

中国中西部地区年降水量空间分布数据集 (2010)

蒋育昊^{1,2}, 李宝林^{1,2*}, 袁焯城¹, 高锡章¹, 张涛^{1,2}, 刘岩^{1,2}, 李影^{1,2},
李红³, 罗智勇³, 马强³, 王晓敏³, 次仁多吉⁴

1. 中国科学院地理科学与资源研究所资源与环境信息系统国家重点实验室, 北京 100101; 2. 中国科学院大学, 北京 100049; 3. 青海省黄南藏族自治州河南蒙古族自治县气象局, 河南 811599;
4. 西藏自治区拉萨市气象局, 拉萨 850000

摘要: 中国中西部地区年降水量空间分布数据集 (2010) 是采用混合插值法进行高精度曲面建模 (High Accuracy Surface Modeling, HASM) 区域降水模拟的数据结果。该方法包括三个步骤: 首先使用 TRMM 数据 (或者站点观测值空间插值结果) 代表降水在空间中变化平稳的部分, 即趋势面; 然后结合地面站点观测值, 计算去除趋势面后的残差值, 并利用 HASM 对残差进行插值, 得到空间中变化不平稳的残差场; 最后将趋势面与残差场进行加和, 完成插值计算。结果表明, TRMM 作为背景场的 HASM 模拟精度明显高于传统插值方法, 适应性较好。采用平均绝对误差 (Mean Absolute Error, MAE) 和均方根误差 (Root Mean Square Error, RMSE) 作为精度评价指标完成精度验证。在全球尺度上, MAE 和 RMSE 分别为 125.15、155.80 mm; 在局部尺度上, MAE 和 RMSE 分别为 167.53、228.81 mm。与本数据集相关的科学发现论文已发表在《地球信息科学学报》2015 年第 17 卷第 8 期。

关键词: 降水; TRMM 卫星; 高精度曲面建模; 中国中西部地区; 地球信息科学学报

DOI: 10.3974/geodp.2019.04.05

1 前言

降水是重要的环境变量, 在地表径流、大气运动、农业资源等方面都扮演着不可忽视的角色。大范围降水数据通常是以离散站点的观测结果为基础, 使用空间插值的方式获得。传统的空间插值方法并未考虑空间曲面自身的特点, 在建模过程中忽略了曲面的内蕴因素对曲面重建的约束作用, 而高精度曲面建模方法 (High Accuracy Surface Modeling, HASM) 却可以很好地解决传统插值方法难以避免的峰值削平和边界动荡等问题^[1], 但是, 在站点较少的地区, HASM 模型的适用性受到很大限制。卫星遥感数据具有大范围、同步观测等优势, 可以弥补地面站点观测范围有限的不足。因此, 在利用 HASM 进行降水模拟过程中融合卫星降水产品的信息, 可以提高区域降水空间模拟精度^[2]。

本产品^[3]基于目前可以方便获取的 0.25°×0.25°分辨率 TRMM 3B43 V7 数据产品^[4-5]

收稿日期: 2019-11-07; 修订日期: 2019-12-09; 出版日期: 2019-12-24

基金项目: 中华人民共和国科学技术部 (2016YFC0500205, 2015SCB954103)

*通讯作者: 李宝林 N-8884-2019, 中国科学院地理科学与资源研究所资源与环境信息系统国家重点实验室, libl@lreis.ac.cn

数据引用方式: [1] 蒋育昊, 李宝林, 袁焯城等. 中国中西部地区年降水量空间分布数据集(2010) [J]. 全球变化数据学报, 2019, 3(4): 343-348. DOI: 10.3974/geodp.2019.04.05.

[2] 蒋育昊, 李宝林, 袁焯城等. 中国中西部地区年降水量空间分布数据集(2010) [DB/OL]. 全球变化科学研究数据出版系统, 2019. DOI: 10.3974/geodb.2019.05.18.V1.

进行研发。针对 HASM 方法在站点较少地区精度较差的问题,使用混合插值法将 TRMM 数据作为背景场(趋势面),利用 HASM 修正残差场(去除趋势后),从而提高 HASM 对区域降水空间细节的刻画能力。将订正后的产品与基于地面站点插值方法生成的结果进行对比,验证其有效性。

2 数据集元数据简介

《中国中西部地区年降水量空间分布数据集(2010)》^[3]的名称、作者、地理区域、数据年代、时间分辨率、空间分辨率、数据集组成、数据出版与共享服务平台、数据共享政策等信息见表1。

表1 《中西部地区年降水量空间分布数据集(2010)》元数据简表

条 目	描 述
数据集名称	中国中西部地区年降水量空间分布数据集(2010)
数据集短名	PrecipMidwesternChina2010
作者信息	蒋育昆 N-8765-2019, 中国科学院地理科学与资源研究所资源与环境信息系统国家重点实验室, jiangyh@lreis.ac.cn 李宝林 N-8884-2019, 中国科学院地理科学与资源研究所资源与环境信息系统国家重点实验室, libl@lreis.ac.cn 袁焯城 N-9047-2019, 中国科学院地理科学与资源研究所资源与环境信息系统国家重点实验室, yuanyc@lreis.ac.cn 高锡章 N-1655-2019, 中国科学院地理科学与资源研究所资源与环境信息系统国家重点实验室, gaoxz@lreis.ac.cn 张涛 N-8690-2019, 中国科学院地理科学与资源研究所资源与环境信息系统国家重点实验室, zhangtao@lreis.ac.cn 刘岩 N-8844-2019, 中国科学院地理科学与资源研究所资源与环境信息系统国家重点实验室, liuy.18b@igsnr.an.c.cn 李影 Y-4384-2019, 中国科学院地理科学与资源研究所资源与环境信息系统国家重点实验室, liying9391@126.com 李红, 青海省黄南藏族自治州河南蒙古族自治县气象局, lh691208@163.com 罗智勇, 青海省黄南藏族自治州河南蒙古族自治县气象局, 393352158@qq.com 马强, 青海省黄南藏族自治州河南蒙古族自治县气象局, 107082968@qq.com 王晓敏, 青海省黄南藏族自治州河南蒙古族自治县气象局, 2444869807@qq.com 次仁多吉, 西藏自治区拉萨市气象局, LSNSE111@126.com
地理区域	25°N–35°N、105°E–115°E, 包括重庆、贵州、湖南、湖北、河南大部和四川东部及陕西南部 and 甘肃的东南部, 总面积约 $1.06 \times 10^6 \text{ km}^2$
数据年代	2010
时间分辨率	月
空间分辨率	0.25°×0.25°
数据格式	.tif
数据量	48.8 KB (压缩后)
数据集组成	由 20 个数据文件组成

续表

条 目	描 述
基金项目	中华人民共和国科学技术部（2016YFC0500205, 2015CB954103）
数据计算环境	Matlab 2011b；中国科学院地理科学与资源研究所 ArcGIS 所内计算平台
出版与共享服务平台	全球变化科学研究数据出版系统 http://www.geodoi.ac.cn
地址	北京市朝阳区大屯路甲 11 号 100101，中国科学院地理科学与资源研究所
数据共享政策	全球变化科学研究数据出版系统的“数据”包括元数据（中英文）、实体数据（中英文）和通过《全球变化数据学报》（中英文）发表的数据论文。其共享政策如下：（1）“数据”以最便利的方式通过互联网系统免费向全社会开放，用户免费浏览、免费下载；（2）最终用户使用“数据”需要按照引用格式在参考文献或适当的位置标注数据来源；（3）增值服务用户或以任何形式散发和传播（包括通过计算机服务器）“数据”的用户需要与《全球变化数据学报》（中英文）编辑部签署书面协议，获得许可；（4）摘取“数据”中的部分记录创作新数据的作者需要遵循 10% 引用原则，即从本数据集中摘取的数据记录少于新数据集总记录量的 10%，同时需要对摘取的数据记录标注数据来源 ^[6]
数据和论文检索系统	DOI, DCI, CSCD, WDS/ISC, GEOSS, China GEOSS

3 数据研发方法

3.1 算法原理

稀疏站点以及地表变化较大区域的降水插值结果不确定性较大，而遥感降水产品可以较好地体现降水空间分布模式。因此，本数据集利用混合插值法，结合 HASM 模型和遥感降水产品对区域降水进行模拟。该方法主要包括三个步骤：首先使用 TRMM 数据（或者对照方法中的空间插值结果）代表降水在空间中变化平稳的部分，即趋势面；然后结合地面站点观测值，计算去除趋势面后的残差值，并利用 HASM 对残差进行插值，得到空间中变化不平稳的残差场；最后将趋势面与残差场进行加和，得到最终插值结果。

3.2 技术路线

本文在混合插值中，利用 HASM 模拟残差场，根据反距离加权（IDW）、样条函数（Spline）和克里金（Kriging）插值结果以及 TRMM 数据来计算趋势面，其对应的插值结果分别表示为 HASM_I、HASM_S、HASM_K 和 HASM_T。HASM 模型方程组解的求取是利用预处理共轭梯度法（PCG）^[7]迭代进行（图 1）。

使用随机选取的标准气象站点作为建模点，其余标准气象站点作为全局验证点，用来检验大尺度范围的模型模拟精度；在不改变建模点的前提下，将全部地方气象站点作为局部验证点，用

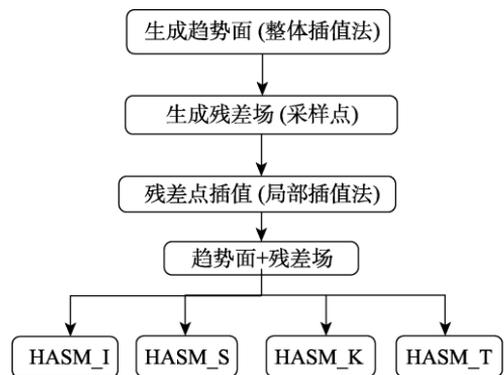


图 1 中国中西部地区年降水量空间分布数据集（2010）研发技术路线

于检验小尺度范围的模型模拟效果。利用 TRMM 数据产品、IDW、Spline 和 Kriging 插值结果,以 $0.25^{\circ} \times 0.25^{\circ}$ 为分析单元,进行背景场的 HASM 模拟,并比较结果精度。

4 数据结果与验证

4.1 数据集组成

数据集由 20 个数据文件组成,其中 hasm_idw2010.tif、hasm_kriging2010.tif、hasm_spline2010.tif 和 hasm_trmm2010.tif 代表研究区范围内分别以反距离权重插值结果 (IDW)、克里金插值结果 (Kriging)、样条函数插值结果 (Spline) 以及 TRMM 数据作为 HASM 驱动场,利用混合插值法计算得到的 2010 年降水数据,其空间分辨率为 0.25° 。

4.2 数据结果

从图 2 可以看出,四种来源的降水数据具有相似的空间分布格局,各种背景场下 HASM 所模拟的降水空间分布格局差异较小,降水量整体呈现出自东南向西北逐渐减少的趋势。而 TRMM 作为驱动场的 HASM 模拟结果在没有站点观测的区域展现出了更多的降水空间变化信息,优于其他以传统插值方法生成驱动场的 HASM 模拟结果。图 2 中 A 地位于我国湿润半湿润地区向半干旱区急剧过渡区,降水应为从东南向西北呈现迅速减少的趋势^[8],在四种方法中,只有 HASM_T 方法模拟得到了这种变化。B 地位于渭河平原, HASM_T 方法模拟得到的降水分布与前人研究结果相一致。C 处为武陵山脉北支的澧水上游,为湖南省主要降水中心之一, D 处为衡阳-邵阳丘陵区,是湖南省主要“干旱区”^[9], HASM_T 方法清楚地反映了这两个降水极值地区。

4.3 数据结果验证

由表 2 可知,以 TRMM 数据作为驱动场的 HASM 模拟结果精度优于以传统插值方法为驱动场的 HASM 模拟结果。全局验证结果中, HASM_T 结果的 MAE 和 RMSE 分别为 125 和 156 mm, HASM_I、HASM_S 和 HASM_K 的 MAE 分别为 212、234 和 192 mm,比 HASM_T 的结果高 70%、87% 和 54%; RMSE 分别为 260、328 和 241 mm,比 HASM_T 的结果高 67%、110% 和 54%。局部验证结果中, HASM_T 结果的 MAE 和 RMSE 分别为 168、229 mm; HASM_I、HASM_S 和 HASM_K 的 MAE 则分别为 196、198 和 197 mm,比 HASM_T 的结果高 17%、18% 和 17%; RMSE 分别为 263、260 和 256 mm,比 HASM_T 的结果高 15%、14% 和 12%。

表 2 不同背景场的 HASM 计算精度 (mm)

验证点	HSAM_I		HSAM_S		HSAM_K		HSAM_T	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
全局	212	260	234	328	192	241	125	156
局部	196	263	198	260	197	256	168	229

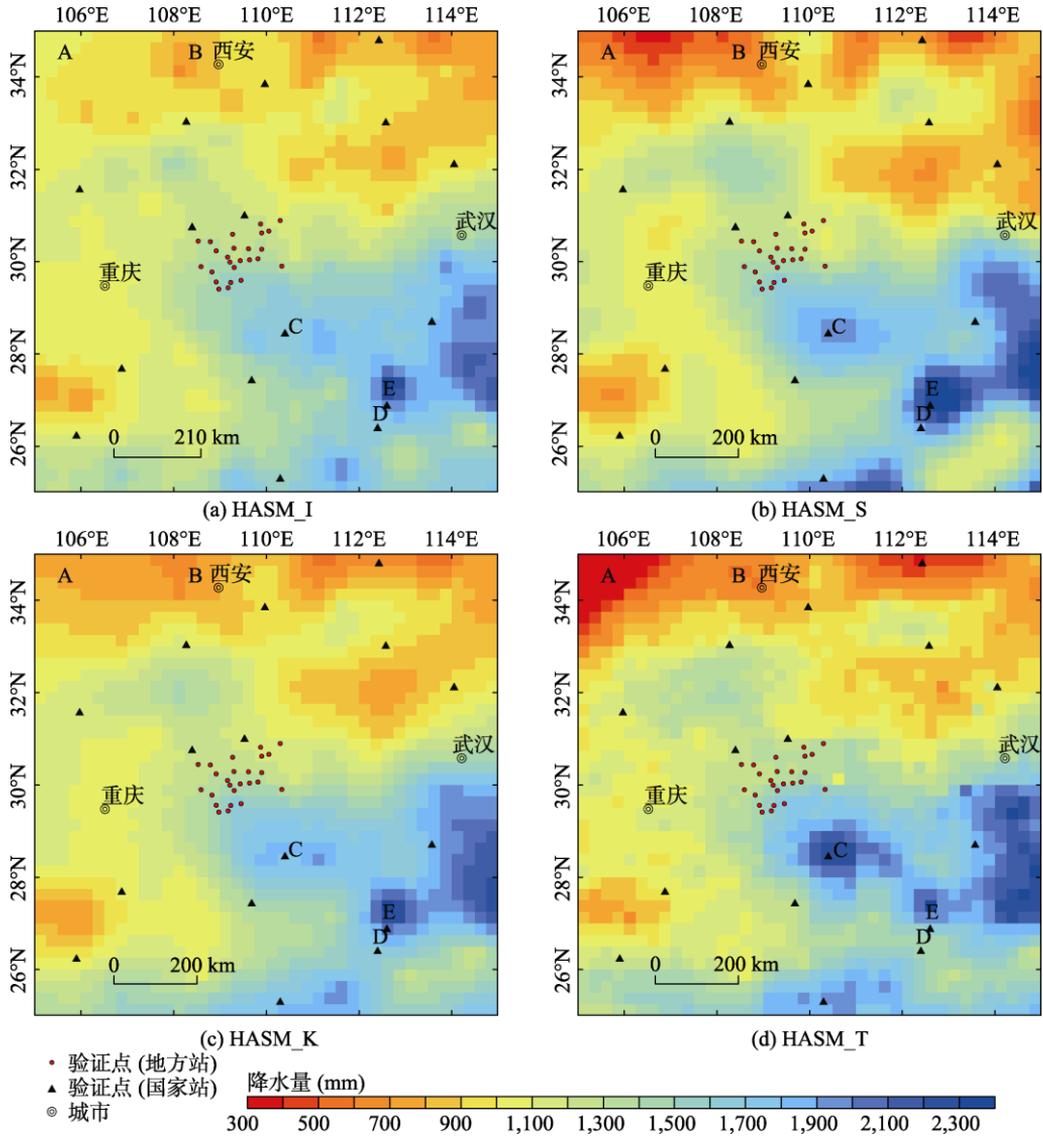


图2 不同背景场 HASM 模拟的降水量分布

5 讨论与结论

本数据集利用 TRMM 卫星降水数据作为背景场，结合 HASM 模型修正残差场，以中国中西部地区为例，模拟了 2010 年降水量空间分布；同时利用传统插值方法得到同区域降水空间分布数据。TRMM 作为背景场的 HASM 模拟精度明显高于传统插值方法对应的模拟精度。在全球尺度上，其 MAE 和 RMSE 分别为 125.15、155.80 mm；在局部尺度上，其 MAE 和 RMSE 分别为 167.53、228.81 mm。在局部区域 TRMM 作为背景场的 HASM 模拟结果更能反映降水的基本空间分布规律，模拟结果不仅可以显示降水的迅速变化趋势，反映降水极值地区，更能有效地避免传统空间插值算法中，由个别

站点观测高值导致的突出现象,并且在不同子区域内,以 TRMM 为背景场的 HASM 模拟误差较低,适应性较好。目前,在气象要素空间插值的相关应用中,不可避免地会出现由于站点获取难度较大而带来的实测数据不足等方面的问题,造成了气象要素模拟结果的不确定。因此,将 HASM 模型与其他能够反映待模拟气象要素空间分布结果的其他辅助数据相结合,以弥补实际可利用采样点数量的不足,是 HASM 在实际应用过程中应深入研究的方向。

作者分工: 李宝林、蒋育昊对数据集的开发做了总体设计;刘岩、李影处理了 TRMM 降水估值数据;张涛、袁焯城设计了模型和算法;高锡章做了数据验证;李红、罗智勇、马强、王晓敏和次仁多吉提供了部分气象站点数据;蒋育昊、李宝林撰写了数据论文。

参考文献

- [1] 赵娜, 焦毅蒙. 基于 TRMM 降水数据的空间降尺度模拟[J]. 地球信息科学学报, 2018, 20(10): 20–27.
- [2] 张涛, 李宝林, 赵娜等. 结合 TRMM 数据的区域降水高精度曲面建模研究[J]. 地球信息科学学报, 2015, 17(8): 895–901
- [3] 张涛, 李宝林, 袁焯城等. 中国中西部地区年降水量空间分布数据集(2010) [DB/OL]. 全球变化科学研究数据出版系统, 2019. DOI: 10.3974/geodb.2019.05.18.V1.
- [4] Kummerow, C., Barnes, W. The tropical rainfall measuring mission (TRMM) sensor package [J]. *Journal of Atmospheric and Oceanic Technology*, 1998, 15: 809–817.
- [5] 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.
- [6] 全球变化科学研究数据出版系统. 全球变化科学研究数据共享政策 [OL]. DOI: 10.3974/dp.policy.2014.05 (2017 年更新).
- [7] Golub, G. H., Van Loan, C. F. *Matrix Computations* [M]. Beijing: Posts & Telecom Press, 2009.
- [8] 赵鑫, 马金珠, 聂中青等. 天水地区降水时空特征及其对印度洋偶极子的响应[J]. 干旱区研究, 2009, 26(6): 917–922
- [9] 游兴, 李昌, 喻妮厚. 湖南省降水特性分析[J]. 湖南水利水电, 2004(6): 41–43.