

杨荃,杨晓春,王一格,等. 西安市气象-水文雨量信息融合试验分析[J]. 陕西气象,2025(5):49-55.

文章编号:1006-4354(2025)05-0049-07

西安市气象-水文雨量信息融合试验分析

杨 荃^{1,2},杨晓春^{1,2},王一格³,魏俊涛³

(1. 陕西省气象服务中心,西安 710014;

2. 中国气象局秦岭和黄土高原生态环境气象重点开放实验室,西安 710016);

3. 西安市气象局,西安 710016)

摘 要:基于 2023 年 5—10 月西安市水文站降水观测数据和中国气象局多源降水同化系统(CMPAS)降水产品,应用最优插值方法开展逐小时降水数据融合试验,生成融合水文站降水数据后的网格数据产品,并对其进行精度评估。结果表明:水文站点与邻近气象站点的降水观测数据具有较高的一致性,数据可靠性满足融合及独立检验需求。融合水文站降水数据显著提升了网格降水产品的精度;融合后产品的偏差统计指标明显降低,均方根误差和平均偏差均值分别降低了 87% 和 70%,这两个指标的样本分布也更为合理,集中分布于 0 附近;该产品在西安区域的精度均得到显著改善,尤其在秦岭山区和骊山山区有效抑制了原有产品的偏差峰值;同时,融合产品在单站点尺度上能更准确地捕捉强降水过程的降水量峰值。

关键词:小时降水;融合试验;水文站;最优插值

中图分类号:P468

文献标识码:A

降水是地球水循环的基本组成部分,同时也是天气、气候研究的重点^[1-2]。为有效利用站点观测、卫星遥感、雷达估算降水等多种观测降水资料的优势,将多源降水数据融合产出高质量的降水网格实况数据是目前研发的趋势^[3-5]。面向中国区域的多源降水融合研究工作已经取得了一定进展。早期的融合主要是基于卫星反演降水产品和站点观测数据^[6]。后期,国家气象信息中心引入雷达定量降水估测产品,研制了站点降水观测、卫星、雷达等多源融合(CMA multi-source precipitation assimilation system, CMPAS)降水产品,产品的空间分辨率也从 5 km 提高至 1 km^[7]。已有应用研究检验表明,CMPAS 降水产品在全国范围质量较优^[8-12],但是在有些地形复杂地区、站点稀少地区会存在较大偏差。该产品在陕西的检

验^[13]也存在类似的问题。针对此类问题,李显凤等在评估江西省水文站降水观测数据质量的基础上,将该数据作为独立检验样本对 CMPAS 降水产品精度进行了检验^[8];并进一步采用多重网格变分分析方法将江西省水文站观测降水作为多源降水数据的一个来源,开展数据融合试验,结果表明,由于水文站多位于山区,此类数据的加入对于对不同地形条件下的降水融合效果均有改善,特别是对于海拔高度在 800 m 以上的地区改善最为显著^[14]。

本文以西安及周边区域的 CMPAS 降水产品为背景场,采用最优插值方法融入质量控制后的水文站点降水监测数据,提高多源降水融合产品的精度,为防灾减灾、气象服务提供更精准的降水网格实况数据。

收稿日期:2025-04-12

作者简介:杨荃(1980—),男,汉族,陕西耀州人,本科,工程师,从事应用气象研究工作。

通信作者:杨晓春(1984—),女,汉族,陕西安塞人,博士,正研级高工,主要从事数值预报和应用气象研究。

基金项目:秦岭和黄土高原生态环境气象重点实验室开放基金课题(2022G-14;2022Y-20;2021G-26)

1 数据和方法

1.1 数据来源

CMPAS 数据是国家气象信息中心基于“概率密度函数+贝叶斯模型平均+降尺度+最优插值”方法研发的三源融合降水产品^[7,15-16],时间分辨率为 1 h。该降水数据产品的精度高于任何单一来源降水产品的精度^[16]。研究采用 2023 年 5—10 月西安及周边区域 CMPAS 降水产品(33.44°~34.86°N,107.18°~110.19°E)为背景场,数据从气象大数据云平台·天擎获取。

水文站的布设重点在中小河流重要河段、山区的山洪沟及水库周边,其降水数据能有效弥补气象部门资料探测空间的不足。目前,西安区域气象站降水观测主要采用上海气象仪器厂生产的 SL3-1 型双翻斗雨量计,其采集分辨率为 0.1 mm;

西安区域水文站主要采用南京水利水文自动化研究所、江苏赛立科技有限公司生产的雨量计,仪器分辨率为 0.5 mm、1 mm,即水文站雨量监测仅传输实际发生降水的站点数据,均为有效降水数据。将水文站观测数据重新计算至与气象观测数据相同的小时分辨率后,采用气象部门逐小时降水资料的质量控制方法开展气候学界限值检查、时空一致性检查,剔除质量控制异常站点及异常时次数据^[17],得到质控后 2023 年 5—10 月水文站点逐小时观测数据。

评估中所用到的水文站邻近的气象站降水观测数据来自气象大数据云平台·天擎,时间分辨率为 1 h,时效范围与上述两类数据一致。融合范围内气象和水文站点分布如图 1 所示,其中,气象监测站 670 个,水文站 212 个。

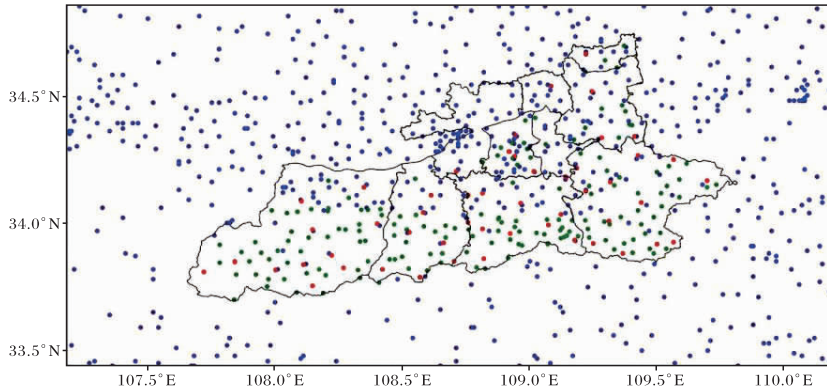


图 1 融合分析区域内站点分布示意图(蓝色圆点为气象监测站点,绿色圆点为融合水文站点,红色圆点为独立检验水文站点,审图号 GS(2020)4619 号)

1.2 研究方法

本文采用最优插值方法(optimal interpolation, OI),以 CMPAS 数据为初估场,将水文站观测数据融入。计算过程中,当某网格中有水文站降水观测值时,观测值即为网格分析值;当某网格中没有水文站点观测值时,网格上降水分析值 A_k 等于该点的初估值 F_k 加上该格点上观测值与初估值的偏差,偏差由规定距离范围内 n 个格点上已知的观测值 O_i 与初估值 F_i 的偏差加权估算而得,即:

$$A_k = F_k + \sum_{i=1}^n W_i (O_i - F_i) \quad (1)$$

式(1)中, k 为分析格点, i 为有效格点,有效

格点的范围可根据具体计算情况而定。 W_i 为权重函数,表示 i 点上观测值与初估值的偏差在估计时分配的权重。详细计算方法参考潘旻等^[18]的方法。以 CMPAS 为初估场,利用最优插值方法对西安区域水文站降水数据进行融合后所得到的网格降水数据记为 XIAN。其与 CMPAS 具有相同的时间、空间分辨率。

融合试验前需要先确定水文站观测降水数据的可靠性。本文选取邻近的气象站与水文站开展降水数据评估分析,统计指标为:相关系数(R)、均方根误差(root mean square error, RMSE)、相对偏差(relative deviation, RE)。对融合后降水产品的效果评估则采用独立检验方法,即均匀选

择西安范围内不同地理位置的水文站作为独立检验站点,不参与融合计算,仅对融合后的多源降水数据进行精度检验^[8],如图1所示,独立检验水文站点有52个。基于降水过程检验所采用的统计指标有平均偏差(mean bias, MB)和均方根误差。所有检验指标具体公式如下:

$$R = \frac{\sum_{i=1}^n (S_i - \bar{S})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^n (S_i - \bar{S})^2 \sum_{i=1}^n (O_i - \bar{O})^2}}, \quad (2)$$

$$X_{RE} = \frac{1}{n} \sum_{i=1}^n \frac{S_i - O_i}{O_i}, \quad (3)$$

$$X_{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (S_i - O_i)^2}, \quad (4)$$

$$X_{MB} = \frac{1}{n} \sum_{i=1}^n (S_i - O_i). \quad (5)$$

融合前的水文站降水数据可靠性检验指标,如公式(2)~(4), R 为相关系数, X_{RE} 为相对偏差, X_{RMSE} 为均方误差(mm), S_i 为水文站观测降水量(mm), O_i 为气象站点观测降水量(mm)。融

合后的格点降水检验指标,如公式(4)~(5), X_{MB} 为平均偏差, S_i 为独立检验水文站点所在位置对应格点的融合后降水量(mm), O_i 为独立检验水文站点的观测降水量(mm)。 n 为样本数。

2 结果与分析

2.1 水文与气象站点观测降水对比

为了检验气象与水文监测数据的一致性,参考李显风等^[8]的方法,选取24组气象站与水文站距离小于1 km的邻近站点(表1所示),对2023年5—10月的观测降水进行对比分析。从气象和水文邻近站点逐小时降水量散点图可以看出(图2),两套观测降水数据具有较好的一致性,相关系数达到0.81。由于降水具有强空间异质性,邻近站点并非相同的地理位置,两者的均方根误差为1.29 mm/h。对比结果表明:水文站的总体降水量级大于气象站,相对偏差为11%。前人研究认为,小时降水量理想误差为20%之内^[19],目前的两组数据满足标准,也表明了水文数据在检验和融合试验中的可靠性。

表1 邻近的气象、水文站点统计表

编号	气象站	水文站	站点间距离/km	编号	气象站	水文站	站点间距离/km
1	57047	80085349	0.985	13	V8523	80085459	0.497
2	V8188	80085540	0.248	14	V8527	80085461	0.418
3	V8153	80085036	0.208	15	V8721	80085483	0.923
4	V8166	80085004	0.239	16	V8781	80085412	0.102
5	V8167	80085011	0.386	17	V8810	80085247	0.726
6	V8256	80085119	0.333	18	V8823	80085406	0.825
7	V8235	80085122	0.921	19	V8870	80085520	0.956
8	V8335	80085218	0.631	20	V8808	80085525	0.029
9	V8488	80085330	0.727	21	V8872	80085473	0.384
10	V8452	80085348	0.497	22	V8864	80085522	0.867
11	V8458	80085321	0.384	23	V8862	80085521	0.022
12	V8586	80085485	0.910	24	V8863	80085514	0.227

2.2 融合前后的统计指标特征

本文选取了2023年夏季7次降水过程开展融合水文站观测数据前、后的网格降水产品精度评估。评估中分别提取降水过程中52个独立检验水文站的逐小时降水数据及相对应的XIAN

和CMPAS网格数据进行配对,计算每个站点MB和RMSE,并绘制分布图(图3)。由MB的分布可以发现,XIAN降水数据偏差分布为扁平图形,数值集中在0 mm附近,正负偏差都较小。而CMPAS降水数据的偏差分布图形较长,显示偏

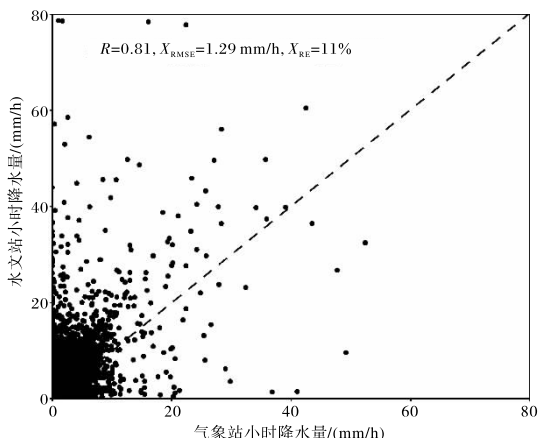


图2 2023年5—10月24组邻近的气象、水文站点小时降水量散点图

差较为分散,以负偏差为主, $-2\sim 0$ mm 为偏差的集中区域,最大值为 -4.3 mm。融合前后 RMSE 的分布特征与 MB 相似。XIAN 的 RMSE 呈现出数据集中的特征, $0\sim 2$ mm 为数据集中区域,而 CMPAS 的 RMSE 则数值较大,分布更分散。从数值的分布表现来看,水文站降水数据的融合不仅大大降低了网格降水产品的 MB 和 RMSE 数值,二者的均值分别降低 87% 和 70%,中位数也分别下降 85% 和 70%,而且使得两个统计量的分布更为合理,表明融合后产品较融合前更接近站点观测实况。

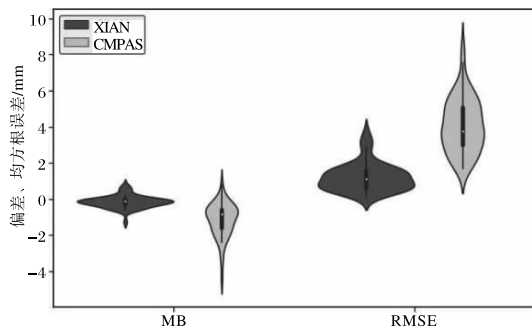


图3 水文站降水观测数据融合前(CMPAS)和融合后(XIAN)的网格降水数据平均偏差(MB)和均方根误差(RMSE)分布图

进一步分析融合前后独立检验站统计指标的空间分布特征。由图 4a、4b 的 MB 空间分布可见,融合前的 CMPAS 数据整体表现为负偏差,其中西安南部秦岭山区及东南部骊山山区为负偏差大值区,该现象与前人研究结论基本一致^[13,20-21]。融合后的 XIAN 产品负偏差整体显著减小,特别是 CMPAS 负偏差大值区改善尤为明显,7 次降水过程站点降水量的平均负偏差最大减少值达 4.5 mm。这种改善主要归因于气象站点在上述区域分布稀疏,而水文站点分布相对密集的状况^[14]。RMSE 的空间分布特征与 MB 相似(图 4c 和 4d)。融合后,全市范围内的 RMSE 均有所降

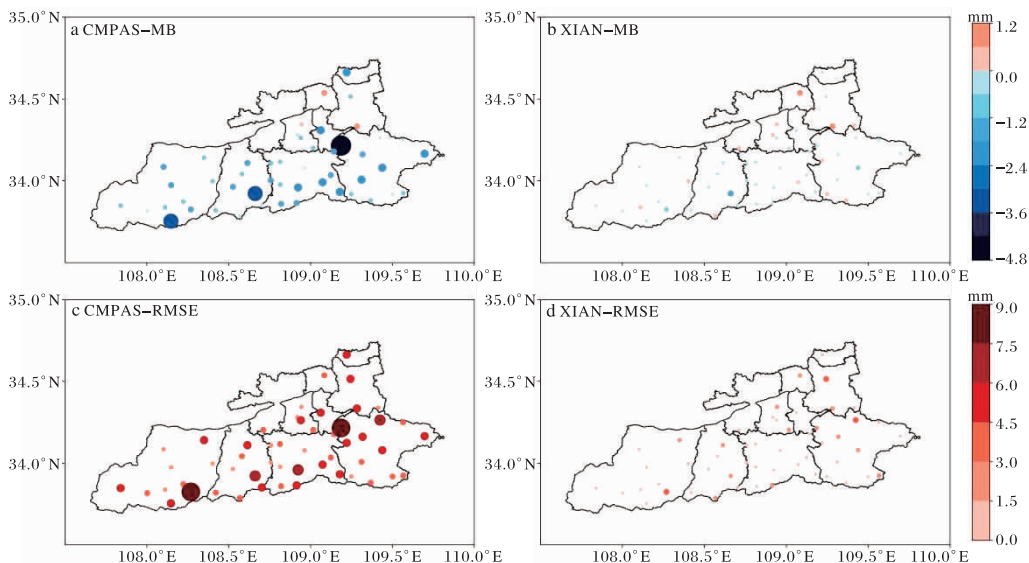


图4 水文站降水数据融合前(CMPAS)和融合后(XIAN)的网格降水数据偏差(MB)和均方根误差(RMSE)空间分布图(审图号为 GS(2020)4619 号)

低。以骊山山区为例,融合前 RMSE 最大可达 9.0 mm,融合后则降至 2.4 mm 以下。融合水文站降水数据有效减小了 CMPAS 在秦岭山区和骊山山区因气象观测站点不足而产生的显著负偏差。

2.3 融合前后单站精度分析

为了更细致地分析融合水文站降水数据对网格降水产品精度的影响,选取秦岭山区 3 个独立检验水文站(80085510、88085504、88085048),对 2023 年 6—7 月 3 次短时强降水过程(雨强 $> 10 \text{ mm/h}$)中融合水文数据前后的网格降水数据

进行对比分析(图 5)。3 次降水过程中,融合后数据 XIAN 比融合前 CMPAS 更接近网格内水文站降水实况,特别是当 7 月 12 日 21 时小时降水量达到 31 mm 时,由于山区气象站点稀少,CMPAS 的降水量仅为 4.9 mm,而 XIAN 在 CMPAS 基础上融合了周边的水文站观测数据,降水量为 27.7 mm,大大减少了强降水下的偏差(图 5a)。类似的情况也出现在 6 月 2 日 10 时、7 月 31 日 17 时的强降水中,如图 5b、图 5c 所示。同时发现,当水文站实况降水出现间歇和降水减弱时,如图 5a 中的 7 月 12 日 23 时—13 日 02 时,XIAN 和

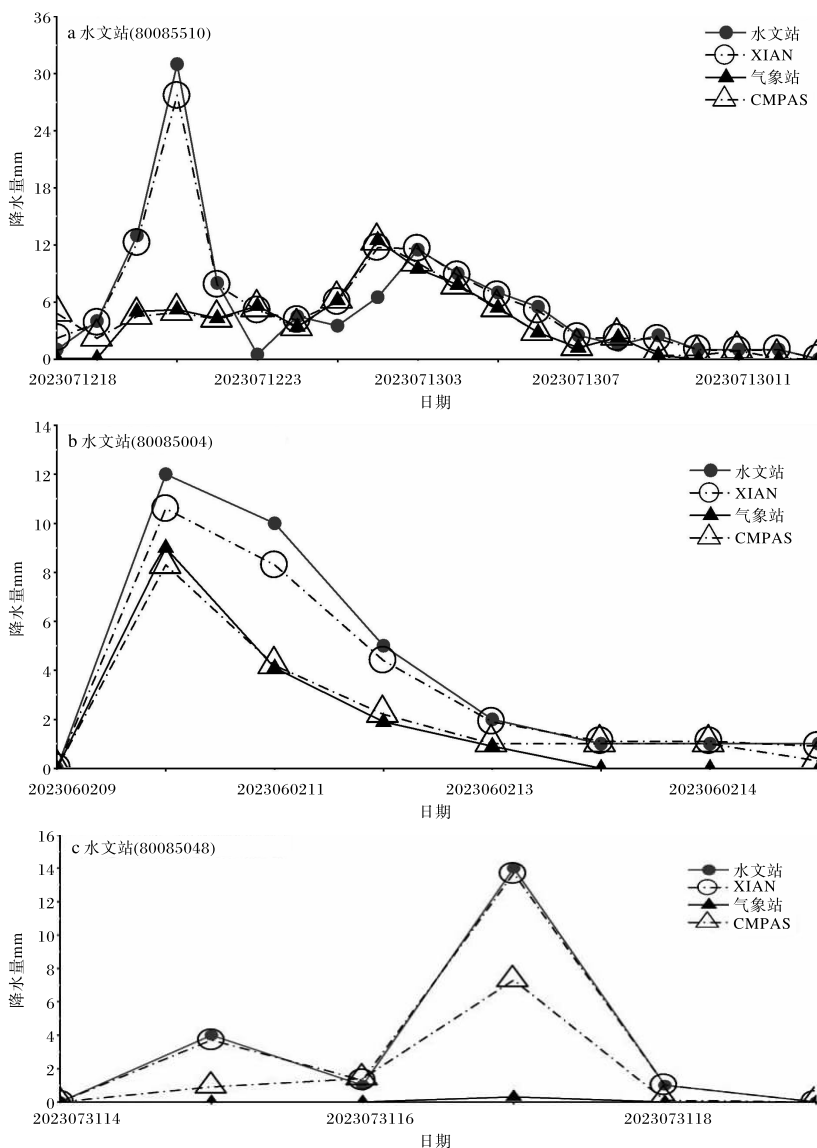


图 5 3 次降水过程中水文站降水量与 XIAN、CMPAS 对应网格降水量以及距离最近气象站降水量对比图

CMPAS 均出现大于水文站实况降水的现象。这可能是 CMPAS 的降水在当前网格点没有气象站降水信息时,更依赖周边气象站降水数据,与水文站降水量有一定的偏差,而 XIAN 是基于 CMPAS 数据对周边水文站点数据进行融合计算,虽然量值较未融合有所减小,但仍较站点水文站降水量值偏大。总体来说,在强降水发生时,融合后数据可以有效改善 CMPAS 降水量低估现象。

3 结论与讨论

本文采用最优插值法在 $1\text{ km}\times 1\text{ km}$ 逐小时的时空尺度上开展了水文站降水与 CMPAS 降水产品的融合试验,利用 2023 年 5—10 月 7 次降水资料,从偏差统计指标数值分布、空间分布、单站时序变化等方面对融合前后的数据进行综合评估分析,主要结论如下。

(1)质量控制后的水文站观测降水数据与距离相近的气象站具有较好的一致性,相关系数达到 0.81。水文站降水量总体大于气象站,小时降水量的相对偏差为 11%。检验结果表明水文站降水数据具有一定可靠性,满足参与融合算法和作为独立检验数据的需求。

(2)独立检验表明,融合水文站降水数据后的 XIAN 比 CMAPS 降低了负偏差和均方根误差,MB 和 RMSE 均值分别降低 87% 和 70%,同时 XIAN 的 MB 和 RMSE 的样本数据分布更为合理,MB 分布集中在 0 mm 附近。

(3)偏差统计指标的空间分布显示,融合水文站降水数据后的 XIAN 比 CMAPS 在整个西安区域内的精度均明显改善,气象站点稀疏的山区显著降低了偏差。骊山山区在融合前的 RMSE 达到 9.0 mm,融合后降至 2.4 mm 以下。

(4)通过对独立检验站点中的单站降水过程数据分析表明,融合水文站降水数据后的 XIAN 比 CMAPS 更准确反映降水峰值,减少强降水时段的偏差。

参考文献:

- [1] 刘元波,傅巧妮,宋平,等. 卫星遥感反演降水研究综述[J]. 地球科学进展,2011,26(11):1162-1172.
- [2] 赵煜飞,朱江. 近 50 年中国降水格点日值数据集精

度及评估[J]. 高原气象,2015,34(1):50-58.

- [3] HUFFMAN G J, BOLVIN D T, NELKIN E J, et al. The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales[J]. Journal of Hydrometeorology, 2007, 8(1): 38-55.
- [4] 李林,范雪波,崔炜,等. 称重与人工观测降水量的差异[J]. 应用气象学报,2015,26(6):688-694.
- [5] TANG G Q, MA Y, LONG D, et al. Evaluation of GPM day-1 IMERG and TMPA version-7 legacy products over mainland china at multiple spatiotemporal scales [J]. Journal of Hydrology, 2016, 533: 152-167.
- [6] LU N, YOU R, ZHANG W J. A fusing technique with satellite precipitation estimate and raingauge data[J]. Acta Meteorolog Sinica, 2004, 18(2): 141-146.
- [7] 潘旸,沈艳,宇婧婧,等. 基于贝叶斯融合方法的高分辨率地面-卫星-雷达三源降水融合试验[J]. 气象学报,2015,73(1):177-186.
- [8] 李显凤,周自江,李志鹏,等. 基于江西省水文资料对中国融合降水产品的质量评估 [J]. 气象,2017, 43(12):1534-1546.
- [9] 龙柯吉,谷军霞,师春香,等. 多种降水实况融合产品在四川一次强降水过程中的评估[J]. 高原山地气象研究,2020,40(2):31-37.
- [10] 吴薇,黄晓龙,徐晓莉,等. 融合降水实况分析产品在四川地区的适用性评估[J]. 沙漠与绿洲气象,2021,15(4):1-8.
- [11] 刘菊菊,陈小婷,肖贻青,等. 降水融合格点产品在陕西 2019 年暴雨过程中的检验[J]. 陕西气象, 2022(2):1-9.
- [12] 高亮书,余鹏. CLDAS 地面要素产品在陕西的适用性评估[J]. 陕西气象,2023(1):67-72.
- [13] 贺音,张雅斌,樊丹丹. 两套降水融合实况产品在陕西省的质量检验[J]. 陕西气象,2023(5):60-66.
- [14] 李显凤,张玮,黄少平,等. 江西省气象-水文-雷达小时雨量信息融合试验结果分析[J]. 暴雨灾害,2020,39(3):276-284.
- [15] 师春香,潘旸,谷军霞,等. 多源气象数据融合格点实况产品研制进展[J]. 气象学报,2019,77(4): 774-783.
- [16] 潘旸,谷军霞,宇婧婧,等. 中国区域高分辨率多

- 源降水观测产品的融合方法试验[J]. 气象学报, 2018, 76(5): 755-766.
- [17] 任芝花, 赵平, 张强, 等. 适用于全国自动站小时降水资料的质量控制方法[J]. 气象, 2010, 36(7): 123-132.
- [18] 潘旻, 沈艳, 宇婧婧, 等. 基于最优插值方法分析的中国区域地面观测与卫星反演逐时降水融合试验[J]. 气象学报, 2012, 70(6): 1381-1389.
- [19] HONG Y, HSH KL, SOROSHIAN S, et al. Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system[J]. J Appl Meteor, 2004, 43(12): 1834-1852.
- [20] 吴薇, 黄晓龙, 徐晓莉, 等. 融合降水实况分析产品在四川地区的适用性评估[J]. 沙漠与绿洲气象, 2021, 15(4): 1-8.
- [21] 俞剑蔚, 李聪, 蔡凝昊, 等. 国家级格点实况分析产品在江苏地区的适用性评估分析[J]. 气象, 2019, 45(9): 1288-1298.