

杨轩,曾燕,邱新法,等.2023.基于机器学习算法的多源月尺度融合降水产品在中国区域的检验评估[J].暴雨灾害,42(5):595-605.
YANG Xuan, ZENG Yan, QIU Xinfu, et al. Examination and evaluation of multi-source monthly scale fusion precipitation product in China based on machine learning algorithm [J]. *Torrential Rain and Disasters*,42(5):595-605 (in Chinese). doi:10.12406/byzh.2023-006

基于机器学习算法的多源月尺度融合降水产品 在中国区域的检验评估

杨轩¹,曾燕^{2,3,4},邱新法¹,朱晓晨⁵

(1. 南京信息工程大学地理科学学院,南京 210044; 2. 中国气象局交通气象重点开放实验室,南京 210041;
3. 江苏省气象科学研究所,南京 210041; 4. 南京气象科技创新研究院,南京 210041;
5. 南京信息工程大学应用气象学院,南京 210044)

摘要: 栅格格式降水产品相对于地面气象站观测资料有更好的空间监测能力,但是不同产品性能存在显著差异。本文评估了9种月尺度栅格降水产品 TRMM、GPM、CMORPH、CHIRPS、ERA5、ERA5-Land、PERSIANN、PERSIANN-CDR、PERSIANN-CCS 在中国的精度,从中择优选取5种较好的降水产品,利用XGBoost、随机森林和多元线性回归3种机器学习算法分别进行数据融合。研究发现,TRMM、GPM、CMORPH、CHIRPS、PERSIANN-CDR 5种产品具有相对较好的精度;在高海拔与干旱区域,降水产品的误差均明显增大。经过机器学习算法融合后,最优的XGBoost算法模型产品相关系数明显提升,均方根误差和偏差明显降低。3种算法各月均表现较高精度,其中XGBoost算法模型产品在夏季表现较好,而随机森林算法模型产品在冬季表现较好,且3种算法模型产品在不同区域均表现较高精度。和融合之前的5种原始产品比较,3种算法模型产品的精度均有提升。经过XGBoost算法融合后的产品在空间上相比较最优的原始GPM产品与气象站点插值产品具有更多的变化和局部降水细节信息。

关键词: 月尺度降水产品;机器学习;评估指标;数据融合

中图分类号:P426.6

文献标志码:A

DOI: 10.12406/byzh.2023-006

Examination and evaluation of multi-source monthly scale fusion precipitation product in China based on machine learning algorithm

YANG Xuan¹, ZENG Yan^{2,3,4}, QIU Xinfu¹, ZHU Xiaochen⁵

(1. School of Geographical Sciences, Nanjing University of Information and Science & Technology, Nanjing 210044; 2. Key Laboratory of Transportation Meteorology, China Meteorological Administration, Nanjing 210041; 3. Jiangsu Institute of Meteorological Sciences, Nanjing 210041; 4. Nanjing Joint Institute for Atmospheric Sciences, Nanjing 210041; 5. School of Applied Meteorology, Nanjing University of Information Technology, Nanjing 210044)

Abstract: Grid format precipitation products have better spatial monitoring capabilities compared to ground meteorological station observations, but there are significant differences in performance among different products. This article evaluates the accuracy of nine monthly scale precipitation products TRMM, GPM, CMORPH, CHIRPS, ERA5, ERA5 Land, PERSIANN, PERSIANN-CDR, PERSIANN-CCS in China, and selects five better precipitation products from them. XGBoost is used to select the best precipitation products. Three machine learning algorithms, random forest and multiple linear regression, were used for data fusion. Research has found that TRMM, GPM, CMORPH, CHIRPS, and PERSIANN-CDR products have relatively good accuracy. In high altitude and arid regions, the error of precipitation products significantly increases. After machine learning algorithm fusion, the optimal XGBoost algorithm model significantly improves product correlation coefficient, and significantly reduces root mean square error and bias. The three algorithms have shown good accuracy in each month, with XGBoost algorithm model products performing better in summer and random forest algorithm model products performing better in winter. Moreover, the three algorithm model products have shown high accuracy in different regions. Compared with the five original products before the fusion, the accuracy of the three algorithm model products has improved. The product fused with XGBoost algorithm has more variation and local precipitation details compared to the optimal original GPM product and meteorological station interpolation product in space.

Key words: monthly scale precipitation products; machine learning; evaluation indicators; data fusion

收稿日期: 2023-01-12; 定稿日期: 2023-07-11

资助项目: 国家重点研发计划项目(2019YFB2102003); 国家自然科学基金项目(41805049, 42075118)

第一作者: 杨轩, 主要从事GIS在气象领域的应用研究。E-mail: yxzj0519@163.com

通信作者: 曾燕, 主要从事水文气象学与气候资源利用的研究。E-mail: jlzengyan@sina.com

引言

降水是最重要的气象要素之一,是全球水循环的重要组成部分(熊立华等,2021)。由于降水的空间异质性,准确测量降水量仍然是一项具有挑战性的任务(Guo et al.,2016)。虽然地面气象站测量的降水数据十分准确,但是气象站分布不均,使用传统气象站点降水数据插值的方法难以真实反映降水实况(曲学斌等,2020)。近些年随着技术发展,利用卫星遥感产品进行算法反演可以提供大范围和连续的降水观测,因此基于栅格格式的高时空分辨率卫星遥感降水产品迅速发展(徐淑琴等,2017)。但不同的降水产品在不同的区域精度不同,Duan等(2016)通过综合评价意大利阿迪杰盆地8个高空间分辨率网格降水产品,发现CHIRPS(Climate Hazards Group Infrared Precipitation with Station Data)、TRMM (Tropical Rainfall Measuring Mission)和CMORPH (Climate Prediction Center Morphing Technique)3种产品表现较好。Wu等(2019)通过比较青藏高原的TRMM、CHIRPS和CMFD (China Meteorological Forcing Dataset)3种降水产品,发现对青藏高原东南部(湿润区)降水的估算优于西北部(干燥区)降水的估算。Fang等(2019)通过评估GPM (Global Precipitation Measurement)与TRMM数据,揭示了GPM的优势。吕爱锋和亓珊珊(2022)等对遥感和再分析降水产品在缺测干旱内陆地区进行评估和水文模型分析,对缺乏资料的干旱区后续水资源科学研究提供依据。虽然目前国内外的降水产品评估取得一定进展,但是局限于较小区域,且大多数为单一产品的研究或者少数产品的对比(Xu et al.,2019;杜懿等,2021;丁光旭等,2022;甘富万等,2022)。中国地区地形复杂,降水产品在不同区域精度不同,研究多种降水产品在中国不同区域的适用性十分重要(刘少军等,2021)。

不同的降水产品具有不同的优势,因此展开多源降水数据融合,取长补短以获得最优精度栅格降水产品是必要的,目前已有许多学者进行研究。李运龙等(2017)结合GWR (Geographically Weighted Regression)与克里金插值方法,在赣江流域进行降水数据融合试验,并将试验结果应用到水文领域。Xie和Xiong (2011)利用CMORPH数据结合中国降水格点数据采用两步融合法建立CMPA (China Merged Precipitation Analysis)数据。该数据集发挥了卫星观测与地面观测的优势,具有较高的精度与合理的降水分布(宇婧婧等,2013;2015)。但随着遥感数据量的加大,以及融合过程中降水与其他变量的非线性关系,常用的统计方法并不能很好地描述降水与其他影响因子

的关系(潘旻等,2018)。因此,机器学习方法逐渐被应用于降水融合中。张均民等(2021)利用XGBoost (eXtreme Gradient Boosting)算法融合雷达与地面站点降水观测资料,结果表明该算法可以获得较高精度的降水。Alizadeh和Nikoo (2018)利用机器学习算法在伊朗地区进行降水产品融合,并将数据应用到干旱监测中取得了良好的效果。虽然多源降水产品融合取得一定进展,但以上研究多针对较小区域,且并未比较多种机器学习算法的优劣,另外参与融合的降水产品数偏少。

本文选择国际上常用的9种卫星反演和再分析降水产品,基于气象站点的观测降水评估各产品的精度,从中选择较好的产品,采用3种机器学习算法进行降水产品的融合并对比评估,以获得更高精度的中国栅格月降水产品。

1 资料与方法

1.1 研究区域

本文研究区域选择中国大陆区域。中国大陆地势西高东低,呈三级阶梯状,地貌形态多样,气候类别多样,是最典型的季风区,降水空间与季节差异明显,降水量自东南向西北呈逐步减少的趋势。

1.2 资料说明

1.2.1 降水产品

本文选用目前国际上具有代表性的TRMM、GPM、CMORPH、CHIRPS、ERA5、ERA5-Land、PERSIANN、PERSIANN-CDR、PERSIANN-CCS 9种降水产品。ERA5系列(The fifth Generation ECMWF Reanalysis)气候数据集由欧洲中期天气预报中心(European Centre for Medium-Range Weather Forecasts,ECMWF)研发的第五代大气再分析产品。TRMM数据是世界上首颗热带降水测量卫星,携带的测雨雷达与微波成像仪相结合,极大改善了降水反演精度(曾琪铨和金鑫,2022)。卫星降水数据IMERG是GPM计划发布的新一代多卫星联合反演降水数据,其融合了多种微波、红外和雷达数据,充分实现多源传感器的优势互补(曾岁康和雍斌,2019)。PERSIANN (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks)系列数据由加利福尼亚大学尔湾分校的水文气象遥感中心开发,主要是基于红外亮温图像通过对云特征分类来估计降水量(刘松楠等,2022)。CMORPH是由美国国家海洋大气局(National Oceanic and Atmospheric Administration,NOAA)气候预测中心(Climatology Prediction Center,CPC)整合全球多卫星被动遥感数据生成(李强宇等,2021)。CHIRPS提供1981年至今横跨全球50°S—50°N空间分辨率为0.05°的降水

产品,该数据融合了卫星降水与地面气象站降水,可以提供卫星网格点降水用于趋势分析与季节尺度的干旱监测(刘冀等,2020)。各数据时段均为2003年1月—2017年12月,基本信息见表1。

表1 9种原始降水产品基本信息介绍

Table 1 Basic information of nine precipitation products

降水产品	空间分辨率/(°)	时间分辨率	数据来源
TRMM	0.25	月	https://disc.gsfc.nasa.gov/
GPM	0.1	月	https://disc.gsfc.nasa.gov/
PERSIANN	0.25	月	https://chrsdata.eng.uci.edu
PERSIANN-CDR	0.25	月	https://chrsdata.eng.uci.edu
PERSIANN-CCS	0.04	月	https://chrsdata.eng.uci.edu
CMORPH	0.25	天	https://rda.ucar.edu/
ERA5-Land	0.1	月	https://cds.climate.copernicus.eu/
ERA5	0.25	月	https://cds.climate.copernicus.eu/
CHIRPS	0.05	月	https://data.chc.ucsb.edu/products/CHIRPS-2.0/

1.2.2 其他数据

本文还使用中国气象数据网发布的中国地面气象站基本气象要素日值数据集V3.0,该数据集经过严格的质量控制,经处理将日数据累加到月数据,参与评估检验的总计650个站点,时间跨度与降水产品一致,从2003年1月—2017年12月,总计180个月数据。DEM数据来自于中科院资源环境科学与数据中心,该数据集基于最新的SRTM V4.1数据生成。

1.2.3 数据预处理

针对原始降水产品不同的产品数据格式(nc格式、tif格式等),采用相应的软件进行处理。为减少插值带来额外误差和保持数据的原始性及空间连续性,利用ArcGIS软件提取站点位置所在的格点降水值作为降水产品在该点的数值。批量提取降水产品值至地面站点,然后利用Python对相关系数、偏差和均方根误差等指标进行计算。

1.3 方法介绍

1.3.1 机器学习算法

为了比较多种机器学习算法产品的精度,本文选取多元线性回归、随机森林和XGBoost 3种算法进行降水产品的融合。

多元线性回归(Multiple Linear Regression, MLR)是线性回归分析中的重要统计方法,被广泛应用于地理、经济、社会等众多领域研究中。在实际研究中,多元线性回归算法模型不能局限于单变量,因此本文建立多种降水产品和辅助变量的多元线性回归算法模型。

随机森林(Random Forest, RF)模型将多棵分类与回归树组合,继承和发展了分类回归树的优点。由于

随机森林在构建过程中采用了自举法重采样技术,因而在运算量没有显著提高的情况下提高了预测精度(沈润平等,2017)。相比较单棵回归树一定程度上的不稳定问题,随机森林具有更稳定和更强的预测能力。

XGBoost算法在2016年提出,XGBoost为一种改进集成树算法,从梯度下降树算法改进与扩展而来。其特点在于对目标函数的损失函数增加了额外正则化项,有助于平滑学习权重以避免过拟合。此外,还使用一阶、二阶梯度函数来优化损失函数,以降低算法复杂度(王浩钰等,2021)。

在进行机器学习训练时,不同算法模型对应的参数对结果有很大的影响,选取最适合的参数组合进行模型构建具有重要意义。因为TRMM和CHIRPS数据不能覆盖50°N以北的区域,因此本文只针对50°N以南区域进行数据融合。首先对样本数据进行读取和划分,将研究时间段内50°N以南气象站点根据空间均匀采样算法划分为训练站和测试站,划分比例为7:3,并以此分别构建训练集和测试集样本;其中训练集只参与融合模型建立、训练以及参数调优,测试集不参与训练并作为独立样本对经过参数调优后的模型进行检验。其次基于GridSearchCV网格搜索函数结合训练集数据并采用十折交叉验证方法寻找最优的模型参数组合,以替代传统的将样本划分训练集和验证集进行训练和验证,从而调校及选取模型参数的方法。然后,利用独立样本构建的测试集对训练出的机器学习算法的最佳参数与变量组合进行检验对比。

通过以上算法进行降水产品融合,详细流程见图1。

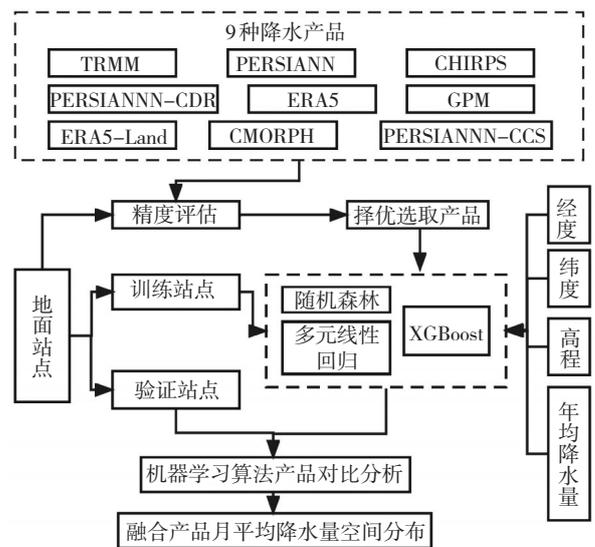


图1 多源降水产品融合流程图

Fig.1 Multi-source precipitation product fusion flow chart

1.3.2 评估方法

降水产品的精度采用相关系数R、偏差BIAS(BIAS)、均方根误差RMSE(R_{MSE})三个指标评估。

相关系数 R 是描述降水产品数据和地面站点实测值之间线性关系的指标, 相关系数越高表明降水站点与站点实测值相关性越强, 其计算公式如下

$$R = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (1)$$

均方根误差 $RMSE (R_{MSE})$ 反映降水产品与地面站点实测值之间的偏离程度, 较低的均方根误差代表数据越准确, 其计算公式如下

$$R_{MSE} = \sqrt{\frac{\sum_{i=1}^n (X_i - Y_i)^2}{n}} \quad (2)$$

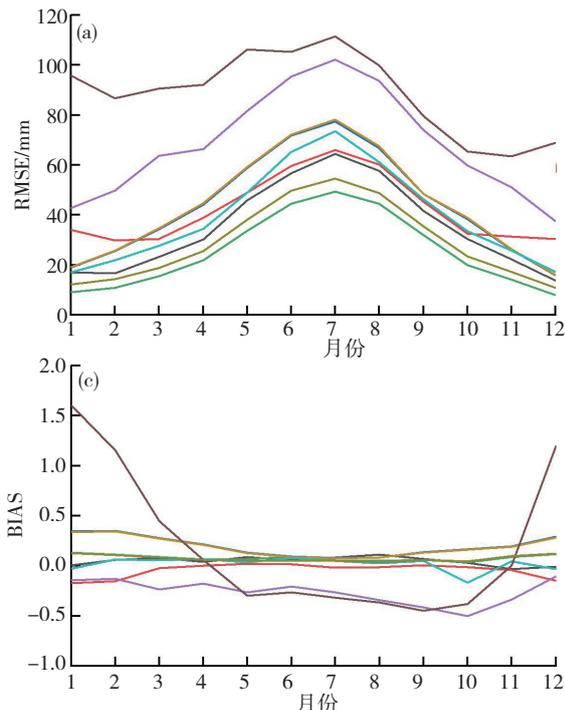
偏差 $BIAS (B_{IAS})$ 用于评价降水产品与站点测值的偏离程度, 值越接近于 0 代表降水产品反演更精确, 其计算公式如下

$$B_{IAS} = \frac{\sum_{i=1}^n (X_i - Y_i)}{\sum_{i=1}^n (Y_i)} \quad (3)$$

式(1)、(2)、(3)中, X_i, Y_i 分别为降水产品估计值和地面站点实测值, n 为样本个数。 \bar{X} 和 \bar{Y} 分别是降水产品估计的均值和地面站点实测值均值。

2 原始月尺度降水产品精度评估

利用地面站点观测降水数据, 对 9 种降水产品在



月尺度上进行精度检验, 结果如表 2 所示。除 PERSIANN 和 PERSIANN-CCS 外, 其余产品相关系数普遍在 0.8 以上, 说明降水产品与实测数据具有良好的一致性, 并且其中以 GPM 数据为最优。而在均方根误差方面, 除 PERSIANN 和 PERSIANN-CCS 外, 其余产品误差均低于 50 mm, 其中 GPM 数据表现出最优。在偏差方面, 九种产品呈现高估或低估状态; 其中 TRMM、GPM、PERSIANN-CDR、PERSIANN-CCS、ERA5、ERA5-Land、CHIRPS 整体高估了降水, 而 PERSIANN、PERSIANN-CCS 与 CMORPH 对降水存在一定的低估。

表 2 不同降水产品的精度评估指标

Table 2 The accuracy evaluation indexes of different precipitation products

降水产品	R	RMSE/mm	BIAS
TRMM	0.937 5	32.72	0.057 9
GPM	0.950 9	29.04	0.065 0
PERSIANN	0.676 7	71.18	-0.278 8
PERSIANN-CDR	0.883 9	43.54	0.043 7
PERSIANN-CCS	0.389 2	90.17	-0.123 5
CMORPH	0.883 3	44.19	-0.013 5
ERA5-Land	0.867 6	48.26	0.144 1
ERA5	0.864 9	48.67	0.139 1
CHIRPS	0.912 5	38.92	0.067 0

注: R 为相关系数, $BIAS$ 为偏差, $RMSE$ 为均方根误差, 下同

各降水产品逐月均方根误差呈正态分布(图 2), 误差最大出现在 7 月; PERSIANN 和 PERSIANN-CCS 数据各月的均方根误差明显高于其余产品。从相关

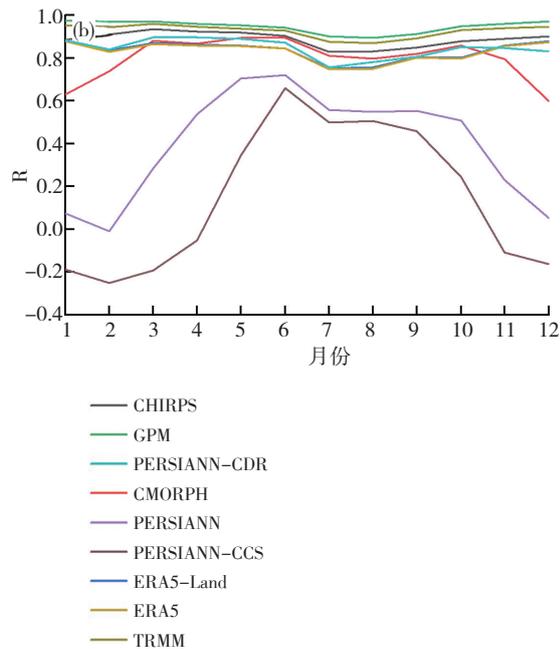


图 2 不同降水产品逐月均方根误差 RMSE(a)、相关系数 R(b)、偏差 BIAS(c)演变图

Fig.2 Monthly evolution chart of (a) root mean square error, (b) correlation coefficients, (c) relative deviation of different precipitation products

系数R来看,除PERSIANN和PERSIANN-CCS数据,其余降水产品均表现出很好的相关性;而PERSIANN和PERSIANN-CCS数据在部分月份出现了负相关,说明这两套产品精度相比较其他产品很差。在BIAS方面,PERSIANN和PERSIANN-CCS产品相对于其他产品,在不同月份波动较大,且偏离程度较高;其余产品相对偏差范围在-0.5~0.5之间。通过对9种降水产品的精度分析,发现相对于其他7种产品,PERSIANN与PERSIANN-CCS产品误差明显较大,这可能与该数据固有的算法和系统误差有关,这与慎璐璐(2022)和Zhang(2018)的结论一致。而CMORPH数据是9种降水产品中较低估降水的产品,且在冬季误差较大,这与夏昕然等(2022)、任阳阳(2022)的结论一致。

为了进一步探究海拔高度和干湿分布对月尺度降水产品的影响,以各站点精度指标和对应位置海拔高度进行相关分析,如表3。结果显示,大部分产品相关系数均与高程呈现负相关,其中CHIRPS、GPM降水产品受海拔影响不大。和其他指标比较,均方根误差与海拔关系最敏感,大部分产品呈现强烈的负相关。在偏差方面,ERA5系列受海拔影响最大,其余产品除CHIRPS、CMORPH和TRMM之外,呈较显著的上升趋势。降水产品受海拔高度影响较大,这一现象是合理的,高海拔地区由于地形复杂,降水受到下垫面和局部地形地貌等众多因素的制约,所以对该地区降水精度的影响较大(金晓龙等,2016)。

表3 不同降水产品的精度评估指标和海拔高度的相关性

Table 3 The correlation of accuracy evaluation indexes of different precipitation products and altitude

降水产品	r_{R-H}	r_{RMSE-H}	r_{BIAS-H}
TRMM	-0.116 4**	-0.297 1**	0.070 8
GPM	-0.052 2	-0.313 6**	0.078 2*
PERSIANN	-0.419 5**	-0.163 1**	0.305 1**
PERSIANN-CDR	-0.021 4	-0.188 1**	0.233 4**
PERSIANN-CCS	-0.259 9**	-0.011 1	0.197 4**
CMORPH	-0.177 9**	-0.250 2**	0.064 8
ERA5-Land	0.079 1*	-0.126 6**	0.552 8**
ERA5	0.078 9*	-0.136 4**	0.527 2**
CHIRPS	0.0064	-0.381 2**	0.076 6

注: *表示通过了0.05显著水平的统计检验, **表示通过了0.01显著水平的统计检验,下同。 r_{R-H} 代表相关系数和海拔高度的相关; r_{RMSE-H} 代表均方根误差和海拔高度的相关; r_{BIAS-H} 代表均方根误差和海拔高度的相关

干湿分布与年均降水量关系密切,计算各站点在研究时间段内的年均降水量,以各站点精度指标和对应站点的年均降水量进行相关性分析(表4),可见

产品相关系数均与年均降水量呈现正相关,9种产品的均方根误差呈现强正相关,且相比相关系数和偏差,与年均降水量关系最敏感。偏差方面,因为干旱区降水量偏小,降水产品对低雨量探测误差较大,使得偏差的误差较大。总体来看,随着年均降水量的减少,降水产品的精度逐渐下降,这一现象是合理的。因为干旱地区的大气水汽含量普遍较低,降水事件的持续时间和范围通常小于湿润地区,因此更难被卫星传感器探测到(Fang et al., 2019)。

表4 不同降水产品的精度评估指标和年均降水量的相关性

Table 4 The correlation of accuracy evaluation indexes of different precipitation products and average annual precipitation

降水产品	r_{R-R}	r_{RMSE-R}	r_{BIAS-R}
TRMM	0.312 9**	0.822 3 **	-0.246 6**
GPM	0.226 4**	0.821 4**	-0.345 8**
PERSIANN	0.520 9**	0.920 6**	-0.498 6**
PERSIANN-CDR	0.207 0**	0.866 7**	-0.372 8**
PERSIANN-CCS	0.778 1**	0.576 5**	-0.472 3**
CMORPH	0.387 5**	0.708 0**	-0.264 8**
ERA5-Land	0.067 2	0.815 6**	-0.366 8**
ERA5	0.075 4	0.810 4**	-0.352 3**
CHIRPS	0.267 3**	0.870 7**	-0.115 2**

注: r_{R-R} 代表相关系数和年均降水量的相关; r_{RMSE-R} 代表均方根误差和年均降水量的相关; r_{BIAS-R} 代表均方根误差和年均降水量的相关”。

综上不同降水产品与降水量实测数据存在相关性,但在空间上误差分布较大,受海拔高度和干湿分布影响严重。此外,虽然部分数据产品具有良好的相关性,如GPM数据,但存在高估现象;而部分产品虽然相关性略差,但存在较低的偏差及低估现象,如CMORPH产品。因此展开多源降水产品融合,以此弥补单一降水产品的缺陷并提高降水产品精度十分必要。经过前文的产品评估比较,GPM、TRMM、PERSIANN-CDR、CHIRPS、CMORPH 5种降水产品相比较其余4种产品具有较高的精度。而ERA5系列产品精度较低以及PERSIANN和PERSIANN-CCS两套数据存在较大的误差,因此不参与后文的融合研究。

3 机器学习算法产品分析

3.1 多种机器学习算法模型设计

由于各降水产品存在误差,因此需要开展降水产品融合研究以提高精度,选择精度较好的GPM、TRMM、PERSIANN-CDR、CHIRPS、CMORPH 5种降水产品进行融合算法研究;为了研究降水产品和地理因素的组合是否比仅使用降水产品或部分地理因素能

更好地估计降水量的空间分布,根据已有的研究(傅抱璞,1992;史岚等,2019;吕爱锋和亓珊珊,2022)并结合前文的误差分析,加入观测站所在位置的高程、经度、纬度、年均降水量作为辅助变量。表5设计了5种机器学习变量组合模型。

表5 5种变量组合模型
Table 5 Five variable combination model

模型	变量组合
Model1	GPM、TRMM、CHIRPS、PERSIANN-CDR、CMORPH的降水量及经度、纬度、高程、年均降水量
Model2	GPM、TRMM、CHIRPS、PERSIANN-CDR、CMORPH的降水量
Model3	GPM、TRMM、CHIRPS、PERSIANN-CDR、CMORPH的降水量及经度、纬度
Model4	GPM、TRMM、CHIRPS、PERSIANN-CDR、CMORPH的降水量及高程
Model5	GPM、TRMM、CHIRPS、PERSIANN-CDR、CMORPH的降水量及年均降水量

表6展现了由独立样本构建的测试集对3种算法针对不同的变量组合模型的精度检验对比结果,所有模型均采用由GridSearchCV网格搜索函数得到对应的最优模型参数组合所构建。由表6可知,3种算法不同的变量组合模型中,Model1均拥有最好的精度,说明考虑多重地理因素得出的模型比考虑地理因素较少的Model2—5得出的估算结果更准确,这表明纳入地理因素可提供更准确的预测结果。由表6结合表7比较可知,经过独立样本构建的测试集的检验,3种算法不同变量组合模型的精度指标相对于原始的5套产品均有提升,说明经过模型融合后的降水数据,相对于原始5种产品,具有更显著的相关性;XGBoost算法的Model1变量组合给出了最优的精度,相关系数达到0.9517,均方根误差减少到29.52mm,偏差降低到-0.0005;此外,随机森林算法和多元线性回归算法的Model1变量组合,相关系数分别达到0.9490和0.9498,均方根误差分别减少到30.28mm和30.09mm,偏差分别降低到-0.0004和-0.0023,精度有明显提升,只是效果相比较XGBoost算法略差。

3.2 各算法最优模型产品的检验对比分析

首先比较各月3种算法最优模型与原始产品的精度差异,表8为数据融合的5种原始降水产品与3种机器学习算法最优模型的各月份精度指标对比。从相关系数来看,3种算法相关系数相比较原始产品有较大提升,大部分月均在0.9以上,体现了融合后的数据与实测降水有更好的相关性。从均方根误差来看,呈现夏季大,冬季少的特点。3种算法的均方根误差明

表6 多种机器学习算法结果精度对比

Table 6 Comparison of results of multiple machine learning algorithms				
机器学习算法	变量组合模型	R	RMSE/mm	BIAS
XGBoost	Model1	0.9517	29.52	-0.0005
	Model2	0.9492	30.32	-0.0153
	Model3	0.9488	30.44	-0.0148
	Model4	0.9491	30.34	-0.0132
	Model5	0.9494	30.22	-0.0110
随机森林	Model1	0.9490	30.28	-0.0004
	Model2	0.9472	30.83	-0.0071
	Model3	0.9474	30.78	-0.0081
	Model4	0.9475	30.74	-0.0053
	Model5	0.9487	30.40	-0.0080
MLR	Model1	0.9498	30.09	-0.0023
	Model2	0.9490	30.34	-0.0077
	Model3	0.9490	30.33	-0.0078
	Model4	0.9489	30.34	-0.0065
	Model5	0.9490	30.32	-0.0072

注:每种机器学习算法的Model1是此种算法的最优模型。

表7 各算法最优模型产品与原始产品精度对比

Table 7 Accuracy comparison between the optimal model of each algorithm and the original product

产品类型	产品名称	精度指标		
		R	RMSE/mm	BIAS
原始产品	TRMM	0.9385	33.44	0.0481
	GPM	0.9449	31.74	0.0562
	PERSIANN-CDR	0.8849	44.93	0.0332
	CHIRPS	0.9150	39.73	0.0686
	CMORPH	0.8984	42.56	-0.0313
机器学习	随机森林	0.9490	30.28	-0.0004
	XGBoost	0.9517	29.52	-0.0005
	MLR	0.9498	30.09	-0.0023

显降低,均小于原始产品。在BIAS方面,3种算法在夏季偏差很小几乎为零,冬季最大,算法有效的抑制原始产品高估或低估的现象,且相对于原始产品,12个月都表现十分稳定,说明机器学习算法反演得到的降水更加合理。此外,3种算法中,多元线性回归要次于随机森林和XGBoost,其中在降水丰富的夏季,XGBoost精度最优,而随机森林在降水缺少的冬季,拥有更高的精度,说明不同的算法在各月具有差异。综上所述,在月尺度上,融合产品好于原始产品。

表9进一步比较了三种机器学习算法最优模型产品与参与数据融合的五种原始产品在不同高度区域的精度指标,将海拔高度划分为五个区域,可以看到三种算法均体现出较好的精度,误差较小。三种算法

表8 各月份不同原始降水产品与最优模型融合产品精度对比

Table 8 Comparison of precision between different precipitation products and algorithm optimal models in different months

精度指标	月份	原始降水产品					最优模型融合产品		
		PERSINANN-CDR	CMORPH	CHIRPS	TRMM	GPM	XGBoost	RF	MLR
RMSE/mm	1月	17.12	23.33	17.72	12.75	9.89	9.08	9.12	9.84
	2月	22.48	23.33	16.27	14.07	11.29	10.42	10.19	11.17
	3月	27.42	26.09	22.70	18.76	15.70	15.03	14.68	14.94
	4月	34.81	33.72	29.45	25.17	22.33	21.52	21.34	20.98
	5月	50.78	49.84	44.47	37.69	35.78	33.14	33.44	33.53
	6月	69.25	63.54	59.91	52.76	50.04	47.46	47.92	47.79
	7月	75.48	66.95	67.18	56.85	54.94	50.17	52.93	51.96
	8月	63.47	61.75	60.18	49.98	49.14	45.09	46.80	46.12
	9月	46.65	43.86	41.60	35.16	33.99	31.13	32.27	32.08
	10月	35.21	32.08	29.71	22.76	21.95	21.72	21.59	21.34
	11月	25.77	25.91	22.33	16.88	14.41	13.75	13.75	14.19
	12月	16.90	19.74	13.28	10.02	7.78	6.96	6.78	7.81
R	1月	0.889 7	0.802 5	0.883 5	0.950 9	0.972 8	0.973 2	0.972 7	0.969 3
	2月	0.837 1	0.827 4	0.914 4	0.943 6	0.964 7	0.965 6	0.967 0	0.962 8
	3月	0.898 7	0.909 0	0.936 6	0.958 1	0.969 8	0.970 5	0.971 7	0.971 2
	4月	0.895 3	0.900 5	0.927 0	0.947 5	0.958 3	0.960 1	0.960 8	0.962 2
	5月	0.887 8	0.893 4	0.924 4	0.940 5	0.947 0	0.953 4	0.952 6	0.952 4
	6月	0.871 7	0.892 3	0.905 7	0.927 1	0.936 0	0.941 9	0.940 6	0.941 6
	7月	0.763 4	0.821 0	0.826 6	0.874 1	0.883 5	0.902 3	0.890 6	0.895 2
	8月	0.787 1	0.806 6	0.841 2	0.874 9	0.881 1	0.898 4	0.890 1	0.893 7
	9月	0.820 1	0.846 5	0.867 4	0.903 7	0.910 6	0.923 7	0.917 8	0.919 1
	10月	0.851 9	0.874 5	0.895 8	0.939 5	0.944 4	0.945 1	0.945 6	0.947 7
	11月	0.853 5	0.855 6	0.894 1	0.945 3	0.958 7	0.960 1	0.960 1	0.957 7
	12月	0.836 1	0.784 6	0.903 7	0.953 8	0.972 5	0.974 1	0.975 4	0.969 0
BIAS	1月	-0.0436	-0.269 6	-0.011 7	0.116 9	0.130 6	0.043 8	0.037 4	0.039 4
	2月	0.0654	-0.183 8	0.049 5	0.086 8	0.106 0	0.007 3	0.004 3	-0.006 2
	3月	0.0571	-0.048 6	0.071 5	0.078 8	0.079 4	0.007 3	0.001 7	0.001 8
	4月	0.0637	-0.021 8	0.042 4	0.047 8	0.054 9	-0.009 0	-0.011 1	-0.016 7
	5月	0.0266	0.001 5	0.088 6	0.037 1	0.052 8	0.000 6	-0.000 1	0.003 5
	6月	0.0548	-0.006 4	0.044 4	0.027 0	0.040 1	-0.011 6	-0.009 1	-0.013 7
	7月	0.0378	-0.025 9	0.079 5	0.039 5	0.043 8	0.002 6	0.005 0	-0.002 5
	8月	0.0266	-0.008 3	0.139 7	0.038 6	0.055 9	0.006 3	0.009 8	0.013 4
	9月	0.0344	0.003 9	0.092 7	0.055 2	0.055 3	0.001 2	0.000 2	0.002 3
	10月	-0.0404	-0.040 1	0.015 1	0.031 0	0.025 2	-0.030 9	-0.031 8	-0.039 2
	11月	0.0376	-0.051 0	-0.037 3	0.094 7	0.085 1	0.000 8	-0.000 8	-0.003 0
	12月	-0.0385	-0.214 6	0.003 3	0.121 6	0.122 3	0.023 1	0.012 6	0.017 6

注:颜色越淡,指标越好,加粗斜体为最优,下同。

表9 不同海拔区域降水产品与算法最优模型精度指标对比

Table 9 Comparison of precision indicators for precipitation products and algorithm optimal models in different altitude regions

精度指标	区域划分/m	原始降水产品					最优模型融合产品		
		PERSIANN-CDR	CMORPH	CHIRPS	TRMM	GPM	XGBoost	RF	MLR
RMSE/mm	0~199	52.02	50.70	47.15	39.87	36.06	34.20	34.43	34.18
	200~499	40.68	38.35	36.62	30.63	28.42	26.71	26.98	27.22
	500~999	42.47	37.81	36.66	31.65	31.12	28.26	29.75	30.36
	1000~1499	26.75	27.57	22.22	19.19	17.58	16.86	17.15	17.02
	≥1499	43.54	37.54	35.70	28.84	32.84	28.69	31.21	30.20
	R	0~199	0.881 3	0.888 6	0.906 3	0.932 2	0.946 4	0.949 8	0.949 0
200~499		0.902 8	0.915 9	0.925 3	0.950 0	0.955 3	0.959 3	0.958 5	0.957 7
500~999		0.887 9	0.911 1	0.912 1	0.936 7	0.942 8	0.949 7	0.946 5	0.946 6
1000~1499		0.904 6	0.898 3	0.936 8	0.952 1	0.960 2	0.962 9	0.961 6	0.962 4
≥1499		0.840 3	0.878 6	0.907 3	0.929 5	0.907 8	0.931 0	0.919 3	0.922 4
BIAS		0~199	0.060 2	-0.016 8	0.072 0	0.055 1	0.081 3	0.001 4	0.005 5
	200~499	0.012 4	-0.031 4	0.066 4	0.083 0	0.062 0	-0.001 5	-0.007 3	0.005 7
	500~999	-0.090 4	-0.109 5	0.035 6	-0.015 5	-0.035 8	-0.035 3	-0.062 5	-0.071 7
	1000~1499	-0.017 7	0.013 8	0.052 5	0.045 5	0.054 0	-0.001 5	0.004 3	-0.027 6
	≥1500	0.078 9	-0.052 7	0.092 1	0.018 6	0.022 6	0.019 1	0.029 7	0.028 5

在不同海拔区域相关系数均在0.9以上,偏差较稳定,未出现极高极低值。表10反映了5种原始降水产品与3种机器学习算法最优模型产品在不同年均降水量区域的精度对比,按照年降水量将中国划分为4个区域(干旱区、半干旱区、半湿润区和湿润区);可以看到大部分区域3种机器学习算法产品均能体现出较好

精度,其中在年均降水量0~200 mm区域,虽然相关系数有所降低,但是偏差精度提升明显,有效的降低原始产品的高偏差。综上,和5种原始产品比较,3种机器学习算法最优模型产品在不同区域均体现出更好的精度,3种机器学习算法中又以XGBoost效果最优。

表10 不同年均降水量区域降水产品与算法最优模型精度指标对比

Table 10 Comparison of precipitation products and algorithm optimal model precision indicators in different annual precipitation regions

精度指标	年均降水量 /mm	原始降水产品					最优模型融合产品		
		PERSIANN-CDR	CMORPH	CHIRPS	TRMM	GPM	XGBoost	RF	MLR
RMSE/mm	0~199	17.72	14.31	16.68	15.28	22.08	16.73	19.96	20.06
	200~399	26.91	23.65	17.74	16.42	13.16	11.38	11.71	11.70
	400~799	30.28	29.94	28.44	23.46	23.25	21.27	22.83	21.93
	≥800	57.16	53.96	50.42	42.43	39.61	37.30	37.68	37.72
R	0~199	0.635 7	0.562 1	0.558 2	0.665 4	0.628 7	0.597 0	0.556 2	0.617 8
	200~399	0.786 1	0.795 4	0.879 1	0.904 2	0.928 8	0.943 3	0.939 8	0.939 8
	400~799	0.865 4	0.863 9	0.880 4	0.919 1	0.924 4	0.930 6	0.921 1	0.926 8
	≥800	0.863 9	0.884 0	0.900 0	0.927 6	0.937 5	0.9444	0.943 2	0.943 4
BIAS	0~199	0.940 6	0.428 8	0.525 7	0.629 0	0.958 2	0.364 3	0.432 8	0.598 9
	200~399	0.229 6	0.137 7	0.124 4	0.152 9	0.091 4	-0.008 0	0.010 6	0.001 8
	400~799	0.134 3	0.027 2	0.079 0	0.084 5	0.110 9	0.016 7	0.021 8	0.005 2
	≥800	-0.012 2	-0.059 6	0.058 9	0.027 8	0.031 1	-0.008 4	-0.011 2	-0.010 4

3.3 特征重要性分析

机器学习的特征重要性分析一定程度上可以反映输入变量对模型的影响大小,进一步增加模型的可解释性,因此利用XGBoost算法的Model1组合进行模型特征重要性分析。图3反映了各特征变量在模型中的F分数情况,F分数反映模型中变量对于模型的贡献度大小,分值越大,说明该变量对于模型影响越大。F变量总体排名依次为:GPM,CHIRPS,PERSIANN-CDR,TRMM,高程,CMORPH,年均降水量,纬度,经度。在5种降水产品中,GPM的重要性最高,说明GPM数据对模型的影响较大,这与前文GPM数据具有较好的精度具有一定的关系。在辅助变量中,高程与年均降水量对模型贡献较大,说明地形和干湿区域的分布对降水量具有一定的影响,未来可以对其进行进一步的研究;相比较而言,经度与纬度的重要性较低。

3.4 月平均降水量空间分布

由前文分析可知,XGBoost算法拥有最好的精度,因此以此算法为基础,将模型输入变量统一为0.05°分辨率进行降水产品的融合,生成0.05°分辨率的降水空间分布。图4展示了采用XGBoost算法融合的月平均降水量中国大陆区域和闽浙赣地区的空间分布、原始

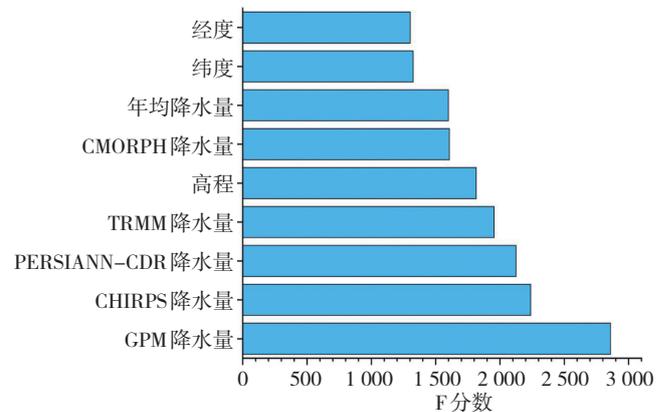


图3 XGBoost模型中特征变量F分数

Fig.3 F-score of characteristic variable in XGBoost model

GPM数据的月平均降水量的空间分布和利用气象站点数据进行IDW插值的月平均降水量的空间分布。从空间分布规律来看,我国各月降水量由东南逐渐向西北递减,这与海陆位置有关;而在青藏高原东南部,由于受西南暖湿气流影响,降水量偏多,3种数据其他区域均呈现出大致相似的降水分布。在局部细节上,因为算法考虑了地理要素,在地形复杂的地区如闽浙赣地区(图4d,e,f);相比较IDW插值带来的点状性与原始GPM数据的低分辨率,经过XGBoost算法融合后的数据具有更多的降水空间变化和局部降水细节信息。

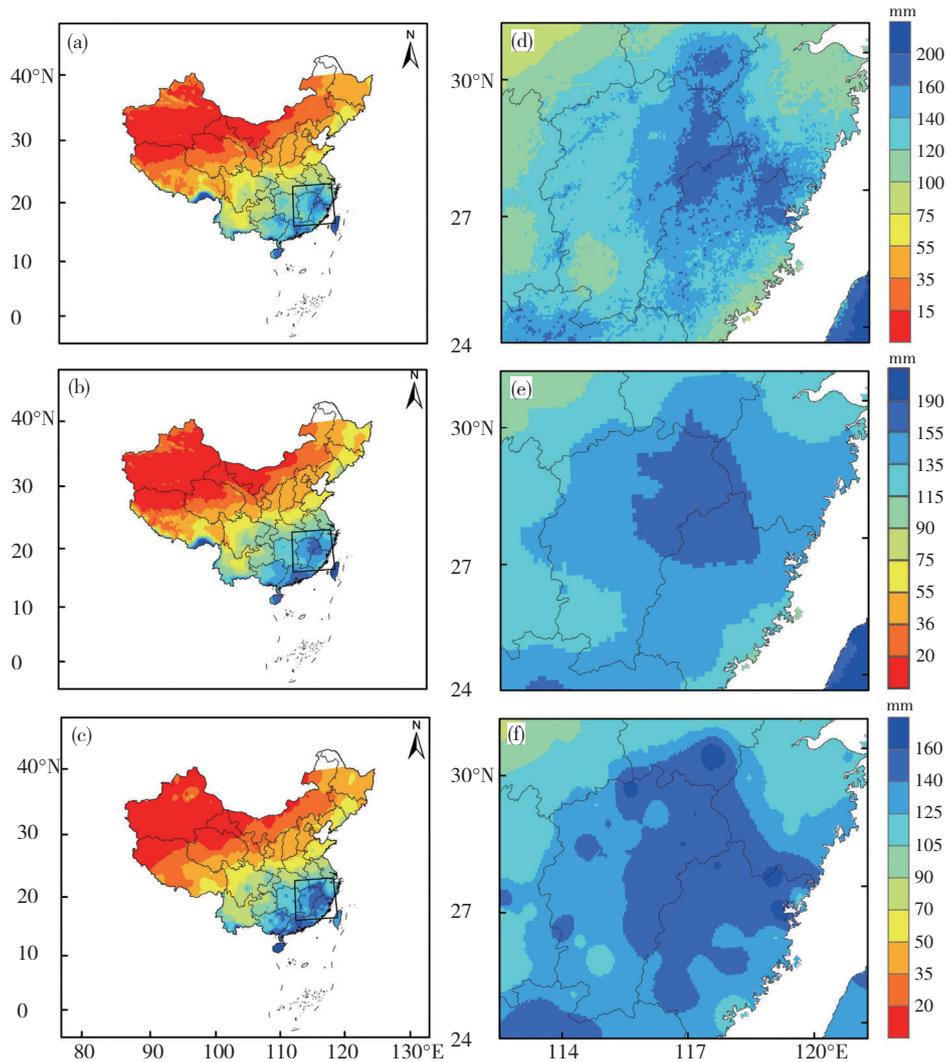


图4 2003—2017年中国大陆区域(a—c)和闽浙赣地区(d—f,见图 a—c 中黑框) XGBoost算法(a,d)、原始GPM (b,e)和气象站点IDW插值(c,f)月平均降水量空间分布图

Fig.4 Spatial distribution map of monthly average precipitation of (a, d) XGBoost algorithm, (b, e) original GPM and (c, f) meteorological station IDW interpolation in (a—c) Chinese mainland region and (d—f, the black flame in map a—c) Fujian Zhejiang Jiangxi region from 2003 to 2017

4 结论与讨论

本文选取2003—2017年的气象站点降水观测数据,评估中国大陆地区9种栅格降水产品的精度并进行误差对比分析,然后利用机器学习算法进行数据融合。得出主要结论如下:

(1) 中国大陆地区的9种降水产品均能很好地表现降水的空间分布特征,但是精度相差较大。其中GPM、TRMM、CMORPH、PERSIANN-CDR、CHIRPS这五种产品较好。ERA5系列次之,PERSIANN和PERSIANN-CCS产品精度最差。

(2) 不同的月份,除PERSIANN和PERSIANN-CCS产品外,其余产品与观测降水的相关性大部分能达到0.8以上;均方根误差则呈现夏季大,冬季小的特点;而所有的降水产品在冬季普遍出现偏差较大的情况。

此外,降水产品受高程和干湿分布影响较大,平原和湿润地区误差较小,高海拔和干旱地区误差增大。

(3) 经过3种机器学习算法融合后,最优的XGBoost算法模型相关系数提升,均方根误差和偏差降低。在不同月份,3种算法产品也明显优于原始产品;其中,XGBoost算法在夏季较好,而随机森林算法在冬季较优。在不同的区域,3种算法也明显优于原始产品,XGBoost算法在不同的海拔区域相关系数均最优,在不同的干湿区域也呈现较好的效果;随机森林与多元线性回归相比较于XGBoost算法则精度较差。

(4) 3种算法有效的消除了各种降水产品之间高估和低估的现象,并以XGBoost算法最优。机器学习算法考虑了其中辅助变量的非线性关系,使机器学习融合产品结果更加符合实际。融合产品在空间上与气象站点插值产品空间趋势一致,且在地形复杂区域

(如闽浙赣地区),具有更多的空间变化并且能够捕捉更多的降水细节信息。

经机器学习算法融合后的产品精度高于原始产品,但是仍然存在一定的误差,总体呈现微弱低估现象。未来研究将依据以下2个方面,首先综合考虑地理环境的影响,可尝试加入云量,水汽等地理要素,其次开展降水相关的不确定性分析,将不确定性分析引入到后处理研究中能增加后处理产品的可信度(熊立华等,2021)。未来将对以上问题进行更加深入的研究,将生成的月降水数据进行业务应用,并尝试日甚至小时数据的融合。

参考文献(References):

- 丁光旭,郭家力,汤正阳,等.2022.多种降水再分析数据在长江流域的适用性对比[J].人民长江,53(9):72-79. Ding G X, Guo J L, Tang Z Y, et al. 2022. Evaluation on applicability of various reanalysis datasets of precipitation in Changjiang River Basin [J]. Yangtze River,53(9):72-79 (in Chinese). doi:10.16232/j.cnki.1001-4179.2022.09.012
- 杜懿,王大洋,张智,等.2021. GPM IMERG 降水产品在珠江流域的适用性分析[J].水文,41(6):1-6. Du Y, Wang D Y, Zhang Z, et al. 2021. Applicability of GPM IMERG satellite precipitation product in Pearl River Basin [J]. Journal of China Hydrology,41(6):1-6 (in Chinese). doi: 10.19797/j.cnki.1000-0852.20200318
- 傅抱璞.1992.地形和海拔高度对降水的影响[J].地理学报,47(4):302-314. Fu B P. 1992. The effects of topography and elevation on precipitation [J]. Acta Geographica Sinica,47(4):302-314 (in Chinese). doi:10.11821/xb.199204002
- 甘富万,陈帅,胡秀英,等.2022. IMERG 及 TMPA 降水卫星产品在红河流域的适用性评估[J].水电能源科学,40(7):5-9. Gan F W, Chen S, Zhang Z, et al. 2022. Applicability evaluation of IMERG and TMPA satellite precipitation products in Honghe River Basin [J]. Water Resources and Power,40(7):5-9 (in Chinese). doi:10.20040/j.cnki.1000-7709.2022.20211937
- 金晓龙,邵华,张弛,等.2016. GPM 卫星降水数据在天山山区的适用性分析[J].自然资源学报,31(12):2074-2085. Jin X L, Shao H, Zhang C, et al. 2016. The applicability evaluation of three satellite products in Tianshan Mountains [J]. Journal of Natural Resources,31(12):2074-2085 (in Chinese). doi:10.11849/zrzyxb.20160057
- 李傲,高媛,崔春光. 2022. 暴雨灾害预警信息传播渠道及社会管理机制分析[J]. 暴雨灾害,41(3):348-354. Li A, Gao Y, Cui C G. 2022. Analysis on the dissemination way and social management strategy of early warning information of rainstorm disaster [J]. Torrential Rain and Disasters,41(3):348-354(in Chinese). doi:10.3969/j.issn.1004-9045.2022.03.011
- 李强宇,楚岱蔚,朱晓晨.2021.长三角地区常用遥感降水产品质量评估[J].科学技术与工程,21(12):4801-4810. Li Q Y, Chu D W, Zhu X C. 2021. Study on the quality evaluation of precipitation products commonly used by remote sensing in the Yangtze River Delta [J]. Science Technology and Engineering,21(12):4801-4810 (in Chinese)
- 黎跃勇,周威,李好,等. 2022. 基于优势分析法的暴雨综合致灾指数及阈值研究——以张家界市保险业为例[J]. 暴雨灾害,41(2):232-239. Li Y Y, Zhou W, Li H, et al. 2022. Study on torrential rain comprehensive disaster-causing index and its threshold based on Dominance Analysis Method: A case of insurance industry in Zhangjiajie City [J]. Torrential Rain and Disasters,41(2):232-239 (in Chinese). doi:10.3969/j.issn.1004-9045.2022.02.014
- 李运龙,熊立华,闫磊.2017.基于地理加权回归克里金的降水数据融合及其在水文预报中的应用[J].长江流域资源与环境,26(9):1359-1368. Li Y L, Xiong L H, Yan L. 2017. A geographically weighted regression kriging approach for TRMM-rain gauge data merging and its application in hydrological forecasting [J]. Resources and Environment in the Yangtze Basin,26(9):1359-1368 (in Chinese). doi:10.11870/cjlyzyy-hj201709008
- 刘冀,魏榕,张特,等.2020.基于CHIRPS卫星降水的雅砻江流域干湿时空演变分析[J].长江流域资源与环境,29(8):1800-1811. Liu J, Wei R, Zhang T, et al. 2020. Spatial and temporal evolution characteristics of dry and wet condition in Yalongjiang River Basin based on the CHIRPS satellite precipitation [J]. Resources and Environment in the Yangtze Basin,29(8):1800-1811 (in Chinese). doi:10.11870/cjlyzyy-hj202008012
- 刘少军,蔡大鑫,韩静,等.2021.卫星遥感反演降水研究进展综述[J].气象科技进展,11(1):28-33. Liu S J, Cai D X, Han J, et al. 2021. Progress of the satellite remote sensing retrieval of precipitation [J]. Advances in Meteorological Science and Technology,11(1):28-33 (in Chinese). doi: 10.3969/j.issn.2095-1973.2021.01.005
- 刘松楠,汪君,王会军.2022.高分辨率卫星对“21·7”河南特大暴雨监测能力分析[J].气象学报,80(5):765-776. Liu S N, Wang J, Wang H J. 2022. Analysis of the monitoring ability of high-resolution satellites for the "21·7" heavy rain in Henan [J]. Acta Meteorologica Sinica,80(5):765-776 (in Chinese). doi:10.11676/qxxb2022.053
- 吕爱锋,尹珊珊.2022.遥感及再分析降水产品在缺资料干旱内陆盆地的适用性评估[J].地球信息科学学报,24(9):1817-1834. Lv A F, Qi S S. 2022. Applicability analysis of satellite-based and reanalysis precipitation products in poorly-gauged arid inland basins [J]. Journal of Geo-information Science,24(9):1817-1834 (in Chinese). doi:10.12082/dqxxkx.2022.220356
- 罗潇,廖春花,罗菁,等. 2021.两套新型高分辨率再分析降水资料在湖南年季尺度上的适用性评估[J].暴雨灾害,40(2):167-173. Luo X, Liao C H, Luo Q, et al.2021. Applicability evaluation of two sets of new high-resolution reanalysis precipitation data on annual and seasonal scale in Hunan [J]. Torrential Rain and Disasters,40(2):167-173 (in Chinese). doi:10.3969/j.issn.1004-9045.2021.02.008
- 潘旸,谷军霞,徐宾,等.2018.多源降水数据融合研究及应用进展[J].气象科技进展,8(1):143-152. Pan Y, Gu J X, Xu B, et al.2018. Advances in multi-source precipitation merging research [J]. Advances in Meteorological Science and Technology,8(1):143-152 (in Chinese). doi: 10.3969/j.issn.2095-1973.2018.01.019
- 曲学斌,付亚男,袁秀芝,等.2020. GPM-IMERG 日降水数据在内蒙古地区的适用性分析[J].暴雨灾害,39(3):293-299. Qu X B, Fu Y N, Yuan X Z, et al. 2020. Applicability analysis of GPM-IMERG daily precipitation data in Inner Mongolia [J]. Torrential Rain and Disasters,39(3):293-299 (in Chinese). doi:10.3969/j.issn.1004-9045.2020.03.010

- 任阳阳.2022.多种降水产品在中国大陆地区的适用性评估[D].邯郸:河北工程大学. Ren Y Y. 2022. Applicability of multi-source satellite precipitation products in Chinese mainland [D]. Handan: Hebei University of Engineering (in Chinese)
- 沈润平,郭佳,张婧娴,等.2017.基于随机森林的遥感干旱监测模型的构建[J].地球信息科学学报,19(1):125-133. Shen R P, Guo J, Zhang J X, et al.2017. Construction of a drought monitoring model using the random forest based remote sensing [J]. Journal of Geo-information Science,19(1):125-133 (in Chinese). doi:10.3724/SP.J.1047.2017.00125
- 慎璐璐.2022.黄河流域多源卫星降水产品精度评估[D].咸阳:西北农林科技大学. Shen L L. 2022. Evaluation of Multi-source satellite precipitation products in the Yellow River Basin [D]. Xianyang: Institute of Soil and Water Conservation Northwest A & F University (in Chinese)
- 史岚,何其全,杨娇,等.2019.闽浙赣地区 GPM IMERG 降水产品降尺度建模与比较分析[J].地球信息科学学报,21(10):1642-1652. Shi L, He Q Q, Yang J, et al. 2019. Downscaling modeling of the GPM IMERG precipitation product and comparative analysis in the Fujian-Zhejiang-Jiangxi region [J]. Journal of Geo-information Science,21(10):1642-1652 (in Chinese). doi:10.12082/dqxxkx.2019.180603
- 王浩钰,付国永,李黎,等.2021.一种基于 Xgboost+GWR 算法的中国地区降水量空间插值模型[J].水电能源科学,39(12):7-11. Wang H Y, Fu G Y, Li L, et al. 2021. A spatial precipitation distribution prediction model for Chinese regions based on Xgboost+GWR algorithm [J]. Water Resources and Power,39(12):7-11 (in Chinese).
- 夏昕然,田焯,谭伟丽,等.2022.多种卫星降水产品在中国的精度评估[J].水利水电技术(中英文),53(8):29-40. Xia X R, Tian Y, Tan W L, et al. 2022. Evaluation of satellite precipitation products in China [J]. Water Resources and Hydropower Engineering,53(8):29-40 (in Chinese). doi:10.13928/j.cnki.wrahe.2022.08.003
- 熊立华,刘成凯,陈石磊,等.2021.遥感降水资料后处理研究综述[J].水科学进展,32(4):627-637. Xiong L H, Liu C K, Chen S L, et al. 2021. Review of post-processing research for remote-sensing precipitation products [J]. Advance in Water Science,32(4):627-637 (in Chinese). doi:10.14042/j.cnki.32.1309.2021.04.014
- 徐淑婷,丁星臣,王斌.2017.多卫星遥感降水产品寒区误差时空特征分析[J].东北农业大学学报,48(11):63-73. Xu S Q, Ding X C, Wang B.2017. Alternative observation of ground stations by multiple satellite remote sensing precipitation in alpine region [J]. Journal of Northeast Agricultural University,48(11):63-73 (in Chinese). doi:10.19720/j.cnki.issn.1005-9369.2017.11.008
- 宇婧婧,沈艳,潘畅,等.2013.概率密度匹配法对中国区域卫星降水资料的改进[J].应用气象学报,24(5):544-553. Yu J J, Shen Y, Pan Y, et al. 2013. Improvement of satellite-based precipitation estimates over China based on probability density function matching method [J].Journal of Applied Meteorological Science,24(5):544-553 (in Chinese). doi:10.11898/1001-7313.20130504
- 宇婧婧,沈艳,潘畅,等.2015.中国区域逐日融合降水数据集与国际降水产品的对比评估[J].气象学报,73(2):394-410. Yu J J, Shen Y, Pan Y, et al. 2015. Comparative assessment between the daily merged precipitation dataset over China and the world's popular counterparts [J]. Acta Meteorologica Sinica,73(2):394-410 (in Chinese). doi:10.11676/qxxb2015.033
- 曾琪铖,金鑫.2022. GPM 与 TRMM 降水数据在柴达木盆地的精度评估[J].人民黄河,44(3):16-20. Zeng Q C, Jin X.2022. Accuracy evaluation of the TRMM and GPM precipitation data in Qaidam Basin [J]. Yellow River,44(3):16-20 (in Chinese). doi:10.3969/j.issn.1000-1379.2022.03.004
- 曾岁康,雍斌.2019.全球降水计划 IMERG 和 GSMaP 反演降水在四川地区的精度评估[J].地理学报,74(7):1305-1318. Zeng S K, Yong B. 2019. Evaluation of the GPM-based IMERG and GSMaP precipitation estimates over the Sichuan region [J]. Acta Geographica Sinica,74(7):1305-1318 (in Chinese). doi:10.11821/dlxb201907003
- 张钧民,阮惠华,许剑辉,等.2021.基于 XGBoost 的多源降水数据融合方法研究[J].热带地理,41(4):845-856. Zhang J M, Ruan H H, Xu J H, et al. 2021. An XGBoost-merging method for high-Resolution daily precipitation estimation for a regional rainstorm event [J]. Tropical Geography,41(4):845-856 (in Chinese). doi:10.13284/j.cnki.rddl.003372
- Alizadeh M R, Nikoo M R. 2018. A fusion-based methodology for meteorological drought estimation using remote sensing data [J]. Remote Sensing of Environment,211:229-47. doi:10.1016/j.rse.2018.04.001
- Duan Z, Liu J, Tuo Y, et al. 2016. Evaluation of eight high spatial resolution gridded precipitation products in Adige Basin (Italy) at multiple temporal and spatial scales [J]. Science of The Total Environment,573:1536-53. doi:10.1016/j.scitotenv.2016.08.213
- Fang J, Yang W, Luan Y, et al. 2019. Evaluation of the TRMM 3B42 and GPM IMERG products for extreme precipitation analysis over China [J]. Atmospheric Research,223:24-38. doi:10.1016/j.atmosres.2019.03.001
- Guo H, Chen S, Bao A, et al.2016. Early assessment of integrated multi-satellite retrievals for global precipitation measurement over China [J]. Atmospheric Research,176-177:121-133. doi:10.1016/j.atmosres.2016.02.020
- Wu Y, Guo L, Zheng H, et al. 2019. Hydroclimate assessment of gridded precipitation products for the Tibetan Plateau [J]. Science of the Total Environment,660:1555-64. doi:10.1016/j.scitotenv.2019.01.119
- Xie P, Xiong A. 2011. A conceptual model for constructing high-resolution gauge-satellite merged precipitation analyses [J]. Journal of Geophysical Research-Atmospheres,116: D21106. doi:10.1029/2011JD016118
- Xu S, Shen Y, Niu Z. 2019. Evaluation of the IMERG version 05B precipitation product and comparison with IMERG version 04A over mainland China at hourly and daily scales [J]. Advances in Space Research,63(8):2387-98. doi:10.1016/j.asr.2019.01.014
- Zhang T, Li B, Yuan Y, et al.2018. Spatial downscaling of TRMM precipitation data considering the impacts of macro-geographical factors and local elevation in the Three-River Headwaters Region [J]. Remote Sensing of Environment,215:109-27. doi:10.1016/j.rse.2018.06.004

(责任编辑 张文)