

引用格式:胡顺石,张辰璐,彭雨龙,等.湖南省不同云状态的时空特征分析[J].地球信息科学,2019,21(5):688-698. [Hu S S, Zhang C L, Peng Y L, et al. Spatiotemporal patterns of the different cloud states in Hunan province[J]. Journal of Geo-information Science, 2019,21(5):688-698.] DOI:10.12082/dqxxkx.2019.180588

湖南省不同云状态的时空特征分析

胡顺石^{1,2*}, 张辰璐¹, 彭雨龙¹, 谭子芳¹

1. 湖南师范大学资源与环境科学学院,长沙 410081;2. 湖南师范大学地理空间大数据挖掘与应用湖南省重点实验室,长沙 410081

Spatiotemporal Patterns of the Different Cloud States in Hunan Province

HU Shunshi^{1,2*}, ZHANG Chenlu¹, PENG Yulong¹, TAN Zifang¹

1. College of Resources and Environmental Sciences, Hunan Normal University, Changsha 410081, China; 2. Key Laboratory of Geospatial Big Data Mining and Application, Changsha 410081, China

Abstract: The presence of clouds hinders remote sensing of earth surface and causes information loss. Understanding the spatiotemporal patterns of different cloud states over a given region can improve the pertinence and accuracy of cloud noise elimination. In this study, the MODIS surface reflectance products, MOD09A1, were used to extract the cloud state information for each pixel of Hunan Province, spanning from 2001 to 2017. Then, geo-statistics and spatial analysis methods were used to detect Hunan's spatiotemporal patterns of the clear, cloudy, and mixed states. The examined patterns included the spatial distribution, seasonal distribution, gap length duration, and cloud interference for the different cloud states. The results show that: (1) Hunan Province in general was quite severely affected by clouds, with obvious spatial heterogeneity, i.e., the cloudy state when imagery are shadowed by clouds mainly occurred over the western and southern mountain areas of Hunan, while the mixed state when imagery are distorted by clouds existed over the gentle plains and hilly regions of the northern, central, and southern Hunan; (2) the cloudy state was dominant among the cloudy and mixed states, with the former more likely to occur between January and February, November and December, late May and early July, while the latter maintained 10% of cloud cover during the whole year and up to 18% from June to October; (3) cloud gap lengths of 8-days and 16-days were the primary situation for the cloudy and mixed states; (4) increasing the composite window size, cloud interference effects declined dramatically and could be neglected if monthly composite size (4 collections) was utilized; (5) by Principal Component Analysis (PCA) of the different cloudy state data, the first two PCA components were derived which indicated different spatial distribution patterns and divided Hunan into four sub regions with distinctive cloud characteristics; and (6) there are significant relationships between the different cloud states and elevation variation, and except for the cloudy state, they are all negatively correlated with elevation variation. This study can provide technical support for the selection of remote sensing data, the removal method of cloud noises, and the reconstruction of vegetation indices time series for Hunan Province.

收稿日期:2018-11-15;修回日期:2019-01-31.

基金项目:湖南省自然科学基金项目(2018JJ3348);湖南省教育厅科学研究项目(17C0952);中国国家留学基金项目(201806725009)。

[**Foundation items:** Hunan Provincial Natural Science Foundation of China, No.2018JJ3348; Research Foundation of Education Bureau of Hunan Province, China, No.17C0952; Foundation of China Scholarship Council, No.201806725009.]

作者简介:胡顺石(1984-),男,湖南永州人,讲师,博士,主要从事资源环境遥感与灾害遥感、遥感大数据应用研究。

E-mail: hufrank@163.com

Key words: cloud states; spatiotemporal patterns; Hunan Province; MODIS; remote sensing

***Corresponding author:** HU Shunshi, E-mail: huf Frank@163.com

摘要:云的存在使得遥感对地观测地表信息受损或缺失,特定区域的云状态时空分布特征有助于提升遥感数据云噪声去除的针对性和准确性。本文采用2001–2017年湖南省MOD09A1地表反射率产品像素级产品质量数据集,逐像素解析并提取云状态信息,利用地统计学方法和空间分析方法等,从云状态空间分布概率、云状态季节分布概率、云状态持续时长和云干扰率4个方面分析了湖南省晴朗无云、云污染和云混合3种云状态时空分布特征。研究结果表明:①湖南省整体受云影响较为严重,不同云状态空间分布呈现较为明显的差异性,云污染状态主要分布在湘西、湘南山区,云混合状态主要分布在湘北、湘中、湘南连接的地势相对平缓的平原、丘陵地区;②云污染状态在云影响中起主导作用,其主要分布在1–2月、11–12月以及5月下旬–7月上旬之间,云混合状态平均出现概率全年维持在10%左右,在6–10月增加至18%左右;③云状态持续时长为8、16 d是受云影响的主要情形;④随着合成窗口的增加,云干扰率迅速下降,采用月合成(4期)后,可以忽略云的影响;⑤云状态数据经主成分变换后,前2个主分量可表征不同云状态空间分布模式,全省可划分为4个具有显著云状态特征的区域;⑥不同云状态与高程变化具有显著的关系模型,除云污染状态外,均与高程呈负相关关系。本研究可为湖南省遥感数据选择、云去除方法选择、植被指数时间序列重构等提供技术支持。

关键词:云状态;时空特征;湖南省;MODIS;遥感

1 引言

光学遥感对地观测过程中易受到大气和云的影响。大气中的气体分子、气溶胶、水汽等会对太阳辐射产生吸收、反射、散射等作用,造成遥感图像质量衰减;这类问题可通过大气校正模型和方法进行恢复,如暗目标法^[1]、经验线性法^[1]、MODTRAN模型^[2]、6S模型^[3]等。而云在陆地上空广泛存在,有研究表明陆地云量约为55%,并且有明显的季节循环特征^[4];不同于大气对光学遥感的影响,云的存在会使得遥感传感器不能接收到被云遮挡的地表信息,这在一定程度上限制了光学遥感的应用。

云的时空分布特征和规律有助于人们了解天气过程和气候特征,国内外相关学者利用MODIS云产品、地面监测资料开展了一系列相关研究。King等^[4]利用MODIS云产品数据,采用数学统计分析、直方图、联合直方图分析法对全球范围云量及云物理参数等时空分布规律进行了研究,结果表明MODIS云产品可以较好地描述全球云及其相关物理特征参数;张雪芹等^[5]利用青藏高原地面监测资料,采用正交经验函数和线性倾向估计等方法,分析了该区域1971–2004年总云量的时空变化特征,结果表明该区域总云量空间分布呈自东南向西北减少的态势,且年与四季总云量变化都存在显著下降趋势;李慧晶等^[6]利用西南地区地面测站1956–2005年月平均云量资料,采用经验正交函数分析和Mann-Kendall突变检验方法,研究分析了西南地区云量的时空分布特征,结果表明整个西南地区总云

量的变化趋势一致且存在着明显的年际变化特征,总云量的分布在一定程度上受地形和区域气候的影响;Ma等^[7]利用地面监测数据和MODIS云产品对中国境内的云量分布进行了详细比较和验证,表明中国境内云量分布具有明显的时空差异性,两类数据具有较高的相关性,MODIS云产品可以用于气象领域相关研究;Ghasemifar等^[8]利用MODIS云产品数据、地面监测数据对伊朗境内云空间变化及其影响因素进行了研究,结果表明伊朗境内云量空间分布时空差异性较大,纬度位置、植被、地形是影响云量空间分布的3个重要因素。

上述研究中,不同学者通过利用地面站监测数据和遥感数据对云量空间分布进行定量分析以及年际、季节性趋势分析,多集中在云气候学相关领域,对气候变化研究具有重要意义;然而采用地面站监测数据进行分析时,会存在一定风险且不利于了解云量的综合客观特征^[7]。MODIS云产品可为大空间尺度云量、云物理参数的特征研究提供有效数据源,在一定程度上克服了地面监测数据的局限性,但是对于地表遥感监测而言,云层阻碍了遥感传感器获取地表信息,特别是对于长时间序列高时间分辨率的MODIS产品而言,云噪声的影响需要格外引起重视。深入了解云状态特征信息有助于消除云噪声影响,提升数据质量,但这方面研究相对较少。钟洪麟等^[9]在MODIS元数据基础上提出了局地云量概念,用以反映局部区域云在原始条带影像中的空间分布状况,但其仅能提供研究区域云覆盖率数据,主要服务于高质量遥感图像检索,不

能反映云状态在时间维的出现概率及其空间分布特征;彭秋志等^[10]以MOD13Q1植被指数产品中的云标记信息为数据源,利用地理信息系统空间分析方法,分析了云对东江流域2001–2015年植被指数产品质量的干扰情况,但对16 d内合成得到的植被指数产品而言,其整体云干扰程度已有所减轻,对于综合反映区域性不同云状态出现概率的时空特征具有一定的限制性。

目前,利用遥感长时间观测序列数据进行区域资源环境监测成为研究热点。长时间序列遥感数据易受云的影响,需要采用一定的模型和方法去除云的影响,提升遥感数据质量及数据可用性,获得能反映地表正确信息的遥感数据,这类方法主要包括合成法和数据平滑与拟合方法。合成法以最大值合成法MVC^[11]为代表,这类方法可以很大程度上减轻云带来的影响,但是它降低了数据的时间分辨率;数据平滑与拟合方法也可以减轻云的影响,如Savitzky-Golay滤波方法^[12]、时间序列谐波分析法^[13]、非对称高斯模型^[14]、双逻辑斯蒂法^[14]等,但这些方法对于中国多阴雨季节、连续多期云干扰区域,可能导致云影响误判^[10]。

由此可见,在进行长时间序列对地观测以获取地表变化信息时,云作为时空多变要素,其影响不可忽视。对于单景影像而言,可以通过光谱及空间关系等多特征耦合检测单景影像中的薄云、厚云及云阴影^[15],从而采用相应方法予以消除;对长时间序列遥感影像而言,可以采用不同的云噪声去除方法,重构长时间序列产品,提升产品质量^[16]。但是云在不同区域、不同季节存在一定的时空差异性,而采用统一的算法模型及参数难免会造成误差^[10]。对于特定研究区域,研究人员首先应该根据研究区域的不同云状态影响特点,选用合适的云去除方法,并调整数据处理流程与模型相关参数,以提高数据质量和可靠性。最新的MODIS C6产品提供了高质量的像素级质量控制状态参数,可以用于不同云状态时空分布特征研究。

本研究在此基础上,选择2001–2017年湖南省MODIS反射率产品MOD09A1中的云状态信息为数据源,通过统计分析、空间分析方法等从云的空间分布特征、云季节分布特征、云持续干扰时长等方面,探究分析云对该区域遥感对地观测的影响,较为全面地揭示该区域不同云状态时空分布特征及规律,为该区域遥感对地观测数据选择、云去除方法选择、植被指数时间序列重构等提供技术支撑。

2 研究区概况、数据源及研究方法

2.1 研究区概况

湖南省位于24°38'N–30°08'N, 108°47'E–114°15'E之间,国土面积为21.18万km²,辖14个地州市,如图1所示。湖南省东、南、西三面山地环绕,西北有武陵山脉,西南有雪峰山脉,南部为南岭山脉,东面为湘赣交界的罗霄山脉,中部和北部地势低平。湖南省丘陵、盆地、平原地貌呈马蹄形分布,整体地形地貌较为复杂。湖南省亚热带季风气候明显,气候温暖湿润,阴雨季节持续时间较长,云量较多。

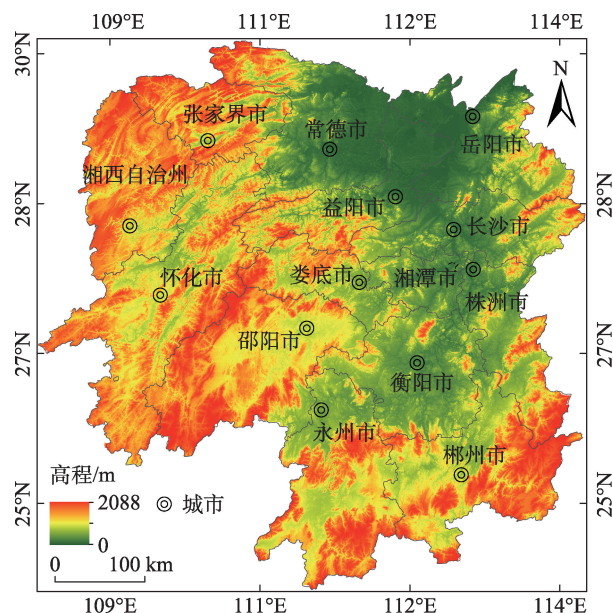


图1 湖南省行政区划及高程分布

Fig. 1 Administrative division and elevation map of Hunan Province

2.2 数据源及数据预处理

MODIS提供全球范围内辐射平衡、生态系统、土地覆盖特征3大类近157种产品供用户免费使用^[17],采用HDF格式、以瓦片形式和正弦投影方式进行存储。MODIS C6版产品不但消除了传感器性能衰减的影响^[18],还在一些产品集成算法等方面进行了改进^[19]。MOD09A1产品^[20]是8 d合成地表反射率产品,一年46期,空间分辨率为500 m,除第1–7波段地表反射率产品外,还提供了相应的质量控制参数。本研究采用MOD09A1产品中的sur_refl_state_500 m数据集进行云量信息提取和分析。

覆盖整个湖南省区域的MOD09A1产品标识号为h27v05、h27v06、h28v06,但是h27v05仅包含湖

南省北部极少部分区域,为了减少数据预处理工作量,本研究仅选取了h27v06、h28v06两景影像。借助于DAAC2Disk工具(<https://lpdaac.usgs.gov/tools/daac2diskscripts/>), Earthdata 注册用户可以进行MODIS产品批量下载。本研究下载了2001–2017年间h27v06、h28v06共1564景原始影像。

除此外,本文还使用了GADM提供的湖南省行政边界矢量数据(<https://gadm.org/>),用于MOD09A1数据产品掩膜,提取研究区数据。

下载后的MOD09A1反射率产品需要经过重投影、镶嵌、裁剪等工作,才能进行后续处理与分析。MODIS重投影工具MRT可以完成重投影和镶嵌工作。在使用MRT进行重投影和镶嵌时,设置需要输出数据集为sur_refl_state_500 m,输出文件格式为tif,投影方式为WGS84地理坐标。需要注意的是,用户需要将数据输出格式设置为tif格式,才能获得地理位置正确的输出结果,否则输出结果会存在一定地理位置偏差。由于MRT工具不能实现影像的不规则裁剪,本文在IDL 8.5平台上编写批处理程序,利用湖南省行政边界矢量数据对重投影后的镶嵌结果进行不规则裁剪,并输出为ENVI标准数据格式,方便数据后续批量处理和读取。

sur_refl_state_500 m数据集是16位无符号整型数据,其对应的二进制位描述了每个像素的云、阴影、雪等11种状态^[20]。二进制位编码的最右端两位表示不同的云状态,具体含义^[19]为:00表示晴朗无云干净像元;01表示云污染像元;10表示云混合像元;11表示没有设置,假定为无云像元。将该数据集转换成相应二进制编码,取最右端2位,进行云状态解析,提取每个像素的晴朗无云状态(00)、云污染状态(01)、云混合状态(10)共3类数据集。其中,晴朗无云状态表示不受云影响,其对应产品进行了大气校正,像素值真实有效且质量很高,用户可放心使用;云污染状态表示该像元受云影响严重,被云完全遮挡或覆盖,其对应产品的像素值无效;云混合状态表示像素受薄云、气溶胶、大气等影响,尽管对其进行了大气校正,但其像元值的质量可靠性及可用性还需要参考其他质量控制参数进行综合判断。由于假定无云状态(11)在影像中的比例极小,故而并未输出此状态。此项工作在MATLAB 2017a中编码实现,考虑到计算机性能限制,本研究采用逻辑值表示不同云状态,以减轻存储压力,1表示为对应状态为真,0表示对应状态为假。

2.3 研究方法

不同云状态对数据质量及数据可靠性具有不同的影响,图2为某像素利用MOD09A1地表反射率计算得到的NDVI序列(图中未标明的序列为晴朗无云状态对应的值),可以发现云污染和云混合状态一般对应于NDVI值中异常离群值,需要采用相关方法予以校正,以提高数据的质量和可用性。

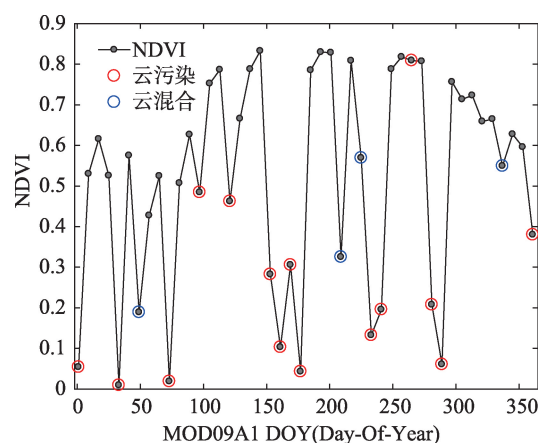


图2 不同云状态对NDVI时间序列的影响示例

Fig. 2 An example of the effects of different cloud states on NDVI time series

本文定义了如下概念用于分析不同云状态的时空分布特征和规律:

(1) 云状态空间分布概率。将2001–2017年间每一像素分别出现晴朗无云、云污染、云混合3种状态的概率定义为云状态空间分布概率。分别统计每个像素的不同云状态的出现概率,获得湖南省不同云状态出现概率空间分布图,分析不同云状态的空间分布特征,揭示云状态空间分布规律。

(2) 云状态季节分布概率。通过计算湖南省2001–2017年同一时期内的不同云状态平均出现概率,即可获得一年内不同季节不同云状态的变化规律。尽管MOD09A1产品采用8 d合后可在一定程度上消除大气、云等带来的噪声,但是受天气、气候等多种因素影响,8 d合成产品依然受云影响严重。为了揭示不同云状态在一年之内不同季节的变化趋势,本文对云状态平均出现概率进行Savitzky-Golay滤波,以获得其季节变化趋势和特征。Savitzky-Golay滤波在消除噪声、获取长时间序列的变化趋势以及局部突变信息具有较好的效果^[12,21]。平滑窗口大小和平滑多项式次数是Savitzky-Golay滤波过程中需要确定的2个重要参数。本研究经多次试验后,将平滑窗口大小设置为9,多项式次数设

置为2,滤波结果可以较好地反映不同云状态的季节变化趋势。

(3) 云干扰持续时长。将同一像素连续出现相同云状态的期数称为云干扰持续时长。如果某区域云干扰持续时长过大、数量过多,则表明该区域受云连续干扰严重。获取云持续干扰时长的过程如下:首先,挑选云状态为真的值,记录其对应的索引号;然后对索引号进行逐个甄别,分别记录连续云状态的数量、起始索引号;最后,通过起始索引号计算云状态持续时长。

(4) 云干扰率。在设定的合成时长内(滑动窗口),将连续有云次数与窗口滑动次数之比定义为某像素在该合成时段内的云干扰率^[10]。其操作过程是采用一个长度为 m 的卷积核对长度为 n 的云状态时间序列进行卷积操作,在滑动过程中,如果在卷积核窗口内均为相同云状态,则定义为一次连续有云状态,共卷积操作 $n-m+1$ 次,最终计算连续有云状态占卷积操作的比值。云干扰率在在一定程度上可以指示数据合成后的数据质量及其可靠性。

采用主成分分析法(PCA)对2001–2017年湖南省不同云状态平均出现概率数据进行变换分析,得到PCA各主分量,然后通过直方图分析法、空间叠置分析法对不同PCA主分量进行分析,探索不同云状态之间的关系。最后,以湖南省高程数据进行区间划分,对不同云状态数据进行重分类,计算重分类后的不同云状态平均出现概率和平均持续时长,对其进行回归建模分析,探索不同云状态与高程之间的关系。

3 结果及分析

3.1 云状态空间分布概率特征

通过对2001–2017年共782期MOD09A1产品云状态进行解析,分别获得了晴朗无云、云污染和云混合状态出现概率空间分布图(图3)。

从图3可以看出,出现晴朗无云状态概率较大的区域主要分布在湘北洞庭湖平原区,包括岳阳、常德、怀化东部、益阳西部及娄底北部区域;云污染和云混合状态在空间分布上呈现差异性,云污染状态在常德、岳阳、益阳等湘北区域出现概率相对较低,湘中低山丘陵、平原区域次之,而湘西(湘西州和怀化)、湘南山区(永州)地形复杂、海拔较高、森林覆盖率高,受天气条件、蒸散发影响等,更容易受到云的影响,云污染出现概率较高,这与文献^[8]提出的良好的植被条件容易产生云研究结论一致;云混合状态主要出现在洞庭湖中部区域、长沙、湘潭、衡阳组成的湘北、湘中、湘南连接的地势相对平缓的平原、丘陵地区,而这些区域既是湖南省主要人口聚集区,同时也是农业、工业等生产活动主要聚集区域;湘西、湘东区域受云混合状态影响较低。

图4为全省晴朗无云、云污染及云混合状态平均出现概率及方差图。2001–2017年晴朗无云状态的平均出现概率为57.32%,方差为5.70%;云污染状态的平均出现概率为27.59%,方差为5.08%;云混合状态的平均出现概率为12.26%,方差为3.43%。这说明全省范围内出现晴朗无云状态的概率不高,全省受云影响较大,云污染和云混合状态

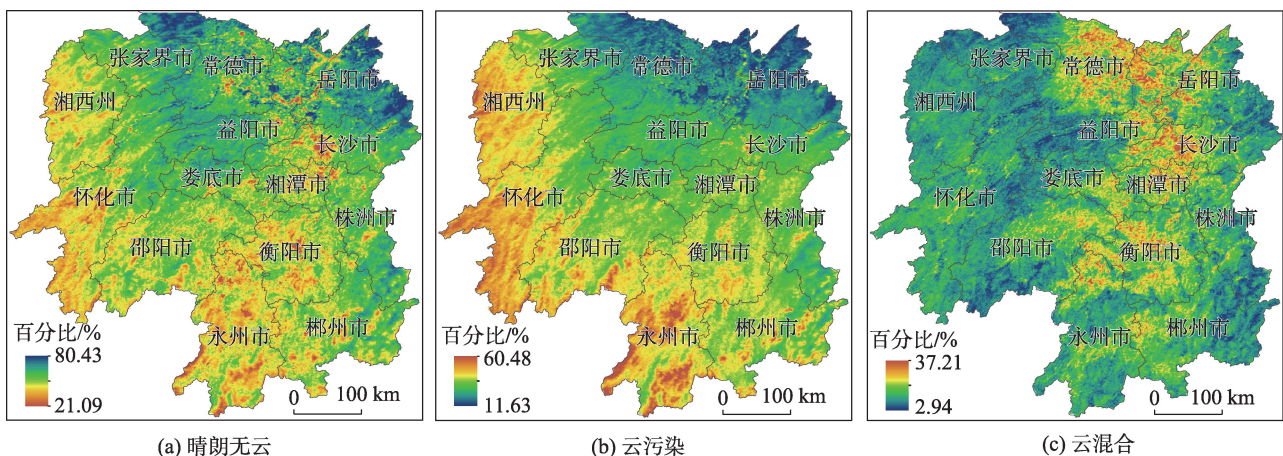


图3 2001–2017年湖南省晴朗无云、云污染及云混合状态平均出现概率空间分布
Fig. 3 Spatial distribution of average occurrence probability of cloud free, cloud shadowed, and cloud distorted states in Hunan Province from 2001 to 2017

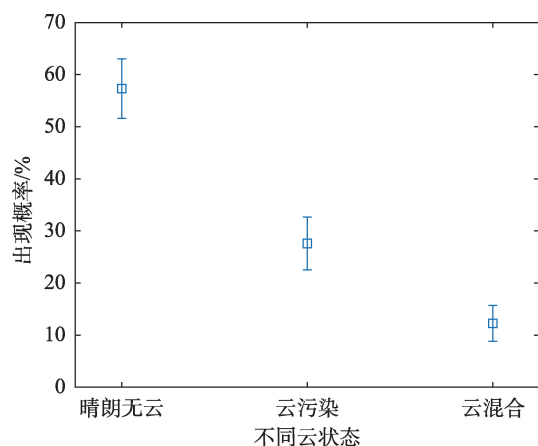


图4 2001–2017年湖南省晴朗无云、云污染及云混合状态平均出现概率及方差

Fig. 4 Average occurrence probability and variances of cloud free, cloud shadowed, and cloud distorted states in Hunan Province from 2001 to 2017

是重要影响因素,云污染状态是主导因素;在两者共同作用下,全省一年期内平均约40%的时段数据是受云影响的,图2中的云污染和云混合状态数量也说明了此现象。

3.2 云状态季节分布率特征

湖南省2001–2017年不同云状态在各月份的平均出现概率如图5所示。图5中的不同颜色的虚线表示不同云状态的分布概率,绿色粗实线表示经Savitzky-Golay滤波处理后的相对应云状态分布趋势。

从图5可以看出,全省晴朗无云状态在4月–5月上旬、7月下旬–8月初、9月中旬、11月下旬出现概率较高,约为60%~70%;全省1–2月出现晴朗无云状态概率较低,并与云污染状态出现概率相当;在6月下旬、8月中下旬、10月中下旬全省出现晴朗

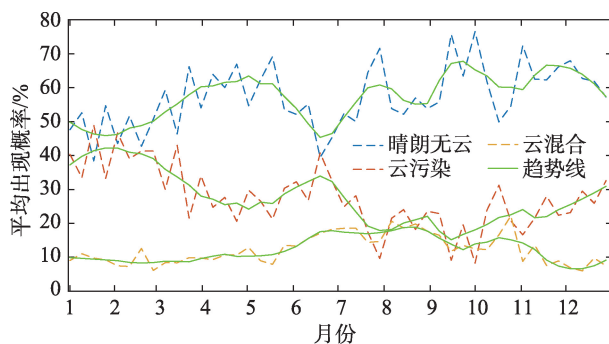


图5 2001–2017年湖南省不同云状态在各月份平均出现概率

Fig. 5 Average occurrence probability of different cloud states in different months in Hunan Province from 2001 to 2017

无云状态的概率相对邻近月份会出现低谷;全省云污染和云混合状态出现概率之和与晴朗无云状态出现概率呈相反的趋势,这说明云污染和云混合状态是影响全省数据质量的重要因素,而云污染状态又起着主导作用。云污染状态主要分布在1–2月、11–12月之间的春、冬季阴雨季节、5月下旬–7月上旬之间的雨季,在9月中旬以后全省云污染状态开始趋于上升趋势;云混合状态全年基本维持在10%~20%水平之间,但在6–10月之间增加至18%左右水平,这可能与夏秋季节天气状况、人类活动等因素有关,从11月开始,云混合状态的影响又递减至10%水平。

3.3 云状态持续时长特征

云持续时长以期数为单位进行表征,8 d为1期。本文分别统计全省云污染和云混合状态持续时长为1–5期平均持续时长和方差,即连续8、16、24、32、40 d,结果如图6所示。从图6可以看出,无论是云污染还是云混合状态,持续时长为1、2期受云影响是主要情形,其累计百分比超过90%;云持续时长为1、2期的云污染状态比率分别为68.67%、20.90%,而云混合状态分别为85.42%、12.07%。这说明云混合状态在时间维上分布不连续,而云污染更容易出现连续2期云污染影响现象。

全省不同云状态平均持续时长空间分布如图7所示。从图7可以看出,云污染与云混合状态平均持续时长空间分上表现出不同的特征。全省云污染平均持续时长在空间上可以分成4个区域,即湘

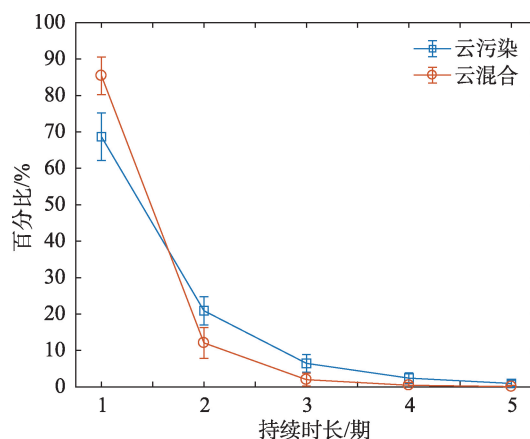


图6 2001–2017年湖南省云污染、云混合状态平均持续时长及方差

Fig. 6 Average gap length duration and variance of cloud shadowed and cloud distorted states in Hunan Province from 2001 to 2017

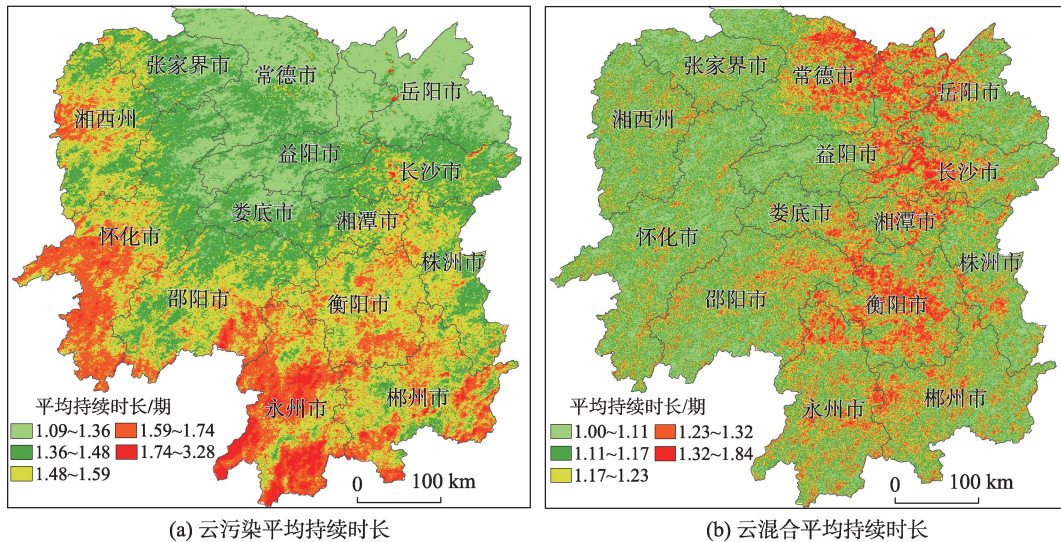


图7 2001–2017年湖南省云污染、云混合状态平均持续时长空间分布

Fig. 7 Average gap length duration spatial distribution of cloud shadowed and cloud distorted states in Hunan Province from 2001 to 2017

北平原区域(平均持续时长为1.09~1.36期),以益阳、常德、怀化三市交界处为中心呈环状分布区域(平均持续时长为1.36~1.48期),以湘西山区、湘南山区为主的高云污染持续时长区域(平均持续时长为1.59~3.28期),湘中丘陵、平原交叠区域(平均持续时长为1.48~1.74期);云混合平均持续时长在空间分布相对较为集中,主要分布在洞庭湖区域,长株潭城市群区域,以及衡阳、娄底、邵阳、永州城镇较为集中的区域,其平均持续时长为1.23~1.84期;而其他区域的平均持续时长相对偏低,为1~1.17期,云混合状态在这些区域属于非长时间持续性影响。

3.4 云干扰率及其空间分布

2001–2017年湖南省不同云状态的云干扰率随不同合成窗口(即合成期数)变化情况如图8所示,随着合成窗口增加,云干扰率迅速下降;当合成窗口为1期时,云污染、云混合状态造成的云干扰率分别为27.59%、12.26%,这意味着约40%的数据受云影响;通过2期合成后,云污染和云混合的干扰率已分别下降至9.12%、1.94%;当合成窗口期数为4期时,全省云污染和云混合状态的云平均干扰率已经分别下降至1.26%、0.08%,即如果是月合成产品,则可以忽略云的影响。

云污染是影响数据质量的主要因素,当采用不同窗口(1、2、3期)进行合成时,云干扰率空间分布如图9所示。当合成期窗口为1期时,也即是原始

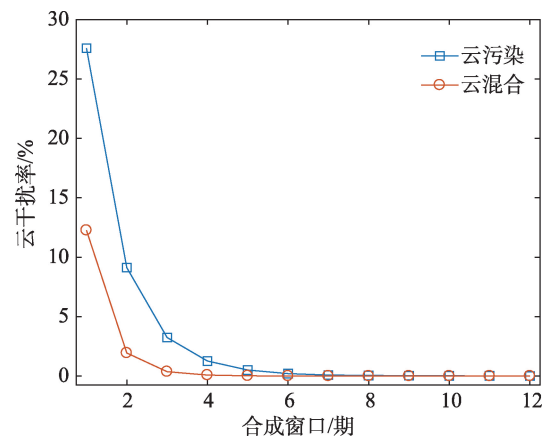


图8 不同合成窗口下云污染和云混合状态的整体平均云干扰率

Fig. 8 Overall cloudy interference ratio of cloud shadowed and cloud distorted states by using different composite window sizes

数据,云污染状态的干扰率特征与云污染的空间分布概率一致;当采用2期合成后,云干扰率空间分布呈现较为明显的空间差异性,其中,湘北洞庭湖区域云的干扰率已经下降至5%,湘中区域的干扰率在5%~10%之间,而湘西山区、湘南山区以及湘中衡阳、长株湘城市群区域的云干扰率在10%~15%之间,部分区域略高;而当采用3期合成后,全省绝大多数区域的云干扰率已经下降至5%;只有湘西山区、湘南部分区域及永州境内的云干扰率为5%~10%。因此,通过3期合后,云对数据质量的影响可以在很大程度上予以消除和减轻。

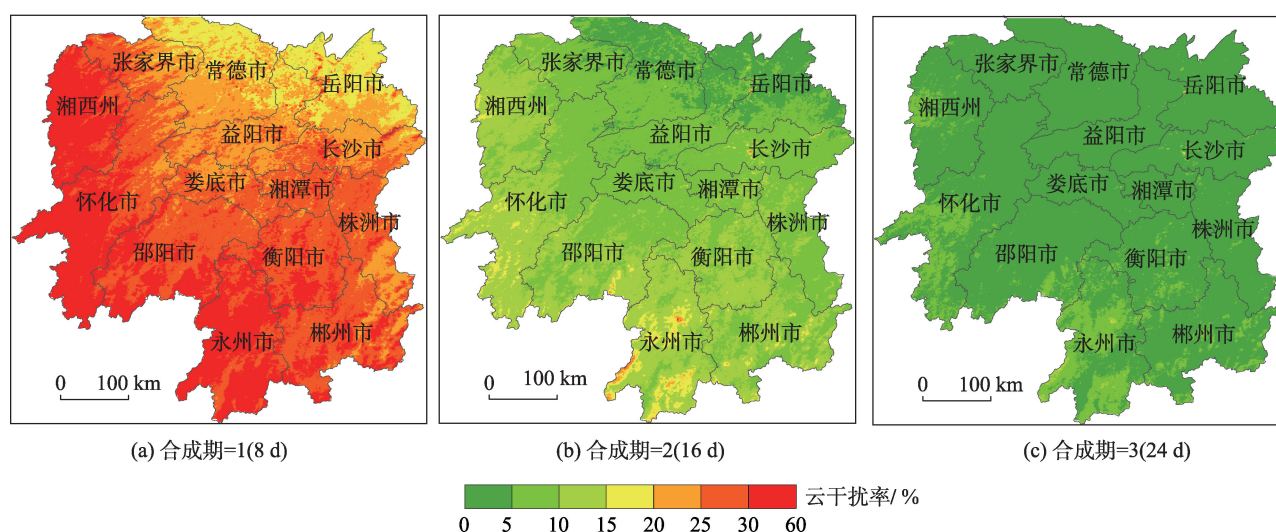


图9 2001–2017年湖南省不同合成期平均云干扰率空间分布

Fig. 9 Cloudy interference ratio spatial distribution for different composite window sizes in Hunan Province from 2001 to 2017

3.5 不同云状态及其与高程关系

对2001–2017年湖南省晴朗无云、云污染和云混合状态进行主成分变换,得到PCA1-3分量,各分量对应的特征值分别为52.83、15.67和1.34;根据PCA原理可知,前2个分量所占的信息量之和为98.08%,几乎包含所有云状态信息;PCA1分量的信息量为75.64%,PCA2分量的信息量为22.44%;PCA1分量反映的是晴朗无云和云污染状态综合空间分布模式,PCA2分量反映的是云混合状态空间分布模式。

图10(a)为PCA1分量直方图,以零值为界线,将PCA1分量划分成“正负”2个区域,其对应空间范围如图10(d)所示。计算不同云状态的均值,结果显示:对于PCA1分量为负值的区域,晴朗无云状态的平均出现概率为62.00%,高于全省平均值(57.32%),云污染状态的平均概率为23.60%,低于全省平均值(27.59%);对于PCA1分量为正值的区域,晴朗无云状态的平均出现概率为53.09%,低于全省平均值(57.32%),云污染状态的平均概率为31.18%,高于全省平均值(27.59%)。因此将PCA1分量为负值的区域称为“高晴朗无云区”,占全省面积47.52%,将PCA1分量为正值的区域称为“高云污染区”,占全省面积52.48%。采用同样的方法对PCA2分量进行分析,结果如图10(b)、(e)所示,PCA2分量为负值的区域云混合状态的平均出现概率为10.07%,低于全省平均值(12.26%);正值区域的云混合状态平均概率为15.48%,高于全省平均值

(12.26%)。因此,将PCA2分量负值的区域称为“低云混合区”,正值区域称为“高云混合区”。

对PCA1和PCA2两分量空间分布模式进行空间叠置分析,结果如图10(f)所示,叠置结果分成四个区域,每个区域3种不同云状态平均出现概率如图10(c)所示。综合图10(c)、(f)可以看出,这四个区域体现出不同的云状态特征:Ⅰ区、Ⅲ区出现晴朗无云状态的平均概率较高(分别为62.55%、61.20%),超过全省平均值(57.32%),云污染状态出现平均概率较低(分别为24.86%、21.80%),低于全省平均值(27.59%),但Ⅲ区较Ⅰ区更容易出现云混合状态,Ⅲ区为14.73%,Ⅰ区为9.44%;Ⅱ区、Ⅳ区更容易出现云污染状态(分别为32.57%、29.10%),超过全省平均值(27.59%),其出现晴朗无云状态平均概率较Ⅰ区、Ⅲ区偏低,且Ⅳ区比Ⅱ区更容易出现云混合状态影响。总体上来说,Ⅰ区更容易出现晴朗无云状态,Ⅱ区更容易出现云污染状态,Ⅲ区出现晴朗无云状态概率较高,但其受云混合影响较大,Ⅳ区易同时受云污染、云混合影响,晴朗无云状态平均出现概率最低。

以100 m高程间距对湖南省高程进行区间划分,对每个区间内不同云状态的值进行重分类,获取其平均值;再利用不同回归模型分别对其与高程之间的关系进行回归建模分析,结果如图11所示。从图11(a)–11(e)可以看出,晴朗无云、云污染和云混合3种云状态平均出现概率与高程变化分别呈线性负相关、非线性正相关(对数关系)和非线性反比

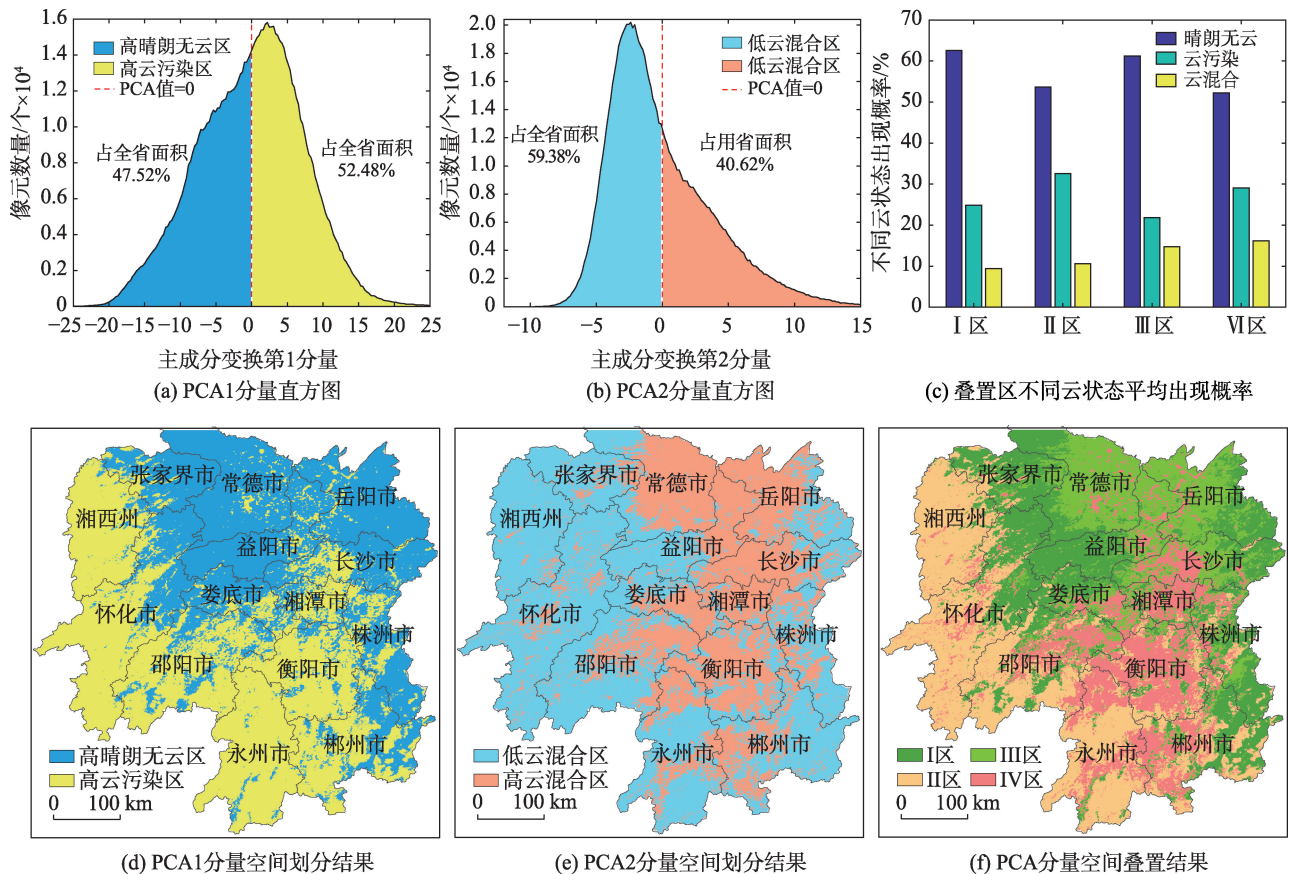


图10 2001–2017年湖南省不同云状态第1、2主分量直方图及空间分布模式

Fig. 10 Histograms and spatial distribution patterns of the first and second principal components for different cloud states in Hunan Province from 2001 to 2017

例关系；云污染、云混合平均持续时长与地表高程变化分别呈较强的对数关系和弱反比例关系；所有模型 R^2 值较大（云混合平均持续时长除外），模型假设检验 P 值几乎为0，模型通过显著性 $\alpha=0.01$ 检验，具有统计意义；当地表高程超过1200 m后，3种云状态平均出现概率在一定范围内波动，晴朗无云状态平均出现概率为 $54\%\pm 1\%$ ，云污染状态平均出现概率为 $32\%\sim 33\%$ ，云混合状态平均出现概率为 $10\%\sim 11\%$ 。

从全省范围来看，不同云状态平均出现概率、持续时长与高程的相关性如表1所示。可以看出，不同云状态与高程变化有不同程度的弱相关关系，相关系数较小，但检验 P 值为0，显著性极高。

4 结论与讨论

本文利用MOD09A1产品中的像素级质量状态标记数据集，解析并获取了湖南省2001–2017年云状态信息，从云状态空间分布概率、季节分布概率、云状态持续时长以及云干扰率4个方面分析了湖南省

不同云状态时空分布特征，得出了如下结论：

(1) 湖南省受云影响较为严重，不同云状态空间分布呈现差异性，云污染和云混合状态是云影响的主要因素。全省云污染状态的平均出现概率为27.59%，云混合状态为12.26%；云污染状态主要分布在湘西山区、湘南山区，而云混合状态主要分布在湘北、中、南连接的主要地势低缓的平原、丘陵地区。

(2) 不同云状态的季节分布概率呈现差异性。云污染状态在云影响中起着主导作用，云污染状态主要分布在1–2月、11–12月之间以及5月下旬–7月上旬之间；云混合状态平均出现概率全年维持在10%左右，但在6–10月增加至18%左右。

(3) 不同云状态的平均持续时长空间分布呈现差异性，无论是云污染还是云混合状态，持续时长为1、2期（即连续8、16 d内）受云影响是主要情形，其占所有受云影响的累计比率超过90%。

(4) 云污染和云混合状态的云干扰率随着合成期数增加，云干扰率迅速下降，如果采用月合成，则云干扰的影响已经很微弱了；且不同区域体现了差异性。

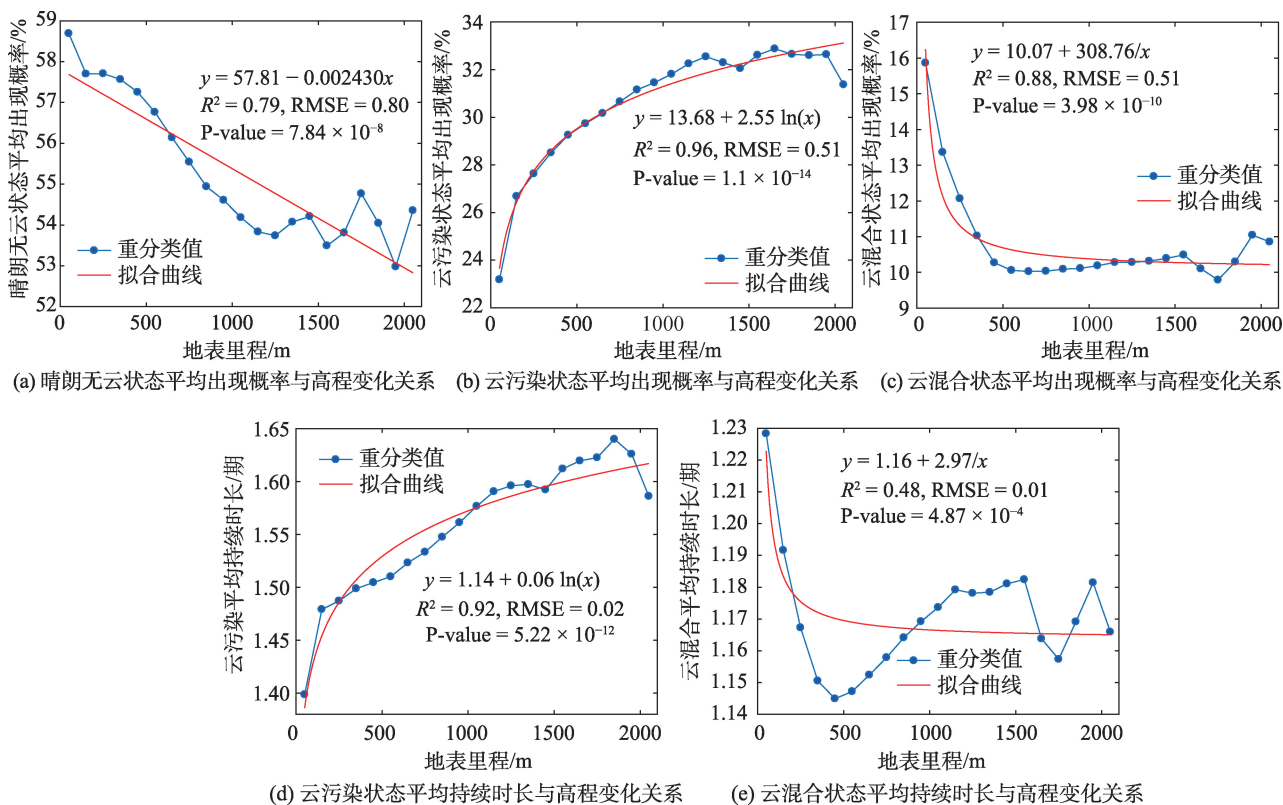


图 11 2001–2017年湖南省不同云状态平均出现概率、平均持续时长与高程变化之间的关系

Fig. 11 Relationships between average occurrence probabilities, average durations and elevation variations for different cloud states in Hunan Province from 2001 to 2017

表 1 2001–2017年湖南省不同云状态平均出现概率、持续时长与高程的相关系数

Tab. 1 Correlation coefficients between average occurrence probability, duration and elevation variations for different cloud states in Hunan Province from 2001 to 2017

云状态出现概率/持续时长	相关系数	P-value	$\alpha=0.01$ 置信区间
晴朗无云平均出现概率	-0.2026	0.000	[-0.2051 -0.2001]
云污染平均出现概率	0.4700	0.000	[0.4679 0.4720]
云混合平均出现概率	-0.5024	0.000	[-0.5043 -0.5004]
云污染平均持续时长	0.3346	0.000	[0.3322 0.3369]
云混合平均持续时长	-0.2479	0.000	[-0.2504 -0.2454]

(5) 对不同云状态数据进行主成分变换后,得到的前2个PCA分量占总信息量的98.08%,几乎包含所有云状态信息;PCA1分量反映的是晴朗无云和云污染状态综合空间分布模式;PCA2分量反映的是云混合状态空间分布模式。通过直方图分析和空间叠置分析,可以将湖南省云状态影响分成四个区域:Ⅰ区为易出现晴朗无云状态,Ⅱ区为易出现云污染状态,Ⅲ区易出现晴朗无云状态,但其受

云混合影响较大,Ⅳ区易同时受云污染、云混合影响,晴朗无云状态平均出现概率偏低。

(6) 以100 m高程间距对不同云状态数据进行重分类,然后进行回归建模分析,结果显示:晴朗无云、云污染和云混合3种云状态平均出现概率与高程变化分别呈线性负相关、非线性正相关(对数关系)和非线性反比例关系;云污染、云混合平均持续时长与地表高程变化分别呈较强的对数关系和弱反比例关系;所有模型相关系数较高,均通过显著性检验。

综上所述,在使用长时间序列遥感产品时,需要考虑到不同云状态的时空分布特征和规律,并采取相应策略和方法对云噪声进行处理。8 d合成的MOD09A1反射率产品中的云状态标记信息是合成期内某一天对应的云状态,并不能完全反映这8 d的云状态,但其合成后的地表反射率产品是8 d内获得的最佳遥感观测值,这对光学遥感监测具有一定的指示意义,同时也能在一定程度上揭示湖南省不同云状态的时空分布特征和规律。如果是研究云气候学或区域性云量时空分布特征,MODIS云产品则是最佳选择^[7]。另外,云的出现受天气、气候

影响较为严重,随机性也较大,采用17年8 d合成的MOD09A1产品进行研究,可在一定程度上反映全省不同云状态基本情况,但是17年的云状态标记数据时间尺度较短,且MODIS仅对云光学厚度大于等于0.4的云层具有较高的探测精度^[4],这也会对云观测精度带来一定影响。

参考文献(References):

- [1] 童庆禧,张兵,郑兰芬.高光谱遥感——原理、技术与应用[M].北京:高等教育出版社,2006:83-86. [Tong Q X, Zhang B, Zheng L F. Hyperspectral remote sensing: Principle, technology and applications[M]. Beijing: Higher Education Press, 2006:83-86.]
- [2] Berk A, Conforti P, Kennett R, et al. MODTRAN6: A major upgrade of the MODTRAN radiative transfer code[J]. 2014 6th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, 2014:1-4.
- [3] Vermote E F, Tanre D, Deuze J L, et al. Second simulation of the satellite signal in the solar spectrum, 6s: An overview[J]. Geoscience & Remote Sensing IEEE Transactions on, 2002,35(3):675-686.
- [4] King M D, Platnick S, Menzel W P, et al. Spatial and temporal distribution of clouds observed by MODIS onboard the terra and aqua satellites[J]. IEEE Transactions on Geoscience and Remote Sensing, 2013,51(7):3826-3852.
- [5] 张雪芹,彭莉莉,郑度,等.1971-2004年青藏高原总云量时空变化及其影响因子[J].地理学报,2007,62(9):959-969. [Zhang X Q, Peng L L, Zheng D, et al. Cloudiness variations over the Qinghai-Tibet Plateau during 1971-2004 [J]. Journal of Geographical Science, 2007,62(9):959-969.]
- [6] 李慧晶,刘建西,刘东升,等.西南地区云量变化特征[J].干旱气象,2014,32(2):194-200,219. [Li H J, Liu J X, Liu D S, et al. Variation characteristics of cloud cover over Southwestern China[J]. Journal of Arid Meteorology, 2014,32(2):194-200,219.]
- [7] Ma J J, Wu H, Wang C, et al. Multiyear satellite and surface observations of cloud fraction over China[J]. Journal of Geophysical Research: Atmospheres, 2014,119(12): 7655-7666.
- [8] Ghasemifar E, Farajzadeh M, Perry M C, et al. Analysis of spatiotemporal variations of cloud fraction based on geographic characteristics over Iran[J]. Theoretical and Applied Climatology, 2018,134(3-4):1429-1445.
- [9] 钟洪麟,施润和,曲培青,等.MODIS影像的局地云量信息元数据提取算法与应用[J].地球信息科学学报,2010,12(4):587-592. [Zhong H L, Shi R H, Qu P Q, Zhang H F, et al. The regional cloud-cover metadata extraction based on MODIS image[J]. Journal of Geo-information Science, 2010,12(4):587-592.]
- [10] 彭秋志,秦国玲,吕乐婷,等.基于MODIS数据的东江流域云干扰时空特征分析[J].国土资源遥感,2018,30(1):109-115. [Peng Q Z, Qin G L, Lv L T, et al. Spatio-temporal patterns of cloud interference in Dongjiang River basin from MODIS data[J]. Remote Sensing for Land & Resources, 2018,30(1):109-115.]
- [11] Huete A, Didan K, Miura T, et al. Overview of the radiometric and biophysical performance of the MODIS vegetation indices[J]. Remote Sensing of Environment, 2002, 83(1):195-213.
- [12] Chen J, Jönsson P, Tamura M, et al. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter[J]. Remote Sensing of Environment, 2004,91(3):332-344.
- [13] Petitjean F, Ketterlin A, Gançarski P. A global averaging method for dynamic time warping, with applications to clustering[J]. Pattern Recognition, 2011,44(3):678-693.
- [14] Jönsson P, Eklundh L. Seasonality extraction by function fitting to time-series of satellite sensor data[J]. Geoscience & Remote Sensing IEEE Transactions on, 2002,40(8):1824-1832.
- [15] 沈金祥,季璇.遥感影像云及云影多特征协同检测方法[J].地球信息科学学报,2016,18(5):599-605. [Shen J X, Ji X. Cloud and cloud shadow multi-feature collaborative detection from remote sensing image[J]. Journal of Geo-information Science, 2016,18(5):599-605.]
- [16] 周惠慧,王楠,黄瑶,等.不同时间间隔下的遥感时间序列重构模型比较分析[J].地球信息科学学报,2016,18(10): 1410-1417. [Zhou H H, Wang N, Huang Y, et al. Comparison and analysis of remotely sensed time series of reconstruction models at various intervals[J]. Journal of Geo-information Science, 2016,18(10):1410-1417.]
- [17] LP DAAC. MODIS products table[OL]: https://lpdaac.usgs.gov/product_search/?view=list&collection_version_id=6&collections=Aqua+MODIS&collections=Combined+MODIS&collections=Terra+MODIS.
- [18] Lyapustin A, Wang Y, Xiong X, et al. Scientific impact of MODIS C5 calibration degradation and C6 + improvements[J]. Atmospheric Measurement Techniques, 2014,7(12):4353-4365.
- [19] Roger J C, Vermote E F, Ray J P. MODIS surface reflectance user's guide, Collection 6[R]. 2015.
- [20] Vermote E F. MOD09A1 MODIS/Terra surface reflectance 8-day 13 global 500m sin grid V006 [DB]. 2015. doi:10.5067/MODIS/MOD09A1.006
- [21] 边金虎,李爱农,宋孟强,等. MODIS植被指数时间序列 Savitzky-Golay 滤波算法重构[J].遥感学报,2010,14(4): 725-741. [Bian J H, Li A N, Song M Q, et al. Reconstruction of NDVI time-series datasets of MODIS based on Savitzky-Golay filter[J]. Journal of Remote Sensing, 2010,14(4):725-741.]