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

Major Comments:

1 *The introduction lacks a comprehensive review of snow depth retrieval algorithm development. Please expand this section to include a detailed discussion of recent progress in the field, citing both foundational and current works to contextualize your research.*

Reply: Thank you for taking the time to point out this deficiency amidst your busy schedule.

We have carefully read domestic and foreign literature on passive microwave remote sensing snow depth inversion. It was found that there is a lack of data assimilation method in the snow depth inversion method in the original text. There is a lack of review on the progress and shortcomings of semi empirical methods in foreign research. At the same time, the main line of the review is not clear enough.

We have supplemented the review of the data assimilation method and expanded the content on the international progress and shortcomings of the semi-empirical statistical method. Meanwhile, we have optimized the description of the downscaling method and integrated it into the review of the machine learning (ML) method. The revised content is as follows:

(Lines: 35-76)

Research on SD inversion based on passive microwave (PMW) remote sensing has been conducted for more than 40 years. Multiple mature inversion algorithms have been developed, and various SD products have been released. Currently, there are three main methods for using PMW remote sensing to invert SD: physical model method, data assimilation method, semi-empirical statistical method, and machine learning (ML). Among them, the physical model simulates the scattering and absorption characteristics of snow in microwave bands, fully considering snow properties such as snow density and snow grain size. However, due to the complexity of the microwave radiation transmission model and the difficulty in accurately obtaining these snow characteristic parameters, the reliability of SD physical model is reduced (Kwon et al., 2017; Wainwright, et al., 2017). The data assimilation approach enhances the precision of SD estimation through the optimal integration of PMW data and other auxiliary information prior to the estimation of model fluxes. The quality of observational data is a critical factor affecting the accuracy of data assimilation methods (Cortés et al., 2017; Aalstad et al., 2018). The scarcity of observational data over the QTP imposes limitations on the implementation of data assimilation methods. (Li et al., 2022b). The SD inversion of semi-empirical statistical method primarily utilizes the correlation between the difference in the snow scattering characteristics of different frequency brightness temperature (BT) and SD. The ‘brightness temperature gradient method’, initially proposed by

Chang et al. (1987) (Chang et al., 1987), has been widely used. and numerous scholars have subsequently improved SD inversion algorithms based on Chang algorithm (Cao et al., 1993; Che et al., 2008; Foster et al., 1997; Jiang et al., 2014; Kelly, 2009). Among them, Che et al. (2008) improved the Chang algorithm based on SD measurements from Chinese meteorological stations in response to the low snow density in China, and released two long-term time series SD datasets in China: Che_SSMI/S product and Che_SMSR2 product. 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). Nevertheless, the inversion products obtained by the semi-empirical statistical method are associated with two drawbacks in mountainous areas. Firstly, the relationship between BT and SD is nonlinear, while semi-empirical methods often treat it as a linear relationship, resulting in significant errors (Tanniru and Ramsankaran, 2023). Secondly, owing to the low spatial resolution (10-25 km) of these PMW SD products, the accuracy of SD inversion is significantly limited in mountainous areas (Yan et al., 2022; Tanniru and Ramsankaran, 2023).

In recent years, ML has become a significant means of SD inversion. By training ML models, such as Support Vector Machine (SVM), Random Forest Method (RF), and Artificial Neural Network (ANN), a nonlinear relationship between microwave radiation BT and SD is established, and the SD inversion accuracy is improved by integrating multisource remote sensing data (Xiao et al., 2018; Zhong et al., 2021). Yang et al. (2020) proposed a SD inversion algorithm based on RF that considers multiple factors (BT at different frequencies, geographic location information, and land cover types), the SD results obtained were superior in accuracy compared to the Che_SSMI/S and Che_SMSR2 products. Hu et al. (2021b) compared three ML methods (ANN, SVM, and RF) using five SD products and ground-based SD measurements as prior data, and found that RF had the highest accuracy. Meanwell, ML demonstrate remarkable efficacy in enhancing the spatial resolution of PMW SD products. In general, scholars frequently employ downscaling methodologies that are founded upon empirical fusion rules and snowmelt regression curves, thereby obtaining higher resolution SD products (Tang et al., 2016; Hu et al., 2021a; Xu et al., 2024). In contradistinction to SD downscaling methodologies, ML approaches boast the advantage of being able to integrate PMW, optical remote sensing data and terrain factors, as well as automatically extracting complex nonlinear relationships (Wei et al., 2022; Tanniru and Ramsankaran, 2023). The aforementioned studies indicated that the SD estimation method based on ML models exhibits significant advantages, however, they still have shortcoming in mountainous area.

2 *The manuscript does not clearly articulate the innovation of the proposed method or its contribution to snow remote sensing research. Compared to existing snow depth retrieval studies, the work appears incremental. Please revise to explicitly highlight the novelty of your approach and its unique contributions to the field.*

Reply: Thank you very much for your constructive feedback, which is crucial for improving the quality of this manuscript. The unclear articulation of innovation stemmed from insufficient description in the original review section. We have adjusted the introduction to clearly present the novelty of our method and its contributions, adding key content in the final paragraph of the introduction. The revisions are as follows:

(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 regions 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.

3 The motivation for this work is not well-defined. The stated aim in Line 74, "to address the issues in ML models mentioned above," is insufficient, as these issues have been partially addressed in prior publications. Clearly define the specific research gap your snow depth retrieval algorithm addresses to justify the study. Without a compelling motivation, the manuscript risks rejection.

Reply: Thank you very much for your feedback, which is critical for enhancing the overall logic of the manuscript. After reviewing literature on ML and AutoML, we identified a key research gap: traditional ML heavily relies on researchers' experiential choices (in data processing, feature engineering, parameter tuning, and model selection), leading to overlooked risks of low robustness and overfitting (due to prioritizing high accuracy). In contrast, AutoML avoids human bias through algorithm-driven processes (e.g., Bayesian optimization), improving model generalizability. We have revised the AutoML section to clarify this motivation:

(Lines: 77-85)

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).

4 Section 2.2.2 introduces the Che_AMSR2_NSD snow depth product at 500-m resolution derived from AMSR-2 data. It is unclear why a new method was developed when this product exists. Additionally, if Che_AMSR2_NSD is used as a reference dataset (i.e., "true" snow depth), discuss its uncertainties and their potential impact on the stability and reliability of your retrieval model.

Reply: Thank you very much for your concern about this issue. It is indeed a deficiency in our article.

Firstly, the snow depth observation data in the Qinghai Tibet Plateau mainly comes from meteorological stations, but the distribution of meteorological stations is sparse and mostly concentrated in the east. Therefore, we have to use this data as a supplement to the meteorological station data (reference dataset) for automatic machine learning training.

Secondly, although there is already a 500 meter resolution Che_SMSR2-NSD snow depth product, it is mainly obtained through downscaling of snow accumulation experience curves. Although it is already one of the better 500m resolution microwave remote sensing snow depth products in the existing Qinghai Tibet Plateau region, there is still a lot of room for accuracy improvement. Therefore, we designed this study in order to obtain a higher precision dataset and make a certain contribution to the microwave remote sensing snow depth research in the Qinghai Tibet Plateau region. We have found that the product has improved accuracy compared to the Che_SMSR2RNSD snow depth product.

We have supplemented explanations for two core points: (1) the necessity of developing a new method (Che_AMSR2_NSD still has accuracy room for improvement); (2) the uncertainty of Che_AMSR2_NSD and its impact. The revisions are as follows:

(Lines: 136-142)

2.2.2 Che_AMSR2_NSD

Due to the sparse and uneven distribution of meteorological stations in the QTP, this study utilised a downscaled Che_AMSR2_NSD SD data as the reference dataset. Che_AMSR2_NSD is a 500 m downscaled Che_AMSR2 dataset, which was obtained from the results of a published study that utilised empirical fusion rules and snowmelt regression curves (Xu et al., 2024). In comparison to the SD data from meteorological stations, it exhibits a higher degree of concordance with measured SD, with an R of 0.72 and a root mean square error (RMSE) of 3.21 cm (Xu et al., 2024). Therefore, the Che_AMSR2_NSD, in conjunction with ground-based SD observations, was utilised as a training sample for the AutoML.

(Lines: 430-439)

5 Discussion

The QTP is characterised by complex terrain, comprising a mosaic of mountains, plateaus, and basins. The distribution of meteorological stations is characterised by sparsity and heterogeneity. In order to address this issue, the approach involved the introduction of downscaled SD data (Che_AMSR2_NSD), which was utilised in conjunction with ground SD observations as training data for AutoML models. Despite the fact that this approach enhanced the precision of SD estimation in the QTP region to a certain degree, the downscaled SD data itself is inherently uncertain. Consequently, this has led to inconsistencies between the downscaled and observed SDs. In this study, meteorological station data on SD observations were utilised to analyse the SD results obtained from AutoML estimations, considering different SDs, different snow periods and land cover type. A comparative analysis was conducted on the accuracy of Che_SMSR2_NSD and Auto_NSD data from varying snow depths and snow periods, and the factors affecting snow depth inversion were discussed.

5 *The daily cloud-free snow cover dataset, combining MODIS and passive microwave-derived snow depth data, is mentioned. However, the manuscript does not evaluate the uncertainty in snow cover identification due to Huang's snow cover data. Please provide a quantitative assessment and discussion of this uncertainty and its implications for your results.*

Reply: Thank you very much for pointing out this deficiency. The original manuscript failed to clarify the purpose of using Huang's dataset (it is not a model variable, but a tool to identify snow presence/absence).

This study uses the snow cover product as a test method to identify whether there is no snow in the pixels, while for pixels with snow, the snow depth value obtained by the snow depth calculation method in this study is used, and this data is not used, nor is it used as the independent or dependent variable of the model algorithm.

In addition, the product is based on MODIS products and uses Hidden Markov Random Field modeling technology to obtain a 500m resolution daily snow depth coverage dataset. Huang pointed out in the manuscript that the snow cover dataset he obtained was validated with field observation data and snow cover data from Landsat-8 OLI images, and found that the accuracy of these new snow cover products reached 98.29% and 91.36%, respectively. Therefore, it can be considered that this dataset meets the accuracy requirements for direct use.

We have supplemented the dataset's background, accuracy, and application logic:

(Lines: 143-151)

2.2.3 The daily cloud-free snow cover dataset

This Dataset based on long-term series MODIS snow cover products, daily snow cover products without data gaps at 500 m spatial resolution from 2002 to 2024 over the QTP. The author's findings indicate that the validation process involving in situ observations and snow cover derived from Landsat-8 OLI images has demonstrated that these new snow cover products achieve an accuracy of 98.29% and 91.36%, respectively (Huang et al., 2018). The dataset is freely available on the Big Earth Data Platform for Three Poles at https://poles.tpdc.ac.cn/zh-hans. The present study sought to ascertain whether there is a distribution of snow in pixels by downloading daily snow cover data from 2012 to 2021 over the QTP during the snow cover periods. The value assigned to pixels devoid of snow is set to 0, whilst the value calculated for pixels containing snow is determined by the algorithm proposed in this paper.

(Lines: 199-201)

In conclusion, the data concerning snow depth obtained in the present study were evaluated using a daily cloud-free snow cover dataset, in order to ascertain the presence of snow at each pixel. The snow depth of pixels that were snow-free was assigned a value of 0. The technical roadmap of this study is illustrated in Figure 2.

Minors

1 ***Lines 35–40:*** *Provide references for claims about SD retrieval challenges.*

Reply: Thank you very much for your attention to detail—this was an oversight in the original

manuscript. We have added relevant references to support claims about SD retrieval challenges:
(Lines: 35-42)

Research on SD inversion based on passive microwave remote sensing has been conducted for more than 40 years. Multiple mature inversion algorithms have been developed, and various SD products have been released. Currently, there are three main methods for using passive microwave remote sensing to invert SD: physical model method, data assimilation method, semi-empirical statistical method, and machine learning (ML). Among them, the physical model simulates the scattering and absorption characteristics of snow in microwave bands, fully considering snow properties such as snow density and snow grain size. However, due to the complexity of the microwave radiation transmission model and the difficulty in accurately obtaining these snow characteristic parameters, the reliability of SD physical model is reduced (Kwon et al., 2017; Wainwright, et al., 2017).

Kwon, Y., Yang, Z.L., Hoar, T.J., Toure, A.M.: Improving the Radiance Assimilation Performance in Estimating Snow Water Storage across Snow and Land-Cover Types in North America. J. Hydrometeor. 18, 651–668. https://doi.org/10.1175/JHM-D-16-0102.1, 2017.

Wainwright, H. M., Liljedahl, A. K., Dafflon, B., Ulrich, C., Peterson, J. E., Gusmeroli, A., & Hubbard, S. S., Hubbard, S.S.: Mapping Snow Depth within a Tundra Ecosystem Using Multiscale Observations and Bayesian Methods. Cryosphere 11, 857–875. https://doi.org/10.5194/tc-11-857-2017, 2017.

2 ***Lines 41–73:*** *Expand the literature review to include key international studies.*

Reply: Thank you for pointing out this shortcoming—international studies are essential to contextualize our work. We have supplemented key international studies on semi-empirical SD inversion algorithms:
(Lines: 45-57)

The SD inversion of semi-empirical statistical method primarily utilizes the correlation between the difference in the snow scattering characteristics of different frequency brightness temperature (BT) and SD. The 'brightness temperature gradient method', initially proposed by Chang et al. (1987) (Chang et al., 1987), has been widely used. and numerous scholars have subsequently improved SD inversion algorithms based on Chang algorithm (Cao et al., 1993; Che et al., 2008; Foster et al., 1997; Jiang et al., 2014; Kelly, 2009). Among them, Che et al. (2008) improved the Chang algorithm based on SD measurements from Chinese meteorological stations in response to the low snow density in China, and released two long-term time series SD datasets in China: Che_SSMI/S product and Che_SMSR2 product. 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).

3 ***Lines 83:*** *Line 83: Remove "in" for grammatical correctness.*

Reply: Thank you for correcting this grammatical issue. After adjusting the original text, we have deleted this sentence and checked the entire manuscript for similar errors. Thank you again for pointing out this issue.

*4 **Section 3.1.1:** Justify the use of the 23-GHz band given its sensitivity to water vapor.*

Reply 4: Thank you for raising this professional question—it has deepened our understanding of microwave remote sensing for SD inversion.

After careful review of literature on the 23GHZ frequency band, we found that 23GHZ is indeed sensitive to water vapor. However, due to its cancellation of negative sensitivity in the upper atmosphere and positive sensitivity in the lower atmosphere, the sensitivity of 23GHZ to TPW is very low on land. Meanwhile, studies have shown that its combined effect with other frequency bands can increase the accuracy of identifying shallow snow.

We have supplemented the justification for using the 23-GHz band, emphasizing its low water vapor sensitivity on land and its role in improving shallow snow identification:

(Lines: 230-240)

The SD inversion is affected by multiple factors, and initial research focused on the sensitivity of various microwave frequencies to snow cover. The SD inversion was carried out by using the BT values of each microwave frequency. Chang's algorithm is chiefly reliant on BT data from 18 GHz and 36 GHz in order to derive SD. Nevertheless, in regions characterised by shallow snow cover, the SD inversion results obtained using this algorithm demonstrate poor performance (Chang et al., 1987). While the 23 GHz frequency demonstrates high sensitivity to water vapour in the boundary layer, its sensitivity to water vapour in terrestrial regions is comparatively low. (Liu et al., 2021; Xing et al., 2022). Meanwhile, Kelly et al. (2009) indicated that a 23 GHz channel can be utilised for the identification of shallow snow. Furthermore, the capacity of different bands to express snow characteristics varies, and the combination of these bands is more conducive to the acquisition of more comprehensive snow information (Liu et al., 2021; Xing et al., 2022). Consequently, a number of scholars have employed the BT data supplied at 89 GHz, 23 GHz and 10 GHz in the context of SD inversion studies (Jiang et al., 2014; Kelly, 2009; Yang et al., 2020a).

*5 **Line 213-215:** Clarify why all AMSR-2 bands and band differences were selected as input features.*

Reply 5: We carefully reviewed relevant literature and found that different bands often exhibit different characteristics in terms of snow cover. The combination of these bands with other bands helps to obtain more comprehensive snow cover information. In addition, we did not use all bands and band differences for training. In the study, we used the correlation coefficient method to eliminate collinearity and performed feature selection. Regarding the 10V, 18V, 23V, 36V, and 89V in BT and the 10H23H, 10H36H, 18H36H, and 36H89H in BT difference, they were ultimately discarded. The final selected BTs are 10H, 18H, 23H, 36H, and 89H, with BT differences of 18V23H, 18H23H, 10V23V, 10V23H, 23V23H, 10V36H, 36V89V, and 18V36V, as well as a total of 19 influencing factors including lon, lat, slope, roughness, dem, and aspect.

Thank you very much for raising such a professional question, which is of great help to our in-depth understanding of snow depth inversion. This issue stems from our unclear description in the article. We have clarified the screening process and final selected features:

(Lines: 307-311)

Finally, 5 sets of BTs (10H, 18H, 23H, 36H, and 89H), 8 sets of BT differences (18V23H, 18H23H, 10V23V, 10V23H, 23V23H, 10V36H, 36V89V, and 18V36V), as well as longitude, latitude, slope, roughness, DEM, and aspect were selected. In conclusion, a total of 19 independent variables were selected for utilisation as input data for the AutoML model. The dependent variable, SD data, was applied in conjunction with the model during the training process.

6 ***Section 4.1.1:*** *The correlation coefficient is not a robust metric for variable selection. Consider alternative methods (e.g., feature importance from ML models).*

Reply 6: We solved the problem of collinearity by using a more traditional correlation coefficient method, without using feature importance or variance inflation factor methods. When designing this study, the main consideration was that the automatic machine learning framework actually performs feature selection again based on the input data. Using statistical methods or importance indicators of models to select the most helpful features for prediction. Therefore, in order to enable automatic machine learning to mine and preserve useful features as much as possible during operation, the most traditional correlation coefficient method is adopted for data preprocessing and feature selection.

Of course, the issue you pointed out is indeed a more unique insight, and we will consider more robust methods to eliminate collinearity in future work, rather than leaving it entirely to the black box of automatic machine learning.

Thank you for raising this crucial issue. We have supplemented the rationale for using the correlation coefficient and committed to adopting more robust methods in future work:

(Lines: 527-530)

(3) There are still areas that require enhancement in the research methodology of this study. In the context of feature selection, a more conventional approach, predicated on the calculation of the correlation coefficient, was utilised for the identification of influential factors. In the future, we will continue to refine our existing methods, including the adoption of more robust approaches for feature selection (feature importance, and variance inflation factor, among others).

7 ***Figure 9:*** *Revise captions for clarity (subfigures c and d are unclear).*

Reply 7: Thank you very much for pointing out this issue. The relevant parts are now modified as follows:

(Lines: 386-389)

Figure 9: Spatial error distribution between Auto_NSD data and observed SD at meteorological stations: (a) average SD at meteorological stations; (b) the average SD of Auto_NSD data; (c) Auto_NSD data and the BIAS of SD at meteorological stations; (d) Auto_NSD data and RMSE of SD at meteorological stations.

8 ***Section 4.3:*** *The SD-temperature relationship is well-known. Emphasize new insights (e.g., regional variability, model sensitivity) rather than restating basics.*

Reply 8: Thank you for your constructive suggestion—this helps us refocus on novel contributions. As the purpose of this article is to develop a more accurate, high-quality microwave remote sensing snow depth inversion model based on previous research, the relationship between snow depth and temperature has not been thoroughly investigated. Secondly, because the relationship between snow depth and temperature is well known and closely related, we use the relationship between snow depth and temperature to verify the rationality of the snow depth inversion model. However, implementing this research direction is currently relatively difficult. Thank you very much for your feedback. We will take it into account in our future work.

We have revised the discussion to acknowledge the limitations of our current analysis and propose future research directions for new insights:

(Lines: 528-531)

(4) This study merely conducted a cursory exploration of the relationship between SD and temperature. In future research, we will consider further exploring the relationship between SD and temperature in the QTP, such as seasonal patterns, spatial heterogeneity, sensitivity, and dependence.