

# 江苏省太湖流域 MSWEP 数据 与地表站网雨量资料融合研究

云兆得<sup>1</sup>, 杜 飞<sup>1</sup>, 胡庆芳<sup>1,2</sup>, 王银堂<sup>1,2</sup>

(1. 南京水利科学研究院 水文水资源与水利工程科学国家重点实验室, 江苏 南京 210029;

2. 长江保护与绿色发展研究院, 江苏 南京 210098)

**摘要:**以江苏省太湖流域采用空间降尺度处理后的 MSWEP 日降水数据为基础,开展基于地理加权回归方法的 MSWEP 与地表站网资料的降水融合研究,解析了降水融合数据较原始 MSWEP 数据的精度改善情况。通过对 1979—2019 年的长系列研究结果发现,MSWEP 日降水与站网降水融合后,对研究区域日雨量场空间变异性的刻画能力明显提高,其平均绝对误差、相关系数等各项定量精度指标均有所改善。

**关键词:**降水融合; MSWEP 数据; 太湖流域

中图分类号:TV125

文献标识码:A

文章编号:1007-7839(2023)01-0054-0005

## Study on the integration of MSWEP data and rainfall information from surface station network in Taihu Lake Basin, Jiangsu Province

YUN Zhaode<sup>1</sup>, DU Fei<sup>2</sup>, HU Qingfang<sup>1,2</sup>, WANG Yintang<sup>1,2</sup>

(1. State Key Laboratory of Hydrology-Water Resource and Hydraulic Engineering,

Nanjing Hydraulic Research Institute, Nanjing 210029, China;

2. Yangtze Institute for Conservation and Development, Nanjing 210098, China)

**Abstract:** Using daily downscaled MSWEP rainfall data as the background field, and merging with rainfall data from gauge observation based on Geographically Weighted Regression in the Taihu Lake Basin within Jiangsu Province. The result from long series of 1979~2019 shows that, after rainfall merging of MSWEP daily rainfall data and gauge observation, the ability of describe the characterize spatial variability of precipitation have a remarkable enhancement, meanwhile the mean absolute error, correlation coefficient and other quantitative accuracy index are significantly improved.

**Key words:** rainfall merging; MSWEP Data; Taihu Lake Basin

收稿日期: 2022-07-16

基金项目: 江苏省水利科技项目(2020034); 国家自然科学基金面上项目(51479118)

作者简介: 云兆得(1996—),男,硕士,主要从事遥感水文的研究工作。E-mail: zd\_yun@foxmail.com

## 1 概述

降水是最基本的气象水文要素,同时也是水循环和能量循环过程的关键要素,具有强烈的时空变异性<sup>[1]</sup>。然而基于现有观测方式中所获取的降水信息并不能完全准确地展现降水时空分布特征,如地表雨量站网监测,受限于雨量站空间分布和自身监测情况等,无法表征降水空间的连续性和异质性。天气雷达、卫星遥感反演等空间雨量监测,尽管时空连续性更强,覆盖范围更广,但局地误差相对明显,无法直接运用于流域尺度研究中<sup>[2]</sup>。因此,国际上广泛开展了降水融合研究,将遥感反演降水与地表站网雨量信息相融合。

目前,降水融合按照基本原理包括全局和局部融合2种方法。全局方法通常假定地表雨量与栅格降水之间的关系不随空间位置的变化而改变,故其计算相对简单,但融合结果较为平滑,在低时间尺度的降水融合中易丢失强降水信息,代表性方法主要有全局回归法和平均偏差校正法<sup>[3]</sup>。然而受下垫面、降水类型等多重因素的影响,不同空间位置的遥感反演估计降水与地表雨量之间存在非平稳空间关系,因此考虑空间非平稳性的局部方法逐步发展成为主流融合技术。这类融合技术主要涉及了地质统计学<sup>[4]</sup>、最优插值<sup>[5]</sup>等方法,通过考虑多元影响变量(经纬度、高程等)对“真实”空间降水进行估计<sup>[6]</sup>。

江苏省太湖流域受梅雨、台风和局地性强对流天气影响,汛期暴雨多发,获取长时间序列高精度的空间降水信息有利于对流域降水过程的认识。本文采用普通克里金法(Ordinary Kriging, OK)<sup>[7]</sup>空间降尺度处理后的MSWEP日降水数据作为背景场,开展了基于地理加权回归(Geographically Weighted Regression, GWR)<sup>[8]</sup>方法的MSWEP与地表站网雨量资料的降水融合研究。解析降水融合数据相对于原始MSWEP数据的精度改善情况,分析降水融合数据对于降水空间变异性的刻画能力。

## 2 研究方法

### 2.1 降水融合方法

本文采用OK方法将MSWEP的1979—2019年日降水数据的空间分辨率由 $0.1^{\circ} \times 0.1^{\circ}$ 降尺度至 $0.01^{\circ} \times 0.01^{\circ}$ 。OK方法假定空间降尺度栅格估计降水量与原始栅格降水量满足一定关系,通过拟合计算求解权重系数,得到MSWEP降尺度雨量场。

以降尺度雨量场作为背景场,基于GWR方法构建“加法模式”,开展了MSWEP与52个雨量站实测资料的降水融合,即对降水误差场进行估计,本文中选取经度和纬度作为影响变量。后续计算中,采用AIC信息准则法(Akaike Information Criterion)对带宽进行优选,并通过Gauss截尾函数作为自适应空间权重函数计算权重。

### 2.2 精度评价方法

以地表雨量站观测数据为基准,采用分类指标和定量指标评价降水融合数据的精度。分类指标用于反映数据对日降水事件是否发生的判别能力,定量指标用于反映数据对降水量值的准确性。

分类指标<sup>[9]</sup>包括探测率POD(Probability of Detection)、误报率FAR(False Alarm Ratio)和Heidke技巧评分HSS(Heidke Skill Score),取0.1 mm/d作为有雨或无雨的判断阈值。定量指标包括平均误差ME(Mean Error)、平均绝对误差MAE(Mean Absolute Error)、相对偏差RBIAS(Relative Bias)、绝对偏差ARBIAS(Absolute Relative Bias)、均方根误差RMSE(Root Mean Square Error)和相关系数CC(Correlation Coefficient)。

## 3 研究区概况

### 3.1 研究区域

太湖流域位于长三角核心地带,是著名的江南水乡,地势低洼,受梅雨、台风和局地性强对流天气影响,汛期暴雨多发。本文研究区域为太湖流域的江苏省部分,即“江苏省太湖流域”。研究范围包括苏州、无锡、常州、镇江市全部及南京市部分区域,面积约19 399 km<sup>2</sup>,约占整个太湖流域面积的52.6%。研究区域内地形以平原为主,山丘区主要分布在茅山和宜溧山区一带。根据地形和水系特征,研究区域又分为湖西区、武澄锡虞区、阳澄淀泖区和太湖区4个水利分区。

### 3.2 数据资料

(1)站点雨量。为保证地表逐日雨量资料的数据质量,对数据连续性较差或日降水资料缺测较多的站点予以剔除,最后筛选了1979—2019年期间具有比较完整的日降水量记录的52个雨量站点,单站平均控制面积约373 km<sup>2</sup>。

(2)MSWEP雨量。本文采用的MSWEP数据为2.0版本<sup>[10]</sup>。该数据时间分辨率为3 h,空间分辨率为 $0.1^{\circ} \times 0.1^{\circ}$ 。由于地表站网观测资料为日降水数据,先将MSWEP降水数据累加为日降水资料,然后

再通过江苏省太湖流域的流域边界进行掩膜提取得到。

## 4 结果分析

### 4.1 点尺度精度

为便于表述,本文将空间分辨率为 $0.1^{\circ} \times 0.1^{\circ}$ 的MSWEP原始数据记作 $MSWEP_{Origin}$ ,将降水融合数据记作 $MSWEP_{OK-GWR}$ 。首先从点尺度开展了 $MSWEP_{Origin}$ 和 $MSWEP_{OK-GWR}$ 的精度评价。图1和图2分别为1979—2019年江苏省太湖流域52个雨量站实测日降水量与其相应空间位置的 $MSWEP_{Origin}$ 和 $MSWEP_{OK-GWR}$ 栅格降水量及误差散点图。

由图1(a)可知, $MSWEP_{Origin}$ 在江苏省太湖流域显著低估了降水,平均误差、平均绝对误差、均方根误差分别为 $-0.2\text{ mm}$ 、 $2.3\text{ mm}$ 、 $7.0\text{ mm}$ ,相对偏差和绝对偏差分别为 $-7.0\%$ 、 $71.0\%$ 。且 $MSWEP_{Origin}$ 和站点雨量相关关系一般,方差解释相对能力较弱,相关系数为 $0.7$ , $R^2$ 仅为 $0.5$ 。由图2(a)可知,当雨量站实测日降水量超过一定幅度时, $MSWEP_{Origin}$ 误差与日降水量呈显著负相关关系。而当地表日降水量为零时,由于 $MSWEP_{Origin}$ 存在一定误报,其误差值可能也较大。故在江苏省太湖流域, $MSWEP_{Origin}$ 对地表降水具备一定的解释能力,但其误差较为明显,可能源于数据自身精度有限,也可能源于

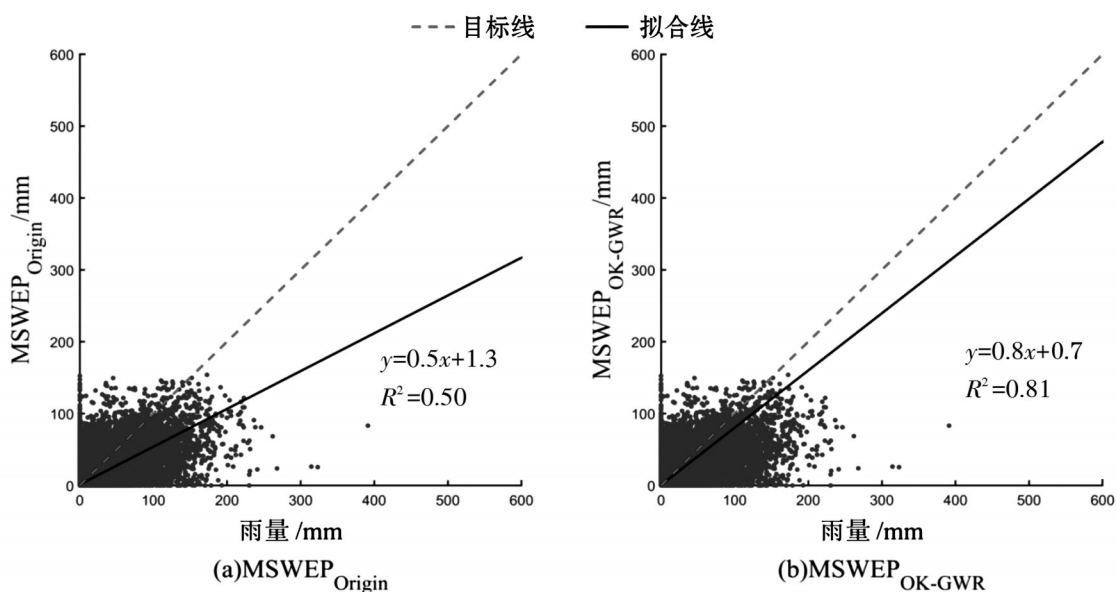


图1  $MSWEP_{Origin}$ 和 $MSWEP_{OK-GWR}$ 与地表雨量的日降水量散点分布

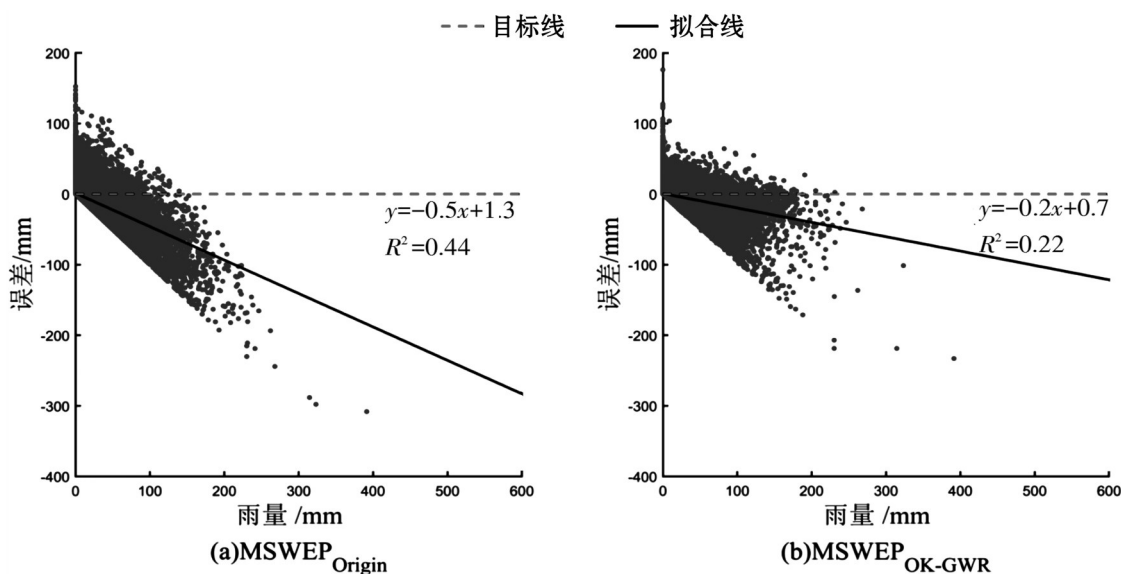


图2  $MSWEP_{Origin}$ 和 $MSWEP_{OK-GWR}$ 与地表雨量的日降水量误差散点分布

MSWEP栅格单元降水量和站点雨量空间尺度的不匹配性。

相较于图1(a),图1(b)中经融合后的MSWEP<sub>OK-GWR</sub>的定量精度得到了明显的提升,总体上降低了对日降水量的低估程度,其中平均误差、平均绝对误差和均方根误差指标分别改善了0.2 mm、1.0 mm、2.7 mm,相对偏差和绝对偏差分别改善了8.0%、31.5%。可见随着栅格和站点降水的匹配性增强,MSWEP<sub>OK-GWR</sub>的误差削减显著,同时MSWEP<sub>Origin</sub>和实测日降水量的相关性和方差解释能力也进一步提升,相关系数和 $R^2$ 均达到0.80以上。尽管图2(b)中两者的负相关关系已不甚明显,但是降水事件的误报仍然存在。

进一步研究雨量站相应MSWEP<sub>Origin</sub>和MSWEP<sub>OK-GWR</sub>栅格时序降水的分类和定量精度指标,从中可知除异常值外,MSWEP<sub>Origin</sub>相应栅格的探测率为0.9~1.0,误报率为0.3~0.4,技巧评分为0.5~0.7,可见MSWEP<sub>Origin</sub>在江苏省太湖流域对有雨事件的漏报率较小,然而对有雨事件的误报率相对较高。相对于MSWEP<sub>Origin</sub>,MSWEP<sub>OK-GWR</sub>相应栅格的3种分类指标并未得到明显改善,其3种分类指标分别介于0.9~1.0、0.2~0.5、0.4~0.8之间,说明采用这一降水融合方法可能无法有效改善降水分类指标的精度。

定量精度指标则恰恰相反,MSWEP<sub>Origin</sub>的平均误差、平均绝对误差、均方根误差分别为-0.4~0.0 mm、1.8~2.3 mm、5.5~7.7 mm,而MSWEP<sub>OK-GWR</sub>的3种指标分别介于-0.2~0.2 mm、0.7~1.7 mm、2.7~4.6 mm,可见相较于前者,后者相应栅格降水的定量误差得到了不同程度的削减,整体水平提高较为显著。MSWEP<sub>Origin</sub>的相对偏差和绝对偏差分别为-12.9%~1.3%、57.4%~71.3%,而MSWEP<sub>OK-GWR</sub>分别为-5.5%~7.7%、21.6%~54.6%,且MSWEP<sub>OK-GWR</sub>的相对偏差和绝对偏差和中位数相较前者更接近于0,可见后者相应栅格降水的偏差也得到一定的纠正。同时,相较于MSWEP<sub>Origin</sub>,MSWEP<sub>OK-GWR</sub>的相关系数也由原始的0.70以上提升至0.87,相关程度进一步增强。综上,尽管MSWEP<sub>OK-GWR</sub>的分类指标精度并未显著提升,但实现了定量误差的大幅削减和相关性提升。

#### 4.2 雨量场对比

本文进一步以1987年汛期单日降水的降水空间分布为例,解析降水融合数据对降水空间变异性的刻画效果。由分析可知,1987年6月1日江苏省太湖流域总体呈现降水由中部湖西区、武澄锡虞

区和太湖区交界地区向四周递减的空间格局。然而由于丢失部分强降水信息,降水空间分布无法表征这一特性,反而呈现了降水由西部湖西区到东部阳澄淀泖区逐渐降低的空间格局,这说明MSWEP<sub>Origin</sub>对雨量场空间变异性的刻画能力严重不足。在对站网雨量资料进行融合过程中发现,MSWEP<sub>OK-GWR</sub>可有效融合两者的信息,所刻画的雨量场更加接近真实情况,突显了降水空间分布特性,并且相对于MSWEP<sub>Origin</sub>,MSWEP<sub>OK-GWR</sub>的雨量场更为连续和细化。

## 5 结 语

通过开展1979—2019年江苏省太湖流域的MSWEP与地表站网日降水资料的融合研究,解析了降水融合数据较MSWEP原始数据的精度改善情况。

(1)MSWEP对江苏省太湖流域日降水量总体上存在低估现象,雨量场空间变异性的刻画能力严重不足,平均绝对误差、相关系数等各项定量精度指标较差,方差解释能力相对较弱。

(2)MSWEP日降水与站网降水融合后,雨量场更为连续和细化,空间变异性刻画能力明显增强,误差大幅减小,与地表雨量的相关性也显著增强,各项定量指标精度均得到显著提升。

(3)降水融合数据的探测率和误报率等分类精度指标,特别是日降水事件的误报率情况还需进一步改善。

本文仅开展考虑经度和纬度2种影响变量的非连续性日降水融合研究,今后有必要在此基础上继续开展考虑下垫面等影响要素的单日及小时以下尺度的连续性降水融合研究,以进一步提升降水融合数据的精度水平。

#### 参考文献:

- [1] DONOHUE A. Atmospheric science: energy and precipitation[J]. Nature Geoscience, 2016, 9(12): 861-862.
- [2] LU X Y, TANG G Q, WANG X Q. The development of a two-step merging and downscaling method for satellite precipitation products[J]. Remote Sensing, 2020, 12(3): 398.
- [3] NERINI D, ZULKAFLI Z, WANG L P, et al. A comparative analysis of TRMM-rain gauge data merging techniques at the daily time scale for distributed rainfall: runoff modeling applications[J]. Journal of Hydrometeorology, 2015, 16(5): 2153-2168.



- [4] LIU Y X, ZHUO L, MARIA P, et al. An assessment of statistical interpolation methods suited for gridded rainfall datasets[J]. International Journal of Climatology, 2021, 42(5):5.
- [5] YU J J, LI X F, LEWIS E, et al. UKGrHP: a UK high-resolution gauge-radar-satellite merged hourly precipitation analysis dataset[J]. Climate Dynamics: Observational, Theoretical and Computational Research on the Climate System, 2020, 54(5):2919-2940.
- [6] 熊立华, 刘成凯, 陈石磊, 等. 遥感降水资料后处理研究综述[J]. 水科学进展, 2021, 32(4):627-637.
- [7] ANDIE S, BASARUDDIN T, ANIATI M A. Minimax approach for semi variogram fitting in ordinary kriging[J]. IEEE Access, 2020(8):82054-82065.
- [8] 胡庆芳. 基于多源信息的降水空间估计及其水文应用研究[D]. 北京:清华大学, 2013.
- [9] 黄琦, 覃光华, 王瑞敏, 等. 基于MSWEP的祁连山地区降水空间分布特性解析[J]. 水利学报, 2020, 51(2):232-244.
- [10] BECK H E, VAN D A I J M, VINCENZO L, et al. MSWEP: 3-hourly 0.25° global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data[J]. Hydrology and Earth System Sciences, 2017, 21(1):589-615.

(上接第53页)

数计算,充分利用以往灌溉水利用系数研究成果,针对渠系组成与灌溉农田的分布,推算各级渠道的渠道水利用系数,对于不同灌区不同级渠道的具体情况,采用修正系数反映具体特征,最终得到完整的各级渠道水利用系数公式。利用分布式模型可以计算不同灌区渠系水利用系数,可以反映各个灌区渠系组成、灌溉面积分布、渠道衬砌情况等,相对于只分大、中、小灌区得到的灌溉水利用系数精度更高。但该方法需要计算的数据比较多,需要灌区渠系水利用系数实测资料进行模型参数确定,要求调查灌区渠系基本情况与农田分布情况,随着灌区信息化建设的不断推进,所需基础资料是可以得到

保证的,故此方法具有一定的应用前景。

#### 参考文献:

- [1] 中华人民共和国水利部. 灌溉与排水工程设计标准: GB 50288—2018[S]. 北京:中国计划出版社, 2018.
- [2] 白美健, 谢崇宝, 许迪, 等. 渠道输水损失计算公式中用平均流量代替净流量的误差分析[J]. 中国农村水利水电, 2001(6):12-13.
- [3] 任可, 王红雨. 灌区输水渠道渗漏损失测算与分析[J]. 中国农村水利水电, 2006(12):16-20.
- [4] 荣丰涛. 议论考斯加可夫经验公式[J]. 山西水利科技, 2004(1):17-19.