

文章编号: 1672-8785(2021)05-0039-06

风云四号红外高光谱 GIIRS 中波 通道亮温偏差订正

王 根^{1,2} 陈 娇¹ 戴 娟³ 王 悦¹

(1. 安徽省气象台 大气科学与卫星遥感安徽省重点实验室, 安徽 合肥 230031;

2. 中亚大气科学研究中心, 新疆 乌鲁木齐 830002;

3. 安徽省气候中心, 安徽 合肥 230031)

摘 要: 变分同化风云四号干涉式大气垂直探测仪(Geostationary Interferometric Infrared Sounder, GIIRS)中波通道亮温偏差要求满足高斯分布, 因此需进行 GIIRS 资料偏差订正。在 Harris B A 等人提出的“离线”法的基础上, 发展了基于随机森林(Random Forest, RF)的 GIIRS 偏差订正方法。在具体执行过程中, 基于风云四号多通道扫描成像辐射计(Advanced Geosynchronous Radiation Imager, AGRI)云产品对 GIIRS 资料进行了云检测。试验结果表明, 经过偏差订正的 GIIRS 亮温偏差满足高斯分布的假定。与“离线”法相比, RF 法的订正效果更好。

关键词: 高光谱 GIIRS; 偏差订正; “离线”法; 随机森林; 云检测

中图分类号: P407 文献标志码: A DOI: 10.3969/j.issn.1672-8785.2021.05.007

Bias Correction of Brightness Temperatures in Medium Wave Channel of FY-4A Infrared Hyperspectral GIIRS

WANG Gen^{1,2}, CHEN Jiao¹, DAI Juan³, WANG Yue¹

(1. Anhui Key Lab of Atmospheric Science and Satellite Remote Sensing, Anhui Meteorological

Observatory, Hefei 230031, China; 2. Center of Central Asia Atmospheric Science

Research, Urumqi 830002, China; 3. Anhui Climate Center, Hefei 230031, China)

Abstract: The brightness temperature bias of the medium wave channel of the variational assimilation geostationary interferometric infrared sounder (GIIRS) of FY-4 is required to meet the Gaussian distribution, so the bias correction of GIIRS data is necessary. Based on Harris B A and Kelly G's "off-line" method, a method for GIIRS bias correction based on the random forest is developed in this paper. In the specific implementation process, the cloud detection of GIIRS data is carried out based on the advanced geosynchronous radiation imager (AGRI) cloud products of FY-4. The experimental results show that the brightness temperature bias of GIIRS satisfies the assumption of Gaussian distribution after the bias correction. Compared with "off-line" method, random forest method has a better correction effect.

收稿日期: 2021-01-07

基金项目: 国家自然科学基金项目(41805080); 中亚大气科学研究基金项目(CAAS202003); 安徽省气象台自立项目(AHMO202007; AHMO202004)

作者简介: 王根(1983-), 男, 江苏泰州人, 高级工程师, 博士, 主要从事卫星资料同化、正则化反问题与人工智能应用等方面的研究。E-mail: 203wanggen@163.com

Key words: hyperspectral GIIRS; bias correction; "off-line" method; random forest; cloud detection

0 引言

数值天气预报是一个初/边值问题。星载高光谱红外探测器通道主要覆盖 CO₂ 和 H₂O 光谱区域。CO₂ 和 H₂O 吸收带提供的温度和湿度信息是数值预报的模式变量。同化该吸收带的通道亮温能够改进模式的初值, 进而提高预报精度^[1]。

我国于 2016 年发射的风云四号 A 星(FY-4A)携带的 GIIRS 是国际上首台在静止卫星上装载的高光谱大气红外探测器^[2]。1650 个通道覆盖 700~2250 cm⁻¹ 光谱区域。其中, 中波通道有 961 个^[2]。GIIRS 能从高空遥测地球的温度、湿度以及其他气体成分的垂直分布, 可实现快速和大范围的连续观测, 为区域和全球数值天气模式预报提供了资料服务^[3]。高光谱红外探测器能提供高分辨率大气结构信息, 并丰富对温度和水汽信息的描述^[3]。特别是静止卫星上的高光谱红外探测中波通道, 为实现高时间分辨率和高垂直分辨率大气/水汽探测提供了可能。尤其在台风和暴雨等与湿过程密切相关的高影响天气预报中, 水汽信息很重要。

变分同化是基于误差服从高斯分布的假定理论而构建的目标泛函。实际上偏差包含了多种系统性误差。在同化卫星资料时需进行资料的偏差订正, 使得订正后的偏差满足高斯分布^[4]。本文中的卫星通道亮温偏差是其观测和模拟亮温的差值^[5]。关于卫星资料的偏差订正, 不少学者开展了富有成效的研究。Eyre J R^[6]首先提出卫星通道亮温偏差订正分为扫描和气团偏差订正两步。Harris B A 等人^[7]在 Eyre J R 研究的基础上提出了分纬度带的扫描偏差订正法。以上方法统称为“离线”偏差订正法。“离线”法统计某个时间段(如两周、一个月或三个月等)的资料, 并将统计后的订正系数用于后面时间段的资料偏差订正。Dee D

P^[8]提出的变分(也称为“在线”)偏差订正法将气团偏差订正的系数加入到变分同化总目标(或能量)的泛函中。在极小化迭代过程中, 该系数和其它模式控制变量(如温度和湿度等)一起更新优化。

总之, 目前偏差订正主要有“离线”和“在线”两种方法。它们又可具体细分为静态偏差、回归、自适应^[9]、基于辐射传输模式^[10]、基于卡尔曼滤波^[11]、变分偏差、静态与变分结合的动态偏差订正更新^[12]、带约束项^[13]、泰勒级数订正^[14]以及基于机器学习^[15]等方法。这类订正方法各有优缺点, 具体论述可参考王根等^[5]的综述论文。此外, Yin R Y 等人^[16]发现 FY-4A/GIIRS 观测亮温存在阵列偏差。

由于 RF^[17-18]能够度量因变量与自变量的非线性关系, 且基于大数据训练样本, 因此它可能会挖掘出数据内部的深层次关系。本文发展了基于 RF 的 GIIRS 偏差订正方法, 然后将其与“离线”方法进行了比较, 并分析了气团预报因子对偏差订正结果的重要性。

1 基于 RF 的 GIIRS 中波通道亮温偏差订正

极轨卫星资料偏差订正分为扫描和气团偏差订正两步^[5]。极轨卫星的扫描偏差是临边测量相对于星下点测量的系统偏差。考虑到静止卫星视场点恒定不变, 故无需进行扫描偏差订正, 而仅需进行气团偏差订正^[19]。其关键是选取预报因子。针对 Harris B A 等人在气团偏差订正中采用的线性回归“离线”法, 本文基于 RF^[17-18]发展了静止卫星 FY-4A/GIIRS 中波水汽通道亮温偏差订正的方法。

1.1 RF

Breiman L^[17]提出的 RF 模型是基于一定规则分类与回归树算法的集合。RF 中的每个

独立树都是由训练样本和输入变量随机选择子集而创建的。具体执行过程如下: 首先合并从原始训练样本库中抽取的多个样本, 并将其记为训练集; 然后根据样本集生成多个决策树, 并将它们组成 RF; 最后, 依据决策树的投票分数确定分类或回归模型结果^[4]。在将 RF 用于 GIIRS 中波水汽通道亮温偏差订正时, 该模型能够处理亮温偏差与预报因子的非线性关系。

RF 需要调整“树数量”以及在每个节点处分割时随机采样的“变量数量”两个基本模型参数^[18]。本文中的“树数量”和“变量数量”分别设定为 500 和 7。RF 模型可评价变量的重要性以优选变量。评价原理如下: 随机改变某解释变量的取值, 然后利用生成的 RF 模型进行袋外数据(Out-Of-Bag, OOB)误差拟合; OOB 误差增加越多, 该变量就越重要^[20]。

1.2 离线和 RF 偏差订正

Harris B A 等人的“离线”法在进行气团偏差订正时使用一组偏差预报因子 $X_i (i=1, 2, \dots, n-1, n)$ 。每个通道 j 的亮温偏差 B_j 为^[21]

$$B_j = \sum_{i=1}^n A_{ji} X_i + C_j \quad (1)$$

式中, 系数 A_{ji} 和 C_j 通过对大量样本进行最小二乘拟合计算得到。本文中的这一步骤采用 RF 进行气团偏差订正。RF 方法是用基于 Python 的 scikit 软件包来实现的^[18]。

1.3 预报因子选取

气团偏差订正的主要问题是预报因子的选取。预报因子有两种分类方法: 第一种基于“观测空间”, 使用通道观测或模拟亮温作为预报因子; 第二种基于“模式空间”, 使用温度和湿度等变量的积分组合值作为预报因子。“离线”偏差订正法从背景场中提取信息作为预报因子。预报因子的获取和不同预报因子的组合会对最终偏差订正效果产生一定的影响。本文采用的预报因子分别为总可降水量、模式的地表温度、50~200 hPa 和 300~1000 hPa

厚度^[21]。

2 FY-4A/GIIRS 通道亮温偏差订正试验与分析

2.1 模式与数据说明

本文采用 Tiros 业务垂直探测器辐射传输(Radiative Transfer for Tiros Operational Vertical Sounder, RTTOV)模式^[22]来模拟 FY-4A/GIIRS 亮温, 使用的光谱系数等文件由中国气象局“风云四号科研 A 星数值预报应用示范”攻关团队提供。

本文的背景场资料来源于 NCEP 的 FNL 资料, 用于亮温模拟和预报因子计算。FY-4A/GIIRS 资料来源于国家卫星气象中心网站(<http://satellite.nsmc.org.cn>)。

由于本文仅分析 GIIRS 晴空视场点的观测资料, 因此使用 AGRI 的 4 km 分辨率全圆盘云检测产品(Cloud Mask, CLM)^[23]进行 GIIRS 云检测。在执行过程中, 采用插值方法将 CLM 匹配至 GIIRS 视场点, 具体可参考 Yin R Y 等^[16]和 Zhang Q 等^[24]的工作。

2.2 初步质量控制

由于通过统计大量样本得到, “离线”偏差订正系数依赖于样本的数量和质量。本文在偏差订正之前先进行资料初步质量控制。考虑到红外探测器易受云的影响, 在质量控制过程中剔除了 GIIRS 的有云视场点。其它质量控制如下:

(1)极值判断。如果任一通道亮温偏差超出给定的阈值(-20~20 K)或者通道的观测/模拟亮温超出了给定的阈值(150~350 K), 则剔除该通道所在视场点的资料。

(2)偏差检验。如果某视场点中某通道的亮温偏差大于该通道亮温偏差的 3 倍标准差, 则剔除该视场点这一通道的资料。

2.3 云检测及区域试验设置

本文研究的时间段为 2019 年 8 月 8 日 20 时至 2019 年 8 月 10 日 11 时。因为 FY-4A 为静止卫星, 且此时间段的数据为“加密数据”

(30 min 全覆盖一次), 因此可将该样本数据用于本文中的统计研究。图 1(b)所示为 2019 年 8 月 9 日 12 时超强台风“利奇马”(Lekima, 国际编号为 1909)的 GIIRS 资料加密区分布。图 1(a)为 AGRI 的 CLM 分布图。其中, 0、1、2 和 3 分别表示云、可能云、可能晴空和晴空。图 1(b)是将 CLM 匹配至 GIIRS 视场点的云量分布图, 其中 0 为绝对晴空视场点。该视场点的资料用于本文中的 GIIRS 偏差订正研究。

图 2 所示为采用“离线”和 RF 订正方法前后 GIIRS 959、963、1024 和 1347 通道的亮温偏差概率密度函数分布。其中, Random forest-4 表示基于总可降水量、模式的地表温度、50~200 hPa 和 300~1000 hPa 厚度四个预报因子且采用 RF 得到的偏差订正结果; Random forest-6 表示在 Random forest-4 四个预报因子的基础上引入视场点经度和纬度信息(用于偏差订正)。

由图 2 可知, 经过偏差订正后, 高光谱 GIIRS 选取的 4 个通道亮温偏差的概率密度接近 0。偏差整体服从高斯分布, 满足变分同化基本理论的要求。与“离线”法相比, RF 法的订正效果更优, 尤其是引入经度和纬度信息后。这可能与 RF 能够度量因变量与自变量的非线性关系有关。

2.4 RF 回归模型变量的重要性分析

图 3 所示为不同预报因子以及经度(标记为“lon”)和纬度(标记为“lat”)信息对偏差订正 RF 回归模型的重要性分布。它是对本文所给个例的整体分析结果。其中, 总可降水量、模式的地表温度、50~200 hPa 和 300~1000 hPa 厚度分别表示为因子 1、2、3 和 4。

从图 3 中可以看出, 经度和纬度信息对 RF 回归模型的贡献率较大(较为重要)。此外, 预报因子 50~200 hPa 的厚度对最终偏差订正结果的影响也较大。

3 总结与展望

在变分同化卫星资料时, 通道亮温偏差需满足高斯分布的假定。因此本文开展了 FY-4A/GIIRS 中波通道亮温偏差订正研究, 并基于 Harris B A 等人提出的“离线”法发展了 RF 偏差订正法。试验结果表明, 经过偏差订正后的 GIIRS 亮温偏差满足高斯分布的假定。与“离线”法相比, RF 法的订正效果更好, 尤其是引入经度和纬度信息后。本文方法对于高光谱遥感数据应用的优化具有一定的参考意义, 也在一定程度上为后续风云系列卫星的偏差订正算法奠定了基础。

本文的不足之处是采用的个例较少。以后将采用 GIIRS 中波亮温同化数值天气预报模式来开展台风和强降水等相关的预报研究, 并进一步优化本文提出的偏差订正算法。

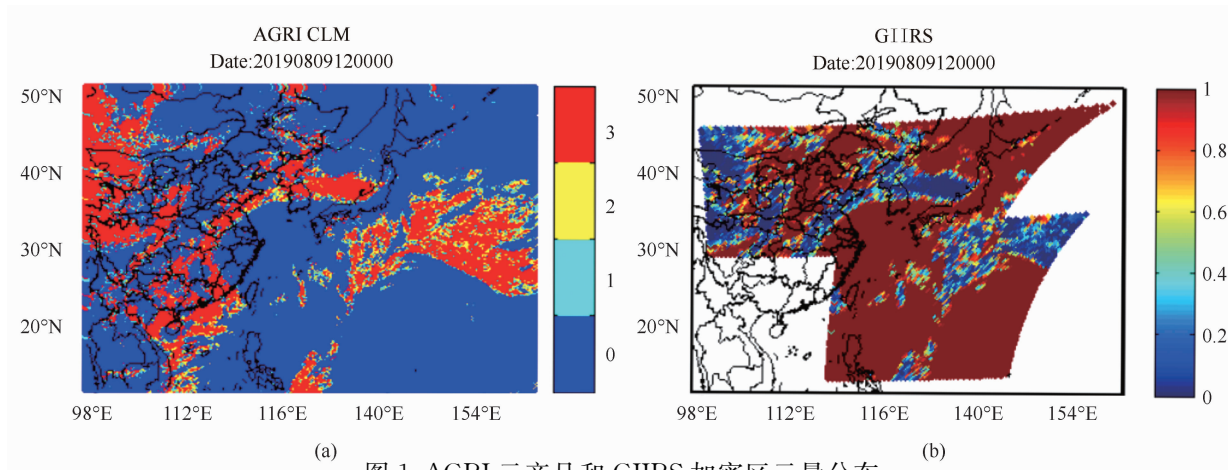


图 1 AGRI 云产品和 GIIRS 加密区云量分布

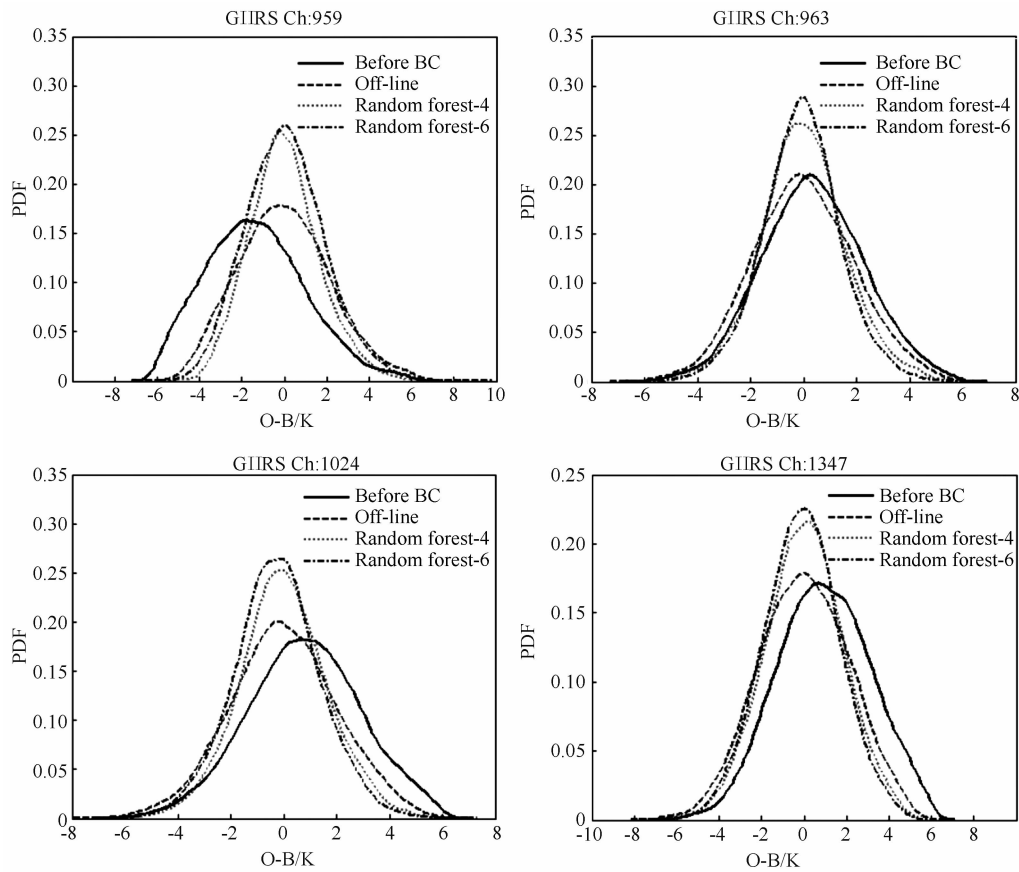


图 2 高光谱 GIRS 通道亮温偏差订正前后的概率密度分布

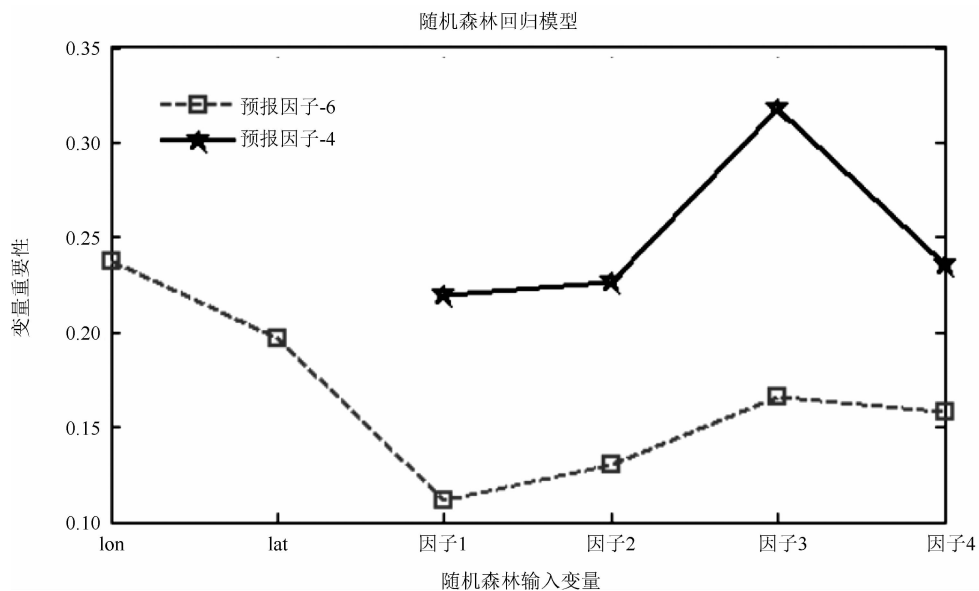


图 3 RF 回归模型 GIRS 偏差订正的不同因子变量的重要性分析

参考文献

[1] Wang G, Zhang J. Generalised variational assimilation of cloud-affected brightness temperature using simulated hyper-spectral atmospheric infrared

sounder data [J]. *Advances in Space Research*, 2014, 54(1): 49-58.

[2] Yang J, Zhang Z, Wei C, et al. Introducing the new generation of Chinese geostationary weather satellites, FengYun 4 [J]. *Bulletin of the Ameri-*

- can Meteorological Society*, 2017, **98**(8): 1637–1658.
- [3] 陆风, 张晓虎, 陈博洋, 等. 风云四号气象卫星成像特性及其应用前景 [J]. *海洋气象学报*, 2017, **37**(2): 1–12.
- [4] 王根. 卫星资料广义同化、大气科学中的数学物理反问题与人工智能应用——学习笔记 [M]. 北京: 气象出版社, 2020.
- [5] 王根, 张华, 杨寅. 高光谱大气红外探测器 AIRS 资料质量控制研究进展 [J]. *地球科学进展*, 2017, **32**(2): 139–150.
- [6] Eyre J R. A bias correction scheme for simulated TOVS brightness temperatures [R]. ECMWF Research Dept Technical Memorandum, 1992.
- [7] Harris B A, Kelly G. A satellite radiance bias correction scheme for data assimilation [J]. *Quarterly Journal of the Royal Meteorological Society*, 2001, **127**: 1453–1468.
- [8] Dee D P. Variational bias correction of radiance data in the ECMWF system [C]. Reading: Proceedings of the ECMWF Workshop on Assimilation of High Spectral Resolution Sounders in NWP, 2004.
- [9] Auligne T, McNally A P, Dee D P. Adaptive bias correction for satellite data in a numerical weather prediction system [J]. *Quarterly Journal of the Royal Meteorological Society*, 2007, **133**: 631–642.
- [10] Watts P D, McNally A P. Identification and correction of radiative transfer modeling errors for atmospheric sounders: AIRS and AMSU-A [C]. Reading: Proceedings of the ECMWF Workshop on Assimilation of High Spectral Resolution Sounders in NWP, 2004.
- [11] Elana J F, Seung J B, Brianr H, et al. Observation bias correction with an ensemble Kalman filter [J]. *Tellus*, 2009, **61**(2): 210–226.
- [12] 王雪曼, 李刚, 张华. 极轨卫星观测偏差订正动态更新技术在 GRAPES-3Dvar 的应用研究 [J]. *气象*, 2015, **41**(7): 863–871.
- [13] Han W, Bormann N. Constrained adaptive bias correction for satellite radiances assimilation in the ECMWF 4D-Var [R]. ECMWF Research Dept Technical Memorandum, 2016.
- [14] Jason A O, Roland P, Amos S L. Nonlinear Bias Correction for Satellite Data Assimilation Using Taylor Series Polynomials [J]. *Monthly Weather Review*, 2018, **146**: 263–285.
- [15] Jin J B, Lin H X, Segers A, et al. Machine learning for observation bias correction with application to dust storm data assimilation [J]. *Atmospheric Chemistry and Physics*, 2019, **19**: 10009–10026.
- [16] Yin R Y, Han W, Gao Z Q, et al. The evaluation of FY4A's Geostationary Interferometric Infrared Sounder (GIIRS) longwave temperature sounding channels using the GRAPES global 4D-Var [J]. *Quarterly Journal of the Royal Meteorological Society*, 2020, **146**(728): 1459–1476.
- [17] Breiman L. Random forests [J]. *Machine Learning*, 2001, **45**(1): 5–32.
- [18] Wang G, Han W, Lu S. Precipitation retrieval by the L1-norm regularization: typhoon Hagibis case [J]. *Quarterly Journal of the Royal Meteorological Society*, 2021, **147**(735): 773–785.
- [19] 郭姿佑, 秦正坤, 杜家铭, 等. 气团订正对静止卫星成像仪资料同化在台风预报中的改进效果研究 [J]. *热带气象学报*, 2019, **35**(4): 539–555.
- [20] Lee Y J, Daehyeon H, Ahn M H, et al. Retrieval of Total Precipitable Water from Himawari-8 AHI Data: A Comparison of Random Forest, Extreme Gradient Boosting, and Deep Neural Network [J]. *Remote Sensing*, 2019, **11**(15): 1741.
- [21] 王根, 陆其峰, 刘辉, 等. FY-3B 红外分光计亮温观测模拟偏差订正的初步研究 [J]. *红外*, 2014, **35**(1): 18–23.
- [22] Saunders R, Hocking J, Turner E, et al. An update on the RTTOV fast radiative transfer model (currently at version 12) [J]. *Geoscientific Model Development*, 2018, **11**: 2717–2737.
- [23] Min M, Wu C Q, Li C, et al. Developing the science product algorithm testbed for Chinese next-generation geostationary meteorological satellites: Fengyun-4 series [J]. *Journal of Meteorological Research*, 2017, **31**(4): 708–719.
- [24] Zhang Q, Yu Y, Zhang W M, et al. Cloud Detection from FY-4A's Geostationary Interferometric Infrared Sounder Using Machine Learning Approaches [J]. *Remote Sensing*, 2019, **11**(24): 3035.