

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^{*1}

李运龙 熊立华 闫磊

(武汉大学水资源与水电工程科学国家重点实验室, 湖北武汉 430072)

【摘要】:地理加权回归克里金(GWRK)是在地理加权回归(GWR)基础上扩展得到的一种既能考虑回归关系的空间非平稳性又能考虑回归变量空间自相关性的降水数据融合方法。以赣江流域为例,在评价 TRMM 卫星数据精度的基础上,分别以 GWRK 和 GWR 方法构建了站点-卫星降水数据融合模型,然后采用降水融合数据驱动 GR4J 水文模型进行水文预报。根据站点尺度降水融合数据精度及水文预报表现,对 GWRK 和 GWR 构建的降水融合模型效果进行评价,结果表明:较之 GWR 方法,GWRK 方法能较明显的提高降水融合数据在站点尺度上的精度,但是由于输入到水文模型中的数据为面降水数据,受空间均化的影响,对水文预报精度的提高不如对站点尺度降水融合数据精度的提高明显。

【关键词】:地理加权回归克里金(GWRK);降水融合;水文模型;TRMM

【中图分类号】:P339 **【文献标识码】**:A **【文章编号】**:1004-8227(2017)09-1359-10

DOI:10.11870/cjlyzyyhj201709008

准确的降水估计是水文模型和环境模型的重要基础工作,同时也是具有挑战性的难题。地面站点是获取降水信息的最主要途径,但实际中的站网往往稀疏、分布不均且易受环境影响,使得估计的降水具有很大的不确定性。卫星反演技术相比于地面站点可以提供较高时空分辨率的区域降水数据,但在现有技术条件下还存在着局地测雨精度不高,定量误差突出等不足^[1]。卫星降水数据的局部精度尚不足,而在站点密度不高时由地面站网估计的降水又具有较高的不确定性,因此通过对站点观测降水数据和卫星反演降水数据的权衡、匹配和优化组合来实现两种数据的融合,可以进行更加精确的降水估计。目前已有多种方法用于站点-卫星降水数据融合,例如地理加权回归^[2,3](Geographically Weighted Regression, GWR)、最优插值法^[4](Optimum Interpolation, OI)、地质统计学法^[5]、自适应核密度估计^[6]等。

GWR 可以在降水估计中同时考虑地理位置、高程及其他辅助信息,并且可以反映降水和影响因素之间的空间非平稳关系^[7],目前已得到广泛应用。许时光等^[8]提出了一种基于 GWR 方法考虑降水与 NDVI 和 DEM 之间非平稳关系的卫星月降水降尺度算法。胡庆芳^[3]采用 GWR 方法在赣江流域进行站点-卫星降水数据融合并对其水文应用效果进行评价。吕爱锋^[2]探讨了基于 GWR 的站点-卫星降水数据融合方法在柴达木盆地的应用。GWR 方法采用空间变回归参数来估计无实测点的降水数据,虽然考虑了降水和影响因素之间的空间非平稳性,但是没有直接考虑降水的空间自相关性。

¹ 收稿日期:2017-03-27; 修回日期:2017-06-28

基金项目:国家自然科学基金资助项目(51525902, 51479139) [National Natural Science Foundation of China(51525902, 51479139)]

作者简介:李运龙(1992~),男,硕士研究生,主要从事水文与水资源研究. E-mail:liyulong@163.com

地理加权回归克里金(Geographically Weighted Regression Kriging, GWRK)是在 GWR 基础上扩展的一种混合方法。GWRK 首先用 GWR 方法建立降水和辅助变量信息之间的回归关系,然后用克里金对 GWR 回归残差进行插值,最后将克里金插值结果和 GWR 计算结果相加获得 GWRK 估计结果。因此,GWRK 方法不但能考虑降水和影响因素之间的空间非平稳性,还能比较好的考虑降水的空间自相关性。GWRK 已经在日平均气温插值、土壤有机碳储量空间分布等研究中得到应用^[9, 10],但在降水数据融合中的研究还比较缺乏。

本文在降水数据融合中引入 GWRK 方法,并对 GWRK 和 GWR 两种方法降水融合数据精度及相应的水文预报效果进行了比较分析。为此,以赣江流域为例,选取具有良好表现且已被广泛应用的热带降雨观测计划(Tropical Rainfall Measuring Mission, TRMM)的实时卫星数据 TRMM3B42RTV7^[11, 12],在评估卫星数据精度的基础上,首先采用上述两种方法构建了地面站点和卫星降水数据融合模型,然后采用获得的降水融合数据驱动 GR4J 水文模型进行水文预报和分析,以期合理利用卫星数据及缺乏地面观测数据地区的降雨空间估计提供参考。

1 方法

1.1 GWRK 基本原理

GWRK 是在 GWR 基础上扩展得到的,因此本文先对 GWR 方法进行简单的介绍。

1.1.1 地理加权回归(GWR)

GWR 方法可以同时考虑多个影响变量,并以最邻近的子样来进行回归参数估计,从而能够直接有效地描述变量间的空间非平稳关系。对于变量 y ,在某区域内有 n 个观测点,则 GWR 方程形式为:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^p \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \quad (1)$$

式中: x_{ik} 是第 i 个观测点的第 k 个影响变量; $\beta_k(u_i, v_i)$ 是第 i 个观测点第 k 个影响变量的回归系数,是空间位置的函数; ε_i 为回归残差; u_i 和 v_i 分别为第 i 个空间点的经度和纬度坐标; p 为影响变量个数。

本文采用和胡庆芳^[3]相似的方法进行局部参数估计,利用自适应空间权函数(近高斯截函数)来反映各邻近点对于 i 点参数估计的重要性程度,采用交叉验证(Cross Validation, CV)的方法来确定最优带宽^[13]。

1.1.2 地理加权回归克里金(GWRK)

GWRK 方法是首先用克里金插值对 GWR 回归残差 ε 进行空间插值,然后将插值结果和 GWR 回归估计值相加,从而获得 GWRK 估计值,可用下式来表示 GWRK 方法:

$$\hat{y}_{GWRK}(u_i, v_i) = \hat{y}_{GWR}(u_i, v_i) + \hat{\varepsilon}_{ok}(u_i, v_i) \quad (2)$$

式中: $\hat{y}_{GWRK}(u_i, v_i)$ 为 GWRK 估计值; $\hat{y}_{GWR}(u_i, v_i)$ 为 GWR 估计值; $\hat{\varepsilon}_{ok}(u_i, v_i)$ 为利用普通克里金 (Ordinary Kriging, OK) 插值获得的 GWR 回归残差插值结果。

针对本文的降水数据, 采用指数变差函数进行 OK 插值。

1.2 降水融合模型

1.2.1 融合模型构建方式

本文在“加法模式”框架下构建降水融合模型。假设 P_o 为由 n 个地表实测值构成的降水实测场, P_s 为降水背景场 (卫星降水数据构成), 在实测点 i 处, 忽略观测误差, 有如下关系:

$$e_{oi} = P_{oi} - P_{si}, \quad i = 1, 2, \dots, n \quad (3)$$

式中: e_{oi} 为实测点 i 处降水背景场误差。

由式 (3) 可知, 在实测点 i 处, 降水背景误差可以表达为实测值与背景值之差。因此, 可以由实测点的降水背景误差 e_{oi} ($i=1, 2, \dots, n$) 来估计无实测点 j 处的降水背景误差 e_{uj} ($j=1, 2, \dots, k$), 即:

$$\hat{e}_{uj} = f [P_{o1} - P_{s1}, P_{o2} - P_{s2}, \dots, P_{on} - P_{sn}] \quad (4)$$

降水融合也就是将背景场误差估计值和降水背景值叠加, 即:

$$P_j = P_{sj} + \hat{e}_{uj} = P_{sj} + f [P_{o1} - P_{s1}, P_{o2} - P_{s2}, \dots, P_{on} - P_{sn}] \quad (5)$$

式中: P_j 为降水融合值。

1.2.2 降水背景场生成

TRMM 卫星数据的空间分辨率为 $0.25^\circ \times 0.25^\circ$, 而区域水文分析中要求的降水空间分辨率一般要高于 $10\text{km} \times 10\text{km}$, 甚至要超过 $5\text{km} \times 5\text{km}$ 。因此, 不适宜直接将原始 TRMM 卫星降水数据作为降水背景场。本文按照如下方法生成了 $5\text{km} \times 5\text{km}$ 空间分辨率的降水背景场:

对任一 $5\text{km} \times 5\text{km}$ 网格中心点 (u_i, v_i), 通过空间搜索找到与其最邻近的 9 个 $0.25^\circ \times 0.25^\circ$ 网格单元中心点, 计算其空间距离, 则该点的降水背景值计算如下:

$$P_{bi} = \sum_{k=1}^{k=9} \omega_k P_{sk} \quad (6)$$

式中: P_{bi} 是 $5\text{km} \times 5\text{km}$ 的网格点 (u_i, v_i) 卫星降水背景值; P_{sk} 是 (u_i, v_i) 的第 k 个最邻近的 $0.25^\circ \times 0.25^\circ$ 卫星降水值; ω_k 是权重值, 按下式计算:

$$\omega_k = \frac{1/d_k^2}{\sum_{k=1}^{k=n} 1/d_k^2} \quad (7)$$

式中: d_k 是 (u_i, v_i) 和第 k 个最邻近点间的距离。

降水融合的关键是无实测点处降水背景误差 e_{ij} 的估计。本文在生成降水背景场的基础上, 计算了各实测点处降水背景误差 e_{oi} , 选择地理坐标 u 、 v 作为影响变量, 分别利用 GWRK 和 GWR 两种方法对无实测点的降水背景误差 e_{ij} 进行估计, 然后将估计的降水背景误差值和降水背景值相加获得降水融合结果。

1.3 面降水量计算

降水融合数据为各网格点数据, 而将降水融合数据应用到水文模型中需要的为流域面降水数据。本文采用如下公式由网格点降水数据计算流域面降水数据:

$$\bar{P} = \frac{\sum_{m=1}^M P_m}{M} \quad (8)$$

式中: \bar{P} 为流域面降水量; P_m 为第 m 个网格的降水融合数据; m 和 M 分别为某一网格点和网格点数。

1.4 GR4J 模型及参数优化

采用 GR4J 模型进行流域水文模拟。GR4J 模型是 Perrin^[14] 在 GR3J 基础上改进的集总式降雨~径流模型, 目前已在包含半干旱区在内的 400 多个区域接受了检验。GR4J 模型采用两个非线性水库进行产、汇流计算, 其中第一个水库为产流水库, 第二个水库为汇流水库, 同时该模型还考虑了流域不闭合所导致的地下水交换问题。GR4J 共包含 4 个模型参数, 分别为产流水库蓄水容量 x_1 、地下水交换系数 x_2 、汇流水库蓄水容量 x_3 和汇流时间 x_4 。模型参数优选采用 SCEMUA^[15] 算法。

1.5 评价指标

从降水融合数据和站点实测降水值的一致性 & 径流预报精度两方面对 GWRK 和 GWR 降水融合效果进行比较分析。

(1) 降水融合数据精度评价指标

从“时序精度”和“空间精度”两方面，评价降水融合数据和站点实测降水数据的一致性。“时序精度”是评价降水融合数据和站点数据在时间上的同步性，是在计算各站点时序精度指标基础上计算的平均值；“空间精度”是评价降水融合数据对降水空间分布的估计能力，是在计算逐日空间精度指标的基础上计算的平均值。采用的评价指标为平均绝对值误差(MAE)、相对绝对值误差(ABIAS)和相关系数(CC)3项定量指标，具体计算公式如下：

①当计算空间精度时：

$$MAE^t = \sum_{i=1}^n \left| G_i^t - S_i^t \right| / n \quad (9)$$

$$ABIAS^t = MAE^t / \bar{G}^t \times 100\% \quad (10)$$

$$CC^t = \frac{\sum_{i=1}^n (G_i^t - \bar{G}^t) (S_i^t - \bar{S}^t)}{\sqrt{\sum_{i=1}^n (G_i^t - \bar{G}^t)^2 \sum_{i=1}^n (S_i^t - \bar{S}^t)^2}} \quad (11)$$

式中： MAE^t 为第 t 天的平均绝对值误差， $ABIAS^t$ 、 CC^t 定义类似； G_i^t 、 S_i^t 分别为第 i 个站点的第 t 天降水量和对应的 $5\text{km} \times 5\text{km}$ 栅格单元融合降水量； \bar{G}^t 、 \bar{S}^t 分别为第 t 天站点降水量和融合降水量空间均值； i 、 n 为某一站点和总站点数。

$$MAE_i = \sum_{t=1}^T \left| G_i^t - S_i^t \right| / T \quad (12)$$

$$ABIAS_i = MAE_i / \bar{G}_i \times 100\% \quad (13)$$

$$CC_i = \frac{\sum_{t=1}^T (G_i^t - \bar{G}_i) (S_i^t - \bar{S}_i)}{\sqrt{\sum_{t=1}^T (G_i^t - \bar{G}_i)^2 \sum_{t=1}^T (S_i^t - \bar{S}_i)^2}} \quad (14)$$

②当计算时间序列精度时：

式中： MAE_i 为第 i 个站点的平均绝对值误差， $ABIAS_i$ 、 CC_i 定义类似； \bar{G}_i 、 \bar{S}_i 分别为第 i 个站点降水量和融合降水量日均值； t 、 T 为某一天和总天数。

(2) 水文预报效果评价指标

采用纳什效率系数(NSE)、相对误差(RB)和均方根误差(RMSE)来评价径流预报精度。NSE反映预报径流和实测径流的一致性, RB反映预报径流的系统偏差程度, RMSE反映洪峰的预报精度, 具体计算公式如下:

$$NSE = 1.0 - \frac{\sum_{t=1}^T (Q_{obs}^t - Q_{sim}^t)^2}{\sum_{t=1}^T (Q_{obs}^t - \bar{Q}_{obs})^2} \quad (15)$$

$$RB = \frac{\sum_{t=1}^T (Q_{obs}^t - Q_{sim}^t)}{\sum_{t=1}^T Q_{obs}^t} \quad (16)$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^T (Q_{obs}^t - Q_{sim}^t)^2}{T}} \quad (17)$$

式中: Q_{obs}^t 、 Q_{sim}^t 分别为第 t 天实测径流值和预报径流值; \bar{Q}_{obs} 为实测径流均值。

为比较某一子流域, GWRK 和 GWR 降水融合数据对应的径流预报精度差别, 定义以下变量:

$$\Delta NSE(k) = NSE_{GWRK}(k) - NSE_{GWR}(k) \quad (18)$$

$$\Delta |RB(k)| = (|RB_{GWRK}(k)| - |RB_{GWR}(k)|) / 100 \quad (19)$$

$$\Delta RMSE(k) = RMSE_{GWRK}(k) - RMSE_{GWR}(k) \quad (20)$$

式中: $NSE_{GWRK}(k)$ 、 $NSE_{GWR}(k)$ 分别为第 k 个子流域 ($k=1, \dots, 14$) 采用 GWRK 和 GWR 降水融合数据预报的日径流量NSE值。 $RB_{GWRK}(k)$ 、 $RB_{GWR}(k)$ 、 $RMSE_{GWRK}(k)$ 和 $RMSE_{GWR}(k)$ 定义类似。显然, 若 $\Delta NSE(k) > 0$, 则说明 GWRK 降水融合数据提高了对实测日径流的方差解释能力; 若 $\Delta |RB(k)| < 0$, 则说明 GWRK 降水融合数据降低了日径流预报中的系统误差; $\Delta RMSE(k) < 0$, 则说明 GWRK 降水融合数据具有更好的洪峰预报能力。

2 应用实例

2.1 实验流域

实验流域采用长江中下游的赣江流域。该流域位于中国东南部(113° 30' ~116° 40' E, 24° 29' ~29° 11' N), 是长江中下游鄱阳湖水系的最大子流域, 其中外洲控制站以上集水面积 80948km²(图 1)。流域内地形复杂, 四周多山, 山间盆地发育, 中部为丘陵与盆地复合体, 下游为冲积平原。赣江流域属亚热带湿润季风气候, 2003~2009 年多年平均降水量约 1449mm, 4~6 月梅雨季降水集中, 7~9 月易受台风影响。

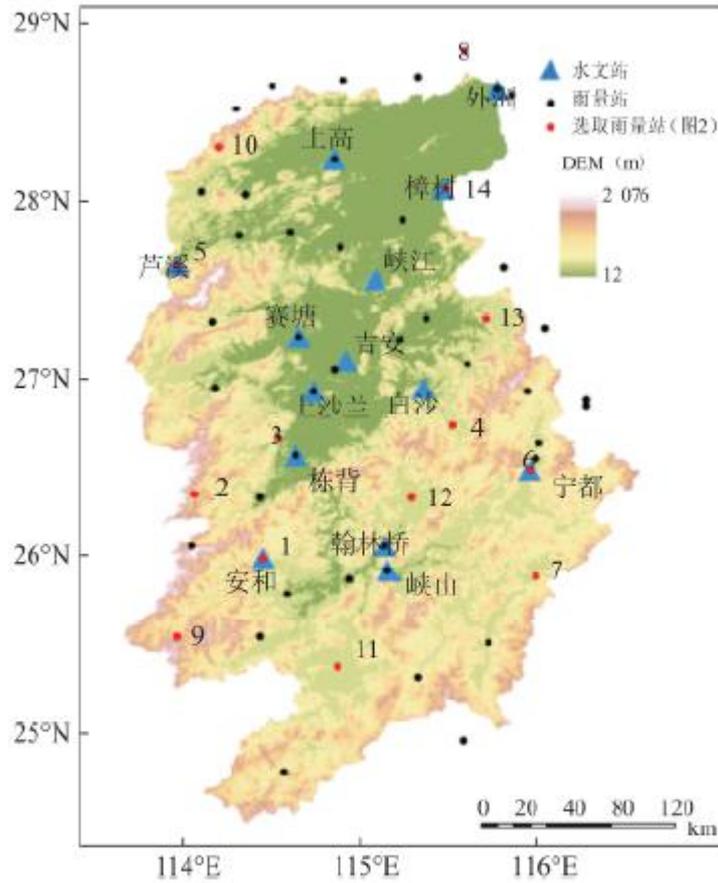


图 1 赣江流域及水文站、雨量站分布图
 Fig. 1 Location of the Ganjing Basin and the Distribution of Hydrological Stations, Rainfall Stations

2.2 数据

本文使用的数据资料及处理方法如下: (1) TRMM 卫星数据^[16]: 研究中采用的卫星降水数据为美国 Goddard 中心 (Goddard Space Flight Center, GSFC) 发布的最新 V7 版本 TRMM3B42RT_Daily 数据, 它是 V6 版本的升级数据, 其空间分辨率为 0.25° × 0.25°, 覆盖纬度从 50° N 到 50° S, 时间延迟为 3~9h, 为准实时数据。

(2) 站点观测数据: 降水和蒸发数据来自中国气象数据网, 其中雨量站 54 站, 蒸发站 10 站。流量数据为由水文年鉴获取的赣江流域 14 个水文站日流量数据, 并对数据进行了一致性、代表性和有效性检查。14 个水文站的空间位置、地形特征不一, 可以用来比较降水融合数据在水文分析中的应用效果和差异。14 个子流域对应的逐日面平均降水量数据和逐日平均蒸发能力均采用 GWRK 插值获得。雨量站及水文站分布见图 1。各子流域特性见表 1。

表 1 各子流域主要特性

站点名称	编号	经度 (°)	纬度 (°)	集水面积 (km ²)	多年平均 (mm)	多年平均 (mm)	降水量 (mm)	径流深 (mm)
外洲	1	115.83	28.63	80 948	1 449.4			713.3
樟树	2	115.53	28.07	71 324	1 448.7			663.8
峡江	3	115.15	27.55	62 724	1 449.6			708.3
吉安	4	114.98	27.10	56 223	1 449.8			712.9
栋背	5	114.70	26.57	40 231	1 454.5			711.7
峡山	6	115.22	25.92	16 033	1 459.3			710.6
上沙兰	7	114.80	26.93	5 257	1 470.5			750.6
上高	8	114.92	28.23	4 076	1 578.6			777.9
赛塘	9	114.72	27.23	3004	1 476.7			706.2
翰林桥	10	115.20	26.05	2 689	1 428.2			767.3
宁都	11	116.02	26.48	2 372	1 502.4			814.3
白沙	12	115.42	26.95	1 573	1 500.8			720.9
芦溪	13	114.03	27.63	331	1 550.0			1045.8
安和	14	114.52	25.98	246	1 453.8			802.0

流域地形资料为 SRTM(Shuttle Radar Topography Mission, SRTM)90m×90m 分辨率的 DEM 数据。研究中采用的 TRMM 卫星数据以及站点实测数据的系列长度均为 2003~2009 年。

3 结果与比较

3.1 TRMM 降水数据精度评价

图 2 给出了赣江流域 2003~2009 年雨量站日降水与最邻近的 TRMM 网格日降水散点图, 由于篇幅所限, 本文只绘制了在每个子流域随机选取的一个雨量站点, 共计 14 个站点以及流域平均降水数据散点图, 选取的雨量站位置见图 1。由图 2 可知, 两组数据在各站点均能较好的分布在 1:1 线两侧, 表明两组数据具有较好的线性相关关系。但是散点沿线分布的集中程度不足, 并且各站点具有一定的差别, 说明 TRMM 网格日降水数据定量误差比较明显, 且数据精度在空间上存在差异。对于流域平均数据, 大部分散点位于 1:1 线的下方, 表明 TRMM 卫星数据在整个流域上有低估降水的倾向。

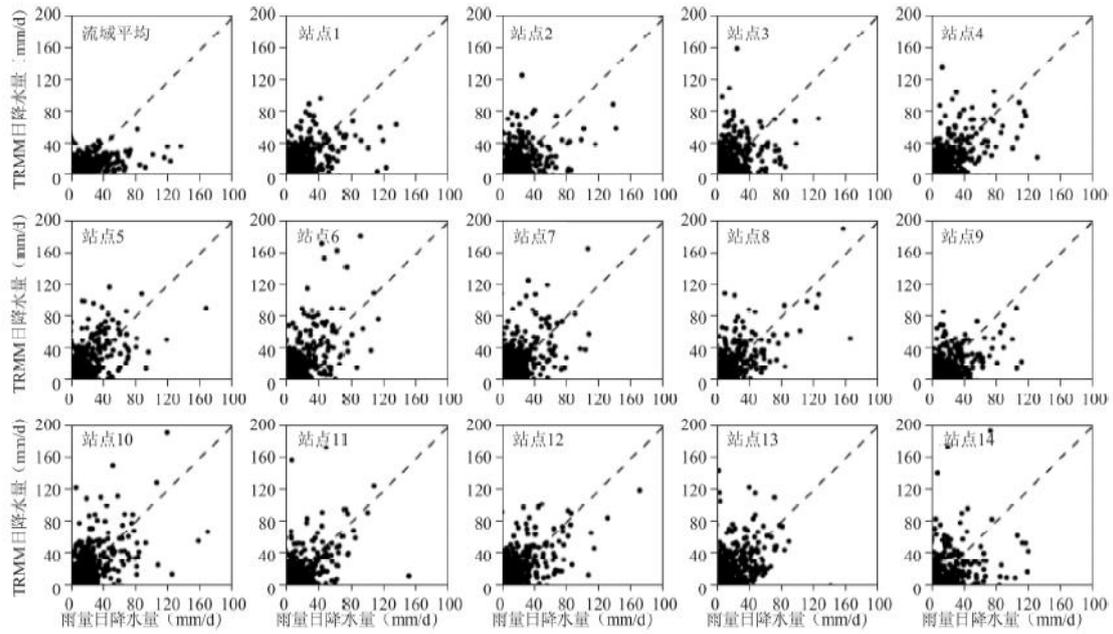


图2 雨量站日降水与邻近 $0.25^{\circ} \times 0.25^{\circ}$ TRMM 网格日降水比较

Fig. 2 Scatter Plots of Daily Rainfall Between Rainfall Station and Nearby $0.25^{\circ} \times 0.25^{\circ}$ TRMM Grid

表2给出了2003~2009年赣江全流域54个雨量站站点日降水与最邻近的TRMM网格日降水数据精度评价指标均值。对于降水量很小的日期，即使较小的卫星降水数据误差也可能导致很差的空问精度评价指标(ABIAS和CC)统计结果，因此表2只对流域平均降水量大于 0.5mm/d 的日期进行统计。由表2可以看出，在时间精度上，CC均值在0.60以上，表明卫星数据与站点数据在时程变化上具有较强的同步性，但是MAE和ABIAS指标相对较大，表明卫星数据存在较大的绝对偏差。在空间精度上，CC均值在0.40左右，MAE和ABIAS较大，说明在日平均意义上，TRMM卫星数据可大致反映流域内降水的空间分布，但是存在很大的绝对偏差。对比两种网格尺度的卫星降水数据评价结果，可以发现对TRMM卫星数据进行重采样后，无论是在空间精度还是时间精度上，均有稍微的提高，表明在降水融合数据中使用重采样后的数据是可行的。

表2 TRMM 网格日降水数据与雨量站日降水数据比较

网格尺度	时序精度评价指标均值			空间精度评价指标均值		
	MAE (mm)	ABIAS (%)	CC	MAE (mm)	ABIAS (%)	CC
$0.25^{\circ} \times 0.25^{\circ}$	3.77	94.2	0.63	4.30	127.7	0.39
5 km \times 5 km	3.56	89.0	0.66	3.84	121.2	0.41

3.2 降水融合数据精度对比

图3为赣江全流域不同日降水量级下GWR降水融合模型残差的指数半变差函数。由图3可以看出在不同的日降水量级下，GWR降水融合模型残差都存在着较强的空问自相关性，说明GWR方法在日降水融合上存在误差。因此，用克里金方法对GWR降水融合模型残差进行插值来降低其降水融合误差是可行的。

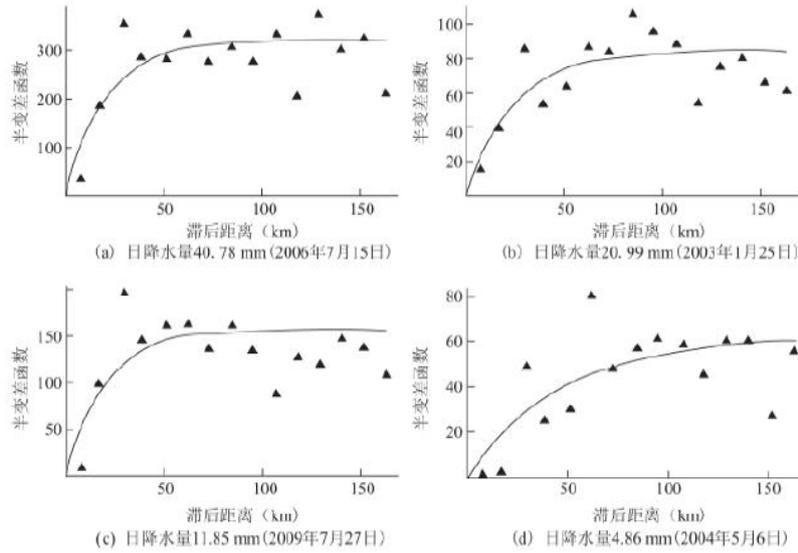


图3 赣江全流域 GWR 日降水融合模型残差指数半变差函数图

Fig. 3 Experimental Semi-Variogram and Fit Models of Residuals from GWR Daily Rainfall Merging Model in Ganjing Basin

表3给出了赣江流域2003~2009年GWRK和GWR日降水融合数据各项评价指标均值,其中空间精度指标是基于实测站点的交叉检验获得,空间精度指标中的ABIAS和CC也是只对流域平均降水量大于0.5mm/d的日期进行统计。为分析在不同的站网密度下,GWRK降水融合数据精度的变化特征,本文采用K均值聚类^[17]的方法来抽取站点,从而组成不同的站网密度。为了将降水融合结果和插值结果进行对比,表3也给出了GWR日降水插值结果,用GWR-I表示。

表3 赣江流域2003~2009年不同站网密度降水融合数据精度对比

站网密度 (站)	模型	评价指标					
		空间精度评价指标均值			时序精度评价指标均值		
		MAE (mm)	ABIAS (%)	CC	MAE (mm)	ABIAS (%)	CC
54	GWRK	3.62	0.90	0.46	2.57	0.65	0.78
	GWR	3.72	0.92	0.46	2.64	0.66	0.78
	GWR-I	3.71	0.91	0.40	2.64	0.66	0.77
30	GWRK	3.87	0.95	0.45	2.69	0.68	0.77
	GWR	3.97	0.97	0.44	2.76	0.70	0.76
	GWR-I	4.00	0.95	0.34	2.78	0.71	0.74
20	GWRK	4.23	0.99	0.43	2.73	0.72	0.75
	GWR	4.32	1.01	0.42	2.79	0.73	0.75
	GWR-I	4.55	0.99	0.26	2.87	0.75	0.71
10	GWRK	4.94	1.16	0.38	3.35	0.85	0.68
	GWR	4.94	1.16	0.38	3.35	0.85	0.69
	GWR-I	5.20	1.12	0.17	3.53	0.89	0.59

由表 3 可以看出, GWRK 相比于 GWR 对降水融合数据的 ABIAS 和 CC 改进作用不大, 但是对降低日降水融合数据的 MAE 有明显效果。但是在站网密度为 10 站时, GWRK 相比于 GWR 融合方法并不能表现出优越性, 可能是因为站点距离较大超出了半变差函数中变程的范围, 导致对 GWR 降水融合模型残差进行插值失去了效果。相比于时间精度指标, GWRK 方法对于空间精度指标的改进更加明显, 这可能是由于克里金插值方法能更好的捕捉降水之间的空间自相关性。

为比较降水融合数据在空间上的分布差别, 图 4 绘制了站网密度为 20 站时 2003~2009 年 GWRK 和 GWR 日平均降水融合数据空间分布图。由图 4 可以看出, 无论 GWRK 还是 GWR 降水融合数据均在流域的北部和东南部展示出了较高的降水强度, 在流域的中部具有较低的降水强度, 这可能是受赣江流域四周多高山, 中部多丘陵及盆地的影响, 并且图 4 结果与已有研究也具有一致性^[18]。GWRK 降水融合数据相比于 GWR 降水融合数据在流域的北部和东南部具有更高的降水强度, 但在流域的东北部有着更低的降水强度。

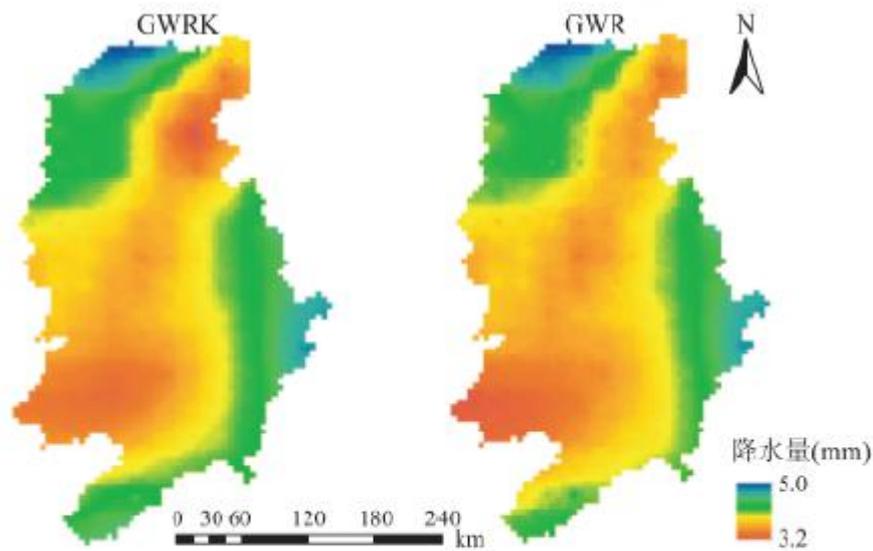


图 4 站网密度为 20 站时 2003~2009 年日平均降水融合数据空间分布

Fig. 4 Spatial Distribution of 2003~2009 Daily Average Rainfall Merging Data When the Rainfall Station Density is 20

3. 3GR4J 模型参数率定及验证

表 4 是以 54 个雨量站 GWRK 插值得到的逐日面降水量作为输入, 以 2003~2007 年(1826d)为率定期, 以 2008~2009 年(731d)为验证期的赣江流域 GR4J 径流模拟结果。由表 4 可以看出, GR4J 在赣江流域日径流模拟中具有较高的模拟精度。对于大多数流域, NSE 接近或者超过 0.85, RB 在 $\pm 5\%$ 以内, CC 值在 0.90 以上, RMSE 值在 1.00 左右。模拟精度较差的上沙兰、芦溪以及安和均位于雨量站分布稀少的地区, 可能是相应区域面雨量精度较低导致模拟精度较差, 可能是相应区域面雨量精度较低导致模拟精度较差, 但也基本在 0.70 以上。总体上, GR4J 模型在赣江流域具有较好的适用性, 可以用来评估降水融合数据在赣江流域的水文应用效果。

表 4 赣江流域 GR4J 模型率定与验证结果

水文站	X1 (mm)	x2 (mm)	x3 (mm)	x4 (d)	率定期 (2003~2007 年)				验证期 (2008~2009 年)			
					NSE	CC	RB %	RMSE (mm)	NSE	CC	RB (%)	RMSE (mm)
					外洲	634.9	1.81	105.6	4.56	0.89	0.94	2.34
樟树	682.9	1.10	89.7	4.26	0.89	0.94	3.60	0.65	0.88	0.94	1.29	0.51
峡江	594.3	1.84	110.9	3.41	0.90	0.95	2.39	0.70	0.87	0.94	-1.93	0.56
吉安	602.0	1.78	107.4	3.20	0.89	0.95	2.57	0.74	0.87	0.94	2.94	0.60
栋背	666.8	1.76	120.6	2.64	0.87	0.93	2.91	0.85	0.83	0.91	2.08	0.76
峡山	717.9	1.09	57.2	2.82	0.89	0.95	0.72	0.94	0.89	0.94	3.21	0.68
上沙兰	236.3	1.12	90.6	2.25	0.80	0.90	7.82	1.26	0.70	0.87	10.80	0.82
上高	697.4	0.46	50.5	2.17	0.88	0.94	4.05	0.88	0.90	0.95	1.89	0.79
赛塘	367.6	0.90	72.3	2.06	0.84	0.92	4.21	1.15	0.80	0.90	-3.72	0.98
翰林桥	570.2	1.68	61.8	2.08	0.90	0.95	3.01	1.01	0.91	0.96	4.16	0.84
宁都	1087.1	1.45	49.2	2.19	0.93	0.96	-0.20	0.92	0.89	0.95	-4.39	0.80
白沙	415.9	0.72	39.8	2.11	0.86	0.93	2.16	1.42	0.90	0.95	-7.67	0.94
芦溪	875.1	1.30	15.6	1.54	0.77	0.88	0.29	1.68	0.75	0.87	-4.75	1.91
安和	308.4	1.59	36.1	1.31	0.67	0.82	-2.27	2.52	0.78	0.89	2.60	1.90

3.4 水文预报表现对比

根据不同站网密度下降水融合精度分析结果, 本文以站网密度为 20 站时为代表进行降水融合数据水文预报效果分析。因此采用 GWRK 和 GWR 降水融合方法计算的站网密度为 20 站时 2008~2009 年赣江全流域 5km×5km 的日降水网格数据集, 计算了同时期 14 个子流域对应的逐日面降水过程线。然后, 输入到已率定好的 GR4J 水文模型中(模型参数见表 4), 分别对 14 个子流域 2008~2009 年日径流进行预报。

表 5 为 GWRK 和 GWR 降水融合数据在 14 个子流域的水文预报结果。GWRK 和 GWR 降水融合数据在赣江 14 个子流域的平均 NSE 值分别为 0.82 和 0.80, 从整体上看都能得到较好的径流预报效果。但是在个别流域两种方法获得的降水融合结果相差较大, 如上沙兰流域 GWRK 和 GWR 降水融合数据的 NSE 值分别为 0.68 和 0.50。

表 5 降水融合数据的水文预报结果

编号	水文站	预报精度					
		GWRK			GWR		
		NSE	RB (%)	RMSE (mm)	NSE	RB (%)	RMSE (mm)
1	外洲	0.89	-2.11	0.52	0.87	-2.89	0.56
2	樟树	0.89	-0.33	0.49	0.87	-0.99	0.52
3	峡江	0.88	-5.42	0.55	0.85	-7.07	0.60
4	吉安	0.89	-0.65	0.56	0.87	-3.06	0.60
5	栋背	0.85	-1.96	0.70	0.85	-4.36	0.72
6	峡山	0.90	-6.24	0.64	0.89	-9.71	0.67
7	上沙兰	0.68	8.64	0.84	0.50	3.73	1.06
8	上高	0.86	3.94	0.94	0.84	-1.63	0.98
9	赛塘	0.71	-13.98	1.19	0.75	-7.54	1.10
10	翰林桥	0.75	3.91	1.38	0.73	2.95	1.43
11	宁都	0.88	-13.62	0.83	0.86	-16.01	0.90
12	白沙	0.87	-6.20	1.06	0.81	-5.20	1.27
13	芦溪	0.69	-13.12	2.11	0.72	-4.59	2.02
14	安和	0.77	-4.36	1.94	0.74	-5.94	2.06
	平均	0.82			0.80		

图 5 绘制了在赣江全流域上 20080609 场次洪水发生时，外洲站、上沙兰站、白沙站和安和站 GWRK 和 GWR 降水融合数据预报流量与实测流量过程线。由图 5 可以看出，在外洲站、白沙站和安和站 GWRK 和 GWR 降水融合数据预报流量和实测流量吻合性较好，且两者的预报效果差别不大。在上沙兰站，GWRK 和 GWR 存在对洪峰流量的严重高估，可能是由于上沙兰站上游南车水库(大 2 型)径流调节的影响，但是相比之下 GWRK 预报径流和实测径流的一致性更好。

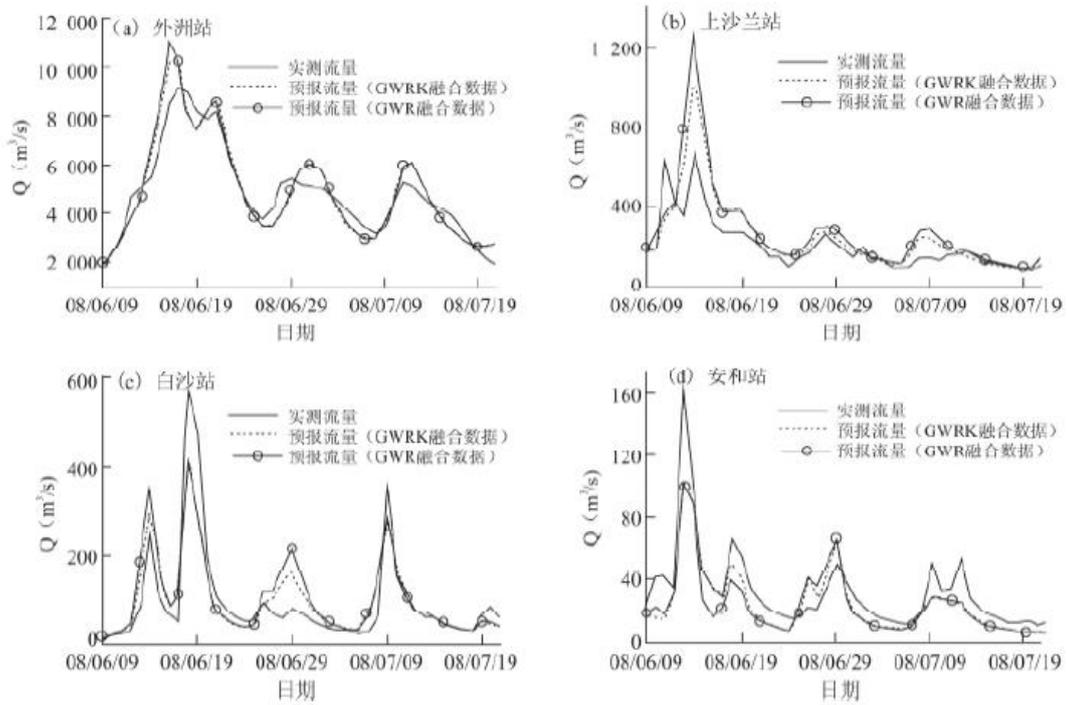


图 5 20080609 场次洪水 GWRK 和 GWR 降水融合数据的径流预报对比

Fig. 5 Comparison of the 20080609 Forecast Flood by the GWRK and GWR Rainfall Merging Data

为进一步比较两种降水融合方法在赣江流域水文预报中的差别，计算了 ΔNSE 、 ΔRB 和 $\Delta RMSE$ ，结果见图 6。由图 6 可以看出，无论对于 ΔNSE 、 ΔRB 还是 $\Delta RMSE$ 指标，相比于 GWR，GWRK 获得的降水融合数据在绝大多数流域的预报精度均得到了改善，但是改善程度不同。在 1~6 号流域（集水面积较大的干流站）各项评价指标均有轻微的改变，但是改善程度不明显，可能是受大流域空间均化现象比较突出的影响。在其他流域， ΔNSE 、 $\Delta RMSE$ 指标在绝大多数流域得到了改善，但是 ΔRB 指标增加的流域多，说明 GWRK 在面积较小的支流集水区，能有效提高预报精度，但是对预报径流系统偏差的降低不如 GWR 有效。总体来说，相比于 GWR 方法，GWRK 方法可以提高赣江流域径流预报精度，但是对径流预报精度的改善作用具有不均匀性。

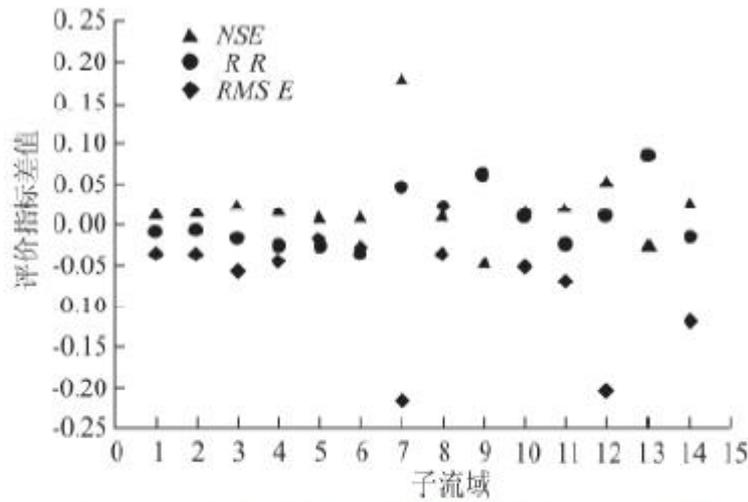


图 6 各子流域水文预报精度改善效果对比

Fig. 6 Comparison of the Improvement of Hydrological Forecast Accuracy in Each Sub Basin

4 结论与讨论

在评估赣江流域 TRMM3B42RTV7 数据精度的基础上，分别采用 GWRK 方法和 GWR 方法构建了站点观测和卫星遥测的降水融合模型，并从与实测站点的交叉检验和径流预报两方面对融合结果进行对比分析，结果表明：

(1) TRMM3B42RTV7 网格日降水数据和站点日降水数据具有较好的相关性，但定量误差较大；

(2) GWR 降水融合模型残差具有较强的空间自相关性，利用克里金对其残差进行插值来提高其降水融合数据精度是可行的；

(3) 在统计平均意义上，GWRK 相比于 GWR 方法能有效的降低站点尺度上降水融合数据的绝对偏差，但是对相关系数和相对误差的改进不够明显；

(4) GWRK 相比于 GWR，总体上可以提高赣江流域的径流预报精度，但在大流域的效果不明显。

降水融合结果要受到多种因素的影响，本文并没有考虑到不同的站点空间分布以及高程信息对融合结果可能的影响。时间尺度不同，降水融合的误差特征也不同，在今后的研究中可以探讨 GWRK 相比于 GWR 方法在日以下尺度以及月尺度上的改进效果。

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.

【LIU Y B, FU Q N, SONG P, et al. Satellite retrieval of precipitation: An overview [J]. Advances in Earth Science, 2011, 26(11) : 1162—1172. 】

-
- [2] LV A, ZHOU L. A rainfall model based on a geographically weighted regression algorithm for rainfall estimations over the arid Qaidam basin in China [J]. *Remote Sensing*, 2016, 8(4) : 311.
- [3] 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]. 清华大学, 2013.
- 【HU Q F. Rainfall spatial estimation using multi-source information and its hydrological application [D]. Tsinghua University, 2013. 】
- [4] 潘昶, 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.
- 【PAN Y, SHEN Y, YU J, et al. Analysis of the combined gauge-satellite hourly precipitation over China based on the Oltechnique [J]. *Acta Meteorologica Sinica*, 2012, 70 (6) :1381—1389. 】
- [5] NERINI D, ZULKAFI 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.
- [6] LI M, SHAO Q. An improved statistical approach to merge satellite rainfall estimates and rain gauge data [J]. *Journal of Hydrology*, 2010, 385: 51—64.
- [7] 覃文忠. 地理加权回归基本理论与应用研究 [D]. 同济大学, 2007.
- 【QIN W Z. The basic theoretics and application research on geographically weighted regression [D]. Tongji University, 2007. 】
- [8] XU S, WU C, WANG L, et al. A new satellite-based monthly precipitation downscaling algorithm with non-stationary relationship between precipitation and land surface characteristics [J]. *Remote Sensing of Environment*, 2015, 162: 119—140.
- [9] 张国峰, 杨立荣, 瞿明凯, 等. 基于地理加权回归克里格的日平均气温插值 [J]. *应用生态学报*, 2015, 26 (5) :1531—1536.
- 【ZHANG G F, YANG L R, QU M K, et al. Interpolation of daily mean temperature by using geographically weighted regression kriging [J]. *Chinese Journal of Applied Ecology*, 2015, 26(5) : 1531—1536. 】
- [10] KUMAR S, LAL R, LIU D. A geographically weighted regression kriging approach for mapping soil organic carbon stock [J]. *Geoderma*, 2012, 189—190(6) : 627—634.
- [11] 费明哲, 张增信, 原立峰, 等. TRMM 降水产品在鄱阳湖流域的精度评价 [J]. *长江流域资源与环境*, 2015, 24(8) :1322—1330.

【FEI M Z, ZHANG Z X, YUAN L F, et al. Accuracy assessment for TRMM in the Poyang lake basin [J]. Resources and Environment in the Yangtze Basin, 2015, 24 (8) : 1322 –1330. 】

[12] 黄钰瀚, 张增信, 费明哲, 等. TRMM3B42 卫星降水数据在赣江流域径流模拟中的应用 [J]. 长江流域资源与环境, 2016, 25(10) : 1618–1625.

【HUANG Y H, ZHANG Z X, FEI M Z, et al. Hydrological evaluation of the TMPA multi-satellite precipitation estimates over the Ganjing basin [J]. Resources and Environment in the Yangtze Basin, 2016, 25(10) : 1618–1625. 】

[13] FORTHERINGHAM A S, BRUNSDON C, CHARLTON M. Geographically weighted regression—the analysis of spatially varying relationships [M]. Wiley, 2002.

[14] PWRRIN C, MICHEL C, ANDREASSIAN V. Improvement of a parsimonious model for streamflow simulation [J]. Journal of Hydrology, 2003, 279(1–4) : 275–289.

[15] VRUGT J A, GUPTA H V, BOUTEN W. A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters [J]. Water Resources Research, 2003, 39(8) : 1201.

[16] HUFFMAN G J, BOLVIN D T. TRMM and other data precipitation data set documentation [EB /OL]. 2014, ftp://precip.gsfc.nasa.gov/pub/trmmdocs/3B42_3B43_doc.pdf.

[17] 谢娟英, 蒋帅, 王春霞, 等. 一种改进的全局 K 均值聚类算法[J]. 陕西师范大学学报(自然科学版), 2010(02) : 18–22.

【XIE J Y, JIANG S, WANG C X, et al. An improved global K-means clustering algorithm [J]. Journal of Shaanxi Normal University (Natural Science Edition) , 2010(02) : 18–22. 】

[18] XIAO Y, ZHANG X, WAN H, et al. Spatial and temporal characteristics of rainfall across Ganjiang River Basin in China [J]. Meteorology and Atmospheric Physics, 2016, 128 (2) :167–179.