

文章编号: 2096-1618(2022)03-0284-06

基于 FY-3D 和 FY-4A 的气温时空融合

陈瑶瑶¹, 邓小波¹, 黄启宏¹, 刘海磊¹, 王乙竹¹, 袁淑杰²

(1. 成都信息工程大学电子工程学院, 四川 成都 610225; 2. 成都信息工程大学大气科学学院, 四川 成都 610225)

摘要: 卫星遥感技术可大范围、连续性地获取数据, 然而静止卫星的时间分辨率高, 空间分辨率低; 极轨卫星的空间分辨率高, 时间分辨率低。多源卫星数据融合方法能结合静止卫星的高时间分辨率特性和极轨卫星的高空间分辨率特性, 从而得到高时空分辨率数据。选取湖南省作为研究区域, 利用静止气象卫星 FY4A 和极轨气象卫星 FY3D 的气温产品, 使用基于字典融合算法进行气温数据融合, 得到时间分辨率为1 h、空间分辨率为250 m的高时空分辨率合成气温数据。从融合结果来看, 此方法可以提高气温数据的空间分辨率, 从得到的逐小时气温图像来看, 此方法可以提高气温数据的时间分辨率。最后, 将融合得到的24 h高时空分辨率的气温数据与台站数据进行时空匹配并验证。结果表明, 相关系数 R 为0.764, 均方根误差 RSME 为2.97 °C, 平均偏差 BIAS 为-1 °C, 平均绝对百分比误差为-2.75%。

关键词: 卫星遥感; 气温产品; FY-3D; FY-4A; 高时空分辨率; 字典融合算法

中图分类号: TP79

文献标志码: A

doi: 10.16836/j.cnki.jcuit.2022.03.008

0 引言

近年, 卫星资料在天气预报、气候监测预测、自然灾害监测等方面都得到广泛的应用。气象卫星具有观测范围广、观测次数多、观测时效快、观测数据质量高、不受自然条件和地域条件限制等优势。它提供的气象信息已广泛应用于日常气象业务、环境监测、防灾减灾、大气科学、海洋学和水文学的研究^[1]。

近地表气温空间分布数据是各种植物生理、水文、气象、环境等模式和陆面过程模型中非常重要的气象参数输入因子之一, 也与人类的日常生活息息相关。目前, 近地表气温数据基本由各个气象观测站采集, 气温数据是监测站点离地面1.5~2 m高度的大气温度。由于气象台站获取的气温数据, 不具有时空连续分布的特点, 况且在偏远地区和人烟稀少地区安置的气象观测站点非常稀疏, 甚至没有^[2]。因此采用卫星遥感技术获取大范围空间连续的气温数据是非常重要的手段。

随着气温相关的应用与研究不断地发展, 高时空分辨率的气温数据也越来越被需要。然而, 受到遥感卫星载荷量的限制及传感器制造上的技术瓶颈, 卫星传感器很难获取具有高时空分辨率的数据, 基于以上问题, 国内外学者提出了多源卫星数据融合算法^[3-4]。

最初时空融合算法是基于反射率数据, 之后随着应用不断深入, 逐步扩大至植被指数、建筑指数、水体

指数及温度等方面。其中 Gao 等^[5]提出的时空自适应反射率融合模型(spatial and temporal adaptive reflectance fusion model, STARFM)典型且应用广泛, 该模型通过邻近光谱相似像元值计算中心像元值, 来获取高时空分辨率的合成数据产品。Yang 等^[6]提出基于稀疏理论的图像超分辨率重建算法, 其算法的基本思想是建立高低分辨率影像之间的对应关系以对预测日期的高空间分辨率影像进行预测。Huang 等^[7]提出基于稀疏表示的时空反射率融合模型(sparse representation based on a spatiotemporal reflectance fusion model, SPST-FM)。该模型利用预测时相的前后时相的 MODIS 和 Landsat 影像对的差值影像训练代表时间变化特征的高低分辨率字典, 然后利用预测时刻的 MODIS 影像生成预测时刻类似 Landsat 的融合影像。

本文使用 FY-3D 和 FY-4A 卫星数据得到的气温产品, 基于字典融合算法进行气温数据融合, 选取湖南省作为研究区域, 最后得到时间分辨率为1 h、空间分辨率为250 m的高时空分辨率合成气温数据。

1 气温产品与研究区域

1.1 气温产品介绍

风云三号 D 星于 2017 年 11 月 15 日 2 时 35 分由长征四号丙运载火箭在太原卫星发射中心成功发射。搭载的中分辨率光谱成像仪 II (MERSI-2) 整合了原有风云三号卫星两台成像仪器(MERSI-1 和 VIRR)的功

能,是世界上首台能够获取全球250 m分辨率红外分裂窗区资料的成像仪器,可每日无缝隙获取全球250 m分辨率真彩色图像,实现云、气溶胶、水汽、陆地表面、海洋水色等大气、陆地和海洋参量的高精度定量估算。

风云四号 A 星于 2016 年 12 月 11 日由长征三号乙改进Ⅲ型运载火箭在西昌卫星发射中心成功发射。搭载的多通道扫描成像辐射计(AGRI)通过精密的双扫描镜机构实现精确和灵活的二维指向,可实现分钟级的区域快速扫描;采用离轴三反主光学系统,高频次获取 14 波段的地球云图。相比风云二号单一可见光通道的限制,风云四号首次制作出彩色卫星云图,最快 1 分钟生成一次区域观测图像^[8]。

表 1 是所使用气温产品的相关信息。具体地,利用 FY-4A 的 12、13 通道亮温数据、云检测(CLM)、全球预报系统(GFS)的气温数据、台站气温数据及其他辅助数据,使用 XGB 模型,获得晴空条件下的 FY-4A 气温产品^[9]。利用 FY-4A 云顶温度(CTT)、云顶高度(CTH)和 GFS 气温数据、台站气温数据及其他辅助数据,基于神经网络,获得云天条件下的 FY-4A 气温产品^[10]。使用 FY-3D 的 24、25 通道亮温数据、GFS 气温数据、台站气温数据及其他辅助数据,使用 XGB 模型获得 FY-3D 全天候气温产品。

表 1 气温产品的信息

数据类别	FY-3D 气温产品	FY-4A 气温产品	气象台站 气温数据
空间分辨率	250 m	4000 m	Site
时间分辨率	2 次/d	1 次/h	1 次/h

其中,FY-4A 和 FY-3D 相关数据均可在国家卫星气象中心风云卫星数据中心下载(<http://satellite.nsmc.org.cn/PortalSite/Data/Satellite.aspx>)。美国国家大气研究中心(NCAR)的研究数据档案(RDA)提供了非常便捷获得 GFS 历史数据下载(<https://rda.ucar.edu/datasets/ds084.1/>),实时的 GFS 气温数据可从 NOAA/NCEP 下载(<https://www.nco.ncep.noaa.gov/pmb/products/gfs/>)。台站数据可在中国气象数据服务中心下载(<http://data.cma.cn>)。

1.2 研究区域

研究地区为湖南省。湖南地处 108°47'E ~ 114°15'E, 24°38'N ~ 30°08'N, 土地面积约 21.18 万 km²,位于云贵高原向江南丘陵和南岭山脉向江汉平原过渡的地带,地势呈三面环山、朝北开口的马蹄形地貌,由平原、盆地、丘陵地、山地、河湖构成,地跨长江、珠江两大水系,属大陆性亚热带季风湿润气

候。湖南地区高程图及湖南省 97 个气象站点的分布情况如图 1 所示。

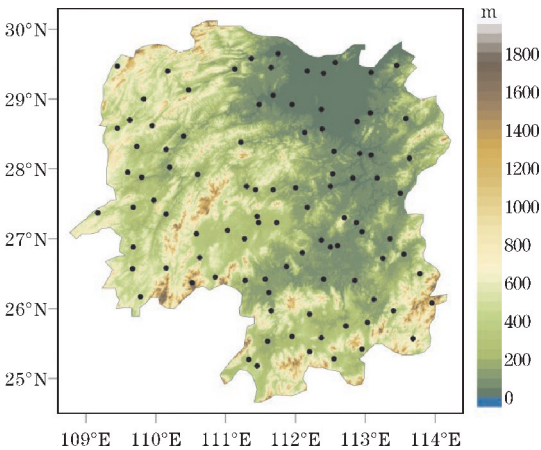


图 1 研究区域及站点分布

2 字典融合算法

字典融合算法核心思想是利用两个时刻的低空间分辨率遥感图像之间的变化信息,基于训练得到的完备字典,利用超分辨率重建方法得到对应的高空间分辨率变化信息后,再结合其中一个时刻观察的高空间分辨率遥感图像的空间信息,得到另一时刻高空间分辨率遥感合成图像^[10-12]。

2.1 算法框架

图 2 为字典融合算法的框架图。使用高空间低时间分辨率(250 m)极轨卫星风云三号 D 星(FY-3D)估算的气温数据和低空间高时间分辨率(4000 m)静止卫星风云四号 A 星(FY-4A)估算的气温数据,采用基于字典融合算法进行数据融合,获得时间分辨率为 1 h,空间分辨率为 250 m 的湖南地区气温数据,并将融合结果与湖南地区的 97 个台站所测气温数据进行验证分析。

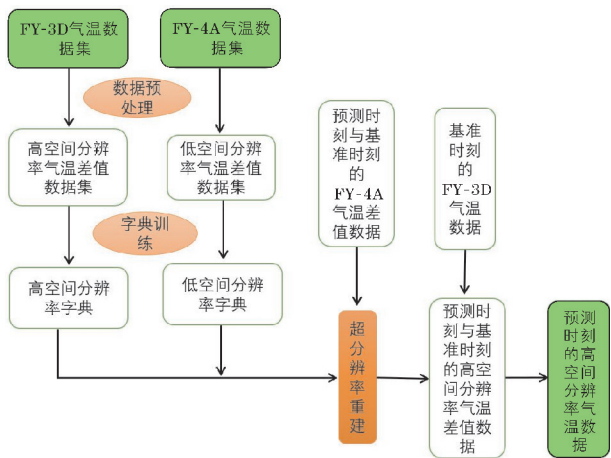


图 2 字典融合算法框架图

利用获得的高低空间分辨率气温差值数据集进行字典的训练,学习并建立两个时刻 FY-4A 气温数据之间的变化差异与对应的 FY-3D 气温数据之间的变化差异的关系,对待融合的变化差异数据采用超分辨率重建,然后再进行时空数据融合。其中,变化差异数据(也称为差值数据)是指两个时刻观察的 FY-4A 气温数据(低空间分辨率)或者 FY-3D 气温数据(高空间分辨率)相减的差值数据。

2.2 算法原理

2.2.1 差值数据集的构建

差值数据集的构建主要是为了冗余字典的构建。对获取的多个时刻的 FY-4A 和 FY-3D 气温数据进行时空匹配,然后按不同时期同一时刻进行差值计算(例如 8 月 1 日 06:00 和 8 月 2 日 06:00 的低空间分辨率气温数据做差值),最后得到 FY-3D 气温差值数据集(即高空间分辨率气温差值数据集)和 FY-4A 卫星气温差值数据集(即低空间分辨率气温差值数据集)。

选用 2019 年夏季 3 个月的气温数据用于字典训练,参数见表 2,对数据集样本进行分块操作,剔除阈值较小的数据块,最后得到需要的高低空间分辨率训练样本。

表 2 设置的参数

参数	变量名	值
字典大小	Dictionary_size	1024
稀疏正则化系数	lambda	0.15
分块大小	Patch_size	5
样本块数量	nSmp	100000
升标因子	upscale	16
阈值	threshold	0.1
迭代次数	num_iters	40

2.2.2 冗余字典的构建

根据得到的高低空间分辨率训练样本,利用 K-SVD 冗余字典构建算法,对高低空间分辨率训练样本进行联合字典构建,得到高空间分辨率字典和低空间分辨率字典。即求解下面的优化问题:

$D, X = \operatorname{argmin} \|Y - DX\|_F^2, \text{s.t. } \forall i, \|x_i\|_0 \leq T_0 \quad (1)$ 其中, Y 为输入的矩阵数据, D 为得到的字典, X 为对应的稀疏表示系数矩阵, T_0 可自行设置数值。

在字典训练过程中,根据经验,将稀疏正则化系数

设置为 0.15,升标因子设置为 16,迭代次数设置为 40。

2.2.3 数据重建与融合

获得高低空间分辨率字典之后,需要使用基准时刻的 FY-3D 和 FY-4A 气温数据及预测时刻 FY-4A 气温数据来预测该时刻对应的 FY-3D 气温数据。首先,需要获取预测时刻 T_p 与基准时刻 T_m 的 FY-4A 差值数据,即 $\Delta M = M^p - M^m$,其中 ΔM 为低空间分辨率气温差值数据。然后求解出 ΔM 在冗余字典 D_l 下的稀疏表示 X_l ,那么对应的高空间分辨率气温差值数据为

$$\Delta L = D_l \cdot X_l \quad (2)$$

其中, D_l 为高空间分辨率字典, ΔL 为得到的高空间分辨率气温差值数据,最后得到预测时刻的高空间分辨率合成气温数据:

$$L^p = L^m + \Delta L \quad (3)$$

其中, L^m 为 T_m 时刻的高空间分辨率气温数据, L^p 为 T_p 时刻得到的高空间分辨率合成气温数据。按照以上所述步骤,根据 FY-4A 气温数据的时间分辨率,可连续生成对应的高时间分辨率的高空间分辨率合成气温数据。

基于以上原理,选用 2019 年夏季的气温数据进行字典训练得到冗余字典,再选用 2020 年夏季其中两个时刻的气温数据进行融合。具体地,因为 FY-3D 只有 06 时和 18 时(协调世界时 UTC)有数据,故基准时刻为 2020 年 7 月 18 日 06 时,预测时刻为 07 时、08 时、09 时、10 时,以此类推,直到下一个基准时刻 18 时,以此可获得逐小时的高空间分辨率合成气温数据。

3 结果与分析

3.1 融合结果对比分析

图 3 是湖南地区气温分布图,图 3(a)和图 3(b)分别是基准时刻(北京时间 2020 年 7 月 19 日 02:00)的 FY-4A 和 FY-3D 气温,图 3(c)为预测时刻(北京时间 2020 年 7 月 19 日 10:00)的 FY-4A 气温,图 3(d)是利用本文方法得到的预测时刻高空间分辨率气温图。气温色柱的单位为℃,对比图 3(c)和图 3(d),气温在空间分布上轮廓基本一致,呈现东南方向的气温高于西北方向的气温,细节部分也较一致,如 25°N ~ 26°N, 111°E ~ 114°E 的区域。此外,28°N, 113°E 附近出现的斑块与图 3(a)有直接关系,可以看出,基准时刻的气温数据对融合结果有一定的影响。

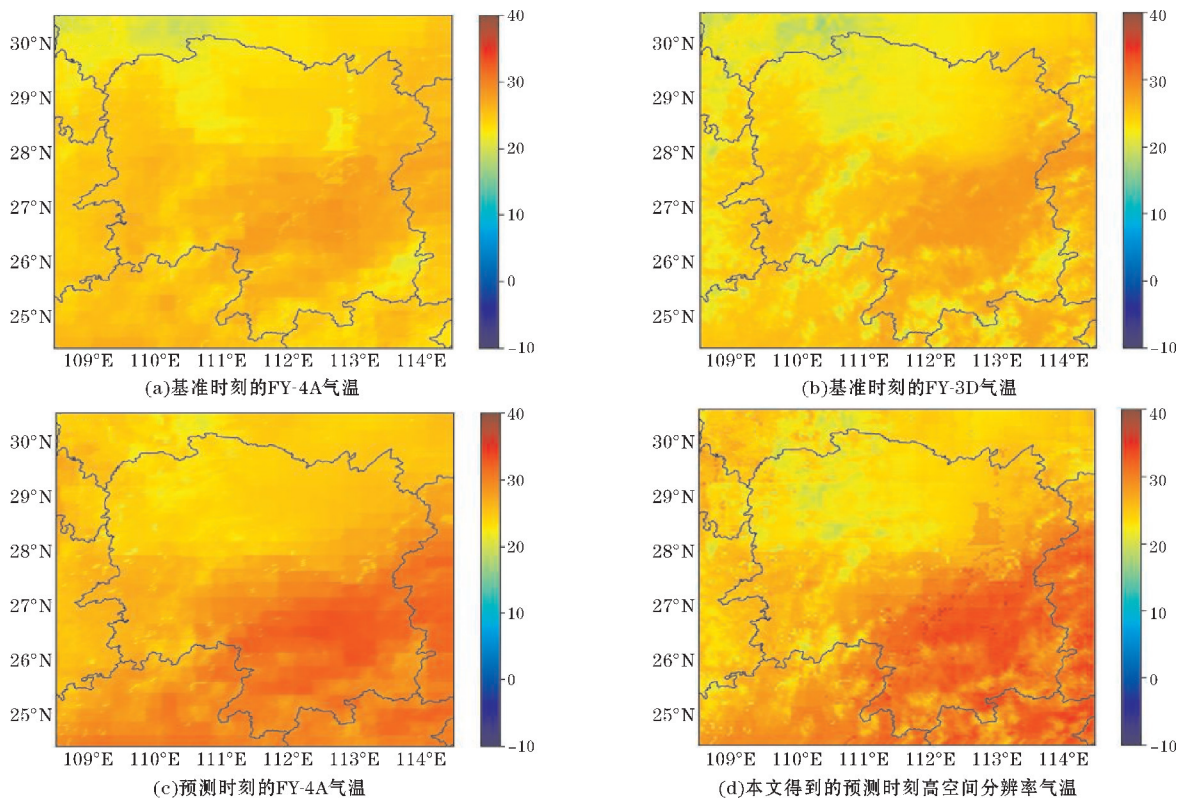


图3 湖南地区气温分布图

图4是预测时刻高空间分辨率气温与台站气温的二维直方图。图中展示了获得的气温与台站实测气温之间的相关系数 R ,均方根误差RMSE,平均偏差BIAS,平均绝对百分比误差RE。如图4所示,色柱单位为个,即重合点的个数,横纵坐标为气温值。图4(a)为湖南地区北京时间2020年7月19日10:00的FY-4A气温与台站气温的二维直方图,图4(b)为融合得

到的2020年7月19日10:00的气温与台站气温的二维直方图。对比可见,相关系数 R 均大于0.9,表明得到的FY-4A气温与台站气温的相关性较强,均方根误差RMSE分别为2.36℃,2.05℃,表明得到的气温与台站气温的误差较小,平均偏差BIAS由2℃下降到1℃,平均绝对百分比误差RE由7.43%下降到6.02%。

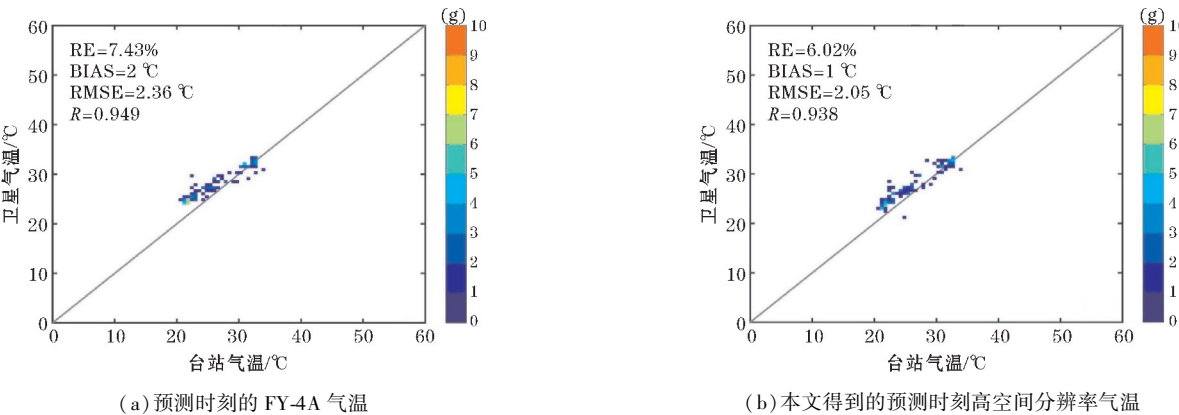


图4 湖南地区气温的二维直方图

3.2 逐小时生成气温图像

利用上述方法,选取UTC时间2020年7月18日06:00,2020年7月18日18:00作为基准时刻,使用两个基准时刻的FY-3D气温数据和此时间段内每小时的FY-4A气温数据进行数据融合,逐小时往后预测气温,可得到24 h

的融合后的气温数据。基于此方法,可以获得每天时间分辨率为1 h,空间分辨率为250 m的气温数据。

图5是北京时间15:00到第二天14:00的24 h的融合后的高空间分辨率合成气温图。从空间尺度上看,湖南地区气温呈现东南方向高于西北方向,这可能是由于西北方向的海拔高于东南方向的海拔。从时间

尺度上看,气温从 09:00 左右(蓝色选框)开始明显升温,18:00 左右(绿色选框)开始明显降温。02:00 气温达到最低,15:00 点左右达到最高,气温的变化趋势较符合实际情况。

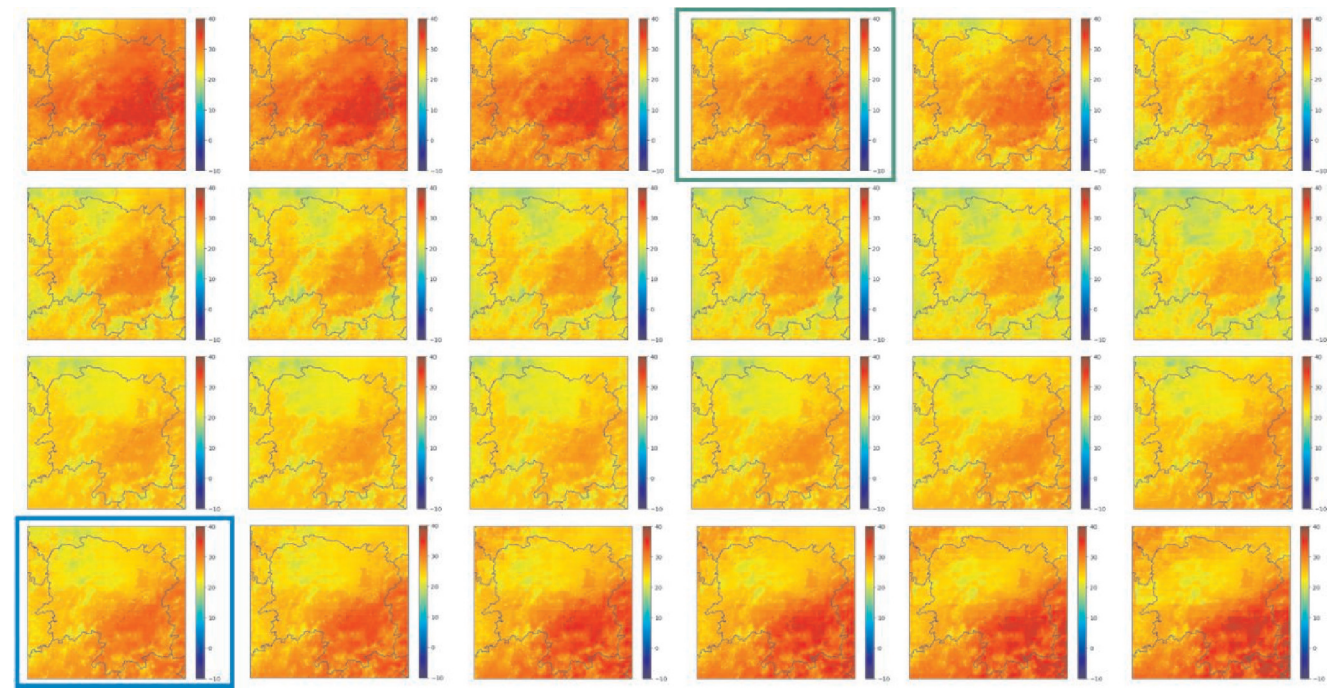


图 5 融合后湖南地区 2020 年 7 月 18-19 日的 24 h 的气温分布图

3.3 精度验证

图 6 是得到的 24 h 气温与台站气温的二维直方图,将融合得到的 2020 年 7 月 18-19 日 24 h 的气温数据与气象台站实测的气温数据进行误差计算,使用平均绝对百分比误差 RE、平均偏差 BIAS、均方根误差 RSME 及相关系数 R 作为评价参数。如图 6 所示,相关系数 R 为 0.764,均方根误差 RSME 为 2.97 $^{\circ}\text{C}$,平均偏差为 -1 $^{\circ}\text{C}$,平均绝对百分比误差为 -2.75%。

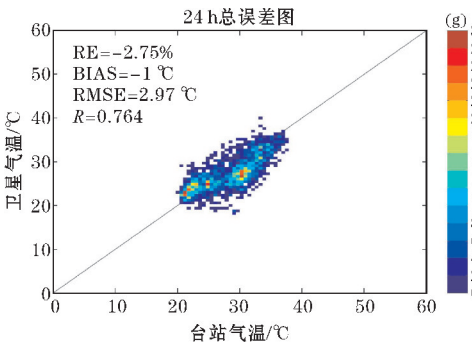


图 6 二维直方分布图

图 7 是 24 h 气温的误差趋势对比图。将融合前 FY-4A 的每小时 RMSE 和融合后的每小时 RMSE 进行对比,画出变化趋势折线图。如图 7 所示,蓝线为融合前的 RMSE 变化趋势,整体分布较平稳,均在 3 $^{\circ}\text{C}$ 以下,红线为融合后的 RMSE 变化趋势,UTC 时间 2020 年 7 月 18 日 07:00-18:00 气温误差较大,高于 3 $^{\circ}\text{C}$,北京时间 2020 年 7 月 18 日 19:00 至 19 日 06:00 气

温误差均低于 FY-4A,分布在 2 $^{\circ}\text{C}$ 左右。融合前 FY-4A 的平均 RMSE 为 2.50 $^{\circ}\text{C}$,融合后的 RMSE 为 2.88 $^{\circ}\text{C}$ 。从变化趋势可以看出,前 12 h 融合效果较差,后 12 h 融合效果稍好,这可能与基准时间 FY-3D 的气温误差有很大关系,也表明此方法有一定适用性和不确定性,若能获得差异较小且误差稳定的基准时刻的气温数据,或许会使结果有较大改善。

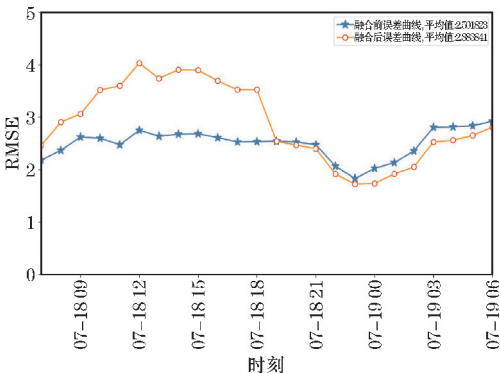


图 7 24 h 的 RMSE 误差趋势对比图

4 结束语

使用高空间低时间分辨率气温数据和低空间高时间分辨率气温数据,采用基于字典融合算法进行气温数据融合。具体地,先利用获得的高低空间分辨率气温差值数据集进行字典的训练,学习并建立两个时刻 FY-4A 气温数据之间的变化差异与对应的 FY-3D 气温数据之

间的变化差异的关系,再对待融合的变化差异数据采用超分辨率重建,最后进行气温数据融合。

据验证,此方法能提高气温的时空分辨率,得到时间分辨率为1 h,空间分辨率为250 m的湖南地区气温数据。融合后的气温与低空间分辨率FY-4A获得的气温分布轮廓基本一致,且获得了更多的细节表现。融合结果比较接近台站气温数据,24 h的总体均方根误差RMSE为2.97℃。相关系数 R 为0.764,平均偏差BIAS为-1℃,平均绝对百分比误差RE为-2.75%。

参考文献:

- [1] 许健民,杨军,张志清.我国气象卫星的发展与应用[J].气象,2010,36(7):94-100.
- [2] 祝善友,张桂欣.近地表气温遥感反演研究进展[J].地球科学进展,2011(7):724-730.
- [3] 董文全,蒙继华.遥感数据时空融合研究进展及展望[J].国土资源遥感,2018,30(2):1-11.
- [4] 刘建波,马勇,武易天,等.遥感高时空融合方法的研究进展及应用现状[J].遥感学报,2016,20(5):1038-1049.
- [5] Gao F, Masek J G, Schwaller M R. On the Blending of the Landsat and MODIS Surface Reflectance: Predicting Daily Landsat Surface Reflectance[J]. IEEE Transactions on Geoscience and Remote Sensing, 2006, 44(8): 2207-2218.
- [6] Yang J, Wright J, Huang T, et al. Image Super-Res-

olution Via Sparse Representation [J]. IEEE Transactions on Image Processing, 2010, 19(11): 2861-2873.

- [7] Huang B, Song H H. Spatiotemporal Reflectance Fusion via Sparse Representation[J]. IEEE Transactions on Geoscience & Remote Sensing, 2012, 50(10): 3707-3716.
- [8] 徐娜,吴荣华,胡秀清.风云三号(02批)气象卫星地面应用系统工程D星中分辨率光谱成像仪数据使用指南[Z].国家卫星气象中心,2019.
- [9] Zhou K, Liu H L, Deng X B, et al. Comparison of Machine-Learning Algorithms for Near-Surface Air-Temperature Estimation from FY-4A AGRI Data [J]. Advances in Meteorology, 2020: 1-14.
- [10] Li H P, Liu H L, Duan M Z, et al. Estimation of Air Temperature under Cloudy Conditions Using Satellite-Based Cloud Products[J]. IEEE Geoscience and Remote Sensing Letters, 2022, 19: 1-5.
- [11] 赵永光.基于稀疏表示超分辨率重建的遥感时空数据融合方法[D].北京:中国科学院,2012.
- [12] Yang J, Wright J, Huang T S. Image super-resolution as sparse representation of raw image patches[J]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition [J]. 2008: 1-8.
- [13] 吴君涵.基于字典学习和多源遥感数据的高时空分辨率地表温度融合研究[D].上海:华东师范大学,2017.

Temporal and Spatial Fusion of Air Temperature based on FY-3D and FY-4A

CHEN Yaoyao¹, DENG Xiaobo¹, HUANG Qihong¹, LIU Hailei¹, WANG Yizhu¹, YUAN Shujie²

(1. College of Electronic Engineering, Chengdu University of Information Technology, Chengdu 601225, China; 2. College of Atmospheric Sciences, Chengdu University of Information Technology, Chengdu 601225, China)

Abstract: Satellite remote sensing technology can obtain data in a large range and continuously. However, geostationary satellites have high temporal resolution but low spatial resolution. Polar-orbiting satellites have high spatial resolution but low temporal resolution. The multi-source satellite data fusion method can combine the high temporal resolution characteristics of geostationary satellite and the high spatial resolution characteristics of polar-orbit satellite to obtain high temporal and spatial resolution data. In this paper, Hunan province is selected as the research area. Temperature products of geostationary meteorological satellite FY-4A and polar-orbit meteorological satellite FY-3D are used to fuse temperature data based on dictionary fusion algorithm, and synthetic temperature data with high temporal and spatial resolution of 1 h and 250 m are obtained. The results show that this method can improve the spatial resolution of air temperature data. According to the hourly temperature images, this method can improve the temporal resolution of temperature data. Finally, the 24 h high spatio-temporal resolution air temperature data obtained by fusion is matched and verified with the station data. The results show that the correlation coefficient R was 0.764, the root mean square error was 2.97℃, the mean bias was -1℃, and the mean absolute percentage error was -2.75%.

Keywords: satellite remote sensing; temperature products; FY-3D; FY-4A; high spatial and temporal resolution; dictionary fusion algorithm