

引用格式: Lei Huajin, Li Hongyi, Wang Jian, *et al.* MODIS Fractional Snow Cover Products Preparing on Tibetan Plateau based on Environmental Information and Regression Model[J]. Remote Sensing Technology and Application, 2020, 35(6): 1303-1311. [雷华锦, 李弘毅, 王建, 等. 基于环境信息和回归模型的青藏高原MODIS积雪面积比例产品制备[J]. 遥感技术与应用, 2020, 35(6): 1303-1311.]
doi: 10.11873/j.issn.1004-0323.2020.6.1303

基于环境信息和回归模型的青藏高原MODIS 积雪面积比例产品制备

雷华锦^{1,2}, 李弘毅^{1,3}, 王 建^{1,4}, 郝晓华^{1,3}, 赵宏宇^{1,2}, 张 娟⁵

(1. 中国科学院西北生态环境资源研究院, 甘肃 兰州 730000;

2. 中国科学院大学, 北京 100049;

3. 中国科学院黑河遥感试验研究站, 甘肃 张掖 734000;

4. 江苏省地理信息资源开发与利用协同创新中心, 江苏 南京 210023;

5. 青海省气象科学研究所, 青海 西宁 810000)

摘要: 积雪面积比例(Fractional Snow Cover, FSC)是定量描述单位像元内积雪覆盖面积(Snow Cover Area, SCA)与像元空间范围的比值, 可为区域气候模拟、水文模型等提供积雪分布的定量信息。MODIS FSC产品是根据经验模型计算得到, 并没有考虑地形、植被和地表温度等环境因素的影响, 在青藏高原的验证精度低。针对此问题, 考虑青藏高原地区环境因素(地形、植被、地表温度)对FSC制备的影响, 基于多元自适应回归模型(Multivariate Adaptive Regression Splines, MARS)和线性回归模型分别建立FSC制备的非参数回归模型和经验回归模型。用Landsat 8地表反射率的数据和SNOMAP算法制备FSC的参考数据集。选取一部分参考数据集作为模型的训练数据集, 另一部分作为模型的检验数据集。研究结果表明: MARS方法估计FSC的精度明显高于线性回归模型和原有的MODIS FSC制备方法。MARS的总体R、RMSE、MAE分别为0.791、0.103、0.058。在线性回归模型中精度最高的总体R、RMSE、MAE分别为0.647、0.128、0.072。MODIS原有FSC制图方法的总体R、RMSE、MAE分别为0.595、0.221、0.170。考虑了环境信息的MARS方法更加适用于青藏高原地区FSC制备。本研究为制备青藏高原地区更高精度的FSC数据提供了新思路。

关键词: 青藏高原; 线性回归模型; 积雪面积比例; MODIS; 多元自适应回归模型

中图分类号: TP75 **文献标志码:** A **文章编号:** 1004-0323(2020)06-1303-09

1 引言

积雪是冰冻圈的重要组成部分, 是气候变化的重要指示器。积雪也是地球表层能量辐射收支平衡、区域水循环的重要影响因子^[1-2]。积雪覆盖

(Snow Cover)是耦合气候模式和水文模型的重要输入参数, 是区域水资源评估的主要指标^[3-5]。因此, 精确估计积雪覆盖对气候模式、水文模型、区域水资源评估具有重要的意义^[6]。积雪覆盖产品是积雪遥感反演的主要产品之一。遥感反演全球尺度

收稿日期: 2019-10-09; 修订日期: 2020-11-18

基金项目: 甘肃省科技计划(17JR5RA296), 科技基础资源调查专项(2017FY100503), 国家自然科学基金项目(41571371、41971399、41861049), 青海省科技厅重点研发与转化项目(2017-SF-131)资助。

作者简介: 雷华锦(1996—), 女, 四川雅安人, 硕士研究生, 主要从事寒区遥感水文研究。E-mail: leihuajin@lzb.ac.cn

通讯作者: 李弘毅(1981—), 男, 四川西充人, 副研究员, 主要从事积雪遥感和寒区水文模拟研究。E-mail: lihongyi@lzb.ac.cn

的积雪覆盖产品主要有 NASA-MODIS、NOAA-AVHRR、SPOT-VEGETATION、GOES-SSM/I 等。MODIS 积雪产品具有时空分辨率高的特点,是目前使用最广泛的积雪覆盖产品^[7]。

二值积雪覆盖产品是将像元判识为雪和非雪,SNOMAP 是 MODIS 二值积雪覆盖产品制图算法^[8]。但是在一个像元内当雪深小于 1cm、积雪面积比例(Fractional Snow Cover, FSC)小于 36% 左右时,SNOMAP 算法不能将该像元识别为雪像元^[9-10]。同时,地形、植被、混合像元、阴影也会影响积雪制图的精度。用二值积雪覆盖产品估计积雪覆盖面积(SCA)存在很大的误差。FSC 产品可以在亚像元尺度上定量描述积雪面积比例。因此,FSC 产品可以替代二值积雪覆盖产品,为气候模式、水文模型、数值模拟提供更精确的输入参数。

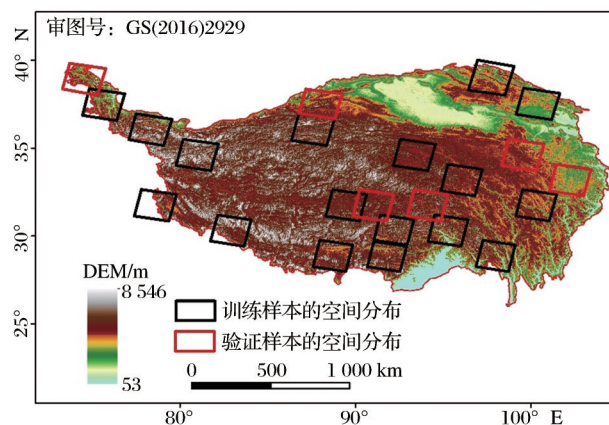
目前,已有大量学者在积雪面积比例提取方面做了很多研究。Salomonson 等^[11]利用 NDSI 与 FSC 之间的线性关系发展了半经验回归模型的方法计算 MODIS FSC,但是该算法严重低估了积雪覆盖的范围,尤其是在植被茂盛和阴影处的积雪区域。为了改进在林区估计 FSC 的精度,研究者们提出了几何光学模型^[12]、冠层辐射传输模型^[13]和混合模型^[14]。然而这些模型主要应用在浓密植被覆盖的积雪区域。Painter 等^[15]利用混合像元分解的方法制备 MODIS FSC。线性光谱混合分析模型是假设混合像元的反射率是混合像元内不同地物反射率的线性组合^[16-17]。混合像元分解的精度很大程度上依赖于端元选择,也有很多研究通过改进端元提取的方法来提高 FSC 估计的精度^[18-20]。Dobrev 等^[21]利用人工神经网络的方法提取 MODIS FSC,也有一些其他机器学习的方法应用在 MODIS FSC 的提取上,例如 SVM、MARS(Multivariate Adaptive Regression Splines)^[22]等。虽然光谱混合分析模型和机器学习的方法在不同下垫面上提取积雪表现较好,但算法复杂度高,不适合做大尺度、长时间序列的 FSC 产品的制备。经验回归模型简单易用,但仅仅只考虑了 NDSI 与 FSC 之间的关系。青藏高原由于其中纬度地区独特的地理位置所形成的自然环境,对区域和全球气候变化都具有很大的影响。青藏高原地形复杂,降雪季积雪覆盖范围广,积雪消融和更替速度快,积雪空间分布破碎化严重是青藏高原区域鲜明的积雪特征。同时,青藏高原植被类型多样,表现出明显的垂直带分布特征和水平分布

特征,并且存在显著的区域差异。在以往的研究中,大多都忽略了地形和植被覆盖对 FSC 提取的影响,并未将地形和植被要素参与 FSC 的制备,这将影响 FSC 的提取精度。

针对以上问题,考虑地形、植被、地表温度对制备 FSC 的影响,通过多元自适应回归模型建立 FSC 与地形(海拔、坡度、坡向)、植被覆盖(NDVI)和 NDSI 之间的关系,与线性回归模型和 MODIS 积雪产品做比较,建立适用于青藏高原区域 FSC 产品的回归模型,从而提高青藏高原区域 FSC 制图的精度。

2 研究区概况

以青藏高原地区为研究区域(图 1)。青藏高原地处我国西南部,位于 73°~105° E, 25°~40° N 之间,南起喜马拉雅山脉,北至昆仑山和祁连山北缘,西起帕米尔高原,东至横断山脉,面积约为 $2.57 \times 10^6 \text{ km}^2$ 。其平均海拔在 4 000 m 以上,是我国主要的季节性积雪区之一。青藏高原是众多河流的发源地,如长江、黄河、澜沧江等^[23],积雪融水是这些河流的主要补给源之一。随着全球气候变暖,青藏高原积雪消融时间会直接影响河川径流^[24]。此外,青藏高原的积雪变化会对东亚地区的大气环流和气候系统产生一定的影响。



(其中黑色框为用于模型训练的样本空间分布,红色框为验证样本的空间分布)

图 1 研究区概况图及样本的空间分布

Fig.1 The location of the study area and the spatial distribution of samples

3 数据及其预处理

采用的数据包括 Landsat 8 SR (Landsat 8 Surface Reflectance Tier 1)数据、DEM 数据、MOD10A1 V6(MODIS 积雪产品)和 MOD11A1 V6(MODIS

地表温度产品)数据。其中 Landsat 8 SR 数据用于制备“真值”FSC、计算 NDSI 和 NDVI。DEM 数据用于提取坡度和坡向。MOD11A1 数据用于提取地表温度。

将“真值”FSC、NDSI、NDVI、地表温度、坡度

和坡向作为 MARS 模型的输入数据。“真值”FSC、NDSI、NDVI 作为线性模型的输入数据。将 MOD10A1 使用 MODIS 原有的经验模型估计 FSC,作为结果之一便于后续与其他模型制备的 FSC 进行对比。数据具体处理流程如图 2 所示。

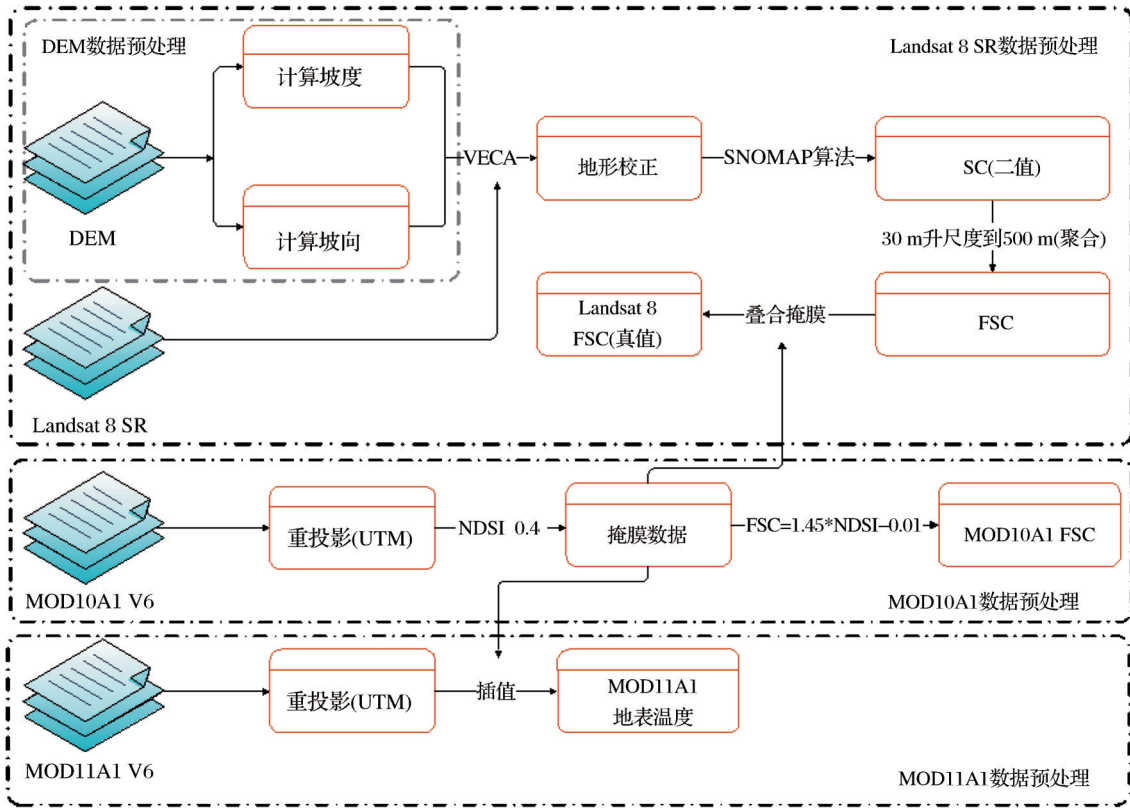


图 2 数据处理流程图
Fig.2 Data processing flow chart

3.1 DEM 数据

DEM 数据是用于计算坡度和坡向。数据来源为 NASA (<https://search.earthdata.nasa.gov>), 空间分辨率为 30 m, 投影为 WGS 84 投影。基于 8 邻域法计算坡度和坡向, 将其用于地形校正以及作为非参数回归的输入数据。

3.2 Landsat 8 SR 数据

该系列数据来源于 USGS。该数据已经去除了大气对地表反射率的影响。其空间分辨率为 30 m, 时间分辨率为 16 d, 投影为 WGS84 UTM 投影。由于太阳方位角、地形的高低起伏对地表反射率的影响较为显著, 因此采用地形校正的方法去除地形对地表反射率的影响。常用的地形校正的方法主要包括经验模型、朗伯体模型、非朗伯体模型。本文采用 VECA^[25] 方法对 Landsat 8 SR 数据进行地形校正。校正公式由式(1)和式(2)给出:

$$\cos i = \cos (\theta_{slop}) \cos (\theta_{sz}) + \sin (\theta_{slop}) \sin (\theta_{sz}) \cos (\theta_{sa} - \theta_{aspect}) \quad (1)$$

$$\frac{R_{corr}}{R_{uncorr}} = \frac{L_a}{m \cdot \cos i + b} \quad (2)$$

其中: θ_{slop} 是坡度、 θ_{aspect} 是坡向、 θ_{sz} 是太阳天顶角、 θ_{sa} 是太阳方位角、 L_a 是未校正图像的平均反射率, m 和 b 是校正系数。

利用经过地形校正后的 Landsat 8 SR 数据计算 NDVI 和制备 FSC“真值”数据集。NDVI 的计算方法由式(3)给出:

$$NDVI = \frac{NIR_{0.86} - Re d_{0.65}}{NIR_{0.86} + Re d_{0.65}} \quad (3)$$

制备 FSC“真值”时首先使用 SNOMAP 算法制备积雪二值图像, 然后将其聚合成 500 m 分辨率的 FSC 数据集, 作为多参数权重回归模型的输入数据。SNOMAP 算法在非林区将 $NDSI \geq 0.4$ 、 $B_5 >$

0.11、 $B_3 > 0.1$ 的像元判识为雪像元,在林区使用动态 NDSI-NDVI 的方法判识积雪。本文在林区将 $NDSI \geq 0.2$ 、 $NDVI > 0.1$ 的像元也被识别为雪像元^[12]。其中 NDSI 的计算公式由式(4)给出:

$$NDSI = \frac{Green_{0.56} - SWIR_{1.61}}{Green_{0.56} + SWIR_{1.61}} \quad (4)$$

随后,采用聚合的方法将 30 m 分辨率的二值积雪数据升尺度到 500 m 分辨的 FSC 数据。聚合的公式由式(5)给出。

$$FSC_i = \frac{s}{n} = \frac{\sum_{k=1}^s 1}{[500/30]^2} \quad (5)$$

其中:[]表示取整, s 表示在 500 m 分辨率的大像元内值为 1 的像元的个数, n 表示 500 m 分辨率内所包含的 30 m 分辨率的像元的个数。本文使用 MOD10A1 V6 掩膜数据去掩膜 500 m 分辨率的 Landsat 8 FSC 数据,最终得到模型训练的“真值”FSC。

3.3 MODIS 数据

MODIS 数据包括 MOD10A1 V6 数据和 MOD11A1 V6 数据,数据均来源于 NASA (<https://search.earthdata.nasa.gov>),其中 MOD10A1 V6 空间分辨率为 500 m、时间分辨率为 1 d、投影为 MODIS Sinusoidal 投影。利用式(6)^[26-27]计算 MOD10A1 V6 FSC 数据,将估计结果作为 MODIS 原有的经验模型得到的 FSC,与其余模型结果进行对比。

$$FSC = 1.45 \cdot NDSI - 0.01 \quad (6)$$

MOD11A1 V6 数据是 MODIS 地表温度的产品。空间分辨率为 1 km、时间分辨率为 1 d、投影为 MODIS Sinusoidal 投影。将其投影转换为 WGS 84 UTM 投影,然后采用最邻近插值的方法重采样成 500 m 分辨率,作为非参数回归的输入数据。

4 研究方法

基于 MARS 模型和线性回归模型制备青藏高原地区的 FSC,其中 MARS 是一种非参数回归的方法。选取了 22 景青藏高原地区 2013 年 10 月到 2019 年的 Landsat 8 SR 数据作为训练数据集和验证数据集,选取原则为积雪季内(上一年 10 月 1 日开始到下一年的 3 月 31 日)云量小于 2%、积雪覆盖率大于 40%。其他数据集都与 Landsat 8 SR 数据集的时间和空间范围一致。其中 75% 的 Landsat 8 SR 数据集(约 16 景影像)作为训练数据集,其余的 25%(6 景影像)作为验证数据集。“真值”FSC、坡度、坡

向、地表温度、NDVI、NDSI 作为 MARS 模型的输入数据。“真值”FSC、NDVI、NDSI 作为线性回归模型的输入数据。

4.1 多元自适应回归模型

MARS(Multivariate Adaptive Regression Splines)是 Friedman 于 1990 年提出^[28],目的是为了解决多变量函数近似、非参数多元回归等问题。响应变量 Y 与多个预测变量 $X = \{x_1, x_2, x_3 \cdots x_n\}$, $x_i \in D$ 之间的回归模型可以表述为式(7):

$$Y = f(X) + \varepsilon \quad (7)$$

其中: ε 表示随机误差,多元非参数回归的目的是找到最优的 $\hat{f}(X)$ 使得 $\hat{f}(X)$ 近似等于 Y 。在 MARS 中使用分段线性基函数来表示 Y 与 X 之间的关系。基函数的形式如式(8)所示:

$$\begin{aligned} (x - \tau)_+ &= \begin{cases} x - \tau, & x > \tau \\ 0, & x \leq \tau \end{cases} \\ (x - \tau)_- &= \begin{cases} x - \tau, & x < \tau \\ 0, & x \geq \tau \end{cases} \end{aligned} \quad (8)$$

其中: τ 是分段函数的结点。每一个基函数就是一个线性样条函数。MARS 的基本思路是为每一个预测变量的值 x_{ij} 构建这样的线性样条函数,则线性样条函数的集合可以表示为 $C = \{(x - \tau)_+, (x - \tau)_-, \tau \in \{x_{1j}, x_{2j}, x_{3j} \cdots x_{nj}\}, j \in \{1, 2, 3 \cdots p\}\}$ 其中 p 表示预测变量的维度。因此 Y 能够被表示成 C 的线性函数。如式(9)所示:

$$Y = \beta_0 + \sum_{m=1}^M \beta_m B_m(X^m) + \varepsilon \quad (9)$$

其中: $B_m(X^m) = \prod_{j=1}^{K_m} (s_{kj}^m \cdot (x_{kj}^m - \tau_{kj}^m))_+$, K_m 表示第 m 个分段线性基函数, x_{kj}^m 表示预测变量 j 第 m 个分段线性基函数的第 k 个截断函数。 τ_{kj}^m 是 x_{kj}^m 的结点。为了增加估计最优模型 $\hat{f}(X)$ 的计算速度, MARS 采用广义交叉验证的方法,如式(10)所示:

$$GCV(\gamma) = \frac{\sum_{i=1}^N (Y_i - \hat{f}_\gamma(X_i))^2}{(1 - Q(\gamma)/N)^2} \quad (10)$$

其中: $Q(\gamma) = \mu + dK$, N 表示观测样本的数量, K 表示前一步当中选择的结点数, μ 表示线性独立函数的数量, d 表示每一个基函数的优化损失。

选取 NDSI、NDVI、坡度、坡向和地表温度作为 MARS 的输入变量进行模型训练,由于坡度和坡向与 FSC 之间的关系不能用明确的数学公式来表达,而 MARS 为非参模型,因此,用 MARS 的方法可以很好地估计 FSC 与各影响因素之间的关系。

4.2 多元线性回归模型

基于 NDSI、NDVI 与 FSC 之间的关系(如图 3 所示)建立多元线性回归模型。根据 NDVI 和 NDSI 对 FSC 的影响建立如式(11)所示的线性回归模型。

$$FSC_a = a_1 \cdot x_1 + a_2 \cdot x_2 + a_3$$
$$FSC_b = \begin{cases} a_1 \cdot x_1 + a_2 \cdot x_2 + a_3, & x_2 > 0 \\ a_1 \cdot x_1 + a_3, & x_2 \leq 0 \end{cases} \quad (11)$$

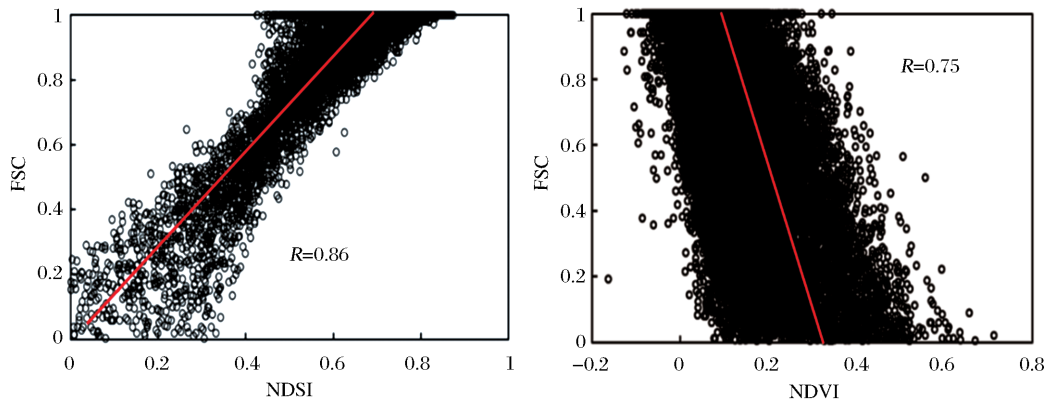


图 3 NDSI、NDVI 与 FSC 之间的关系(其中 R 为参数之间的相关系数)

Fig.3 The relationship between NDSI, NDVI and FSC(R represents the correlation coefficient between two parameters)

其中 $A=(a_1, a_2, a_3)|(a_1, a_3)$, $X=(x_1, x_2, 1)^T|(x_1, 1)^T$ 本文根据选取的 75% 的训练样本建立多元线性回归模型,剔除云像元之后对 86 万个有效的雪像元建立线性回归模型。FSC_a 的解分别为 $a_1 = 0.765$, $a_2 = -0.308$, $a_3 = 0.337$; FSC_b 当 NDVI 大于 0 的解分别为 $a_1 = 0.792$, $a_2 = -0.675$, $a_3 = 0.336$; 当 NDVI 大于 0 的解分别为 $a_1 = 0.402$, $a_3 = 0.648$; FSC_c 的解分别为 $a_1 = 0.882$, $a_3 = 0.252$ 。则式(11)最终可以用式(13)表示为:

$$FSC_a = 0.765 \cdot NDSI - 0.308 \cdot NDVI + 0.337$$
$$FSC_b = \begin{cases} 0.792 \cdot NDSI - 0.675 \cdot NDVI + 0.336, & NDVI > 0 \\ 0.402 \cdot NDSI + 0.648, & NDVI \leq 0 \end{cases} \quad (13)$$
$$FSC_c = 0.882 \cdot NDSI + 0.252$$

4.3 评价方法

选取 25% 的训练数据集(6 景影像)来验证回归模型制备 FSC 的精度。采用相关系数(Correlation Coefficient, R)、均方根误差(Root Mean Square Error, RMSE)、平均绝对误差(Mean absolute error, MAE)作为模型的评价指标,各评价指标是对所有检验样本计算平均值。检验样本的空间分布如图 1 所示,表 1 列举了验证数据集的元数据信息。

$$FSC_c = a_1 \cdot x_1 + a_3$$

其中: a_1, a_2, a_3 为系数, x_1 表示 NDSI, x_2 表示 NDVI。FSC_a 不考虑 NDVI 小于 0 时对 FSC 的影响。FSC_b 根据 NDVI 进行分段建模。FSC_c 根据 NDSI 与 FSC 之间的关系建立模型。本文使用线性最小二乘法解线性回归模型,则式(11)可以表示为:

$$\min_x \frac{1}{2} \|FSC - AX\|^2 \quad (12)$$

表 1 检验数据集的元数据信息

Table 1 The metadata information of the test dataset				
样本编号	云量/%	有效像元个数	行/列号	日期
T _a	1.98	57 457	131/37	2016-03-25
T _b	1.05	39 310	133/36	2018-11-08
T _c	0.74	44 610	136/38	2014-02-19
T _d	0.82	1 039	138/38	2014-12-18
T _e	1.36	20 523	141/34	2013-12-20
T _f	1.26	52 320	150/33	2017-01-12

5 结果与讨论

在验证结果中,使用 FSC_{MARS} 表示 MARS 模型制备的 FSC, FSC_a、FSC_b、FSC_c 分别表示不同线性回归模型计算的 FSC, FSC_{MOD} 表示 MODIS FSC 制图的线性回归模型估计的 FSC。不同模型制备 FSC 精度的评价结果如表 2 所示。

从表 2 中可以看出 MARS 回归模型制备 FSC 的精度明显优于线性回归模型,根据 NDVI 建立分段函数的线性回归模型估计 FSC 要优于其它的线性回归模型。而仅根据 NDSI 与 FSC 之间建立的线性回归模型制备 FSC 的精度要比考虑植被对 FSC 的影响建立的二元线性回归模型的精度要低。在 3 个线性模型的验证结果中,基于 MODIS 制备 FSC

表 2 不同模型制备 FSC 的精度

Table 2 The accuracy of FSC prepared by different models

M	有效像元个数	R	RMSE	MAE
FSC _{MARS}	35 876.50	0.791	0.103	0.058
FSC _a	35 876.50	0.553	0.144	0.091
FSC _b	35 876.50	0.647	0.128	0.072
FSC _c	35 876.50	0.595	0.139	0.090
FSC _{MOD}	35 876.50	0.595	0.221	0.170

的方法的精度最低,改进的模型制备青藏高原区域 FSC 的精度显著高于 MODIS 积雪产品的精度。这说明除 NDSI 之外,植被对 FSC 的计算也有很大的影响。

不同模型应用到不同的检验数据制备 FSC 的精度存在差异。图 4 展示了每一个检验数据对应不同的模型制备 FSC 的精度验证结果。从图 4 可以看出检验数据 T_d 、 T_e 、 T_f 在不同模型中制备 FSC 的精度存在显著差异。用 FSC_{MARS} 方法制备 FSC 的精度优于其它回归模型,在线性回归模型中 FSC_b 的验证精度最高, FSC_{MOD} 的精度最差。在精度差异较大的 3 个检验数据集中 T_d 在不同模型中的精度检验差异较大,主要原因在于有效像元的个数较少,使得结果并不稳定。线性模型在青藏高原制备 FSC 时会受到混合像元和地形的影响从而降低 FSC 估计的精度,并且在线性模型中没有对地形影响的修正,只加入了植被对 FSC 的影响。此外, SNOMAP 在判识积雪的时候雪深越大精度越高,但是青藏高原的积雪薄而变化速度快的特性会造成“真值”的误差。因此,除模型估计 FSC 精度的能力局限之外,青藏高原地区的积雪特性、地形和混合像元也会造成模型估计 FSC 的精度误差。

图 5 是展示了某一天的检验数据 T_d (时间为 2016 年 3 月 25 日)用不同的回归模型制备 FSC 的结果。L-FSC 表示 Landsat 8 SR 用 SNOMAP 制备的 FSC。图中标出了每类结果的评价精度,可以看出 FSC_{MARS} 制备 FSC 的精度最高,其 R、RMSE、MAE 分别为 0.912、0.054、0.034,而 MODIS 线性回归模型制备 FSC 的精度较差。

FSC_{MARS} 制备 FSC 的精度显著高于其它线性回归模型。在线性回归模型中, NDVI 分段线性回归模型 FSC_b 的精度显著高于 FSC_a、FSC_c 和 FSC_{MOD}, FSC_a、FSC_c 和 FSC_{MOD} 在植被覆盖茂盛、地形复杂的区域总体精度差异不大。因此基于 FSC_{MARS} 和 FSC_b 制备青藏高原地区的 FSC 更具可靠性。青

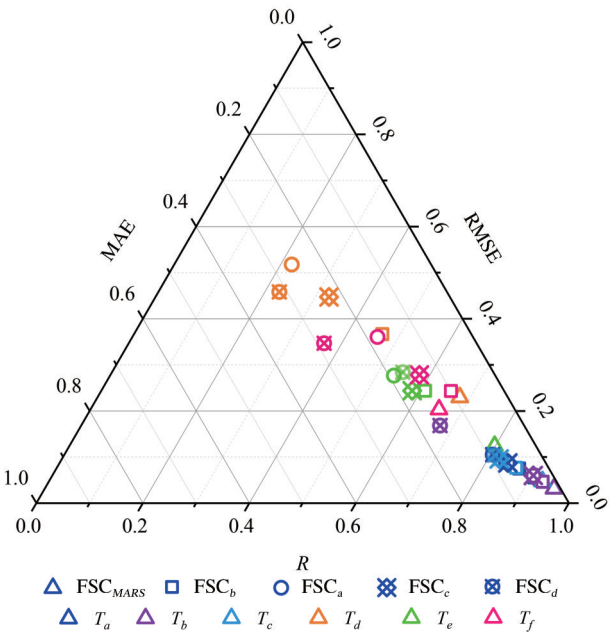


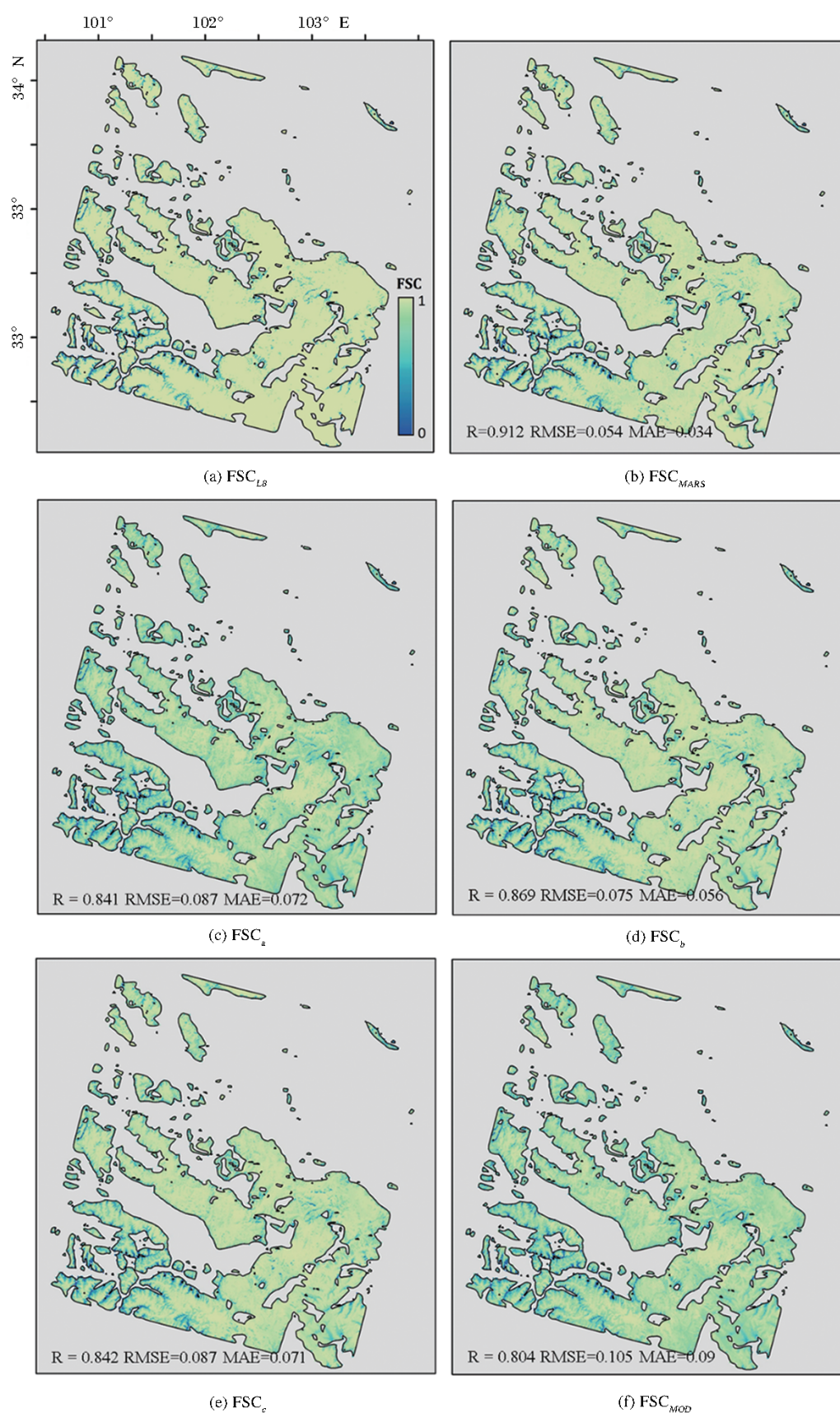
图 4 检验样本的误差图

Fig.4 Error diagram of test samples

高原地区积雪特性的差异(反射率的差异)、混合像元、植被和地形的影响是模型估计 FSC 的主要误差来源。此外,在求解线性回归模型时,最优化的方法会影响模型的系数从而影响 FSC 的估计。本研究发展的 FSC_{MARS} 和线性回归模型在青藏高原地区制备 FSC 的精度显著高于 MODIS 产品的精度。

6 结 论

本文基于 MARS 回归模型和线性回归模型,考虑了地形、植被对 FSC 估计的影响,发展了青藏高原区域 FSC 提取的算法。MARS 回归模型的总体验证精度最高,总体 R、RMSE、MAE 分别为 0.791、0.103、0.058。在线性回归模型中 FSC_b 的精度最高,总体 R、RMSE、MAE 分别为 0.647、0.128、0.072。MODIS 原有 FSC 制图方法的总体 R、RMSE、MAE 分别为 0.595、0.221、0.170。相比原有的 FSC 制图算法的精度, MARS 和 FSC_b 在制备青藏高原地区的 FSC 精度上有了很大的改进。青藏高原这种典型的复杂山区,地形对于积雪识别和 FSC 制备的影响是不容忽视的,本文发展的模型在青藏高原地区制备 FSC 的总体验证精度显著高于 MODIS 产品的总体验证精度,为青藏高原地区制备较高精度的 FSC 提供了可靠的方法。此外,作为水文模型和融雪径流模型的重要输入参数,高精度的积雪面积比例产品可能会改善水文模型的模拟效果。



(a)表示 Landsat 8 的“真值”FSC, (b)~(f)分别表示 MARS 及 4 种不同非线性回归模型的 FSC 结果

图 5 T_a (时间为 2016 年 3 月 25 日) 用不同模型制备 FSC 的结果

Fig.5 T_a results of preparing FSC with different models

参考文献(References):

- [1] Vermote E F, El Saleous N, Justice C O, *et al.* Atmospheric Correction of Visible to Middle-Infrared EOS-MODIS Data over Land Surfaces: Background, Operational Algorithm and Validation[J]. *Journal of Geophysical Research: Atmospheres*, 1997, 102: 17131-17141.
- [2] Xiao Lin, Che Tao, Dai Liyun. Evaluation on the Spatial Characteristics of Multiple Snow Depth Datasets over China [J]. *Remote Sensing Technology and Application*, 2019, 34 (6): 1133-1145.[肖林,车涛,戴礼云.多源雪深数据在中国的空间特征评估[J].*遥感技术与应用*, 2019, 34(6): 1133-1145.]
- [3] Liston G E. Interrelationships among Snow Distribution, Snowmelt, and Snow Cover Depletion: Implications for Atmospheric, Hydrologic, and Ecologic Modeling[J]. *Journal of Applied Meteorology*, 1992, 38(10):1474-1487.
- [4] Ma Lijuan, Qin Dahe. Temporal-spatial Characteristics of Observed Key Parameters of Snow Cover in China during 1957-2009 [J]. *Journal of Glaciology and Geocryology*, 2012, 34 (1):1-11.[马丽娟,秦大河.1957~2009年中国台站观测的关键积雪参数时空变化特征[J].*冰川冻土*, 2012, 34(1):1-11.]
- [5] Hall D K, Riggs G A, Salomonson V V, *et al.* MODIS Snow-cover Products [J]. *Remote Sensing of Environment*, 2002, 83(1):181-194.
- [6] Rittger K, Painter T H, Dozier J. Assessment of Methods for Mapping Snow Cover from MODIS [J]. *Advances in Water Resources*, 2013, 51:367-380.
- [7] Frei A, Tedesco M, Lee S, *et al.* A Review of Global Satellite-derived Snow Products[J]. *Advances in Space Research*, 2012, 50(8): 1007 - 1029.
- [8] Hall D K, Riggs G A, Salomonson V V. Development of Methods for Mapping Global Snow Cover Using Moderate Resolution Imaging Spectroradiometer Data[J]. *Remote Sensing of Environment*, 1995, 54(2):127-140.
- [9] Anttila S, Metsamäki S, Derksen C. A Comparison of Finnish SCAMod Snow Maps and MODIS Snow Maps in Boreal Forests in Finland and in Manitoba, Canada[C]// *IEEE International Conference on Geoscience & Remote Sensing Symposium*. IEEE, 2006. doi:10.1109/igarss.2006.562.
- [10] Klein A G, Barnett A C. Validation of Daily MODIS Snow Cover Maps of the Upper Rio Grande River Basin for the 2000 - 2001 Snow Year [J]. *Remote Sensing of Environment*, 2003, 86(2):162-176.
- [11] Salomonson V V, Appel I. Estimating Fractional Snow Cover from MODIS Using the Normalized Difference Snow Index [J]. *Remote Sensing of Environment*, 2004, 89(3):351-360.
- [12] Klein A G, Hall D K, Riggs G A. Improving Snow Cover Mapping in Forests through the Use of a Canopy Reflectance Model[J].*Hydrological Processes*, 1998, 12(10-11):1723-1744.
- [13] Metsämäki S J, Anttila S T, Markus H J, *et al.* A Feasible Method for Fractional Snow Cover Mapping in Boreal Zone based on a Reflectance Model[J]. *Remote Sensing of Environment*, 2005, 95(1):77-95.
- [14] Liu L, Mishchenko M I, Arnott W P. A Study of Radiative Properties of Fractal Soot Aggregates Using the Superposition T-matrix Method [J]. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 2008, 109(15):2656-2663.
- [15] Painter T H, Rittger K, McKenzie C, *et al.* Retrieval of Sub-pixel Snow Covered Area, Grain Size, and Albedo from MODIS [J]. *Remote Sensing of Environment*, 2009, 113(4): 868-879.
- [16] Wang G, Jiang L, Wu S, *et al.* Fractional Snow Cover Mapping from FY-2 VISSR Imagery of China [J]. *Remote Sensing*, 2017, 9(10):983. DOI: 10.3390/rs9100983.
- [17] Jiancheng S. An Automatic Algorithm on Estimating Sub-pixel Snow Cover from MODIS [J]. *Quaternary Sciences*, 2012, 32(1): 6-15.
- [18] Zhu J, Shi J, Wang Y. Subpixel Snow Mapping of the Qinghai - Tibet Plateau Using MODIS Data [J]. *International Journal of Applied Earth Observations & Geoinformation*, 2012, 18: 251-262.
- [19] Zhang Y, Huang X, Hao X, *et al.* Fractional Snow-cover Mapping Using an Improved Endmember Extraction Algorithm [J]. *Journal of Applied Remote Sensing*, 2014, 8(1): 084691. doi:10.1117/1.JRS.8.084691.
- [20] Zhao Hongyu, Hao Xiaohua, Zheng Zhaojun, *et al.* A New Algorithm of Fractional Snow Cover basing on FY-3D/MERSI- II [J]. *Remote Sensing Technology and Application*, 2018, 33(6):1004-1016.[赵宏宇,郝晓华,郑照军,等.基于FY-3D/MERSI- II的积雪面积比例提取算法[J].*遥感技术与应用*, 2018, 33(6):1004-1016.]
- [21] Dobrev I D, Klein A G. Fractional Snow Cover Mapping through Artificial Neural Network Analysis of MODIS Surface Reflectance [J]. *Remote Sensing of Environment*, 2011, 115(12):3355-3366.
- [22] Kuter S, Akyurek Z, Weber G-W. Retrieval of Fractional Snow Covered Area from MODIS Data by Multivariate Adaptive Regression Splines [J]. *Remote Sensing of Environment*, 2018, 205: 236-252.
- [23] Zhang, Yili, Li, Bingyuan, Zheng Du. A Discussion on the Boundary and Area of the Tibetan Plateau in China [J]. *Geographical Research*, 2002, 21(1):1-8.[张懿铨,李炳元,郑度.论青藏高原范围与面积[J].*地理研究*, 2002, 21(1):1-8.]
- [24] Wang W, Liang T, Huang X, *et al.* Early Warning of Snow-caused Disasters in Pastoral Areas on the Tibetan Plateau [J]. *Natural Hazards and Earth System Sciences*, 2013, 13: 1411-1425.
- [25] Gao Y, Zhang W. A Simple Empirical Topographic Correction Method for ETM+ Imagery [J]. *International Journal of Remote Sensing*, 2009, 30(9):2259-2275.
- [26] Riggs G A, Dorothy K H. MODIS Snow Products Collection 6 User Guide [EB/OL]. 2016. <https://nsidc.org/sites/nsidc.org/files/files/MODIS-snow-user-guide-C6.pdf>.

- [27] Hall D K, Riggs G A. Normalized-Difference Snow Index (NDSI)[J]. Encyclopedia of Snow, Ice and Glaciers, Encyclopedia of Earth Sciences Series, 2011: 779-780. doi: 10.1007/978-90-481-2642-1-376.
- [28] Friedman J H. Multivariate Adaptive Regression Splines[J]. Annals of Statistics, 1991, 19(1): 1-67.

MODIS Fractional Snow Cover Products Preparing on Tibetan Plateau based on Environmental Information and Regression Model

Lei Huajin^{1,2}, Li Hongyi^{1,3}, Wang Jian^{1,4}, Hao Xiaohua^{1,3}, Zhao Hongyu^{1,2},
Zhang Juan⁵

(1. *Northwest Institute of Eco-Environmental Resources, Chinese Academy of Sciences, Lanzhou 730000, China;*

2. *University of Chinese Academy of Sciences, Beijing 100049, China;*

3. *Heihe Remote Sensing Experimental Research Station, Chinese Academy of Sciences, Zhangye 734000, China;*

4. *Geography of Jiangsu Province Collaborative Innovation Center for Information Resources Development and Utilization, Nanjing 210023, China;*

5. *Qinghai Meteorological Science Research Institute, Xining 810000, China)*

Abstract: Fractional Snow Cover (FSC) is the ratio of the Snow Cover Area (SCA) to the spatial area in a unit pixel, which can provide quantitative information of snow cover distribution for regional climate simulation and hydrological model. MODIS FSC products are calculated according to the empirical model, without considering the impact of environmental factors such as topography, vegetation and surface temperature. The accuracy in the Tibetan plateau is low. Therefore, the effects of environmental factors (topography, vegetation, and surface temperature) on FSC preparation were taken into account in Tibetan plateau, based on Multivariate Adaptive Regression Splines (MARS) and linear regression model, and established a non-parametric regression model and an empirical regression model respectively based on Multivariate Adaptive Regression Splines (MARS) and linear regression model. The reference dataset of FSC was prepared with Landsat 8 surface reflectance data and SNOMAP algorithm. A part of reference dataset is selected as the training samples of the model, and the other part as the validation dataset of the model. The results show that the accuracy of the MARS method is significantly higher than that of the linear regression model and the original MODIS FSC preparation method. The total R, RMSE and MAE of MARS were 0.791, 0.103 and 0.058, respectively. In the linear regression model, the overall R, RMSE and MAE with the highest accuracy are 0.647, 0.128 and 0.072, respectively. The overall R, RMSE and MAE of the original MODIS FSC mapping method are 0.595, 0.221 and 0.170 respectively. MARS method with environmental information is more suitable for FSC preparation in Tibetan plateau. This study provides a new idea for preparing FSC data with higher accuracy in Tibetan plateau.

Key words: Tibetan plateau; Regression model; Fractional snow cover; MODIS; Multivariate adaptive regression splines