past should move the Introducion. Figure 4 the color of meteorological stations and downscaled snow depth sample are too close, which are hard to tell from each other.*

Reply: Thank you for your suggestions on manuscript structure and figure readability—these optimize the logical flow and visual clarity.

Regarding the issues you pointed out: We have moved the section introducing the past snow depth (SD) estimation to the Introduction part as suggested, to optimize the logical structure of the manuscript.

(Lines: 52-57)

In addition, some studies have developed distinctive algorithms for SD inversion for different snow underlying surface types (Derksen et al., 2005; Goita et al., 2003; Jiang et al., 2014). For example, Derksen et al. (2005) developed an inversion algorithm for the main land cover types when inverting SD in Canada's forested regions. They then calculated the SD under mixed image elements. Meanwhile, Jiang et al. (2014) combined four frequencies (10 GHz, 18 GHz, 36 GHz, and 89 GHz) BT data to establish a semi-empirical SD inversion algorithm with four snow underlayment cover types (grassland, farmland, bare land, and forest).

Second, we have adjusted the color scheme of Figure 4—specifically, we have increased the color contrast between the meteorological station markers and the downscaled snow depth sample points, making them easily distinguishable and improving the readability of the figure. As shown in Figure 4:

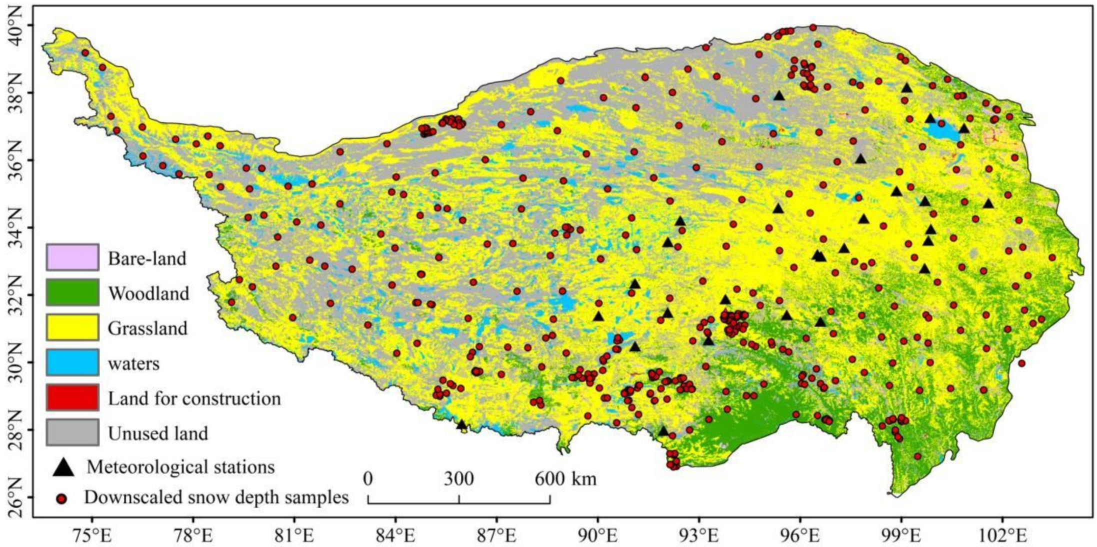


Figure 4: Spatial distribution of input sample data from the AutoML.

13 *The discussion lack the comparison of your work and others and please cite more reference.*

Reply: Thank you very much for pointing out the shortcomings in the discussion section, namely the lack of comparison between our study and others, as well as the need to supplement more references.

For this reason, we have added a comparison with previous research results in the discussion section. Meanwhile, we have added references to sentences that were not cited due to our negligence and errors. The main adjustments are as follows:

(Line: 441-457)

The accuracy of SD retrieval is contingent upon the magnitude of the SD. It is evident that the accuracy of the measurements is enhanced in the case of shallow snow (less than 5 cm), with a concomitant decrease in accuracy as the SD increases. This trend is consistent with the findings of other studies in cold regions (Wang et al., 2022; Wei et al., 2021). The majority of observation stations on the QTP are distributed in shallow snow regions (<10cm). In order to facilitate analysis, this study divides SD into four categories: less than 5 cm, 5-10 cm, 10-20 cm and greater than 20 cm. The accuracy of SD estimation derived from AutoML and downscaled SD methodologies was assessed using meteorological station SD measurements. The results of the study are presented in Table 2. It is evident that when the SD is less than 5 cm, the Auto_NSD data provides the most accurate SD results, exhibiting the minimum RMSE of 1.69 cm. The bias is 2.71, and the MAE is 2.43 cm, with accuracy evaluation indicators that surpass those of the Che_AMSR2_NSD results. When the SD ranges from 5 to cm, the RMSE of the two SD datasets is 2.94 cm and 6.39 cm, respectively. When the SD exceeds 20 cm, the SD result error reaches its maximum. The RMSE of the Auto_NSD data is 6.43 cm higher than when the SD is less than 5 cm. However, the data comparison results indicate that, irrespective of the SD, the accuracy of SD results based on AutoML estimation is superior to the Che_AMSR2_NSD results. Furthermore, the AutoML model developed in this study delivers superior performance to existing ML-based methods across both shallow and deep snow ranges. Compared to the RF model proposed by Yang et al. (2020b), which achieved an RMSE of 4.2 cm for shallow snow (<5 cm) and 10.3 cm for deep snow (>20 cm) on the QTP, the AutoML model reduces these values to 1.69 cm and 8.12 cm, respectively. This improvement is due to AutoML's advantages in model training and selection, reducing the need for manual tuning and model screening.

14 *The conclusion needs the quantitative data to support the conclusion.*

Reply: Thank you very much for your insightful comment pointing out that the conclusion requires quantitative data support. We fully agree with your view—quantitative data is essential to enhance the persuasiveness, objectivity, and scientific rigor of the conclusion, and we highly value this constructive feedback.

In response to your suggestion, we have comprehensively supplemented and refined the conclusion section by integrating key quantitative data from the study. The specific revisions are as follows:

(Lines: 532-547)

6 Conclusion

The present study concentrated on the QTP, employing an ML model with downscaled SD data, ground SD observations, and 19 SD influencing factors as input data. The input data samples were subjected to training under four distinct types of snow cover (forest, grassland, water, and unused land), and the optimal ML model was selected for each type of snow cover using ten-fold cross-validation. Consequently, the SD sequence data for the QTP from 2012 to 2021 were obtained. A thorough investigation was undertaken to evaluate the precision of Auto_NSD and Che_AMSR2_NSD SD data. This investigation encompassed both quantitative and qualitative analyses, with a focus on a comparison with downscaled SD data. The findings suggested that (1) the SD estimates derived from ML techniques exhibited superior accuracy in characterising the snow distribution in the QTP region, closely resembling the ground observations, with an R value of 0.81, RMSE of 3.65 cm, BIAS of 0.26 cm, and MAE of 2.62 cm; (2) The SD estimation accuracy of ML models varies on different underlying surfaces. Unused land (Catboost, R^2=0.82) exhibits the highest accuracy, followed by grassland (Catboost, R^2=0.77, RMSE=3.11cm), water (ET, R^2=0.75, RMSE=2.20cm), and forest (XGBoost, R^2=0.71, RMSE=3.30cm). (3) A comparison with snow cover ranges identified through Landsat-8 imagery demonstrated that both types of SD data were capable of reflecting the detailed spatial features of snow distribution in mountainous regions. However, the Auto_NSD data provided a more consistent description of SD distribution compared to the real SD distribution, fulfilling the monitoring requirements for SD in mountainous regions.

15 *Please check the reference to meet the requirements of magazine.*

Reply: Thank you very much for your valuable feedback on this article. We have checked all the references and found that they lack the DOI number, as well as numerous formatting issues. The changes have been made in the text and highlighted in red font.