

DOI:10.7522/j.issn.1000-0240.2017.0058

QIU Yubao ,ZHANG Huan ,Chuduo ,et al. Cloud removing algorithm for the daily cloud free MODIS-based snow cover product over the Tibetan Plateau[J]. Journal of Glaciology and Geocryology ,2017 ,39(3) : 515 – 526. [邱玉宝,张欢,除多,等. 基于 MODIS 的青藏高原逐日无云积雪产品算法[J]. 冰川冻土,2017 ,39(3) : 515 – 526.]

基于 MODIS 的青藏高原逐日无云积雪产品算法

邱玉宝¹, 张 欢^{1,2}, 除 多³, 张雪成⁴, 于小淇¹, 郑照军⁵

(1. 中国科学院 遥感与数字地球研究所 数字地球重点实验室,北京 100094; 2. 北京旋极伏羲大数据技术有限公司,北京 100081;
3. 青藏高原大气环境科学研究所,西藏 拉萨 850000; 4. 西安科技大学,陕西 西安 710054;
5. 中国气象局国家卫星中心,北京 100081)

摘 要: 青藏高原积雪对高亚洲地区水和能量循环起着重要的反馈和调节作用,其变化影响着融雪性河流流量,对下游水资源和经济活动具有重要影响。中分辨率成像光谱仪(MODIS)具有较高的时空分辨率,被广泛应用于积雪遥感动态监测,然而光学遥感积雪受云层影响严重,且青藏高原地区水汽分布不均,局地对流活跃,积雪的赋存时间变化快,这给高原地区逐日积雪监测及其气候制图带来挑战。在考虑青藏高原地形和积雪分布特征情况下,结合现有的云覆盖下积雪判别算法,采用 8 个不同方法的组合,逐步实现 MODIS 逐日无云积雪算法。选取 2009 年 10 月 1 日–2011 年 4 月 30 日两个积雪季为研究期,并采用 145 个地面台站观测雪深数据对去云算法各步骤过程开展精度验证,结果表明:当积雪深度 > 3 cm 时,逐日无云积雪产品总分类精度达到 96.6%,积雪分类精度达 83%,积雪判对概率(召回率)达到 89.0%,算法可实现青藏高原地区逐日无云积雪动态监测和积雪覆盖气候学数据重建,对高亚洲地区的水、生态和灾害等全球环境变化影响研究具有重要的意义。

关键词: 青藏高原; 积雪遥感; 去云算法; MODIS; 逐日无云雪盖产品;

中图分类号: P426.63+5/P407 文献标志码: A 文章编号: 1000-0240(2017)03-0515-12

0 引言

积雪是冰冻圈的重要组成部分,是地球表面最为活跃的自然要素之一,积雪对区域甚至全球的气候变化、能量平衡以及水循环都有着深刻影响^[1]。积雪反射能力强、热传导性差,通过在融化过程中吸收大量热量改变地表辐射平衡和大气热状况,引起大气环流变化,进而对区域气候产生影响^[2],融雪水文效应对地表径流和大气热状况也产生着深刻的影响^[3]。青藏高原是中纬度地区海拔最高的区域,也是“高亚洲”地区的核心区域,青藏高原积雪对于亚洲季风和南亚、东亚旱涝灾害的影响,一直是中外气候学家的关注焦点^[4–8]。

目前,卫星遥感技术由于重访周期短(半月至日时间尺度)、覆盖范围大等优点逐渐成为积雪参

数监测的主要方式。当前,常用的积雪遥感数据有 Landsat 和 SPOT^[9], AVHRR^[10]、VEGETATION^[11]、MODIS^[12]等,可用于光学积雪覆盖度和二值积雪产品的估算,利用 SMMR、SSM/I^[13–14]和 AMSR-E^[15–17]等微波遥感数据进行雪深监测。其中,MODIS 逐日积雪监测数据(MOD10A1 和 MYD10A1)具有较高的时空分辨率,在北半球大面积积雪制图方面精度可达 93%,可以满足全球气候和环境变化研究的需求^[18]。然而,由于积雪和云的反射特性相似,使得光学遥感产品受天气影响严重,在日时间尺度上,云污染限制了积雪产品的应用。因此,国内外不仅开展云雪分类算法的探索,同时基于 MODIS 积雪产品,开展去云覆盖像元的再分类研究,并提出多种算法,如基于时间连续性的上下午星合成方法^[19–23]、3 d 临近日分析法^[19 21 23]、长时间序列分

收稿日期: 2017-01-13; 修订日期: 2017-03-12

基金项目: 国家自然科学基金项目(41371351); 中国科学院国际合作局对外合作重点项目(NO. 131CIHKYSB20160061); 公益性(气象)行业专项“青藏高原遥感积雪气候数据集建设”(GYHY201206040)资助

作者简介: 邱玉宝(1978–),男,博士,副研究员,2008 年毕业于中国科学院遥感应用研究所,从事微波遥感应用及高山和极地寒区信息服务研究。E-mail: qiuyb@radi.ac.cn.

析^[19-20, 24]；基于空间连续性的临近四像元法^[19-20]、临近八像元法^[19, 25]；基于数字高程模型的雪线方法^[21, 26-27]，以及利用被动微波不受天气制约的优点，结合被动微波数据 AMSR-E 开展的去云方法^[21, 26-29]等。

通过对上述多种方法在青藏高原的适用性进行假设的验证性分析^[30]，结果表明有些方法和步骤精度较高且不受地域限制，可完全用于青藏高原，如临近日分析法、临近像元法；有些方法有其区域局限性，如雪线方法等；而利用被动微波遥感数据帮助去云，由于其空间分辨率较低（0.25° 或 25 km），积雪面积的高估较大，对于积雪分布不均且分散破碎的青藏高原来说可带来较大的误差。

本研究以 MODIS 逐日积雪产品数据 (MOD10A1 和 MYD10A1) 为基础，综合考虑青藏高原的地形特点和积雪分布特征，基于已有的多种去云算法，通过逐步控制剩余云量和积雪的再分类的精度，形成一套新的去云方法流程，并最终获得 MODIS 逐日无云积雪产品，为青藏高原积雪动态监测和历史积雪参数数据的重建提供了参考。

1 研究区概况

青藏高原地处中纬度地区，是中国面积最大、世界海拔最高的高原，是高亚洲地区的核心区，其面积约为 $257 \times 10^4 \text{ km}^2$ ，包括西藏自治区和青海省的全部，四川、云南、新疆和甘肃省部分地区，以

及高原周边的尼泊尔、印度、巴基斯坦以及中亚等国家的一部分毗邻区域。

青藏高原研究区如图 1 所示，平均海拔 4 000 ~ 5 000 m，地形复杂，为中低纬度地区面积较大的低温区，有着丰富的冰雪水资源，是亚洲许多大江大河的发源地，如长江、黄河、雅鲁藏布江等，积雪变化直接影响融雪性河流流量的变化。高原冰雪对气候变化敏感，对东亚大气环流、夏季风降水和全球气候变化等起着重要的调节作用^[31-33]。积雪达到一定厚度可造成气温骤降、交通阻断、大量家畜挨饿受冻而亡，从而引发一系列雪灾，威胁到社会经济的发展及人们的生产生活。该地区的积雪动态监测具有重要的应用价值。

2 研究数据和去云算法

2.1 研究数据与预处理

研究所需要的数据包括 MODIS 逐日积雪数据、数字高程模型以及用于精度对比验证的地面台站雪深观测数据。

其中 MODIS 遥感影像来源于美国冰雪数据中心 (NSIDC)，包括 MODIS/Terra 上午星 (MOD10A1) 和 MODIS/Aqua 下午星 (MYD10A1) 积雪产品数据，时间为 2009 年 10 月 1 日 - 2011 年 4 月 30 日两个雪季，数据从 <ftp://n5eil01u.ecs.nsidc.org/> 免费下载，投影方式为正弦曲线地图投影，转换拼接后，其研究所用为投影 WGS84 坐标系

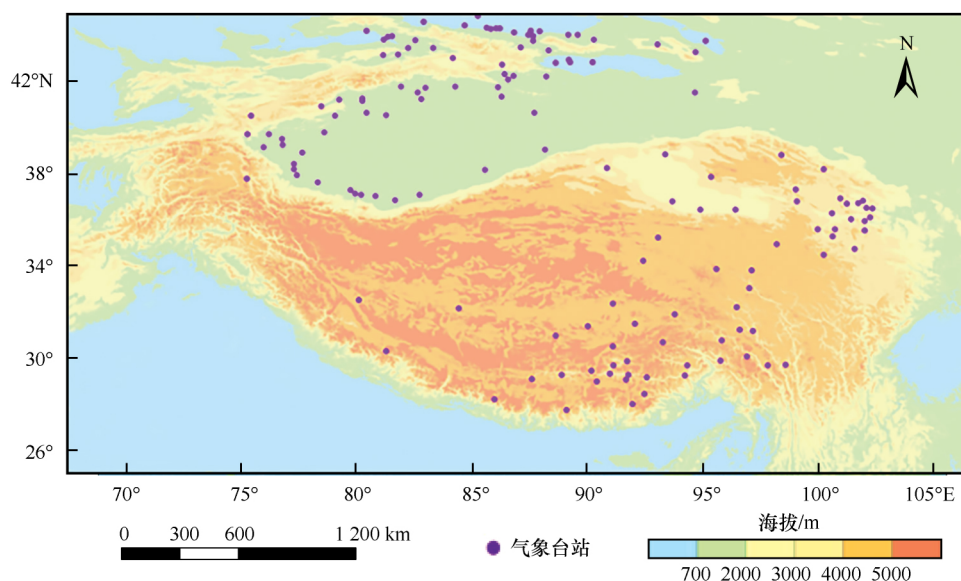


图 1 研究区范围及所采用的积雪地面观测站点分布

Fig. 1 Map showing the location of meteorological stations with snow cover observation in the study area

下的 UTM 投影。

研究中所采用的数字高程数据 (DEM) 来自空间信息协会 CGIAR-CSI 提供的 90 m 分辨率 SRTM 数据集^[34]。将下载的数据进行拼接和裁剪, 重采样成与 MODIS 数据匹配 (约 500 m)。

地面台站观测数据为同时期研究区内 145 个台站逐日雪深数据。当雪深为 0 时记为无雪, 雪深 > 0 时为有雪, 两个积雪季 424 个时相, 共计 61 480 个数据, 用于去云过程中各合成积雪产品分类精度的验证。

2.2 逐步去云算法

青藏高原地区面积广阔, 地形复杂多样, 且积雪空间异质性较大, 这些特点决定了去云方法的研究需考虑区域适用性, 根据对青藏高原积雪分布特征的分区假设检验研究^[30], 发现青藏高原地区云覆盖像元的再分类需要考虑以下方面: 1) 对于高原腹地, 需要考虑对流过程复杂和地形坡度较缓, 积雪出现时间短, 变化速度快等特点; 2) 高原高山区存在常年冰雪 (年度最小雪冰覆盖) 问题, 且积雪受高程的控制明显 (温度较低), 降水的概率受高原西风 and 季风区的影响原因不同, 积雪的分布特征存在差异。因此, 需要考虑时间和空间的特点, 如基于高程依赖的假设算法在青藏高原应采用分区的方式等。

结合前文针对假设的检验分析^[30], 研究采用 8

个步骤对青藏高原 MODIS 积雪产品进行云覆盖条件下的积雪判别, 逐步减少云量, 最终获得逐日无云产品。算法流程图见图 2, 其过程描述和分析如下:

步骤 1: MODIS 上下午星数据合成。

参考 Gao 等^[21]的最大合成算法, 在同一天中, 假设高原地区积雪覆盖在半天内的短时间内不出现大的变化, 对 MODIS 数据进行上下午星积雪产品合成, 公式表达如式 (1):

$$S_{(x,y,t)} = \max(S_{(x,y,t)}^A, S_{(x,y,t)}^T) \quad (1)$$

式中: x, y 分别表示像元 S 所在的行号和列号; t 表示水文年的某一天; A 和 T 分别表示下午星 Aqua 和上午星 Terra; $S_{(x,y,t)}$ 为像素值, 合成优先级顺序定为: 积雪 > 湖冰 > 湖泊 > 陆地, 若合成图像仍有数据缺失或其他无效值, 则赋值为云, 与其他云像素一同进行下一步处理。对 2009–2011 年两个积雪季每天的去云量进行统计, 此步骤平均去除云量约 11.0%, 获得每日最大积雪合成产品。

步骤 2: 3 d 合成。

一般而言, 阴天 (多云) 天气积雪接受太阳辐射较少, 不易融化, 降落在地面上会持续一段时间^[35], 根据临近日的雪盖信息将当天的云像素重新分类。公式表达如式 (2):

$$S_{(x,y,t)} = 1 \quad \text{if} \quad (S_{(x,y,t-1)} = 1 \text{ and } S_{(x,y,t+1)} = 1) \quad (2)$$

式中: x, y, t 含义与式 (1) 中相同, 若当日合成的积雪产品有云像素, 前后日的相应位置同为积雪,

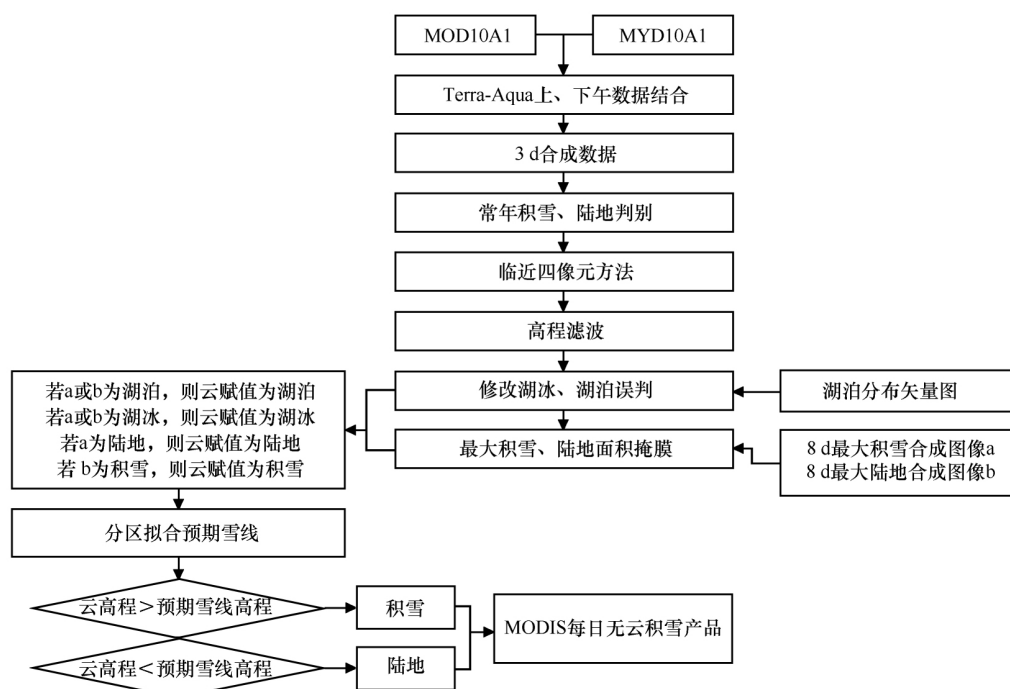


图 2 MODIS 逐日无云积雪产品算法流程

Fig. 2 The algorithm flow of MODIS daily cloudless snow products

则当日云像素分类为积雪,陆地覆盖同理。若前后日的相应位置至少一个为湖冰,则当日云像素赋值为湖冰,湖泊同理,其他情况则不去云处理。3 d 合成在 2009 - 2011 年雪季中平均去除了 7.5% 的云量。

步骤 3: “常年积雪”^[36]和“无雪陆地”判别法。

青藏高原山区分布着丰富的常年冰雪覆盖区,也可认为是年度最小雪冰覆盖区^[36],即无论这些地方是否被云覆盖,都可直接将对应位置的云像素赋值为积雪,而有些地方在某些年份可能从不下雪或极少下雪的区域,则可以通过判别直接赋值为陆地。以 2010 - 2011 年积雪季为例,通过对比分布形态和细节较为丰富的 Google Earth 高清晰度无云卫星图片,并结合已发布的青藏高原“年度最小雪冰覆盖”作为基准参考^[36],基于 DEM 的分布统计分析,分析发现高程在 5 800 m 以上的区域,可认为积雪常年覆盖(5 800 m 以上区域面积包含于“常年积雪”面积);由于积雪上空云层往往多于陆地,且避免 MODIS 积雪识别算法的错判可能,规定只要在一个雪季某像素积雪覆盖和云覆盖的时间综合大于全部时间的 95%,即认为云可分类为积雪。类似地,如果某像素在这段时间内只出现陆地和云,但从未出现积雪,且云覆盖天数较少,则将云分类为陆地。经与 Google Earth 高清晰度雪冰分布图像和已经发布的年度最小雪冰覆盖图对照,云天数小于总天数的 20% 时,可保证精度不产生重大的损失。从而形成经验性的积雪去云的判断条件。

对每个时间段进行“常年积雪”和“无雪陆地”的统计分析,判断条件如下:

①高程 > 5 800 m 的像素全部为“常年积雪”或年度最小雪冰覆盖;②高程在 3 000 ~ 5 800 m 之间的云像素,若满足:云天数 + 积雪天数 > 总天数 × 0.95,则云像素分类为积雪;③若某像素满足:1) 云天数 + 陆地天数 = 总天数;2) 云天数 < 总天数 × 0.2,则云像素分类成陆地。

如图 3 所示,用此图像中的“常年积雪”和“无雪陆地”代替这一雪季每日对应位置的云像素,统计此步骤在 2009 - 2011 年两个雪季的去云量,平均值达 3.1%。

步骤 4: 临近四像元法。

青藏高原在降雪和融雪季节积雪较为破碎,但对于多数情况而言,积雪在空间上连续的概率较大,可根据云像素周围的非云像素信息对其进行重新分类。如果与云像素相邻最近的四个像元中,至少三个为积雪,那么将中心云像素赋值为雪;如果至少三个为陆地,那么将中心云像素赋值为陆地;湖泊、湖冰同理。对于所选的研究数据,该方法去除云量较少,为 5%,但分类精度很高,为整体去云算法的使用减少累积误差^[30]。

步骤 5: 高程滤波法。

基于海拔较高地方温度较低,积雪容易积累的认识,采用高程滤波方法,利用云像素周围信息及数字高程(SRTM DEM)辅助数据,对上一步完成的 MODIS 逐日积雪图像进行去云,算法描述如式(3):

$$S_{(x+i, y+j)} = 1 \quad \text{if} \quad (S_{(x, y)} = 1 \text{ and } H_{(x+i, y+j)} > H_{(x, y)}, \quad i, j = \pm 1) \quad (3)$$

以积雪像素为中心,若与其相邻的八个像元中

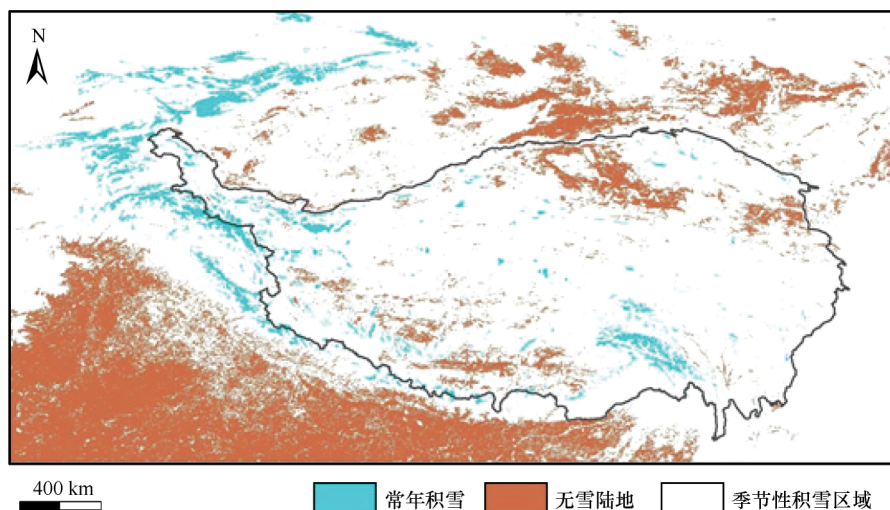


图 3 2010 年 10 月 1 日 - 2011 年 4 月 30 日“常年积雪”和“无雪陆地”分布

Fig. 3 Distributions of permanent snow cover, snow-free land and seasonal snow cover from 2010-10-01 through 2011-04-30

有云,且云像素高程大于中心积雪高程,则将云像素赋值为积雪,对于所选的研究数据,该步骤去量为 0.9%,其精度上在高山区较高^[30]。

步骤 6: 阴影区分类修正。

通过对比世界湖泊矢量边界,发现 MODIS 数据在青藏高原的西部地区,通常把阴影区错误地判断为湖冰或湖泊,这对于后面的去云过程造成误差累计,为减小这种误差,采用国家基础地理信息系统中心的现有的湖泊矢量边界作为辅助数据,对错误的湖泊湖冰进行修正。边界线之内的湖泊或湖冰保留,界线之外错分的湖泊和湖冰重新赋值为云,并进行下一步判断。

步骤 7: 8 d 最大积雪和最大陆地面积掩膜法。

假设 8 d 之内积雪时空分布具有一定的空间随机变化性和时间连续性,那么 8 d 内中,单独每一天的积雪面积应 > 8 d 最小值积雪面积(最大陆地面积),而 < 8 d 最大积雪面积,如此可把每天积雪动态范围控制在以 8 d 为时间窗口的动态变化范围内,从而实现积雪去云过程判识精度的控制。

基于以上高程滤波结果数据,开展 8 d 最大积雪覆盖(类似 MODIS 8 d 积雪产品方法)和 8 d 最小积雪(即最大陆地)覆盖合成,优先级顺序分别为:积雪 > 湖冰 > 湖泊 > 陆地、陆地 > 湖冰 > 湖泊 > 积雪,获得结果分别记为图像 a(8 d 最大积雪合成,也即 8 d 最小陆地合成)和图像 b(8 d 最大陆地合成,即 8 d 最小积雪合成),其算法的理论公式表示如式(4)和式(5),从而获得“最大积雪和最大陆地掩膜”图。据此可以对 8 d 内的逐日积雪覆盖进行掩膜处理,获得去云后的图像。

图 4 显示了 2011 年 1 月 9 日 - 2011 年 1 月 17 日期间 8 d 最小积雪面积与最大积雪面积分布,最小积雪和最小陆地面积之间存在的白色区域为不确定区域,即为 8 d 积雪变化的范围。

$$S_{(x,y)} = 1 \quad \text{if} \quad S_{(x,y)}^b = 1 \quad (4)$$

$$S_{(x,y)} = 0 \quad \text{if} \quad S_{(x,y)}^a = 0 \quad (5)$$

式中: 1 和 0 分别表示积雪和陆地; $S_{(x,y)}^a$ 为图像 a 的 8 d 最大积雪合成图像中第 x 行、第 y 列像素的值; $S_{(x,y)}^b$ 为图像 b 的 8 d 最大积雪合成图像中第 x 行、第 y 列像素的值。如果像素是云,任一最大合成图像中像素是湖冰或湖泊,则给云像素相应赋值为湖冰或湖泊。对于所选的研究数据,这一步骤去除云量最多,约 16.5%。

步骤 8: 拟合预期雪线方法。

由于积雪覆盖受经度、纬度、坡度、坡向以及高程等的因素控制,因此对于剩下的云覆盖区域,用“拟合预期雪线”方法开展最后的再分类处理。算法描述如下:①提取雪线样本。积雪与陆地的分界线即为雪线,提取出雪线之上积雪的高程、坡度、坡向、经度、纬度。②拟合预期雪线高程,以高程为因变量,以其他四个影响因素为自变量,用多元线性回归拟合预期雪线高程。③将云像素重新分类,由云覆盖像素的坡度、坡向、经度、纬度可算出预期雪线高程。将预期雪线高程与云的实际高程比较,若实际高程较高,则将云赋值为积雪,反之赋值为陆地。

其公式表示如下:

$$S_{(x,y)} = 1 \quad \text{if} \quad H_{(x,y)}^c > H_{(x,y)}^c \quad (6)$$

$$S_{(x,y)} = 0 \quad \text{if} \quad H_{(x,y)}^c \leq H_{(x,y)}^c \quad (7)$$

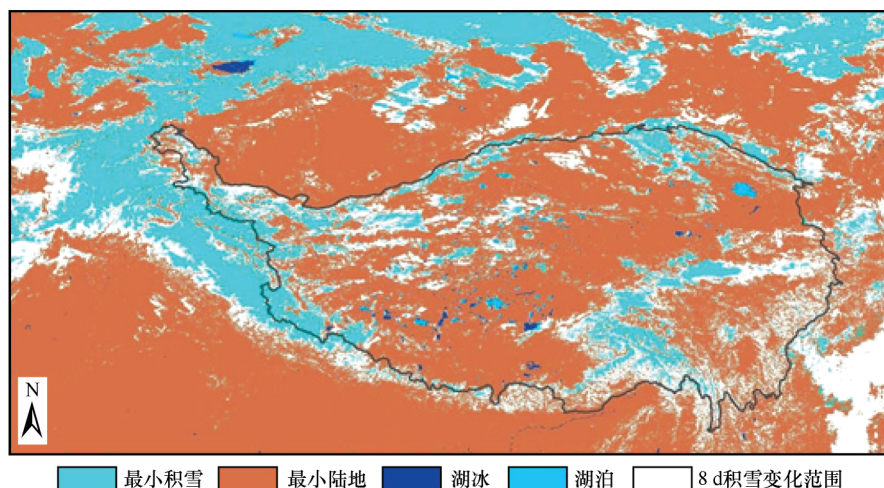


图 4 2011 年 1 月 9 日 - 2011 年 1 月 17 日最小积雪面积、最小陆地面积及积雪变化范围

Fig. 4 Distributions of minimum snow cover, minimum snow-free land and snow cover variation range from 2011-01-09 through 2011-01-17

式中: 1 和 0 分别表示积雪和陆地, $H_{(x,y,t)}^c$ 为云像素对应的高程, $H_{(x,y,t)}^e$ 表示计算出相同位置的预期雪线高程。由于青藏高原面积较大, 且不同地区地形差异很大, 其积雪的依赖特征较大, 按地形和高程将青藏高原分为 7 个子区域(图 5), 每个区域根据不同的时间段分别使用“拟合预期雪线”方法去云。

其多元线性回归模型的一般形式为:

$$H' = a_0 + a_1X_1 + a_2X_2 + a_3X_3 + a_4X_4 \quad (8)$$

采用 2011 年 1 月份拟合出的不同区域的预期高程雪线方程, 根据图 5 中 7 个不同区域, 分别为:

$$Y = -1842.6 + 1.22X_1 - 0.03X_2 + 34.56X_3 + 3.95X_4 \quad (9)$$

区域 2:

$$Y = 0.35X_1 - 0.05X_2 + 15.87X_3 + 36.26X_4 \quad (10)$$

区域 3:

$$Y = 3.63X_1 + 0.05X_2 + 24.7X_3 + 33.44X_4 \quad (11)$$

区域 4:

$$Y = -7.62X_1 + 0.23X_2 + 21.23X_3 + 80.35X_4 \quad (12)$$

区域 5:

$$Y = 14.18X_1 + 0.11X_2 + 39.02X_3 + 51.53X_4 \quad (13)$$

区域 6:

$$Y = 4.38X_1 + 0.09X_2 + 25.41X_3 + 62.43X_4 \quad (14)$$

区域 7:

$$Y = 7.31X_1 + 0.15X_2 + 40.02X_3 + 59.3X_4 \quad (15)$$

通过预期拟合线方法, 可去除最后剩余的云量, 获得结果与相近产品, 如 IMS 等比对发现, 此方法基本将积雪区域和陆地区域正确的分离开来, 整体效果较好。对于试验数据, 这一步骤去除了最

后剩余平均约 5.1% 的云覆盖, 获得 MODIS 逐日无云积雪产品。

经过以上 8 个步骤, 形成基于经验参数的青藏高原逐日无云积雪产品算法流程, 其中每个去云步骤合成图像如图 6 所示。

3 算法测试结果与分析

3.1 各步骤去云效果分析

采用青藏高原地区 2009 - 2011 年两个积雪季的 424 d 数据产品, 分析整个去云过程中云覆盖比率的动态变化, 如图 7 所示, MYD10A1 数据和 MOD10A1 云量较大, 平均分别达到 47.1% 和 41.7%, 其中云量超过 50% 的分别有 186 d 和 121 d, 且多分布在每年的 12 月 - 次年 3 月的主要积雪时期, 10 月、11 月和次年 4 月、5 月云量相对减少, 可见云的大量存在, 特别是冬季积雪期严重制约着积雪动态监测。

在整个去云的步骤和过程中, 不同的步骤对云覆盖下地表的再判断效果也各不相同, 针对测试时间段, 分析发现: “上下午星合成”和“3 d 连续合成”可去除了 11.0% 和 7.5% 的云覆盖, 去云量较多; “常年积雪”和“无雪陆地”判别方法涉及到长时间的积雪覆盖信息, 去除了“常年积雪”和“无雪陆地”上覆盖的云, 约为 3.1%; 而“临近四像元法”去除零星分散约 0.5% 的云像素, 去云量较少但精度较高; “高程滤波方法”将积雪边缘海拔较高的云像素重新分类, 去除约 0.9% 的云量; “8 d 最大积雪、最大陆地”范围掩膜方法利用短时间

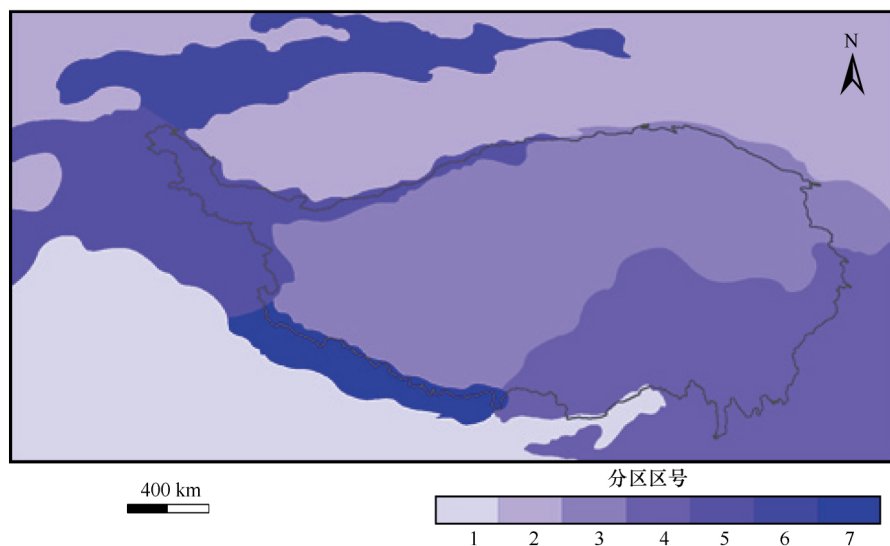


图 5 青藏高原坡度分区情况

Fig.5 Partition of the Tibetan Plateau

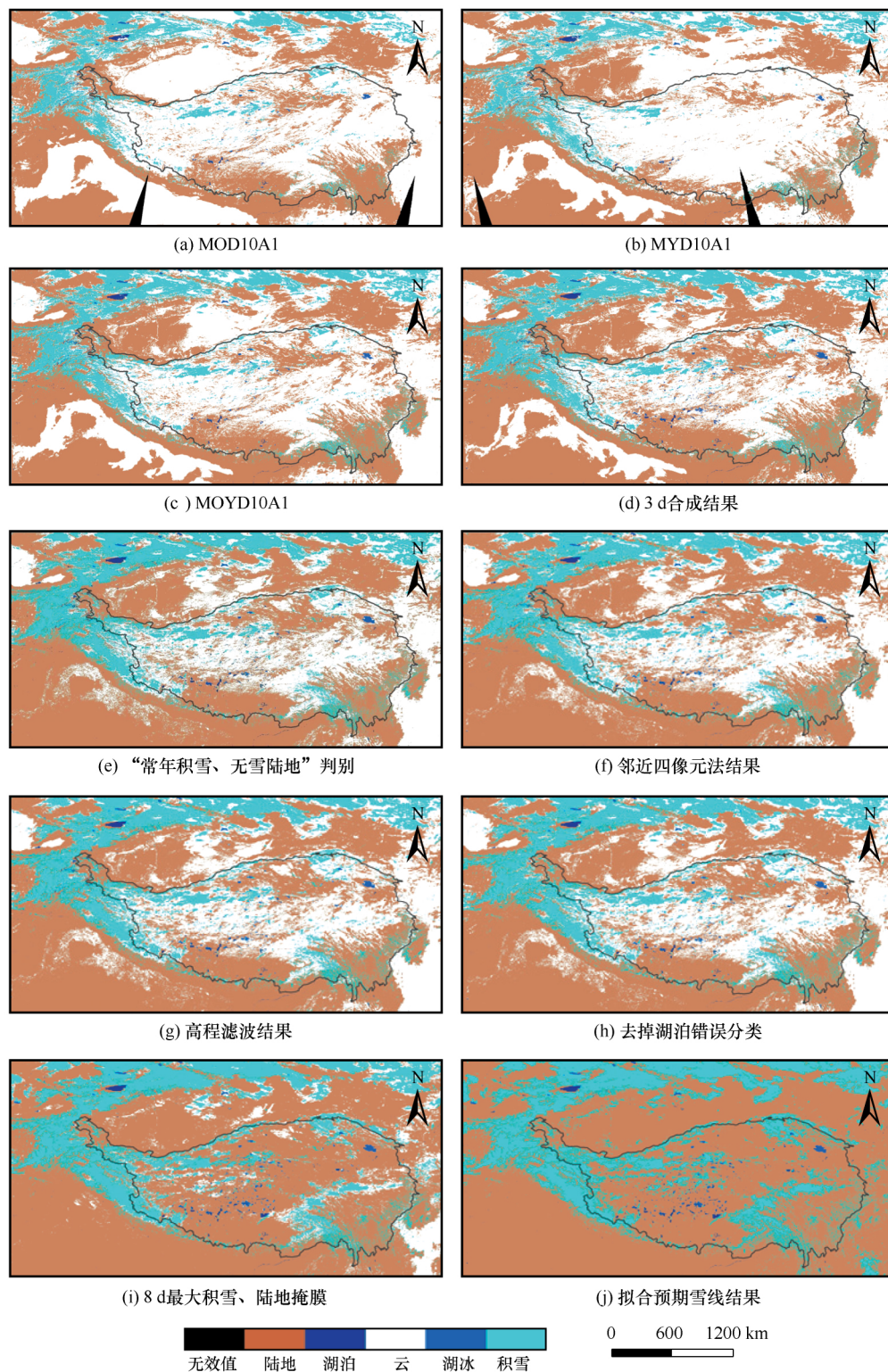


图6 2011年1月9日青藏高原MODIS无云积雪合成图像

Fig.6 MODIS daily cloudless snow product composites on January 9, 2011

内雪盖信息的相关性,将16.4%的云像素重新分类为积雪或陆地像素,去云量最多,其内在的关系是连续多天的积雪的不变性假设为基础,是3 d连续积雪去云方法的进一步的扩展,其潜在的概率扩展

到8 d内积雪的连续性,因此在青藏高原地区对流活动剧烈区,其理论上的可靠性会带来一定的误差。通过以上七个步骤,MODIS积雪合成产品仅剩5.1%的云覆盖,采用具有一定高程和纬度依赖

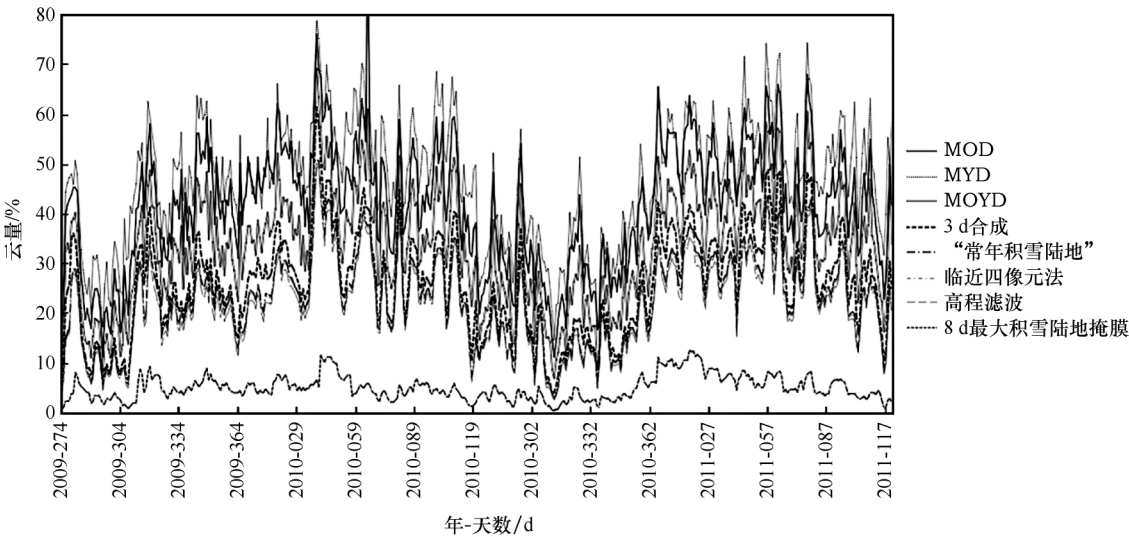


图 7 2009 – 2011 积雪季过程产品云量分析

Fig. 7 Variation of cloud cover percentage during snow cover season of 2009 – 2011

基础的拟合预期雪线算法去除剩余全部的云，从而获得逐日无云积雪产品。

3.2 精度验证

为了评估和验证积雪的分类精度和效果，论文采用积雪面积精度指标来对每一步骤的云像素分类结果做验证评价，指标包括总分类精度 (Accuracy)、召回率 (Recall)、积雪分类精度 (Precision) 以及 f 值^[37]。总精度即准确度，表示一个像素被正确分类的概率 (可能性)，但是当图像中有大面积无雪区域时，这一指标可能会带来误导，如若图像中只有 10% 积雪像元，那么总精度亦会由于大部分甚至全部像元被分为无雪而达到 90%。召回率表示检测到积雪像素的概率，即正确分类出的积雪像元数与实际所有积雪样本像素数的比值。积雪分类精度表示正确分类出的积雪像元数与分类出的所有积雪像元的比值。召回率和分类精度是相互影响的，召回率高表明正确分类的积雪像元数相对于地面站点观测的准确性较高，而精度高则表明正确分类的积雪像元数相对于 MODIS 中总积雪数的准确性较高，召回率高，并不代表分类精度高，因此在评价积雪精度的时候往往采用 f 值 (式 13) 是最合适的指标^[37]，这是召回率和积雪分类精度的调和平均值，能够较好的反映广泛无雪区域的缺失或错误分类的雪的情况。

针对青藏高原 2009 年 10 月 1 日 – 2011 年 4 月 30 日间，采用 145 个地面台站提供每日雪深数据，可以获得共 61 480 对估算和观测值。通过混淆矩阵及精度指标的式 (16)、式 (17)、式 (18)、式 (19)

对积雪分类图像进行评价，分析晴空天气下 MODIS 数据本身 (MOD10A1、MYD10A1) 以及去云过程中各合成产品的精度，在精度评价中考虑以下几个方面样本数：1) 地面台站记录 and 积雪图像均有雪 (雪深 > 0) 的样本数 a ；2) 地面台站记录有雪而图像分为无雪的样本数 b ，即漏测数；3) 地面台站记录无雪而图像分为有雪类型的样本数 c ；4) 地面台站记录 and 图像均无雪的样本数 d (表 1)。

表 1 精度验证混淆矩阵		
Table 1 Confusion matrix for precision validation		
MODIS	积雪	陆地
地面		
积雪	a	b
陆地	c	d

积雪精度指标公式如下：

准确度 (Accuracy) = $\frac{a + d}{a + b + c + d} \times 100\%$ (16)

召回率 (Recall) = $\frac{a}{a + b} \times 100\%$ (17)

精度 (Precision) = $\frac{a}{a + c} \times 100\%$ (18)

$f = 2 \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 100\% = \frac{2a}{2a + b + c} \times 100\%$ (19)

原始数据 MOD10A1、MYD10A1 及各过程产品在混淆矩阵中的样本数统计见表 2，在积雪去云的处理过程中，修正湖泊、湖冰错判改变的像元数少，没有参与统计精度。青藏高原积雪大多较薄，且分布斑驳分散，高原地区下垫面异质性较大，地

表 2 各算法过程产品精度一览
Table 2 The accuracies of various algorithm processes 序号

序号	步骤	a	b	c	d	a3	b3	总分类精度/%		召回率/%		积雪分类精度/%		f 值/%	
								全部	雪深>3 cm	全部	雪深>3 cm	全部	雪深>3 cm	全部	雪深>3 cm
								台站	样本	台站	样本	台站	样本	台站	样本
1	MOD10A1	2 156	419	264	32 424	1 919	104	98.06	98.94	83.73	94.86	89.09	87.91	86.33	91.25
2	MYD10A1	1 708	420	212	29 448	1 596	100	98.01	99.00	80.26	94.10	88.96	88.27	84.39	91.10
3	MOYD10A1	2 853	550	390	36 812	2 564	130	97.69	98.70	83.84	95.17	87.97	86.80	85.86	90.79
4	3 d 合成	3 417	663	420	41 393	3 098	146	97.64	98.74	83.75	95.50	89.05	88.06	86.32	91.63
5	“常年积雪”和“无雪陆地”	3 417	711	420	42 406	3 098	150	97.59	98.76	82.78	95.38	89.05	88.06	85.80	91.58
6	邻近四像元法	3 496	722	425	42 761	3 169	151	97.58	98.76	82.88	95.45	89.16	88.17	85.91	91.67
7	高程滤波法	4 059	722	495	42 761	3 656	151	97.47	98.63	84.90	96.03	89.13	88.08	86.96	91.88
8	最大积雪陆地面积掩膜	6 850	1366	672	50 372	6 187	344	96.56	98.24	83.37	94.73	91.07	90.20	87.05	92.41
9	拟合预期雪线(无云产品)	7 266	2003	1240	50 971	6 406	792	94.73	96.58	78.39	89.00	85.42	83.78	81.76	86.31

面台站雪深数据与 MODIS 积雪图像像元之间空间尺度的差异会造成精度验证的误差。

当融雪期积雪较薄时(如 <3 cm)，可能出现这种情况：包括地面台站所在的小面积区域观测为有雪，而 MODIS 图像 500 m × 500 m 像素判断为无雪，从而产生积雪的低估误差。也可能台站处于城市，由于温度高于四周而较快融雪，此时 MODIS 图像判断为有雪，如此出现积雪高估误差。另外，对于斑驳积雪的深度较小的情况，光学可以做一定的穿透，且混合像元问题较为严重，因此采用一定深度的积雪开展验证，将更好地对积雪去云的判别过程开展评价。若只保留积雪深度 >3 cm 的样本数据，图像和地面观测同时记录为积雪的样本数记为 a3，图像记录为积雪而地面记录为陆地的样本数记录为 b3，仍用上式计算，各合成产品的分类精度较采用全部积雪观测数据有所提高。

对于积雪深度 >3 cm 的情况下，MODIS 数据本身的总分类精度在 98% 以上，而其 f 值超过 90%，在台站所观测的区域精度较高。从整个积雪去云的角度来看，积雪的召回率和积雪分类精度直到少云步骤(步骤七)的过程中，在两个冬季期间的积雪判别中，其精度基本上没有损失，甚至有些提升的情况，主要的原因一方面是由于积雪去除的过程中判别的精度损失较少，也存在着部分在有积雪观测的区域，并未能受到太多的影响，而对于最不稳定的去云的最后一步采用拟合线的时候获得了较大的误差，但是依然可达 86.31% 的 f 值。

总体来看，由表 2 可见，高程滤波之前的去云

步骤对分类精度影响不大，上下午最大积雪合成和高程滤波方法稍微提高了积雪分类精度，最大积雪、陆地面积掩膜方法使合成产品的精度有所降低，但去云量最多。拟合预期雪线方法精度最低，用于去云最后一步，但当雪深 >3 cm 时，无云积雪产品的总分类精度、召回率、积雪分类精度及 f 值仍可达 96.6%、89%、83.8% 和 86.3%，因此积雪分类上依然存在一定的提升空间，但是这种提高需要考虑本身 MODIS 积雪产品在高原地区地表不均一，积雪斑驳和动态变化快等引起的不足，甚至从积雪产品本身的精度开始做提升。

4 讨论与结论

由于积雪和云的反射特性相似，光学积雪遥感数据受到云的严重污染，MODIS 逐日积雪产品(MOD10A1、MYD10A1)往往不能直接进入模型，而是采用八天合成的方式形成少云的积雪数据产品，这在北半球高纬度地区，由于积雪存在的持续性高，八天产品对于每天的积雪状态的代表性强，而对于青藏高原地区的积雪变化快且覆盖较薄等情况下，八天之内的积雪在时空尺度上会发现较大的变化，因此，需要对青藏高原地区的每日无云积雪的开发加以考虑。本项研究在前人 MODIS 积雪产品去云算法的基础上，结合研究区独特的山区地形特点和积雪时空分布的规律，形成了一套适用于青藏高原的 MODIS 逐日无云积雪产品合成算法步骤。

首先考虑到多云天气(阴天)下积雪接受太阳

辐射能量少,不易融化,对 MODIS 逐日积雪图像进行上下午合成和 3 d 合成,利用“常年积雪”和“无雪陆地”的分布范围图像可将每天对应像元的云像素重新分类。这三种方法应用了积雪在时间上的连续性规律,通过两个冬季的测试,可去除 20% 以上的云量。其次,积雪在空间上成片出现的概率较大,根据其空间连续性,采用邻近四像元法进行去云,此方法去云量少但可以获得较高的精度,为后面去云过程减少误差累积。根据山区积雪有一定的垂直分布规律,即积雪海拔高于周围陆地的海拔,采用高程滤波和拟合预期雪线方法,将海拔较高或较低的云像素重新分类。考虑到每天积雪和陆地的面积均大于八天最小积雪和陆地的范围,采用八天最小积雪陆地掩膜方法进行去云,此类方法在时间上的误差不超过 8 d,从统计的结果来看,积雪分类精度较高,但是其潜在的假设依然是积雪的时间连续性,这考虑连续性差的区域其误差和不缺性增加。最后,对于残存的云覆盖区采用具有多元统计拟合预期雪线法方法开展处理,虽然分区处理能降低误差,但其假设在青藏高原存在着较大不确定性,因此带来较大误差。

通过逐步控制云量和精度损失,采用八个步骤去除全部实现云覆盖下再分类,并采用两个冬季的地面台站观测数据对 MODIS 逐日无云积雪产品进行精度验证。结果表明,当积雪深度 > 3 cm 时,无云产品的总分类精度为 96.6%,积雪分类精度为 89.0%,可以认为保持了较高的精度,特别是在冬季条件下,各步骤的假设可靠性较强,精度保持较好,该系列方法步骤的组合可用于青藏高原积雪面积的动态监测。但是,当积雪深度较浅(≤ 3 cm)时,无云产品的总分类精度和积雪分类精度都有所下降,这是由于 MODIS 标准积雪产品本身对薄雪的分类精度不高,且地面台站作为一个单点与 MODIS 像素存在着尺度差异,站点有无积雪不能够准确代表其所在 MODIS 像元的雪盖情况,特别是青藏高原地区斑驳积雪的存在出现了混合像元和光学积雪的穿透性问题。

总结认为,青藏高原地形复杂多样,积雪分布不均,海拔较高的山区积雪分布较多,而高原腹地降雪很少。根据积雪的时间连续性、空间连续性,以及雪盖与高程的关系,采用不同方法进行去云,获得了精度较好的 MODIS 无云积雪产品,并可供一定的模型作为数据输入,当前已形成可供使用的产品数据集发布^[38]。然而,针对青藏高原山区积

雪面积监测精度的进一步提高,还需要在原始观测值的基础之上开展非均匀复杂地形条件下混合像元的分析,并充分考虑高山区的观测和入射角等因素,从根本上提高 MODIS 数据的积雪分类精度。台站分布不均及其代表性也影响着无云产品的精度验证,可考虑采用更多不同的高分辨率数据,开展积雪产品精度的验证,以期实现实时监测青藏高原积雪动态变化、防灾减灾、气候预测等应用。

参考文献(References):

- [1] 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.]
- [2] Liu Yimin, Qian Zheng'an. Influence of sea-land thermal contrast on the climate change in China [M]. Beijing: China Meteorological Press, 2005. [刘屹岷, 钱正安. 海-陆热力差异对我国气候变化的影响[M]. 北京: 气象出版社, 2005.]
- [3] Barnett T P, Dümenil L, Schlese U, et al. The effect of Eurasian snow cover on regional and global climate variations [J]. *Journal of the Atmospheric Sciences*, 1989, 46(5): 661–686.
- [4] Qiu Yubao, Guo Huadong, Bin Chanjia, et al. Comparison on snow depth algorithms over China using AMSR-E passive microwave remote sensing [C]//Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International. IEEE, 2014: 851–854.
- [5] Blanford H F. On the connexion of the Himalaya snowfall with dry winds and seasons of drought in India [J]. *Proceedings of the Royal Society of London*, 1884, 37(232/233/234): 3–22.
- [6] Laba Zhuoma, Qiu Yubao, Cidan Basang, et al. The validation of MODIS daily snow cover products after cloud removal in Tibet Autonomous Region [J]. *Journal of Glaciology and Geocryology*, 2016, 38(1): 159–169. [拉巴卓玛, 邱玉宝, 次旦巴桑, 等. 青藏高原 MODIS 每日积雪产品去云算法过程对比验证研究[J]. 冰川冻土, 2016, 38(1): 159–169.]
- [7] Yan Yuna, Che Tao, Li Hongyi, et al. Using snow remote sensing data to improve the simulation accuracy of spring snow melt runoff: take Babao River basin as an example [J]. *Journal of Glaciology and Geocryology*, 2016, 38(1): 211–221. [闫玉娜, 车涛, 李弘毅, 等. 使用积雪遥感面积数据改善山区春季融雪径流模拟精度[J]. 冰川冻土, 2016, 38(1): 211–221.]
- [8] Chu Duo, Yang Yong, Luobu Jiancan, et al. The variations of snow cover days over the Tibetan Plateau during 1981–2010 [J]. *Journal of Glaciology and Geocryology*, 2015, 37(6): 1461–1472. [除多, 杨勇, 罗布坚参, 等. 1981–2010 年青藏高原积雪日数时空变化特征分析[J]. 冰川冻土, 2015, 37(6): 1461–1472.]
- [9] Dankers R, De Jong S M. Monitoring snow-cover dynamics in Northern Fennoscandia with SPOT VEGETATION images [J]. *International Journal of Remote Sensing*, 2004, 25(15): 2933–2949.
- [10] Hendrix S D. Variation in seed weight and its effects on germination in *Pastinaca sativa* L. (Umbelliferae) [J]. *American Journal of Botany*, 1984: 795–802.
- [11] Xiao X, Zhang Q, Boles S, et al. Mapping snow cover in the

- pan-Arctic zone, using multi-year (1998 – 2001) images from optical VEGETATION sensor[J]. *International Journal of Remote Sensing*, 2004, 25(24): 5731 – 5744.
- [12] 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.
- [13] Grody N C, Basist A N. Global identification of snowcover using SSM/I measurements[J]. *IEEE Transactions on geoscience and remote sensing*, 1996, 34(1): 237 – 249.
- [14] Che Tao, Li Xin, Jin Rui, et al. Snow depth derived from passive microwave remote-sensing data in China[J]. *Annals of Glaciology*, 2008, 49(1): 145 – 154.
- [15] Feng Qisheng, Zhang Xuotong, Liang Tianguang. Dynamic monitoring of snow cover based on MOD10A1 and AMSR-E in the north of Xinjiang Province, China[J]. *Acta Prataculturae Sinica*, 2009, 18(1): 125 – 133. [冯琦胜, 张学通, 梁天刚. 基于 MOD10A1 和 AMSR-E 的北疆牧区积雪动态监测研究[J]. *草业学报*, 2009, 18(1): 125 – 133.]
- [16] 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.]
- [17] Dai Liyun, Che Tao, Wang Jian, et al. 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.
- [18] Hall D K, Riggs G A. Accuracy assessment of the MODIS snow products[J]. *Hydrological Processes*, 2007, 21(12): 1534 – 1547.
- [19] Gafurov A, B rdossy A. Cloud removal methodology from MODIS snow cover product[J]. *Hydrology and Earth System Sciences*, 2009, 13(7): 1361.
- [20] Parajka J, Bl schl G. Spatio-temporal combination of MODIS images – potential for snow cover mapping[J]. *Water Resources Research*, 2008, 44(3).
- [21] Gao Yang, Xie Hongjie, Yao Tandong, et al. Integrated assessment on multi-temporal and multi-sensor combinations for reducing cloud obscuration of MODIS snow cover products of the Pacific Northwest USA[J]. *Remote Sensing of Environment*, 2010, 114(8): 1662 – 1675.
- [22] Paudel K P, Andersen P. Monitoring snow cover variability in an agropastoral area in the Trans Himalayan region of Nepal using MODIS data with improved cloud removal methodology[J]. *Remote Sensing of Environment*, 2011, 115(5): 1234 – 1246.
- [23] Dietz A J, Kuenzer C, Conrad C. Snow cover variability in Central Asia between 2000 and 2011 derived from improved MODIS daily snow cover products[J]. *International Journal of Remote Sensing*, 2013, 34(11): 3879 – 3902.
- [24] Dietz A J, Wohner C, Kuenzer C. European snow cover characteristics between 2000 and 2011 derived from improved MODIS daily snow cover products[J]. *Remote Sensing*, 2012, 4(8): 2432 – 2454.
- [25] López-Burgos V, Gupta H V, Clark M. Reducing cloud obscuration of MODIS snow cover area products by combining spatio-temporal techniques with a probability of snow approach[J]. *Hydrology and Earth System Sciences*, 2013, 17(5): 1809 – 1823.
- [26] Gao Yang, Xie Hongjie, Yao Tandong, et al. Toward advanced daily cloud-free snow cover and snow water equivalent products from Terra Aqua MODIS and Aqua AMSR-E measurements[J]. *Journal of Hydrology*, 2010, 385(1): 23 – 35.
- [27] Thirel G, Salamon P, Burek P, et al. Assimilation of MODIS snow cover area data in a distributed hydrological model[J]. *Hydrology and Earth System Sciences Discussions*, 2011, 8(1): 1329 – 1364.
- [28] Huang Xiaodong, Hao Xiaohua, Wang Wei, et al. Algorithms for cloud removal in MODIS daily snow products[J]. *Journal of Glaciology and Geocryology*, 2012, 34(5): 1118 – 1126. [黄晓东, 郝晓华, 王玮, 等. MODIS 逐日积雪产品去云算法研究[J]. *冰川冻土*, 2012, 34(5): 1118 – 1126.]
- [29] Liang Tianguang, Zhang Xuotong, Xie Hongjie, et al. Toward improved daily snow cover mapping with advanced combination of MODIS and AMSR-E measurements[J]. *Remote Sensing of Environment*, 2008, 112(10): 3750 – 3761.
- [30] Zhang Huan, Qiu Yubao, Zheng Zhaojun, et al. Comparative study of the feasibility of cloud removal methods based on MODIS seasonal snowcover data over the Tibetan Plateau[J]. *Journal of Glaciology and Geocryology*, 2016, 38(3): 714 – 724. [张欢, 邱玉宝, 郑照军, 等. 基于 MODIS 的青藏高原季节性积雪去云方法可行性比较研究[J]. *冰川冻土*, 2016, 38(3): 714 – 724.]
- [31] Li Peiji. Distribution of snow cover over the High Asia[J]. *Journal of Glaciology and Geocryology*, 1995, 17(4): 291 – 298. [李培基. 高亚洲积雪分布[J]. *冰川冻土*, 1995, 17(4): 291 – 298.]
- [32] Verma R K, Subramaniam K, Dugam S S. Interannual and long-term variability of the summer monsoon and its possible link with northern hemispheric surface air temperature[J]. *Proceedings of the Indian Academy of Sciences-Earth and Planetary Sciences*, 1985, 94(3): 187 – 198.
- [33] Stow D A, Hope A, McGuire D, et al. Remote sensing of vegetation and land-cover change in Arctic Tundra Ecosystems[J]. *Remote sensing of environment*, 2004, 89(3): 281 – 308.
- [34] SRTM DEM [EB/OL]. (2016-01-11) [2016-10-15]. <http://srtm.csi.cgiar.org>.
- [35] Dankers R, De Jong S M. Monitoring snow-cover dynamics in Northern Fennoscandia with SPOT VEGETATION images[J]. *International Journal of Remote Sensing*, 2004, 25(15): 2933 – 2949.
- [36] Painter T H, Brodzik M J, Racoviteanu A, et al. Automated mapping of Earth's annual minimum exposed snow and ice with MODIS[J]. *Geophysical Research Letters*, 2012, 39(20): L20501.
- [37] 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.
- [38] Qiu Yubao, Guo Huadong, Chu Duo, et al. MODIS daily cloud-free snow cover product over the Tibetan Plateau (2002 – 2015)[J]. *China Scientific Data*, 2016. DOI:10.11922/cdata.170.2016.0003. [邱玉宝, 郭华东, 除多, 等. 青藏高原 MODIS 逐日无云积雪面积数据集(2002 – 2015 年)[J]. *中国科学数据*, 2016. DOI:10.11922/cdata.170.2016.0003.]

Cloud removing algorithm for the daily cloud free MODIS-based snow cover product over the Tibetan Plateau

QIU Yubao¹, ZHANG Huan^{1,2}, Chuduo³, ZHANG Xuecheng⁴, YU Xiaoqi¹, ZHENG Zhaojun⁵

(1. Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China;

2. Beijing Watertek Fu Xi Big Data Technology Co., Ltd, Beijing 100081, China; 3. Tibetan Institute of Atmospheric Environment

and Science, Lhasa 850000, China; 4. Xi'an University of Science and Technology, Xi'an 710054, China;

5. National Satellite Meteorological Center, Beijing 100081, China)

Abstract: Snow cover on the Tibetan Plateau plays an important role in feedback and regulating to water and energy cycles of the High Asia. Snow melt supplies river runoff, which has an important influence on the social and economic development for the downstream areas. Medium resolution imaging spectrometer (MODIS) is widely used in dynamic remote sensing to monitor snow cover variation, because of its high temporal and spatial resolution. However, optical remote sensing snow cover is seriously affected by cloudy weather. Especially, water vapor over the plateau is low and inhomogeneous, with active convection, resulting in snow cover rapid changing. All the problems bring a challenge to the daily cloud-free snow monitoring and snow climatology mapping in the plateau. Considering the local terrain and snow cover characteristics in the Tibetan Plateau, in this study, based on the existing cloud removal algorithms, gradually controlling cloudage, the daily cloud-free MODIS snow cover algorithm has been carried out. The snow depth records from 145 ground stations for two snow seasons, from 1st October, 2009 through 30th April, 2011, are used to qualify the accuracies of various algorithm processes. The qualifying show that when snow depth is above 3 cm, the total classification accuracy of daily cloud-free snow products is 96.6%, snow classification accuracy is 83.0% and the recall rate is 89%. The new cloud removal process is proved having high precision and is suitable for daily dynamic monitoring snow cover. It provides a new way for reconstructing snow climatology data over the Tibetan Plateau. It is useful to study global environmental change, such as water, ecology, disaster and their impacts on the High Asia.

Key words: Tibetan Plateau; remote sensing of snow cover; MODIS; cloud removing algorithm; daily cloud-free snow cover products

(本文编辑: 周成林)