

DOI:10.7522/j.issn.1000-0240.2020.0080

WANG Jianshun, WANG Yunlong, ZHOU Minqiang, et al. Retrieved snow depth over the Tibetan Plateau using random forest algorithm with AMSR2 passive microwave data[J]. Journal of Glaciology and Geocryology, 2020, 42(3):1077-1086. [王健顺, 王云龙, 周敏强, 等. 基于随机森林算法的青藏高原AMSR2被动微波雪深反演[J]. 冰川冻土, 2020, 42(3):1077-1086.]

## 基于随机森林算法的青藏高原AMSR2 被动微波雪深反演

王健顺<sup>1</sup>, 王云龙<sup>1</sup>, 周敏强<sup>1</sup>, 刘畅宇<sup>1</sup>, 黄晓东<sup>2</sup>

(1. 兰州大学 草地农业生态系统国家重点实验室/草地农业科技学院, 甘肃 兰州 730020; 2. 南京信息工程大学 地理科学学院, 江苏 南京 210044)

**摘要:** 青藏高原因其复杂的地形地势和积雪分布使得多种雪深算法未达到理想的精度。基于新一代被动微波数据 AMSR2 (Advanced Microwave Scanning Radiometer 2), 应用随机森林算法 (Random Forest, RF) 将亮温 (Brightness Temperature, BT) 和亮温差 (Brightness Temperature Difference, BTD) 作为参数输入, 并将高程和纬度参数引入雪深反演模型中, 经过模拟退火算法进行有效反演因子筛选, 构建了基于随机森林算法的青藏高原雪深反演模型。结果表明: 与 AMSR2 全球雪深产品相比, 随机森林算法的拟合优度 ( $R^2$ ) 由 0.41 提升至 0.60, 均方根误差 (Root Mean Square Error, RMSE) 由 7.36 cm 降至 4.88 cm, 偏差 (BIAS) 由 3.24 cm 减小至 -0.16 cm, 随机森林雪深反演模型在青藏高原的精度更高; 青藏高原平均海拔超过 4 000 m, 当海拔大于青藏高原平均海拔时, 随机森林算法的反演效果最差, 但 RMSE 仅为 3.78 cm, BIAS 仅为 -0.09 cm; 高原南部 ( $25^\circ \sim 30^\circ \text{N}$ ) 因其复杂的地势和相对较少的气象站点使得反演效果较差, RMSE 为 5.94 cm, BIAS 为 -0.39 cm; 青藏高原的主要土地覆盖类型为草地, 随机森林算法在草地的 RMSE 约为 3 cm, BIAS 接近 0 cm。

**关键词:** 随机森林算法; 青藏高原; 雪深反演; AMSR2

**中图分类号:** P426.63\*5 **文献标志码:** A **文章编号:** 1000-0240(2020)03-1077-10

### 0 引言

积雪在全球水循环和能量交换中扮演着重要的角色<sup>[1-2]</sup>, 对研究区域气候变化、水资源利用以及自然灾害监测十分重要<sup>[1,3-5]</sup>。积雪深度是描述积雪性质的重要参数之一, 在雪灾监测与预警、水资源评估、气候水文模拟、地表辐射平衡中都发挥着至关重要的作用<sup>[6-10]</sup>。青藏高原对于全球气候变化反应敏感, 是全球气候变化的“驱动力”与“放大器”<sup>[11]</sup>, 对我国乃至亚洲生态安全具有重要的屏障作用<sup>[12]</sup>。青藏高原素有“亚洲水塔”之称, 是我国众多河流的发源地, 而积雪融水是河流的重要补充来

源。在全球气候变化的背景下, 近年来青藏高原冰雪融化加速, 常年积雪范围不断缩小, 积雪累积总量也在减少, 生态环境持续恶化<sup>[13-17]</sup>。青藏高原是中国的三大积雪分布区之一, 也是我国重要的牧区, 积雪的变化与当地农牧业发展密切相关<sup>[18]</sup>。因此, 青藏高原积雪研究对区域生态环境和农牧业可持续发展具有重要意义。

卫星观测是监测积雪的有效方式, 被动微波遥感数据是监测积雪变化的主要手段之一<sup>[19-22]</sup>。微波可以穿透气溶胶层, 云、雾等对低频微波散射较小<sup>[23-24]</sup>, 且可穿透一定深度的地表 (如积雪) 来获取地物信息<sup>[25]</sup>, 是目前获取区域雪深和雪水当量监测

收稿日期: 2019-04-30; 修订日期: 2020-08-11

基金项目: 国家自然科学基金项目 (41691330; 41971293); 科技部基础资源调查专项 (2017FY100501); 教育部长江学者和创新团队发展计划人才项目 (IRT\_17R50) 资助

作者简介: 王健顺 (1994 -), 男, 甘肃秦安人, 2017 年在甘肃农业大学获学士学位, 现为兰州大学在读硕士研究生, 从事积雪遥感研究. E-mail: wangjsh17@lzu.edu.cn

通信作者: 黄晓东, 教授, 从事冰冻圈遥感与地理信息系统研究. E-mail: huangxd@lzu.edu.cn.

的主要数据源<sup>[26-27]</sup>。Chang等<sup>[28]</sup>利用SMMR被动微波数据进行雪深反演,并提出了雪深反演的亮温梯度法,基本满足半球或全球尺度的雪深反演,但是在局部区域误差较大<sup>[28-29]</sup>。Foster等<sup>[21]</sup>以北美大陆和亚欧大陆为研究对象,考虑森林和雪粒径的影响可以提高雪深反演的精度,并在此基础上提出了修正的Chang算法。Tait<sup>[30]</sup>将美国和俄罗斯划分成不同的下垫面类型,利用SMM/I被动微波数据进行雪深反演,发现在湿雪存在的无森林覆盖区反演效果较差。曹梅盛等<sup>[31]</sup>将我国西部划分为不同的地貌单元,基于SMMR被动微波数据发展了订正后的Chang算法,提高了反演精度。车涛等<sup>[7]</sup>使用SMM/I被动微波数据,在剔除了与积雪相似散射特性的降雨、寒漠、冻土等地表类型后,利用修正的Chang算法反演了中国区域的雪深,并获得了较高精度。Dai等<sup>[2]</sup>基于积雪不同的雪粒径、雪密度和温度,利用AMSR-E被动微波数据发展了一套适合于新疆的雪深反演算法。Zhang等<sup>[32]</sup>对北疆地区AMSR2雪深产品的精度进行评价,结果显示其精度较差,并对AMSR2全球算法进行了改进,此改进算法主要利用10 GHz通道差和18 GHz通道差对算法改进。青藏高原海拔高且地形复杂,平均雪深较浅,1979-2010年间的多年平均雪深小于5 cm<sup>[33]</sup>。以上基于经验-半经验雪深反演算法在该地区的积雪深度反演精度均不够理想,存在很大的局限性,验证结果显示仅依赖亮温梯度差进行雪深反演在青藏高原的均方根误差(RMSE)高达9 cm左右<sup>[34]</sup>。

目前,机器学习算法因其精度高的特点已经被广泛运用于遥感定量反演的研究中<sup>[35]</sup>。侯海艳等<sup>[36]</sup>利用人工神经网络(ANN)在北疆地区进行了雪深反演,引入地形等辅助因子对雪深反演精度提升作用显著。Liang等<sup>[37]</sup>将SSM/I和SMMI/S被动微波数据与MODIS的可见光和近红外数据结合,利用支持向量机(SVM)进行了对北疆地区雪深进行反演,此模型的RMSE为6.21 cm。Xiao等<sup>[38]</sup>利用机器学习算法反演了亚欧大陆的雪深,与Chang算法等对比发现机器学习算法反演精度明显提高。目前,全球通用积雪深度反演算法在青藏高原地区的精度均不理想<sup>[7-8,39-40]</sup>,而机器学习以其高效率且模拟精度较高的特点在陆表参数定量反演中日益得到重视,如随机森林能够克服回归中过拟合的问题,对训练集的噪声数据不敏感,且能更好地应用

于大数据集分析<sup>[35]</sup>。

影响积雪分布的因子主要有海拔、坡度、坡向、地表粗糙度、风速和风向等<sup>[41-42]</sup>。Sturm等<sup>[25]</sup>、Xiao等<sup>[38]</sup>和Savoie等<sup>[42]</sup>将地理位置等辅助因子引入到雪深的反演模型中,一定程度上提高了雪深反演的精度。传统被动微波遥感雪深反演算法仅依赖亮度温度差的方法在青藏高原地区存在很大的局限性,因此引入辅助因子,借助机器学习优势,是提高青藏高原雪深反演的有益尝试。针对青藏高原雪深普遍较浅和空间异质性大的特点,本文以青藏高原为研究区,基于AMSR2被动微波亮温数据和地面雪深观测数据,结合地形和地理参数,利用随机森林算法进行雪深反演,以期实现青藏高原地区的高精度雪深反演。

## 1 研究区概况

青藏高原位于我国西南部,东抵横断山脉,西达帕米尔高原,南至喜马拉雅山脉,北到昆仑山-祁连山一线,面积约 $250\times 10^4$  km<sup>2</sup>(图1)。被称为地球的“第三极”和“世界屋脊”,平均海拔在4 000 m以上,是世界上平均海拔最高的高原,也是我国的三大积雪分布中心之一和重要的水源地,我国的众多河流均发源于此。该地区大部分地区年均气温较低,最暖月平均温度在10℃以下<sup>[43]</sup>。青藏高原积雪分布极其不均,雪深较浅,高原东南、西部和南部为雪深分布高值区,自1979年以来,高原大部分地区的雪深呈增加趋势,其中昆仑高寒荒漠地带雪深增加最为明显<sup>[44]</sup>。

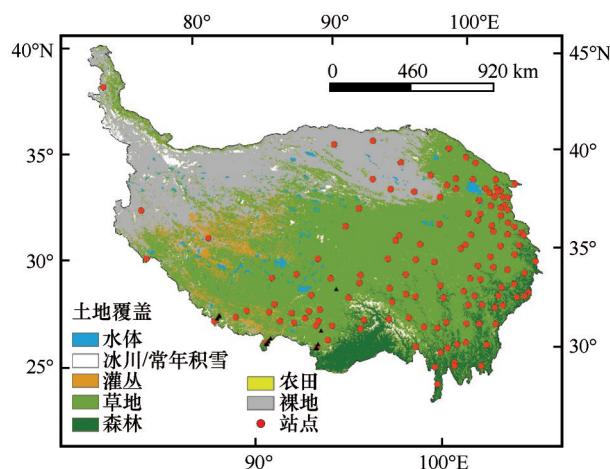


图1 青藏高原气象站点分布及土地覆盖类型 (MCD12Q1-IGBP)

Fig. 1 The distribution of meteorological stations and land cover type over the Tibetan Plateau

## 2 数据来源

### 2.1 AMSR2 微波数据

AMSR2 亮温数据和 AMSR2 全球雪深产品来源于日本航空航天研究开发机构网站 (<https://gcom-w1.jaxa.jp/auth.html>)。AMSR2 是继 AMSR-E 之后又一新型的传感器, 它搭载在日本的 GCOM-W1 (Global Change Observation Mission 1st-Water) “SHIZUKU” 卫星上, 该传感器包括 6.9、7.3、10.65、18.7、23.8、36.5 GHz 和 89 GHz 共 7 个频率, 每个频率有水平和垂直两种极化方式(表 1)。

表 1 AMSR2 传感器参数

Table 1 The sensor Parameter of AMSR2

| 中心频率/GHz     | 6.9   | 7.3   | 10.65 | 18.7  | 23.8  | 36.5  | 89    |
|--------------|-------|-------|-------|-------|-------|-------|-------|
| IFOV/(km×km) | 35×62 | 35×62 | 24×42 | 14×22 | 15×26 | 7×12  | 3×5   |
| 波束宽度/MHz     | 350   | 350   | 100   | 200   | 400   | 1 000 | 3 000 |
| 灵敏度/K        | 0.34  | 0.43  | 0.7   | 0.7   | 0.6   | 0.7   | 1.2   |

该数据的空间分辨率为 10 km, 投影方式为 EQR, 数据格式为 Geo-TIFF。GCOM-W1 卫星一天有两次过境时间, 升轨的过境时间是 13:30(±15 min)(当地时间, 下同), 降轨是 01:30(±15 min)。积雪中液态水的存在会吸收一部分微波辐射能量<sup>[23-24]</sup>, 因此, 升轨数据过境时间存在融雪的影响, 本研究选用降轨数据用于青藏高原雪深反演。

卫星接收到来自地面的上行微波辐射被记录为亮温( $T_b$ ),  $T_b$  可表示为:

$$T_b = (RT_{sky} + (1-R)T_{surf})e^{-\tau} + T_{atm} \quad (1)$$

式中:  $e^{-\tau}$  为大气透过率;  $R$  为表面反射率;  $T_{sky}$  为天空辐射;  $T_{surf}$  为表面发射;  $T_{atm}$  为大气发射。通常情况下  $T_{atm}$  和  $T_{sky}$  非常小, 可以忽略不计。因此, 大气的辐射和散射对微波影响很小, 可以忽略不计<sup>[27]</sup>。研究共选取 10.65、18.7、23.8、36.5、89 GHz(以下简称 10、18、23、36、89 GHz)共 5 个频率的亮温数据, 包括所有频率的垂直和水平通道。因为青藏高原降雪主要集中在 12 月至次年 3 月, 本研究获取的是 2012–2016 年积雪季的 AMSR2 亮温数据。

AMSR2 全球雪深产品的雪深反演算法主要利用 18 GHz 通道差和 36 GHz 通道差, 以及 10 GHz 垂直通道与 18 GHz 垂直通道和 36 GHz 垂直通道的亮温差构建雪深反演模型<sup>[32]</sup>。

### 2.2 气象台站雪深数据与野外调查数据

气象台站雪深数据来源于中国国家气象科学数据共享服务平台 (<http://data.cma.cn/site/index.html>)。

此数据包括台站号、观测日期、经度、纬度、海拔、日积雪深度和日雪压等资料, 气象站点分布如图 1 所示。积雪野外调查数据包括 2018 年 3 月 1 日–3 月 8 日 13 个采样点的数据, 样点测量参数主要包括雪深、雪水当量、积雪密度以及积雪含水量等, 同时记录采样点的经纬度和海拔信息。

### 2.3 辅助数据

积雪的再分配和积雪的微波辐射受不同土地覆盖类型的影响<sup>[21,30]</sup>。土地覆盖类型选取 2014 年 MODIS 土地覆盖产品 MCD12Q1, 分类方案为 IGBP (International Geo-sphere-Biosphere Program), 空间分辨率为 500 m。本研究将 MCD12Q1-IGBP 分类方案重分类为 7 大类: 水体、森林、灌丛、草地、农田、裸地和冰川/常年积雪, 其中草地面积占青藏高原 56.9%, 其次是裸地 (29.6%)。SRTM (Shuttle Radar Topography Mission) DEM (Digital Elevation Model) 数据来源于中国科学院计算机网络信息中心国际科学数据镜像网站 (<http://www.gscloud.cn>), 空间分辨率为 90 m, 数据格式 Geo-TIFF。

## 3 研究方法

本研究利用 AMSR2 被动微波数据和其他辅助遥感数据 (MODIS 和 DEM), 首先剔除与积雪具有相似散射特性的散射体, 然后利用模拟退火算法筛选对雪深变化敏感的反演因子, 最后构建随机森林模型用于青藏高原雪深反演, 与 AMSR2 全球雪深产品进行对比, 并评价不同的土地覆盖、海拔和地理位置因子对雪深模型反演精度的影响。

### 3.1 散射体剔除

寒漠、冻土和降水与积雪有相似的散射特性<sup>[45]</sup>, 因此在建立雪深反演模型前要剔除这些散射体的影响, 从而提高积雪反演算法的精度。在 Grody 等<sup>[46]</sup>的研究基础上, Kelly 等<sup>[45]</sup>利用 AMSR-E 被动微波数据剔除散射体的影响并建立了雪深反演模型, AMSR-E 数据与 AMSR2 数据有相似的频率通道, 因此本研究将利用 Kelly 的方法以剔除散射体的影响。积雪中的存在湿雪同样会影响雪深反演<sup>[47-48]</sup>。因此, 还采用 Walker<sup>[47]</sup>的方法剔除湿雪的影响。上述散射体剔除方法同样被包慧漪等<sup>[48]</sup>与张显峰等<sup>[49]</sup>利用在 AMSR-E 被动微波数据雪深反演中。另外气象站附近存在大型水体时, 会严重影响雪深的反演<sup>[13,32,50]</sup>, 因此在雪深反演算法建立之前要剔除受水体影响的气象站点数据。散射体剔

除方法如下：

- 1) 散射体：Tb18V-Tb36V>0 K；
- 2) 降雨：Tb23V≥259 K 或 254 K≤Tb23V≤258 K 且 Tb18V-Tb36V≤2 K；
- 3) 寒 漠：Tb18V-Tb18H≥18 K 且 Tb18V-Tb36V≤10 K 且 Tb36V-Tb89V≤10 K；
- 4) 冻土：Tb18V-Tb18H≥8 K 且 Tb18V-Tb36V≤2 K 且 Tb23V-Tb89V≤6 K；
- 5) 湿雪：Tb36V-Tb36H≥10 K。

### 3.2 模拟退火算法

多余和无用的变量信息不仅会降低模型运行效率，增加模型运行时间，而且还可能干扰对有效变量的筛选，降低模型的精度<sup>[51]</sup>。模拟退火算法(Simulate Annealing, SA)是一种基于蒙特卡洛思想(Monte-Carlo)的启发式优化算法，该算法的核心思维是避免局部最优解<sup>[52]</sup>。本研究采用的因子包括亮温(BT)、亮温差(BTD)、海拔(Elevation, Ele)、经度(Longitude, Lon)和纬度(Latitude, La)，研究中选用 10、18、23、36 GHz 和 89 GHz 的水平和垂直极化亮温数据产生 45 个亮温差，共组成 58 个反演因子集。利用模拟退火算法在 58 个反演因子中筛选最优反演因子集用于随机森林雪深反演模型构建。

### 3.3 随机森林

随机森林(Random Forest, RF)由 Breiman<sup>[53]</sup>提出，是一种组成式的有监督学习方法。随机森林通过构建无数小决策树来预测，这些小的决策树修剪是基于训练集中另外一个随机的样本子集。随机森林可以克服回归的过拟合问题、噪声数据和大数据集的问题，已经被越来越多的应用于定量参数估算中<sup>[36]</sup>。本研究中的随机森林建模方法通过 R 软件中的“randomForest”数据包实现。随机森林的调优参数是每个切分点上随机选择的预测变量的数目  $k$ ，它通常被记为  $mtry$ 。在使用随机森林进行回归预测时，有两个参数  $ntree$  和  $mtry$  需要预先设定，其中  $ntree$  表示决策树的数量，也就是对原始训练数据集使用 bootstrap 抽样的次数。当  $ntree$  越大时，模拟因变量的结果就越稳定<sup>[51]</sup>；但如果设置太大会增加计算量消耗计算机性能。Breiman<sup>[53]</sup>推荐将  $mtry$  设为总预测变量数目的三分之一，本研究也采用此方法。

### 3.4 精度验证

本研究采用 10 折交叉验证法构建模型并评估

模型的精度。将雪深观测值分为样本量相等的 10 组，9 组用于模型训练，剩余 1 组用于验证模型的精度，循环 10 次，取 10 次循环的平均均方根误差(RMSE)和拟合优度( $R^2$ )作为模型评价标准<sup>[54-55]</sup>。模型的  $R^2$  越高，RMSE 越小，模型的精度越高。

$$R^2 = \frac{\sum_{i=1}^n (\widehat{y}_i - \bar{y})^2 / \sum_{i=1}^n (y_i - \bar{y})^2}{1} \tag{2}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \widehat{y}_i)^2} \tag{3}$$

$$BIAS = \frac{1}{n} \sum_{i=1}^n (y_i - \widehat{y}_i) \tag{4}$$

式中： $y_i$ 、 $\widehat{y}_i$  和  $\bar{y}$  分别为实际雪深值、预测雪深值和平均实际雪深值； $n$  为用于验证的样本数。

## 4 结果与分析

### 4.1 因子筛选结果

如表 2 所示，经过模拟退火算法的因子筛选策略筛选之后，从 58 个反演因子中筛选出 19 个反演

表 2 模拟退火算法变量筛选结果  
Table 2 The results of selected variables with simulate annealing algorithm

| 地形、地理参数 |   | 亮温 (BT) |   | 亮温差 (BTD) |   |        |   |        |   |
|---------|---|---------|---|-----------|---|--------|---|--------|---|
| La      | √ |         |   | 10H10V    | √ | 10H18H | √ | 10H18V | × |
| Lon     | × | 10H     | √ | 10H23H    | × | 10H23V | × | 10H36H | √ |
| Ele     | √ |         |   | 10H36V    | × | 10H89H | × | 10H89V | √ |
|         |   |         |   | 10V18H    | × | 10V18V | × | 10V23H | × |
|         |   | 10V     | √ | 10V23V    | × | 10V36H | × | 10V36V | × |
|         |   |         |   | 10V89H    | × | 10V89V | × |        |   |
|         |   |         |   | 18H18V    | × | 18H23H | √ | 18H23V | × |
|         |   | 18H     | × | 18H36H    | × | 18H36V | × | 18H89H | × |
|         |   |         |   | 18H89V    | √ |        |   |        |   |
|         |   | 18V     | × | 18V23H    | × | 18V23V | √ | 18V36H | × |
|         |   |         |   | 18V36V    | √ | 18V89H | × | 18V89V | × |
|         |   | 23H     | × | 23H23V    | √ | 23H36H | × | 23H36V | √ |
|         |   |         |   | 23H89H    | × | 23H89V | √ |        |   |
|         |   | 23V     | √ | 23V36H    | × | 23V36V | × | 23V89H | × |
|         |   |         |   | 23V89V    | × |        |   |        |   |
|         |   | 36H     | × | 36H36V    | √ | 36H89H | √ | 36H89V | × |
|         |   | 36V     | × | 36V89H    | × | 36V89V | × |        |   |
|         |   | 89H     | × | 89H89V    | × |        |   |        |   |
|         |   | 89V     | √ |           |   |        |   |        |   |

注：“√”表示该变量入选，“×”表示该变量未入选；表中 La、Lon 和 Ele 分别代表纬度 (Latitude)、经度 (Longitude) 和海拔 (Elevation)；10H 和 10V 为 10 GHz 的水平和垂直通道，其他类似；10H18V 为 10 GHz 的水平通道和 18 GHz 的垂直通道的亮温差，其他表述类似。

因子。10 GHz 垂直通道的亮温差组合入选最多, 其次是 18 GHz 和 23 GHz 的垂直通道; 18 GHz 有最多的亮温差组合入选。另外, 10、23 GHz 和 36 GHz 的极化亮温差也通过了模拟退火算法的变量筛选。单通道变量中高频亮温数据仅 89 GHz 的垂直通道通过筛选, 低频通道中 10 GHz 的水平垂直通道全部入选。青藏高原平均海拔较高, 大气稀薄, 更有利于低频微波(如 10 GHz)的传输, 且低频微波更有利于穿透地表覆盖物获取信息。此外, 水平通道比垂直通道对雪深的变化反应更敏感<sup>[56]</sup>, 可能这是水平通道的亮温差组合入选更多的一个原因。

#### 4.2 雪深反演精度验证

基于筛选得到的因子集, 利用随机森林算法构建青藏高原雪深反演模型, 10 折交叉验证结果如表 3 所示, 随机森林雪深算法精度显著地优于 AMSR2 全球雪深产品。相比于 AMSR2 雪深产品, 随机森林反演模型的  $R^2$  由 0.41 提升至 0.60, RMSE 由 7.36 cm 减小至 4.88 cm, BIAS 由 3.24 cm 减小至 -0.16 cm。

图 2 对筛选得到的各反演因子重要性进行评

表 3 随机森林算法与 AMSR2 全球雪深产品的对比

Table 3 Comparison of RF algorithm and AMSR2 global

SD algorithm

|       | 随机森林算法 |       | AMSR2  |
|-------|--------|-------|--------|
|       | 训练集    | 测试集   | 全球雪深产品 |
| $R^2$ | 0.93   | 0.60  | 0.41   |
| RMSE  | 2.13   | 4.88  | 7.36   |
| BIAS  | -0.07  | -0.16 | 3.24   |

分, 结果表明纬度(Latitude, La)和 10 GHz 的水平通道相对重要性最高, 且 10 GHz 的水平通道重要性高于垂直通道。在重要性较小的亮温差组合中, 89 GHz 与其他频率的组合占比较多。虽然高频信号对雪深变化较为敏感, 但高频信号更易受到大气条件的影响<sup>[57-59]</sup>。青藏高原大气稀薄, 对微波的辐射和散射作用较小<sup>[24]</sup>, 但这并不能完全消除这种影响, 这可能是 89 GHz 通道重要性偏低的原因。图 3 表明随机森林算法的拟合优度( $R^2$ )高达 0.88, 而 AMSR2 全球雪深产品仅为 0.41, 该产品对青藏高原雪深低估明显, 且对多数雪深超过 5 cm 的区域严重低估。

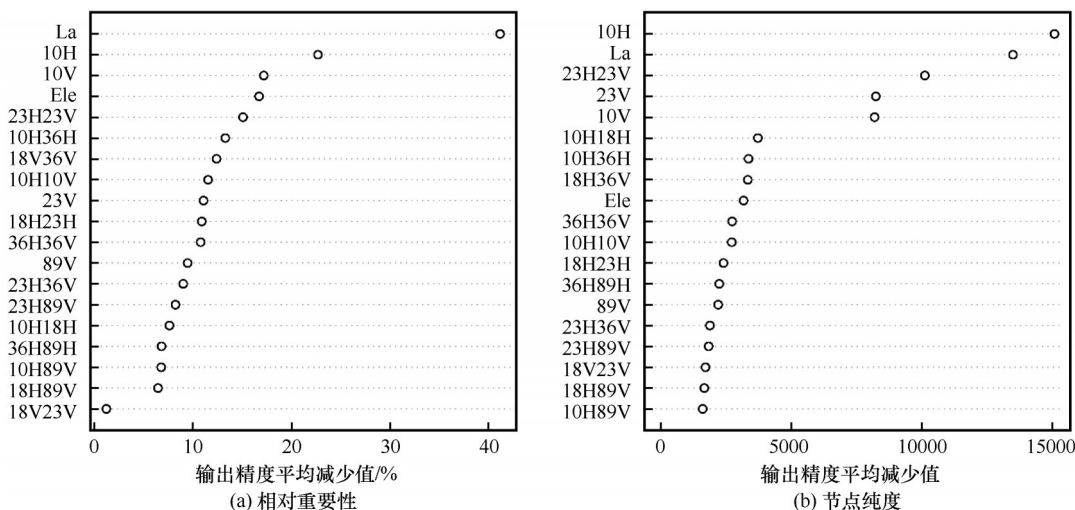


图 2 随机森林雪深反演变量重要性评价

Fig. 2 The importance assessment of selected variables with RF algorithm

#### 4.3 雪深反演影响因素分析

用于本研究的 133 个气象站和 13 个采样点大多分布在海拔 1 100 ~ 4 800 m 之间, 将这些气象站依据海拔划分为三类, 即 <2 800 m、2 800 ~ 3 800 m、>3 800 m; 由南向北划分为三类, 即南部(25° ~ 30° N)、中等纬度部分(30° ~ 35° N)、北部(35° ~ 40° N)。图 4 中, 随机森林算法在不同的海拔和纬度下的精度都优于 AMSR2 全球雪深产

品。如图 4(a), 气象站海拔为 2 800 ~ 3 800 m 之间时, 随机森林雪深反演的误差最小, RMSE 仅为 1 cm 左右, BIAS 仅为 -0.5 cm, 较小的低估了这一区域的雪深。当海拔超过 3 800 m 时, RMSE 相对于 3 800 m 以下区域的误差较大, RMSE 为 3.78 cm, BIAS 为 -0.09 cm。随海拔的上升, 气温逐渐降低, 积雪消融变缓; 海拔越高, 积雪深度越大, 导致青藏高原高海拔地区的雪深反演 RMSE

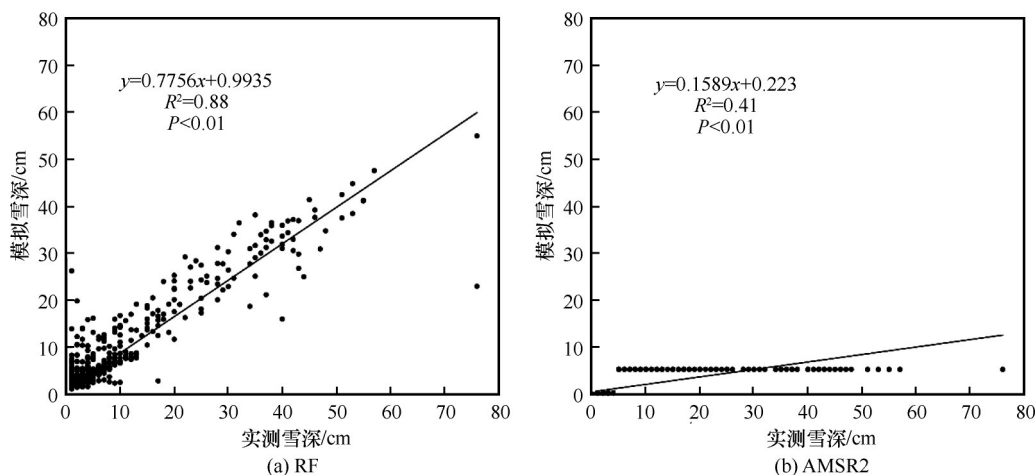


图3 基于气象站点随机森林算法与 AMSR2 全球雪深产品的对比

Fig. 3 Comparison of RF algorithm and AMSR2 global SD algorithm based on the whole meteorological stations

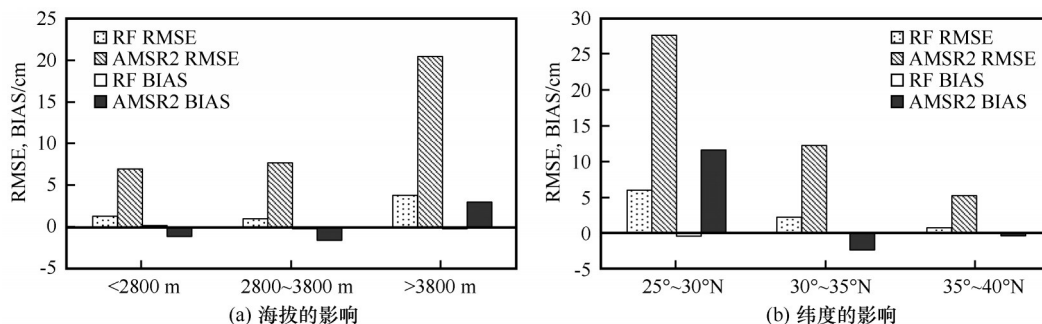


图4 海拔(Elc)和纬度(La)对青藏高原随机森林雪深反演的影响

Fig. 4 The influence of elevation and latitude in RF algorithm SD retrieval over the Tibetan Plateau

增大。图4(b)表明随机森林雪深反演在青藏高原南部地区的误差最大,无论是RMSE还是BIAS都相对较大;中纬度地区的反演误差处于二者之间, RMSE约为2 cm, BIAS接近于0;青藏高原北部地区反演误差最小,这主要是由于青藏高原南部地区各站点雪深普遍较大,此区域的喜马拉雅山脉和念青唐古拉山是青藏高原雪深高值区,导致这些区域的雪深反演误差相对较大,而积雪深度较浅的高原腹地和北部区域,积雪反演相对误差较小。

草地是青藏高原地区主要的土地覆盖类型,其次为裸地,因此,雪深反演模型的精度主要取决于草地区域。通过统计研究期内气象站数据和野外积雪观测数据,草地区域的平均雪深超过10 cm,森林下垫面的平均雪深超过5 cm。图5表明草地的反演误差在所有土地覆盖类型中最高,但RMSE仅为3 cm左右,BIAS更小;随机森林算法对森林区域略有高估,其余土地覆盖类型均有不同程度的低估,对草地覆盖区域低估较小。AMSR2雪深产品

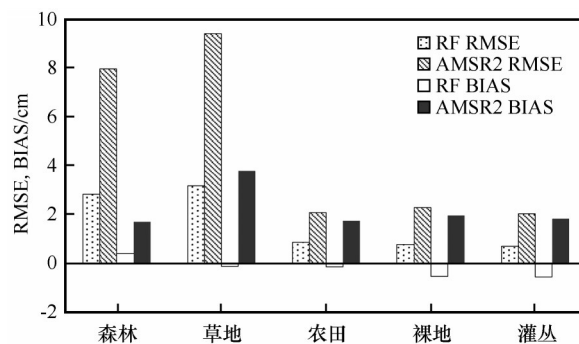


图5 土地覆盖类型对青藏高原随机森林雪深反演的影响

Fig. 5 The influence of land cover in RF algorithm SD retrieval over the Tibetan Plateau

在各个土地覆盖类型中,精度均比随机森林反演模型的精度差。

## 5 讨论

青藏高原是世界上重要的中纬度高海拔积雪区,相比于其他积雪区,其地形更加复杂,这也对雪深的高精度反演造成诸多的限制。机器学习算法已被广泛运用于陆表定量遥感参数的估算中,并

且其反演精度高。青藏高原 5 种雪深算法 (Chang 算法、GSFC96 算法、AMSR-E 算法、青藏高原改进算法和 Savoie 算法) 中 AMSR-E 算法的平均 RMSE 高于  $17\text{ cm}^{[37-38]}$ , 而 AMSR2 雪深产品是主要继承了此算法发展的雪深产品; 可见, 利用随机森林在青藏高原的雪深精度提升较为明显。

青藏高原平均海拔高于  $4\ 000\text{ m}$ , 当对高于平均海拔的地区进行雪深反演时, 其精度明显的低于较低海拔地区。利用随机森林在青藏高原进行雪深反演时, 有必要引入海拔和纬度。青藏高原北部地区 ( $35^{\circ} \sim 40^{\circ}\text{ N}$ ) 主要有柴达木盆地和河湟谷地等, 平均海拔是青藏高原上最低的地区, 地形分布相比中纬度地区 ( $30^{\circ} \sim 35^{\circ}\text{ N}$ ) 和南部地区 ( $25^{\circ} \sim 30^{\circ}\text{ N}$ ) 要缓和简单许多, 因此, 此区域的雪深反演精度较高。而南部地区分布有喜马拉雅山脉、念青唐古拉山和横断山脉等山脉, 此区域地形复杂多变, 且存在林区, 是青藏高原地区雪深反演最困难的区域之一。高原  $90^{\circ}\text{ E}$  以西分布大量水体, 且多为人迹罕至的高海拔无人区, 此区域气象数据获取困难, 使得随机森林模型在该地区的适应性较差, 该区域的反演精度有待进一步的提高。此外, 青藏高原高海拔区域气象站点分布稀少, 且缺乏野外数据采集, 导致该区域的反演精度差。

青藏高原主要的土地覆盖类型为草地, 其气象站点多分布于城镇附近, 因此有较多的气象站点下垫面为农田。草地、农田和裸地均为低矮植被覆盖区域, 其对微波的散射作用较小; 而森林和灌丛地区对微波辐射传输有一定影响, 尤其是下垫面是森林的地区。青藏高原草地相对于其他下垫面类型精度较低, 但其 RMSE 也仅在  $3\text{ cm}$  左右, 绝对偏差 (IBIASI) 未超过  $0.5\text{ cm}$ , 其精度优于整个青藏高原的反演精度。青藏高原林区分布较少, 其较高的精度不能证明随机森林算法适合于林区的雪深反演。此外, 近期的研究显示, 大气条件也会影响被动微波遥感反演地表参数; 因此, 在今后的雪深反演中, 大气条件将是必须考虑的问题<sup>[57-59]</sup>。

## 6 结论

本研究利用机器学习中的随机森林算法用于青藏高原的雪深反演, 为了筛选有效的因子信息和增加模型精度, 使用模拟退火算法进行了因子筛选, 同时在模型中引入了海拔和纬度作为参数用于随机森林雪深反演, 并取得较为理想的反演精度。

基于以上分析, 研究得出以下结论:

(1) 当利用模拟退火算法优化后的随机森林进行青藏高原雪深反演时, 各亮温通道与低频亮温的组合更适合于雪深反演。

(2) 对于地形地势复杂的青藏高原地区, 海拔和纬度引入随机森林雪深反演模型, 可有效提高雪深反演精度。

(3) 对于青藏高原海拔较高和地形、地势复杂的地区反演精度较差, 尤其是气象站点分布较为稀疏的高海拔区域。

综上, 对于地形复杂多变且气象数据获取困难的青藏高原地区, 随机森林算法基本实现了精度较高的雪深反演, 基本适用于青藏高原的大部分地区, 可为类似地区雪深反演模型的建立提供借鉴。

## 参考文献 (References):

- [1] Qin Dahe. Introduction to cryosphere science [M]. Beijing: Science Press, 2017. [秦大河. 冰冻圈科学概论 [M]. 科学出版社, 2017.]
- [2] Dai L Y, Che T, Wang J, Zhang P. Snow depth and snow water equivalent estimation from AMSR-E data based on *a priori* snow characteristics in Xinjiang, China [J]. Remote Sensing of Environment, 2012, 127: 14 - 29.
- [3] Hall D K, Influence of depth hoar on microwave emission from snow in northern Alaska [J]. Cold Region Science Technology, 1987, 13: 225 - 231.
- [4] Foster J L, Hall D K, Eylander J B, et al. A blended global snow product using visible, passive microwave and scatterometer satellite data [J]. International Journal of Remote Sensing, 2011, 32(5): 1371 - 1395.
- [5] Yang Xingguo, Qin Dahe, Qin Xiang. Progress in the study of interaction between ice/snow and atmosphere [J]. Journal of Glaciology and Geocryology, 2012, 34(2): 392 - 402. [杨兴国, 秦大河, 秦翔. 冰川/积雪 - 大气相互作用研究进展 [J]. 冰川冻土, 2012, 34(2): 392 - 402.]
- [6] Li Hongmei, Li Lin, Gao Ge, et al. Snow disaster in Qinghai Plateau: risk division and countermeasure [J]. Journal of Glaciology and Geocryology, 2013, 35(3): 656 - 661. [李红梅, 李林, 高歌, 等. 青海高原雪灾风险区划及对策建议 [J]. 冰川冻土, 2013, 35(3): 656 - 661.]
- [7] Che Tao, Li Xin, Gao Feng. Estimation of snow water equivalent in the Tibetan Plateau using passive microwave remote sensing data (SSM/I) [J]. Journal of Glaciology and Geocryology, 2004, 26(3): 363 - 368. [车涛, 李新, 高峰. 青藏高原积雪深度和雪水当量的被动微波遥感反演 [J]. 冰川冻土, 2004, 26(3): 363 - 368.]
- [8] Che Tao, Li Xin. Retrieval of snow depth in China by passive microwave remote sensing data and its accuracy assessment [J]. Remote Sensing Technology and Application, 2004, 19(5): 301 - 306. [车涛, 李新. 利用被动微波遥感数据反演我国积雪深度及其精度评价 [J]. 遥感技术与应用, 2004, 19(5): 301 - 306.]
- [9] Zhang Y S, Wang S S, Barr A G, et al. Impact of snow cover on soil temperature and its simulation in a boreal aspen forest [J]. Cold Regions Science and Technology, 2008, 52(3):

- 355 – 370.
- [10] Tedesco M, Narvekar P S. Assessment of the NASA AMSR-E SWE product[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2010, 3(1): 141 – 159.
  - [11] Pan Baotian, Li Jijun. Qinghai-Tibetan Plateau: a driver and amplifier of the global climatic change[J]. Journal of Lanzhou University, 1996, 32(1): 108 – 115. [潘保田, 李吉均. 青藏高原: 全球气候变化的驱动力与放大器. III. 青藏高原隆起对气候变化的影响[J]. 兰州大学学报, 1996, 32(1): 108 – 115.]
  - [12] Sun Honglie, Zheng Du, Yao Tandong, et al. Protection and construction of the national ecological security shelter zone on Tibetan Plateau[J]. Acta Geographica Sinica, 2012, 67(1): 3 – 12. [孙鸿烈, 郑度, 姚檀栋, 等. 青藏高原国家生态安全屏障保护与建设[J]. 地理学报, 2012, 67(1): 3 – 12.]
  - [13] Dong J R, Walker J P, Houser P R. Factors affecting remotely sensed snow water equivalent uncertainty[J]. Remote Sensing of Environment, 2005, 97(1): 68 – 82.
  - [14] Gao J, Williams M W, Fu X D, et al. Spatiotemporal distribution of snow in eastern Tibet and the response to climate change[J]. Remote Sensing of Environment, 2012, 121(138): 1 – 9.
  - [15] Xu W F, Ma L J, Ma M N, et al. Spatial-temporal variability of snow cover and depth in the Qinghai-Tibetan Plateau[J]. Journal of Climate, 2015, 30(4): 1521 – 1533.
  - [16] Wang Y L, Huang X D, Liang H, et al. Tracking snow variations in the northern hemisphere using multi-source remote sensing data (2000 – 2015)[J]. Remote Sensing, 2018, 10: 136.
  - [17] Zhang Y S, Ma N. Spatiotemporal variability of snow cover and snow water equivalent in the last three decades over Eurasia[J]. Journal of Hydrology, 2018, 559: 238 – 251.
  - [18] Wang Wei. Snow cover monitoring and early warning of snow-caused disaster based on remote sensing and GIS technologies in pastoral areas of the Tibetan Plateau[D]. Lanzhou: Lanzhou University, 2014. [王玮. 基于遥感和GIS的青藏高原牧区积雪动态监测与雪灾预警研究[D]. 兰州: 兰州大学, 2014.]
  - [19] Derksen C, Toose P, Rees A, et al. Development of a tundra-specific snow water equivalent retrieval algorithm for satellite passive microwave data[J]. Remote Sensing of Environment, 2010, 114(8): 1699 – 1709.
  - [20] Foster J L, Hall D K, Kelly R E J, et al. Seasonal snow extent and snow mass in south America using SMMR and SSM/I passive microwave data (1979 – 2006)[J]. Remote Sensing of Environment, 2009, 113(2): 291 – 305.
  - [21] Foster J L, Chang A T C, Hall D K. Comparison of snow mass estimates from a prototype passive microwave snow algorithm, a revised algorithm and a snow depth climatology[J]. Remote Sensing of Environment, 1997, 62(2): 132 – 142.
  - [22] Che T, Dai L Y, Zheng X M, et al. Estimation of snow depth from passive microwave brightness temperature data in forest regions of northeast China[J]. Remote Sensing of Environment, 2016, 183: 334 – 349.
  - [23] Sun Zhiwen, Shi Jiancheng, Jiang Lingmei, et al. Development of snow depth and snow water equivalent algorithm in western China using passive microwave remote sensing data[J]. Advances in Earth Science, 2006, 21(12): 1363 – 1369. [孙之文, 施建成, 蒋玲梅, 等. 被动微波遥感反演中国西部地区雪深、雪水当量算法初步研究[J]. 地球科学进展, 2006, 21(12): 1363 – 1369.]
  - [24] Xiao Xiongxin, Zhang Tingjun. Passive microwave remote sensing of snow depth and snow water equivalent: overview[J]. Advances in Earth Science, 2018, 33(6): 590 – 605. [肖雄新, 张廷军. 基于被动微波遥感的积雪深度和雪水当量反演研究进展[J]. 地球科学进展, 2018, 33(6): 590 – 605.]
  - [25] Sturm M, Taras B, Liston G E, et al. Estimating snow water equivalent using snow depth data and climate classes[J]. Journal of Hydrometeorology, 2010, 11(11): 1380 – 1394.
  - [26] Jiang L M, Wang P, Zhang L X, et al. Improvement of snow depth retrieval for FY3B-MWRI in China[J]. Science China Earth Sciences, 2014, 57(6): 1278 – 1292.
  - [27] Kelly R. The AMSR-E snow depth algorithm: description and initial results[J]. Journal of the Remote Sensing Society of Japan, 2009, 29(1): 307 – 317.
  - [28] Chang A T C, Foster J L, Hall D K. Nimbus-7 SMMR derived global snow cover parameters[J]. Annals of Glaciology, 1987, 9(71): 39 – 44.
  - [29] Chang A T C, Foster J L, Hall D K, et al. The use of microwave radiometer data for characterizing snow storage in western China[J]. Annals of Glaciology, 1991, 16: 215 – 219.
  - [30] Tait A. Estimation of snow water equivalent using passive microwave radiation data[C]//International Geoscience & Remote Sensing Symposium. IEEE, 1998.
  - [31] Cao Meisheng, Li Peiji, Robinson D A. Evaluation and primary application of microwave remote sensing SMMR-derived snow cover in western China[J]. Remote Sensing of Environment China, 1993, 8(4): 260 – 269. [曹梅盛, 李培基, D. A. Robinson, 等. 中国西部积雪 SMMR 微波遥感的评价与初步应用[J]. 环境遥感, 1993, 8(4): 260 – 269.]
  - [32] Zhang R P, Liang T G, Feng Q S, et al. Evaluation and adjustment of the AMSR2 snow depth algorithm for the northern Xinjiang region, China[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2016, 99: 1 – 12.
  - [33] Bai Shuying, Shi Jianqiao, Shen Weishou, et al. Spatial-temporal variation of snow depth in Tibet and its response to climatic change in the past 30 years[J]. Remote Sensing for Land and Resources, 2014, 26(1): 144 – 151. [白淑英, 史建桥, 沈渭寿, 等. 近30年西藏雪深时空变化及其对气候变化的响应[J]. 国土资源遥感, 2014, 26(1): 144 – 151.]
  - [34] Bin Chanjia. Improvement of passive microwave snow algorithm over Qinghai-Tibet Plateau[D]. Shijiazhuang: Shijiazhuang University of Economics, 2013. [宾婵佳. 青藏高原地区被动微波积雪算法改进研究[D]. 石家庄: 石家庄经济学院, 2013.]
  - [35] Liang Shunlin, Cheng Jie, Jia Kun, et al. Recent progress in land surface quantitative remote sensing[J]. Journal of Remote Sensing, 2016, 20(5). [梁顺林, 程洁, 贾坤, 等. 陆表定量遥感反演方法的发展新动态[J]. 遥感学报, 2016, 20(5).]
  - [36] Hou Haiyan, Hou Jinliang, Huang Chunlin, et al. Retrieve snow depth of north of Xinjiang Region from AMSR2 data based on artificial neural network technology[J]. Remote Sensing Technology and Application, 2018, 33(2): 241 – 251. [侯海艳, 侯金亮, 黄春林, 等. 基于人工神经网络和AMSR2多频微波亮温的北疆地区雪深反演[J]. 遥感技术与应用, 2018, 33(2): 241 – 251.]
  - [37] Liang J Y, Liu X P, Huang K N, et al. Improved snow depth retrieval by integrating microwave brightness temperature and visible/infrared reflectance. Remote Sensing of Environment, 2015, 156: 500 – 509.
  - [38] Xiao X X, Zhang T J, Zhong X Y, et al. Support vector regression snow-depth retrieval algorithm using passive microwave remote sensing data[J]. Remote Sensing of Environment, 2018, 210: 48 – 64.

- [39] Bin Chanjia, Qiu Yubao, Shi Lijuan, et al. Comparative validation of snow depth algorithms using AMSR-E passive microwave data in China[J]. *Journal of Glaciology and Geocryology*, 2013, 35(4): 801 – 813. [宾婵佳, 邱玉宝, 石利娟, 等. 我国主要积雪区 AMSR-E 被动微波雪深算法对比验证研究[J]. *冰川冻土*, 2013, 35(4): 801 – 813.]
- [40] Wang Hongwei, Huang Chunlin, Hao Xiaohua, et al. Analyses of the spatiotemporal variations of snow cover in north Xinjiang[J]. *Journal of Glaciology and Geocryology*, 2014, 36(3): 508 – 516. [王宏伟, 黄春林, 郝晓华, 等. 北疆地区积雪时空变化的影响因素分析[J]. *冰川冻土*, 2014, 36(3): 508 – 516.]
- [41] Tennant C J, Harpold A A, Lohse K A, et al. Regional sensitivities of seasonal snowpack to elevation, aspect, and vegetation cover in western North America[J]. *Water Resources Research*, 2017, 53(8): 1 – 19.
- [42] Savoie M H, Armstrong R L, Brodzik M J, et al. Atmospheric corrections for improved satellite passive microwave snow cover retrievals over the Tibet Plateau[J]. *Remote Sensing of Environment*, 2009, 113(12): 2661 – 2669.
- [43] Yao Yonghui, Zhang Baiping. The spatial pattern of monthly air temperature of the Tibetan Plateau and its implications for the geo-ecology pattern of the Plateau[J]. *Geographical Research*, 2015, 34(11). [姚永慧, 张百平. 青藏高原气温空间分布规律及其生态意义[J]. *地理研究*, 2015, 34(11).]
- [44] Bai Shuying, Shi Jianqiao, Gao Jixi, et al. Analysis of spatial-temporal variations of snow depth over the Qinghai-Tibetan Plateau during 1979 – 2010[J]. *Journal of Geo-Information Science*, 2014, 16(4): 628 – 637. [白淑英, 史建桥, 高吉喜, 等. 1979 – 2010 年青藏高原积雪深度时空变化遥感分析[J]. *地球信息科学学报*, 2014, 16(4): 628 – 637.]
- [45] Kelly R E. A Prototype AMSR-E Global Snow Area and Snow Depth Algorithm[J]. *IEEE Transaction on Geoscience and Remote Sensing*, 2003, 41(2): 230 – 242.
- [46] Grody N C, Basist A N. Global identification of snowcover using SSM/I measurements[J]. *IEEE Transaction on Geoscience and Remote Sensing*, 1996, 34(1): 237 – 249.
- [47] Walker A E, Goodison B E. Discrimination of a wet snow cover using passive microwave satellite data[J]. *Annals of Glaciology*, 1993, 17: 307 – 311.
- [48] Bao Huiyi, Zhang Xianfeng, Liao Chunhua, et al. Co-inversion of snow parameters in Xinjiang based on MODIS and AMSR-E data[J]. *Journal of Natural Disasters*, 2013, 22(4): 41 – 49. [包慧漪, 张显峰, 廖春华, 等. 基于 MODIS 与 AMSR-E 数据的新疆雪情参数协同反演研究[J]. *自然灾害学报*, 2013, 22(4): 41 – 49.]
- [49] Zhang Xianfeng, Bao Huiyi, Liu Yu, et al. Snow parameter estimation from microwave remote sensing data[J]. *Mountain Research*, 2014, 32(3): 307 – 313. [张显峰, 包慧漪, 刘羽, 等. 基于微波遥感数据的雪情参数反演方法[J]. *山地学报*, 2014, 32(3): 307 – 313.]
- [50] Lu Xinyu, Wang Xiuqin, Cui Caixia, et al. Snow depth retrieval based on AMSR-E data in northern Xinjiang region, China[J]. *Journal of Glaciology and Geocryology*, 2013, 35(1): 40 – 47. [卢新玉, 王秀琴, 崔彩霞, 等. 基于 AMSR-E 的北疆地区积雪深度反演[J]. *冰川冻土*, 2013, 35(1): 40 – 47.]
- [51] Max Kuhn. *Applied predictive modeling*[M]. Lin Hui, trans. Beijing: China Machine Press, 2016. [马克斯·库恩. *应用预测建模*[M]. 林荟, 译. 北京: 机械工业出版社, 2016.]
- [52] Bohachevsky I O, Johnson M E, Stein M L, et al. Generalized simulated annealing for function optimization[J]. *Technometrics*, 1986, 28(3): 209 – 217.
- [53] Breiman L. Random Forests[J]. *Machine Learning*, 2001, 45(1): 5–32.
- [54] Meng B, Gao J, Liang T, et al. Modeling of alpine grassland cover based on unmanned aerial vehicle technology and multi-factor methods: a case study in the east of Tibetan Plateau, China[J]. *Remote Sensing*, 2018, 10(34): 1 – 19.
- [55] Liu Y, Bi J W, Fan Z P. Multi-class sentiment classification: The experimental comparisons of feature selection and machine learning algorithms[J]. *Expert Systems with Applications*, 2017, 80: 323 – 339.
- [56] Sayran Wayli, Mao Weiyi. A research on the method of deriving high-precision snow parameters from AMSR2 passive microwave remote sensing data[J]. *Journal of Glaciology and Geocryology*, 2016, 38(1): 145 – 158. [沙依然·外力, 毛炜峰. 基于 AMSR2 被动微波积雪参量高精度反演方法研究[J]. *冰川冻土*, 2016, 38(1): 145 – 158.]
- [57] Liu Jinjun, Fu Yunfei, Li Rui, et al. The influence of atmosphere to passive microwave retrieval of snow depth over Qinghai-Tibetan Plateau[J]. *Plateau Meteorology*, 2018, 37(2): 305 – 316. [刘进军, 傅云飞, 李锐, 等. 青藏高原云和大气对被动微波遥感积雪雪深的影响[J]. *高原气象*, 2018, 37(2): 305 – 316.]
- [58] Qiu Yubao, Shi Lijuan, Shi Jiancheng, et al. Atmospheric influences analysis on the satellite passive microwave remote sensing[J]. *Spectroscopy and Spectral Analysis*, 2016, 36(2): 310 – 315. [邱玉宝, 石利娟, 施建成, 等. 大气对星载被动微波影响分析研究[J]. *光谱学与光谱分析*, 2016, 36(2): 310 – 315.]
- [59] Sun Chuan, Song Xiaoning, Zhou Fangcheng, et al. Effects of cloudy atmosphere on microwave signals in channels of AMSR-E[J]. *Journal of University of Chinese Academy of Sciences*, 2018, 35(1): 42 – 49. [孙川, 宋小宁, 周芳成, 等. 有云大气对 AMSR-E 传感器各通道微波信号影响的分析[J]. *中国科学院大学学报*, 2018, 35(1): 42 – 49.]

## Retrieved snow depth over the Tibetan Plateau using random forest algorithm with AMSR2 passive microwave data

WANG Jianshun<sup>1</sup>, WANG Yunlong<sup>1</sup>, ZHOU Minqiang<sup>1</sup>, LIU Changyu<sup>1</sup>, HUANG Xiaodong<sup>2</sup>

(1. State Key Laboratory of Grassland Agro-ecosystems / College of Pastoral Agriculture Science and Technology, Lanzhou University, Lanzhou 730020, China; 2. School of Geographical Sciences, Nanjing University of

Information Science and Technology, Nanjing 210044, China)

**Abstract:** Due to the complex terrain and snow distribution of the Tibetan Plateau (TP), many snow depth retrieval algorithms have not reached the ideal accuracy. Based on the new generation of AMSR2 passive microwave data, the TP was used as the study area in this study, the Brightness Temperature (BT) and the Brightness Temperature Difference (BTD) were input as parameters, and the elevation and latitude parameters were also introduced into the SD retrieval model. The simulated annealing algorithm (SA) was used to feature selection, and the SD retrieval model of the TP that based on the RF algorithm was constructed. The results showed that the RF algorithm had higher accuracy over the TP compared with the AMSR2 global SD products; the goodness of fit ( $R^2$ ) were increased from 0.41 to 0.60, the root mean square error (RMSE) were decreased from 7.36 cm to 4.88 cm, and the BIAS were decreased from 3.24 cm to -0.16 cm, respectively. The BTD of low frequency are more suitable for SD retrieval with RF algorithm over the TP. When the elevation approached or exceeded the average elevation of the TP by 4 000 m, the retrieval result was not well, but the RMSE was only 3.78 cm, and the BIAS was only -0.09 cm. The southern part of the plateau (25° ~ 30° N) had poor retrieval results due to its complex topography and relatively few meteorological stations, with the RMSE of 5.94 cm and the BIAS of -0.39 cm. The grassland, bare land and farmland with the low vegetation cover area had a maximum RMSE of 3.19 cm, and BIAS of -0.49 cm. The main land cover type on the TP was grassland, and RMSE and BIAS of the random forest algorithm on the grassland underlay were about 3 cm and 0 cm, respectively

**Key words:** random forest algorithm; Tibetan Plateau; snow depth retrieval; AMSR2

(责任编辑: 周成林; 编辑: 周成林)