EVALUATION OF REANALYSIS SNOW DEPTH DATASETS AND RETRIEVAL OF SNOW DEPTH BASED ON PASSIVE MICROWAVE DATA IN SELIN CO AND NAM CO REGIONS

Abstract

Snow cover is an important component of the cryosphere, which is sensitive to climate change and can affect regional energy balance and water cycle process. The Tibetan Plateau is the main snow cover distribution area in the middle and low latitudes of China. The snow cover in this area can affect China and even the global climate system. Snow meltwater also plays an important role in the social and economic development and natural environment stability of downstream countries and regions. Obtaining accurate surface snow depth information is the basis for snow resource assessment, water resource management and climate system simulation, which has important scientific significance and research value. At present, the main ways to obtain surface snow depth information on the Tibetan Plateau are the retrieval of snow depth based on passive microwave and reanalysis snow depth datasets, which have the advantages of wide coverage, high temporal resolution and low acquisition cost. Due to the scarcity of ground snow depth data, there is insufficient understanding of the empirical passive microwave snow-depth retrieval algorithm and the reanalysis snow depth datasets in the central and western Tibetan Plateau regions. In addition, the currently widely used empirical passive microwave snow-depth retrieval algorithms are mostly based on passive microwave brightness temperature difference, and less consideration is given to the influence of environmental factors such as terrain and underlying surface on the retrieval of snow depth, resulting in a large error in the empirical snow-depth retrieval algorithms. Therefore, this study mainly used the measured snow depth data of seven in-situ stations from July 2019 to June 2021 in the Selin Co and Nam Co regions of the central Tibetan Plateau to evaluate the performance

of four common reanalysis snow depths datasets, including ERA5, GLDAS, JRA55 and MERRA2, and five passive microwave snow-depth retrieval algorithms in the study area, including Chang2 algorithm, SPD algorithm, Che algorithm, Jiang algorithm, and AMSR2 algorithm. To recognize the influence of different brightness temperature channels and environment variables such as terrain and land cover types on snow-depth retrieval, and to provide a reference for the study of snow-depth retrieval algorithm based on passive microwave in the Tibetan Plateau, LASSO, SCAD, MCP penalty method, and random forest feature importance method were used to select 61 variables such as slope, aspect, altitude, land cover types, and 10, 18, 23, 37, 89 GHz vertical/horizontal polarization brightness temperature and different frequency brightness temperature difference of AMSR2. Four snow-depth retrieval variables groups were obtained. And all the variables selected by the four methods were used as the experimental control group (ALL method). And then, the five variables groups were combined with random forest regression and support vector regression, respectively, constructing ten models of machine learning for retrieving snow depth. The conclusions of this study are as follows:

(1) The monthly average snow days of the four reanalysis snow depth datasets, ERA5, GLDAS, JRA55, and MERRA2, were all higher than those of the stations' observations (6.59 d). The ERA5 reanalysis snow depth dataset had the best overall performance in the study area, with a correlation coefficient of 0.62. And it overestimated the observed snow depth of the study area overall (2.37 cm), of which the overestimation of deep snow was the largest (12.11 cm). The overall correlation coefficient between the JRA55 reanalysis snow depth dataset and the observed snow depth was 0.45, and the RMSE was the smallest among the four reanalysis snow depth datasets (4.06 cm). The GLDAS and MERRA2 reanalysis snow depth datasets performed the least well in the study area (values of R are 0.23 and 0.14, respectively), and the MERRA2 reanalysis snow depth dataset had the largest overall bias (11.71 cm) among the four analysis snow depth datasets. The correlation coefficient between the four reanalysis snow depth datasets and observed shallow snow was weak (less than 0.20), and the ERA5 and JRA55 reanalysis snow depth datasets had the highest correlation with observed snow depth (values of R are 0.80 and 0.47, respectively).

(2) The Jiang algorithm had the best overall performance among the five algorithms in the study area. The overall correlation coefficient between the snow depth retrieved by the Jiang algorithm and the observed snow depth was the highest (0.65),

and the RMSE was 5.51 cm. The Jiang algorithm outperformed the other four algorithms in snow-depth retrieval for mixed land cover type stations (R = 0.56). The RMSE of the Che algorithm is second only to that of the Jiang algorithm (7.36 cm). The correlation coefficient between snow depth retrieved by the Chang2 algorithm and observed snow depth was high (0.59), but the bias (-12.40 cm) and RMSE (17.00 cm) were large. The SPD algorithm and AMSR2 algorithm performed poorly (values of R are 0.42, 0.45, respectively). Among the five algorithms, the snow depth retrieved by the SPD algorithm overestimated the observed snow depth in the study area (15.80 cm), while the snow depth retrieved by the other four algorithms underestimated the observed snow depth in the study area. And the Chang2 algorithm, Jiang algorithm, and Che algorithm had better retrieval effects on medium snow, and the correlation coefficients between the retrieved snow depth and the observed snow depth were all greater than 0.70.

(3) The determination coefficients of the ten snow depth retrieval models were all high (0.63 ~ 0.82). The distribution of their bias was close to the normal distribution, and their MAE was small (all less than 3.2 cm). Among them, the random forest regression model L-RFR based on the LASSO method had the best performance (R^2 = 0.82). Except for the I-SVR model (R^2 = 0.63), the determination coefficients of the other nine models were all greater than 0.67, their MAE was all less than 3.05 cm, and their overall performances were better than those of the four reanalysis snow depth datasets (the maximum R was 0.62) and the five snow-depth retrieval algorithms based on passive microwave (the maximum R was 0.62). The determination coefficients of the models based on random forest regression (0.79 ~ 0.82) were all higher than those of the models based on support vector regression (0.63 ~ 0.70). Among the random forest regression or support vector regression models, the models based on the variable combination of the LASSO and SCAD methods performed the best, the models based on the variable combination of the MCP and the random forest feature importance method performed poorly, and the performances of models based on the variable combination of the ALL method were between those of the models based on the other four variable combinations.

Keywords: Passive Microwave, Snow Depth, Selin Co, Nam Co, Machine Learning, Snow-depth Retrieval Algorithm, Reanalysis Snow Depth Datasets