

基于广义加性模型估算1979-2014年三峡库区降水及其特征分析

王圆圆, 郭 微, 李贵才, 郭兆迪
(国家卫星气象中心, 北京 100081)

摘要: 高精度、长时间序列、空间连续的降水产品是气候分析、水文模拟等众多研究领域中的重要输入。近期欧洲研究人员融合了3种不同类型的降水数据: 站点观测、遥感产品、再分析资料, 制作了全球1979-2014年3小时一次0.25°空间分辨率的降水产品(Multi-Source Weighted-Ensemble Precipitation, MSWEP), 凭借高时空分辨率及其对多源信息的挖掘和融合, 该产品一经发布即受到广泛关注。本文利用三峡库区及附近地区气象站点的降水资料对MSWEP月降水数据进行评估, 采用广义加性模型算法(General Additive Model, GAM)融合站点降水空间插值结果和MSWEP产品, 对三峡库区融合后降水进行分析。主要结论为: ① 降水估算精度呈现冬春季偏高、夏秋季偏低的特征, MSWEP产品与站点插值方法具有互补性, 前者对夏秋季降水估算精度更高, 后者对冬春季降水估算精度更高; ② GAM算法可以充分发挥站点插值和MSWEP数据各自的优势, 提高区域降水估算精度, 与融合前相比, 均方根误差减少了17%~50%, 相关系数 r 提高了10%~30%; ③ 2003年库区蓄水前后降水变化的主要特征有: 库区中部长江以南地区汛期降水(5-10月)下降, 库区西部干季(1-4月, 11-12月)降水增加, 库区外围西北部(大巴山地区)汛期降水增加, 降水空间格局异质性增加, 干季降水占全年降水比例升高。

关键词: 降水; 三峡库区; MSWEP; GAM

DOI: 10.11821/dlxb201707007

1 引言

由于较高的时空异质性, 降水是最难估算的气象变量之一^[1-3]。目前获取降水信息的主要方法包括: 雨量计观测、地基降水雷达、卫星数据反演、数值预报模式模拟等, 这些方法在观测原理、算法精度、产品时空分辨率和覆盖范围等方面存在巨大差异和互补性^[4]。融合多源信息, 获取高精度、长时间序列、空间连续的降水产品已成为众多研究领域(气候分析、水文模拟、生态监测等)中的热点之一^[4-7]。近期, 欧洲研究人员^[4]制作了1979-2014年全球3小时一次的0.25°分辨率的降水产品(Multi-Source Weighted-Ensemble Precipitation, MSWEP), 该产品根据站网密度、卫星降水反演精度、数值预报降水精度, 适时调整权重, 再通过加权平均的方式融合上述3种降水的结果, 利用全球125个通量塔的降水观测作为独立样本进行检验, 结果表明MSWEP的精度高于其他常用降水产品(WFDEI-CRU, GPCP-1DD, TMPA 3B42, CPC unified)。目前已有多名学者正在利用MSWEP数据开展水文模拟、气候变化等方面的研究。

收稿日期: 2016-11-25; 修订日期: 2017-04-05

基金项目: 国家公益气象行业专项(GYHY201406035), 国家自然科学基金项目(41301461) [Foundation: R&D Special Fund for Public Welfare Industry (Meteorology), No.GYHY201406035; National Natural Science Foundation of China, No.41301461]

作者简介: 王圆圆(1981-), 女, 安徽六安人, 博士, 副研究员, 主要从事生态环境遥感应用研究。

E-mail: wangyuanyan@cma.gov.cn

三峡地区地形复杂,人类活动强度大,景观破碎度高,作为世界最大的水电工程——三峡水库修建后,原本有植被覆盖的陡峭山坡被平坦水域所代替,下垫面的变化如何影响当地及周边地区的降水,一直是众多学者关注和研究的对象^[8-14]。部分学者认为三峡水库建立后,会增加库区气候的海洋性特征,对极端天气事件的影响不明显^[13],对降水的影响也不明显^[8-9]或仅限于近库区^[11];但也有学者认为三峡水库的气候影响不只局限于库区周边数十公里范围,而是可以扩展到数百公里以外的区域,Wu等^[10]利用TRMM数据和MM5模型模拟,发现2003年三峡水库蓄水后,北部秦岭大巴山地区的降水增加,而水库邻近地区降水减少;Xiao等^[12]利用站点长时间序列(1960-2005年)的观测数据,认为Wu等^[10]观测到的现象没有超出降水的自然变化范畴,无法证明是三峡水库的影响。

上述研究中除利用数值模拟技术,主要依赖2种降水数据:站点观测和卫星产品。站点降水虽然精度高,但代表性差,尤其是对于地形复杂的三峡地区;卫星降水产品空间分辨率高,可以进行大面积的整体观测,但时间序列短(1998年至今),精度有限,因此很有必要将站点和卫星数据结合起来,分析三峡库区降水的变化特征。目前常用的方法是利用站点降水对卫星降水进行线性标定,但由于卫星降水误差具有明显的季节性和空间分布特征(例如,对于冬季降水、弱强度降水、地形复杂或高纬度地区,卫星降水误差较大)^[15],简单的线性标定无法取得令人满意的效果。近期,广义加性模型(General Additive Model, GAM)正受到学者日益关注^[16],GAM具有良好的灵活性、非线性以及对空间数据的平滑处理能力,可以自适应的调整站点插值结果和卫星产品的权重,对两者进行有效融合,前人已开展多项研究,如基于遥感数据估算气温^[17]、估算辐射^[18-19]、评估滑坡可能性^[20]等,结果表明GAM精度高、过学习风险小,是值得推广应用的方法。

本文的研究目的是将最新降水产品(MSWEP)用于三峡库区降水特征分析。首先基于气象站点降水观测数据评估了MSWEP在三峡地区的误差,而后利用GAM算法融合了MSWEP和站点插值降水结果,最后基于融合后的高精度区域降水数据,分析三峡库区在2003年水库蓄水前后的降水变化。本文和以往基于卫星降水的分析相比,在保持高空间分辨率(0.25°)的同时大大延长了时间序列(1979-2014年),利用GAM算法充分发挥台站和卫星数据的优势,提高了降水数据精度。本文研究结论不仅为三峡库区降水变化特征提供了新的信息,也为MSWEP降水产品在局地的精度提供参考。

2 研究区域、数据、方法

2.1 研究区域

研究区域如图1所示,地理范围: $33.375^{\circ}\text{N}\sim 26.625^{\circ}\text{N}$ 、 $104.875^{\circ}\text{E}\sim 113.625^{\circ}\text{E}$ (网格中心的经纬度),研究区包含了三峡库区及周边地区,选择如此大的范围一方面是为了增加区域内站点数量,另一方面是为了观察库区和周边区域在降水变化特征方面是否有明显不同。三峡库区的范围大致位于 $32^{\circ}\text{N}\sim 28^{\circ}\text{N}$ 、 $106^{\circ}\text{E}\sim 112^{\circ}\text{E}$,库区内共有34个县(区),靠近库区的沿江县(区)有12个(白色矢量),其余22个为远库区县(区)(黑色矢量)。根据长江两岸地表特征,可以将库区分为两段,重庆至万州为西段,该段区域内长江走向为西南—东北方向,长江北岸地形平坦,以农田覆盖为主,长江南岸地形略有起伏,地表覆盖为农田、森林交错。从云阳至宜昌为东段,该段内长江走向开始水平向东而后向东南倾斜,长江两岸均为浓密的森林,峡谷地形突出。长江离开库区东段以

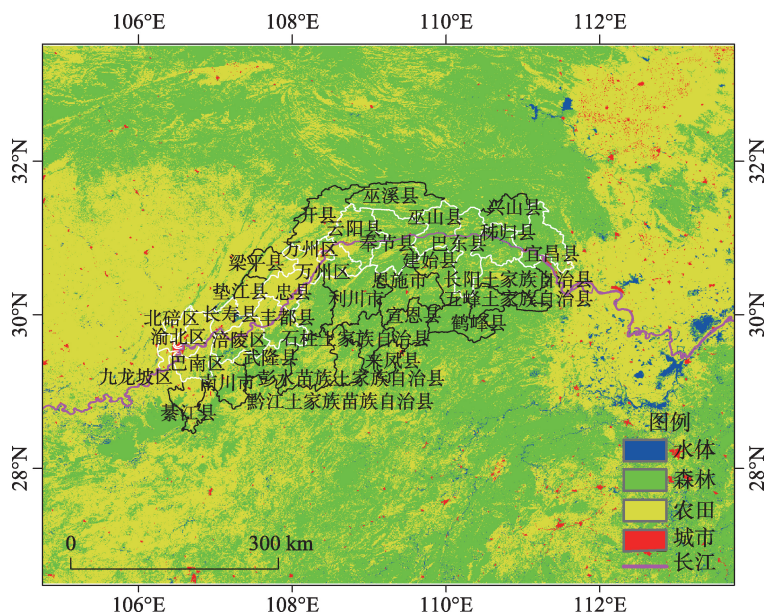


图1 研究区 2006 年 MODIS 土地覆盖图及库区县位置

Fig. 1 Land cover of the study region (MCD12 product in 2006) as well as the boundaries of counties located near the Yangtze River

后, 来到地势平坦地区, 农田为主要土地覆盖。三峡库区在 2003 年 6 月蓄水, 水位由 66 m 涨至 135 m。图 2 显示了研究区的地形图、主要山脉和 220 个气象站点位置。可以看出, 站点的分布比较均匀, 两两站点之间距离的平均值是 380 km。

2.2 研究数据

MSWEP 数据: 1979 年 1 月至 2014 年 12 月 MSWEP 数据产品, 空间分辨率为 0.25° , 时间分辨率为月 (原始数据为 3 小时一次, 处理为月累计降水)。根据研究范围, 提取研究区内格点上 MSWEP 降水数据时间序列。MSWEP 数据的生成过程和下载地址请参考文献[4]。

气象站点数据: 从国家气象信息中心的数据库中获取了三峡库区及周边范围内的

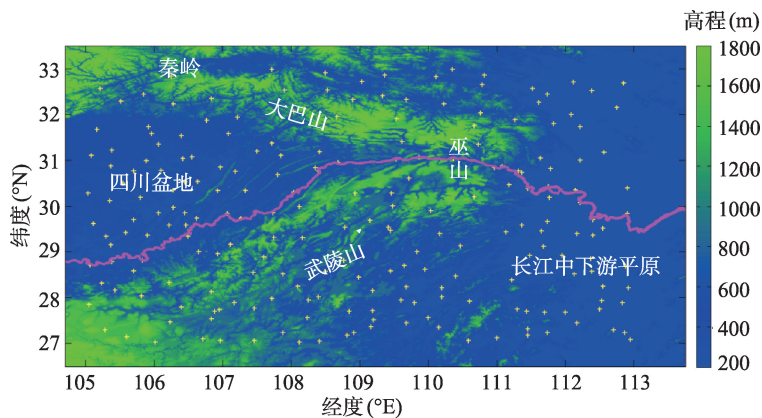


图2 研究区高程、站点分布和主要山脉

Fig. 2 DEM, meteorological stations network, and main mountains of the study region

220个气象站点1979年1月-2014年12月的月降水数据,根据站点经纬度确定其所属的0.25°格点,并提取该格点上MSWEP月降水数据,用于对比分析。

2.3 研究方法

2.3.1 三峡库区站点降水插值和MSWEP数据的精度对比 尽管MSWEP数据已包括了站点的降水观测信息^[4],但只利用了全球的一小部分站点,在局部区域如果有更密集的气象台站(如本研究区有220个气象站),利用插值算法也可以得到空间连续的降水数据,那么MSWEP降水数据和区域降水插值结果比较,谁的精度更高?

为了对此进行分析,本文针对每个站点每个月提取了3种降水数据,分别是:①该站点对应格点上的MSWEP月降水数据 P_{iju} (其中 i 代表站点, j 代表月份, u 代表年份);②利用最近邻插值算法得到的该站点降水数据 T_{iju} ;③该站点实际观测的降水数据 D_{iju} 。前两者代表了假设某站点没有降水观测条件下,可用的两种估算降水的方法(即遥感产品或空间插值),将这两种方法的估算结果和真实降水观测进行对比,则可以评价两种方法的精度。评价的标准主要包括:相关系数(r),均方根误差(rmse),相对均方根误差(relative rmse),以及偏差(bias),为了分析误差随季节的变化,本文首先针对 k 年 j 月的降水数据计算误差参数,而后统计其多年的月均值。

2.3.2 基于GAM方法融合站点观测和MSWEP数据 由于区域插值算法和MSWEP降水数据具有很强的互补性,本文利用GAM算法融合了两种降水信息,模型的形式为:

$$y = s(x_1, x_2) + x_3 \quad (1)$$

式中: x_1 、 x_2 分别为站点经纬度; $s(x_1, x_2)$ 表示利用薄盘样条函数对局部做平滑; x_3 为站点MSWEP降水数据; y 为站点降水观测数据,模型中需要估算的参数是 s 函数的形式,以及 x_3 的系数,通过最小化目标函数(2)实现参数估算。

$$\|y - s\|^2 + \lambda \int [s'']^2 dx \quad (2)$$

目标函数由模型误差和模型复杂度两部分组成, λ 为光滑参数,在误差和复杂度之间起平衡作用,可以用广义交叉验证(Generalized Cross Validation, GCV)或AIC(Akaike's Information Criterion)为标准进行确定,本文使用GCV方法。在拟合薄盘样条时,需要估算的参数个数随着样本量的增加而增加,当样本量大时,计算量很大且容易受到噪音的影响,对此Wood设计了薄盘样条回归函数^[16],该函数需要设定一个基础维数(k)值,来确定模型自由度的上限^[21],如果模型优化后得到的有效自由度(effective degree of freedom, edf)非常接近 k 值,说明先前给定的或缺省的 k 值偏低,模型形式过于简单以至于真实的数据分布无法被很好的模拟。为改善模拟精度,需要增加 k 值(但不能超过用于模型训练的样本个数),当拟合后模型的edf明显小于 k 值(一般要求 $\text{edf} < 0.75 \times k$),并且不随 k 的增加而发生明显变化,此时的 k 值比较适宜,如果再增加,计算量会上升,过拟合的风险也会增加。针对研究数据,本文通过检验 k 值和模型edf的变化,发现更适宜的模型形式是 $k = 70$ 的时候,即:

$$y = s(x_1, x_2, k = 70) + x_3 \quad (3)$$

利用站点每月的降水数据和经纬度可以拟合得到一个GAM模型,将区域的经纬度和该月MSWEP数据输入到模型中,则可以得到区域在该月融合后的降水数据。从模型的结构中可以看出来,融合后的降水数据综合了局部样条平滑插值和MSWEP线性函数两部分信息,两者的权重由他们与站点实测降水的接近程度决定,这种融合策略比单纯基于地理位置的插值或是利用站点降水修订MSWEP的方法要更具有优势。

为了对 GAM 模型精度有客观的评价, 本文采取了留一法 (Leave One Out, LOO) 计算误差。针对某年某月的降水数据, 每次选择一个站点作为检验, 用剩下的站点建立 GAM 模型, 而后用该模型估算检验站点的降水数据, 如此迭代, 直到所有站点的月降水数据都被用作检验样本, 具有 GAM 模型估算的降水数据, 将此估算值和真实值对比, 即可评价 GAM 算法的精度。

2.3.3 三峡库区的降水特征分析 基于 GAM 融合后的降水数据提取降水特征, 为了突出水库蓄水的影响, 将数据分为 2 个时期: 1979-2002 年和 2003-2014 年, 对前期、后期降水的多年平均值、Mann-Kendall 变化斜率等进行统计^[22]。由于研究区域的降水主要集中在夏季, 在不同季节降水的空间格局差异巨大, 且估算误差也明显不同, 因此, 本文将一年分为 2 个时期, 5-10 月为汛期, 1-4 月以及 11 月、12 月为干季, 对汛期和干季的降水特征分别统计。

本研究时间跨度较大, 在近 36 年的时间里, 人类活动、地表覆盖/利用状况都发生很大变化, 很有可能对降水的空间均匀性造成影响, 对此, 本文计算了降水空间均匀性的分布图, 每个格点的空间均匀性定义为该格点年降水时间序列和周边 8 个相邻格点 (3×3 窗口内) 平均年降水时间序列的相关系数的平方 (R^2), 如果 R^2 很大, 说明该格点可以很好的解释周边降水的年际变化, 对周边降水有较好代表性, 即降水的空间均匀性较高; 如果 R^2 很小, 则表明该格点的降水时间序列与周边差异较大, 空间代表性较差, 降水的空间均匀性偏低。该方法曾被用于计算站点观测对周围气象参数的解释方差, 以评估站点的代表性^[23-24]。

3 结果分析

3.1 最近邻插值和 MSWEP 降水的精度对比

MSWEP 和最近邻插值方法估算站点月降水量的精度如图 3 所示, 为了体现精度的季节变化, 本文统计了误差变量在 1979-2014 年期间的月均值, 从图 3a、3b 中可以发现这

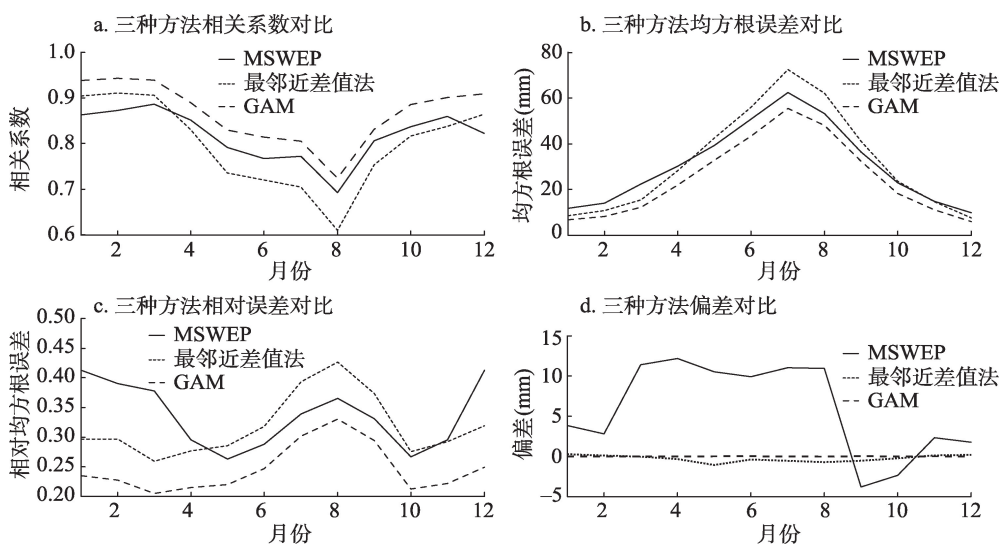


图3 MSWEP降水、最近邻插值和GAM估算精度的对比

Fig. 3 Accuracy comparison among MSWEP, Nearest Neighbour, and GAM estimation results

两种方法都是在冬春季效果更好 ($r = 0.85 \sim 0.9$, $\text{rmse} = 10 \sim 30 \text{ mm}$), 在夏秋季效果更差 ($r = 0.7 \sim 0.8$, $\text{rmse} = 40 \sim 70 \text{ mm}$), 但这两种方法的优势体现在不同季节, 冬春季时插值方法优于 MSWEP, 而夏秋季时却是 MSWEP 优于插值方法, 这和降水的季节分配、空间格局紧密相关。冬春季对流降水偏少, 降水的空间均匀度较高, 有助于提高插值算法的精度, 而受地面上时有冰雪且小雨偏多的影响, 遥感降水产品误差较大, 从图 3c 中可看出, 由于冬春季降水量少, MSWEP 的相对 rmse 可达 40%, 明显高于插值算法 (一般在 30% 以内), 上述结果也表明在 MSWEP 数据的生成中, 虽然冬季降水信息主要依赖于站点资料和再分析资料^[4], 但精度仍然不甚理想。夏秋季对流降水偏多, 降水量大且空间异质性高, 空间插值算法难以准确表达降水的空间格局, 而遥感产品却可以通过对区域的全面观测获取更准确的降水信息。除季节影响外, MSWEP 和地面站点观测的尺度不匹配也是造成误差的主要原因之一。

图 3d 显示了平均偏差的季节变化, 可以看出, 插值算法的平均偏差较小, 并且几乎不随季节发生变化, 相反, MSWEP 具有明显的偏差, 且这种偏差具有季节性, 具体而言, 3-8 月, MSWEP 比实测数据高出 10~12 mm, 在 1-2 月和 11-12 月, MSWEP 比实测数据高出 2~4 mm, 而在 9-10 月, MSWEP 比实测数据平均低 2~4 mm。MSWEP 或遥感降水相对于站点降水的高估现象, 在其他类似研究中也有描述^[14], 但 9-10 月份出现低估, 此前并没有相关研究发现, 这值得进一步探讨, 秋汛是三峡库区的主要气象灾害之一, MSWEP 的低估可能与秋季降水类型、大气条件有关 (秋季发生连阴雨、持续时间长, 但雨强不大)。

3.2 GAM 融合降水精度

图 3 中显示了利用 GAM 方法融合降水的误差, 与 MSWEP 产品和最近邻插值相比, 融合降水的 rmse 减少了 17%~50%, r 提高了 10%~30%, 尽管精度提高明显, 精度的季节特征并没有发生变化, 仍然是夏秋季降水的估算误差较大 ($\text{rmse} = 20 \sim 50 \text{ mm}$, $r = 0.75 \sim 0.85$, 相对 $\text{rmse} = 25\% \sim 33\%$), 而冬春季降水估算误差较小 ($\text{rmse} = 8 \sim 20 \text{ mm}$, $r = 0.9 \sim 0.95$, 相对 $\text{rmse} = 20\% \sim 25\%$)。

融合后降水精度的提高源于 GAM 算法充分地发挥了插值算法和 MSWEP 数据各自的优势, 这一点可以从样条函数的有效自由度 (edf) 以及 MSWEP 的回归系数中可以看出 (图 4), 在冬春季, 样条函数的 edf 较高 (超过 45), 而 MSWEP 系数偏小 (明显小于 1), 表明样条函数拟合的数据分布格局有丰富的细节, MSWEP 产品对融合降水的影响偏小; 在夏秋季, 样条函数的 edf 偏低 (低于 35), MSWEP 偏大 (更接近于 1), 表明样条函数拟合的格局非常概括化, 降水估算信息更多来自于 MSWEP 产品。

3.3 降水估算精度与高程的关系

地形影响降水, 降水估算精度可能和高程有关, 为了体现两者的关系, 本文计算了每个站点年降水观测值与估算值的相关系数, 而后按照高程将站点分成 4 组 ($< 200 \text{ m}$, $200 \sim 500 \text{ m}$, $500 \sim 1000 \text{ m}$, $> 1000 \text{ m}$), 统计每组站点相关系数的平均值 (图 5)。随着高程的增加, 估算精度也略有下降, 这可能和山区降水的时空异质性更高有关。

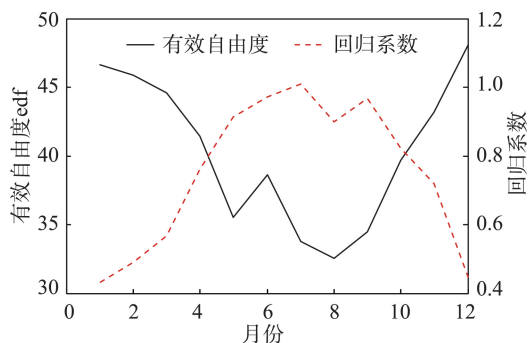


图 4 GAM 模型中样条函数有效自由度(edf)和 MSWEP 的回归系数值的月均值

Fig. 4 Monthly mean values for edf and linear coefficients obtained in fitting GAM models

MSWEP 的效果优于最邻近差值法, 尤其在站点稀少的高海拔地区 ($> 1000\text{ m}$), MSWEP 的优势更加明显, 这说明站点个数是限制最邻近差值法的重要影响因子。当利用 GAM 算法融合 MSWEP 和站点降水时, 精度进一步提高, 与 MSWEP 相比, GAM 的 r 值提高了 4%~7%, 最明显的是在海拔居中的区域 (200~1000 m), 间接说明低海拔和高海拔处降水估算精度改进的余地较小。

3.4 三峡库区降水特征分析

3.4.1 降水均值 降水均值空间格局具有如下特征: 汛期和干季的降水格局明显不同 (图 6a、6d), 干季降水更为均匀, 从东南往西北方向逐渐降低, 有明显的分带性特征; 汛期降水格局更为复杂, 有一条东南—西北方向的降雨带, 雨带上的降水量明显高于两侧, 而且雨带上有两个降雨高值中心, 分布在长江两岸。由于汛期降水量大, 年总降水量的格局基本由汛期降水格局决定。

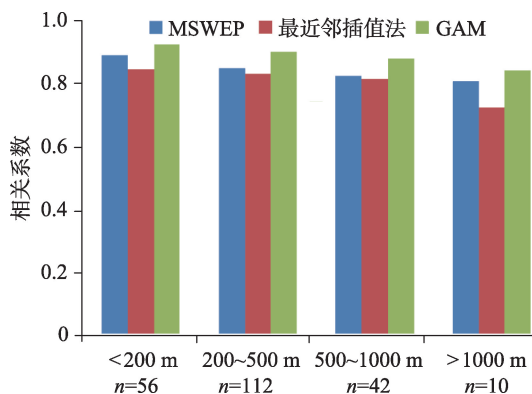


图5 不同海拔区域三种方法估算精度的对比 (n 值为站点个数)

Fig. 5 Accuracy comparison among different elevation groups (n is the number of stations in each group)

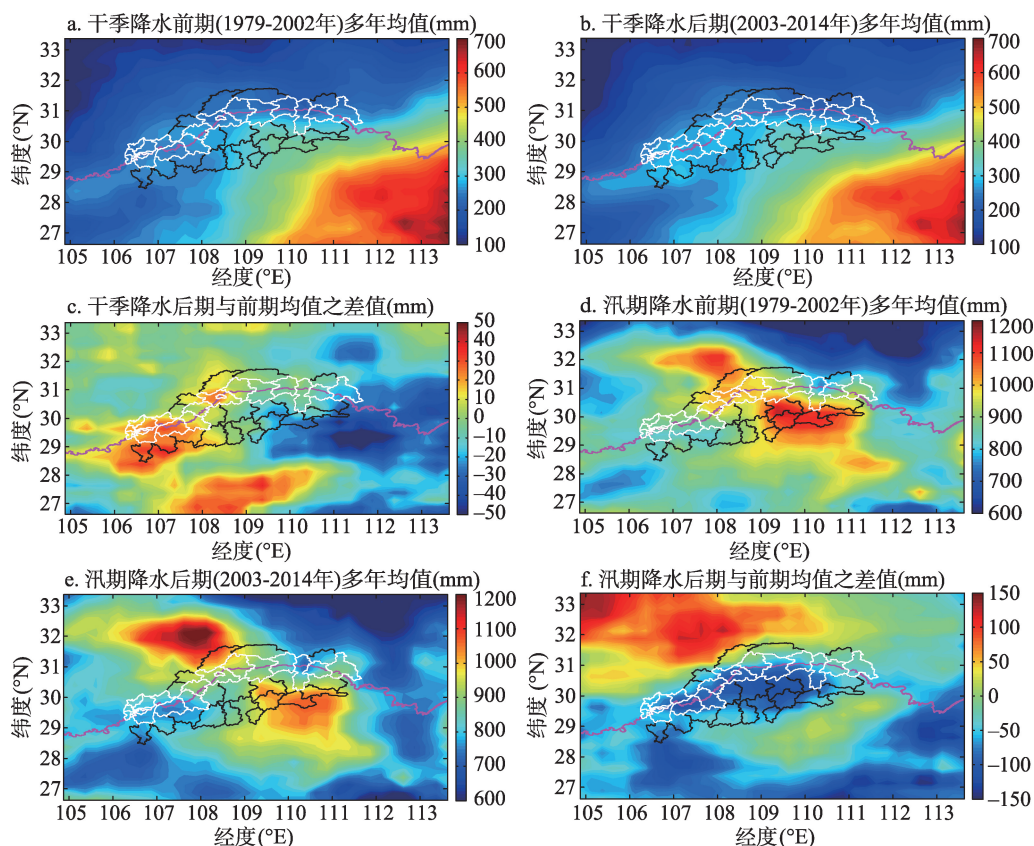


图6 干季和汛期多年平均降水在水库蓄水前后的空间格局及其变化

Fig. 6 Mean precipitation in dry and wet season and the precipitation changes before and after dam water level rose to 135 m

汛期降水变化格局：汛期降水格局在蓄水前后发生了明显变化（图6d、6e、6f），位于长江南岸的降水高值区域空间范围明显缩小、强度减弱，而位于长江北岸的降水高值中心空间范围明显扩大、强度增强，这导致了原本比较连续的东南—西北方向的雨带在靠近长江地区出现了间断。从汛期的降水变化图上可以看出，西北区域（大巴山地区）降水增加，中部、南部区域降水减少。三峡库区内部汛期降水以下降为主，下降最明显的是库区内长江以南地区。汛期降水的减少可能是大量水域在夏季升温缓慢，导致上升对流减少、下沉的大气增多，更容易驱散水汽所造成^[9]。由于汛期降水量明显高于干季，所以汛期降水的变化格局也主导了全年降水的变化格局，受篇幅限制，此处没有显示年降水变化图。

干季降水变化格局：干季降水在蓄水前后也发生了较大变化（图6a、6b、6c），从前后降水差异图上可以看出，蓄水后干季降水在东部地区下降，在南部及西南部地区上升，从变化数值上看，南部的变化幅度高于北部地区（这与南部的干季降水量较大有关）。对于三峡库区内部，除东部小部分地区干季降水略有下降外，大部分区域都表现为上升，尤其是在库区西部，干季降水量明显增多。库区西部为重庆地区，城市快速扩张，干季热岛效应更为明显，可能是导致降水增加的一个原因。

3.4.2 降水变化趋势 图7显示了干季、汛期、全年降水量在水库蓄水前后的MK斜率的

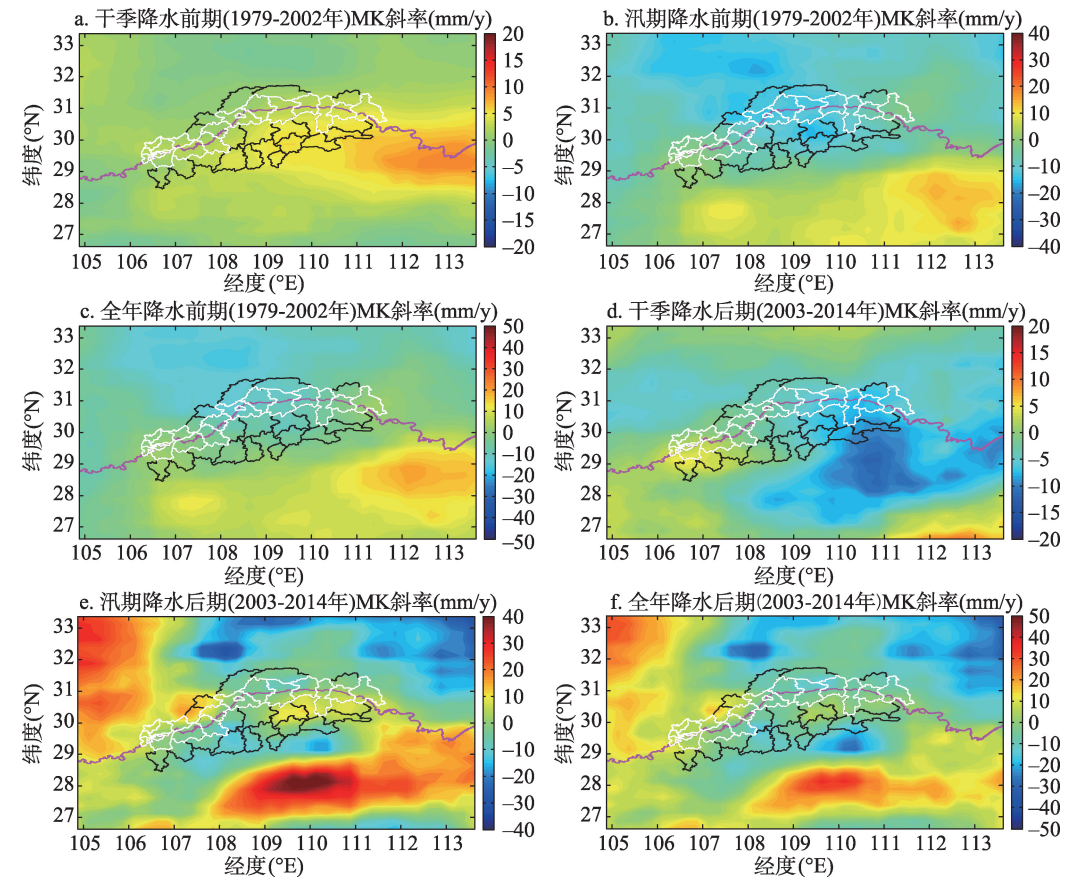


图7 干季、汛期、全年降水变化MK斜率在蓄水前后的分布格局

Fig. 7 MK trends of precipitation in dry season, wet season, and the whole year before and after dam water level rose to 135 m

分布格局,从中可以看出以下特征。

在蓄水前,干季、汛期、全年降水的变化斜率具有较为均匀的空间分布格局(图 7a、7b、7c),而在蓄水后,汛期和全年降水变化斜率的空间分布有明显的异质性,正负斜率在空间分布上有相互交错(图 7d、7e、7f)。从斜率数值变化上看,干季降水由前期的正斜率为主变为后期的负斜率为主(图 7a、7d),说明前期的湿润化趋势转化为后期的干燥化趋势;对库区内部而言,西部干季降水在后期增加明显,东部干季降水由前期的增加转为后期的下降。

汛期降水前期斜率值明显低于后期(图 7b、7e),说明后期变化更为剧烈(这与后期数据序列较短也有关);从格局上看,前期汛期降水为北降南升的格局,后期则更为复杂,西北、东南以上升为主,东北以下降为主;对于三峡库区而言,除西南部汛期降水斜率变化较小外,其他地区的斜率都出现了增加,意味着后期与前期相比,湿润化的趋势更为显著。

全年降水变化斜率格局与汛期降水类似。库区内部的斜率变化整体以平缓上升为主,说明后期与前期相比,全年降水朝着更加湿润的趋势发展。

3.4.3 降水季节分配 由于干季和汛期降水的变化趋势不同,造成了降水的季节分配也出现很大变化,对此计算了干季降水占全年降水的比例,从图 8 中可以看出,蓄水前,干季降水比例从东南向西北呈递减趋势,数值从 0.5 下降至 0.15 左右,库区内干季降水比约为 0.25~0.30。在水库蓄水后,整个库区内部的干季降水比例都有所增加,且库区西南部的增加幅度最大。从降水均值变化中可以看出(图 6),库区西南部干季降水增加、库区其他地区汛期降水减少,是造成整个库区干季降水占全年降水比例提高的主要原因。

3.4.4 降水空间均匀性变化 降水的空间格局与降水类型有关,也会受到地表覆盖、景观格局、地形等因素的影响,图 9 显示了水库蓄水前后降水空间均匀性的分布,对比分析可以看出如下特征。

干季降水相比于汛期降水有更强的空间均匀性,这很好的解释了为何空间插值算法

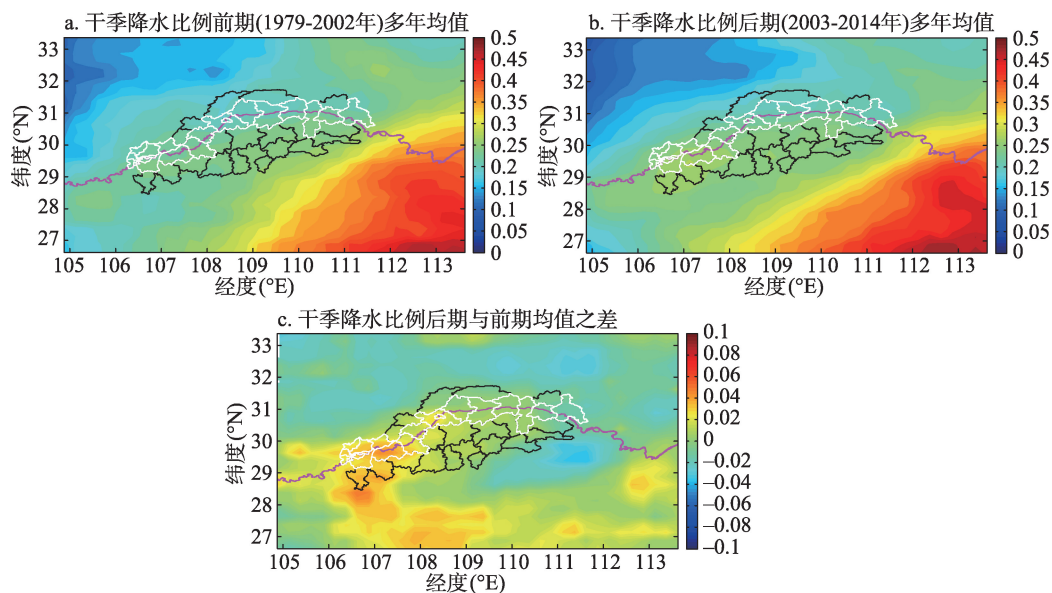


图 8 干季降水比例及其蓄水前后变化

Fig. 8 Proportion of precipitation in dry season before and after dam water level rose to 135 m

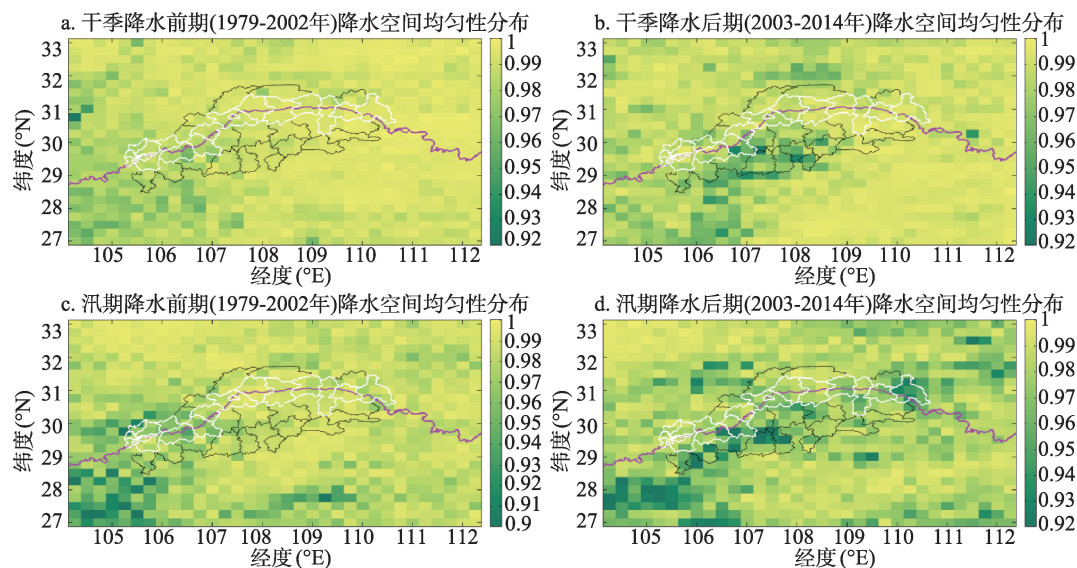


图9 干季、汛期降水空间代表性在蓄水前后的分布

Fig. 9 Spatial homogeneity of precipitation in dry season and wet seasons before and after 2003

在干季的降水估算精度更高。干季绝大部分地区的降水均匀性都非常高，只有西南地区偏低，这与该地区地表覆盖多样、景观上森林和农田交错分布有关。汛期降水的空间均匀性明显偏低，而且格局也更为复杂，和干季一样，西南地区相对其他地区而言，降水异质性更高。

在蓄水以后，干季和汛期的空间均匀性都明显降低，格局也更加复杂化，干季降水空间均匀性下降的区域主要位于中部偏西南，汛期降水均匀性下降明显的区域为西南东北方向穿过库区的带状区域，尤其是在长江以南。与库区外围相比，库区内部的均匀性下降的幅度更大。

4 讨论

4.1 降水产品精度

本文分析的数据是利用GAM算法融合了站点观测和MSWEP降水产品的结果，虽然比直接使用MSWEP或空间插值数据的精度要高，但误差仍是难以避免的，误差基本特征是汛期高于干季、山区高于平原。融合降水数据在汛期主要依赖MSWEP的信息，在干季主要依赖站点降水的观测，这种灵活性是融合降水精度更高的主要原因，同时也体现了MSWEP和站点降水是彼此无法替代的，虽然MSWEP融合了站点、遥感、再分析资料，但当局部地区有比较密集的站点时，充分利用站点观测信息是非常有必要的，而且利用的方式不能仅仅局限于用站点观测去标定MSWEP数据，本文在开始也尝试了这样的标定方法，发现其精度明显低于GAM融合算法。

4.2 库区降水变化特征

大坝蓄水后，汛期的降水空间格局发生显著变化，原本较为连续的东南-西北方向的雨带在三峡库区出现了中断，库区长江以南地区降水明显下降，库区外西北部降水明显增加。对于干季而言，降水空间格局变化不大，仍然保持东南向西北逐渐递减的趋势，

但库区内部西南地区干季降水增加明显。受汛期降水减少或干季降水增加的影响, 库区干季降水占全年降水的比例增加, 即降水的季节分配更为均衡。然而大坝蓄水后, 库区内部降水的空间异质性明显增加, 这可能与地表破碎度增加有关, 也可能与不同地表覆盖对气候变化的响应特征不一致有关, 空间异质性高表明降水的年际变化特征更具局地性, 单个格点对周围降水的代表性下降, 不同区域的降水会朝着更加不同的方向变化, 进一步造成深远的影响。

4.3 三峡水库增加大巴山地区降水?

三峡水库的气候效应一直受到科学界和公众的持续关注, 最初多数研究都认为三峡水库的气候影响只局限在江面附近, 2006年Wu等^[8]首先提出三峡水库的气候影响可以发生在百公里尺度上: 三峡西北部地区(秦岭大巴山地区)的降水在2003年以后明显增加, 而Xiao等^[12]则认为这种增加不超过自然变化范畴, 在水库修建以前, 也发生过这种类似现象。按照Wu等^[8]和Xiao等^[12]研究中的对比方法, 本文计算了三峡北部大巴山地区(中心经纬度范围107.125°E~110.875°E, 33.375°N~30.875°N)区域月平均降水和背景区域(107.125°E~110.875°E, 33.375°N~27.875°N)月平均降水之间差值在蓄水前后的变化, 从图10a中可以看出, 蓄水后, 三峡北部7-9月平均降水与背景值的差异明显增加, 而其他月份平均降水的差异在蓄水前后则无明显变化; 图10b中显示了三峡北部和背景区域7-9月平均降水差异值的年际变化, 可以看出, Wu等研究中对比的两个时段: 1998-2002年和2004-2006年, 其降水差异分别位于历史低位和历史高位, 前后形成强烈反差(降水差异由-11 mm变为24 mm); Xiao研究中对比的两个时段: 1979-1980年和1981-1984年也具有较大反差(由于本文的序列只有1979-2014年, Xiao等研究中提到的1978-1980年只好缩短为1979-1980年), 尤其是1981-1984年, 北部区域比背景区域7-9月平均降水高出的幅度为26 mm。然而这两个研究所用的数据时段都比较短, 虽然差异变化大, 但无法通过显著性检验。

本文以2003年为界, 得出降水差异在蓄水前24年的平均值为-1 mm, 蓄水后12年的平均值为16 mm, 说明蓄水前北部大巴山地区7-9月平均降水非常接近背景值, 蓄水后则是比背景值高出16 mm, 且这一变化通过了显著性检验($P < 5\%$)。如果统计降水差异值最大的5个年份, 发现除1981年以外, 其他年份(2011年, 2005年, 2003年, 2009年)都是在蓄水以后, 说明北部降水高于背景值的情况在近几年发生的比例更高, 笔者认为若要解释其变化原因, 还需要有更多深入研究, 不能简单归结为三峡工程。受数据序列长度限制, 当前蓄水后降水数据只有12年, 随着数据积累时间更长, 可以再检验这种变化是否显著。

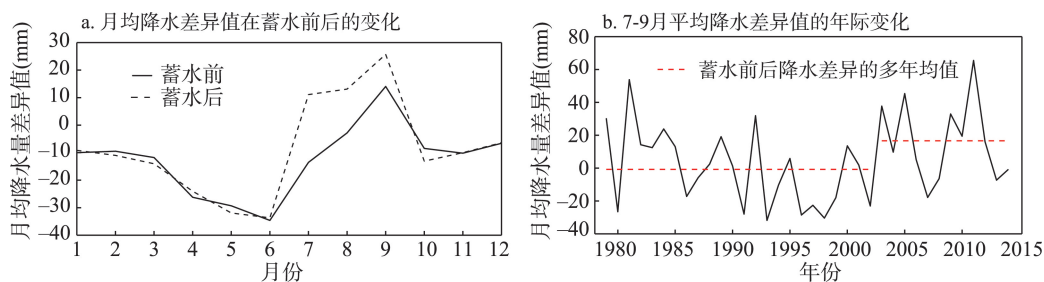


图10 北部区域和背景区域降水比较

Fig. 10 Precipitation contrast of northern region compared against background

5 结论

本文通过融合最新的全球降水产品 (MSWEP) 和站点降水观测数据, 获取了三峡库区及附近地区降水时空数据集, 主要研究结论如下:

虽然 MSWEP 数据在生产过程中综合了 3 方面的降水信息 (站点观测、遥感产品、数值预报), 但其质量仍存在缺陷, 站点插值方法与 MSWEP 产品具有互补性。通过 GAM 模型融合站点降水和 MSWEP 数据, 可以充分发挥两种数据的优势, 提高降水估算精度, 具体而言, 夏秋季降水量大、空间格局复杂, GAM 模型主要依赖于 MSWEP 数据中丰富的空间细节, 冬春季降水量少, 空间均匀性高, GAM 主要依赖于站点空间插值方法的低偏差的估计结果, GAM 融合降水精度明显高于融合前, 但仍然存在明显的季节变化: 冬春季均方根误差 $\text{rmse} = 8\sim 20 \text{ mm/月}$, 夏秋季均方根误差 $\text{rmse} = 20\sim 50 \text{ mm/月}$ 。

与其他关于三峡区域降水变化的研究相比, 本文在保持高空间分辨率的同时, 大大延长了时间序列, 使得降水气候平均值的计算更加准确, 有助于提高变化检测的可靠性。研究结果表明, 2003 年水库蓄水后, 库区内部长江以南地区汛期降水下降, 库区外西北部地区汛期降水上升, 库区内西部地区干季降水上升, 降水空间格局更加复杂化 (空间均匀度下降), 季节分配更具均衡化 (干季降水占全年降水比例上升)。未来如何应对这些变化应值得深入研究。

本文没有涉及到机理的讨论, 在 2003 年前后, 除了水域面积的扩大, 还有其他因素也发生变化, 如各种气候因子、森林覆盖率、植被生物量、景观格局、土地覆被/利用、地表粗糙度等, 如果想把水库的影响分离出来, 还需要整合多方面信息、利用气候模式模拟, 开展深入研究。

参考文献(References)

- [1] Dalya C, Neilson R P, Phillips D L. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meteorology*, 1994, 33: 140-158.
- [2] Herold N, Alexander L V, Donat M G, et al. How much does it rain over land? *Geophysical Research Letter*, 2015, 43: 341-348.
- [3] Prein A F, Gobiet A. Impacts of uncertainties in European gridded precipitation observations on regional climate analysis. *International Journal of Climatology*, 2016. doi: 10.1002/joc.4706.
- [4] Beck H E, van Dijk A I J M, Levizzani V, et al. MSWEP: 3-hourly 0.25° global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data. *Hydrol. Earth Syst. Sci. Discuss.*, 2016. doi:10.5194/hess-2016-236.
- [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, *Journal of Hydrometeorology*, 2007, 8: 38-55.
- [6] Funk C, Verdin A, Michaelsen J, et al. A global satellite-assisted precipitation climatology. *Earth System Science Data*, 2015, 7: 275-287.
- [7] Kidd C, Bauer P, Turk J, et al. Inter comparison of high-resolution precipitation products over Northwest Europe. *Journal of Hydrometeorology*, 2012, 13: 67-83.
- [8] Wu J, Gao X J, Giorgi F, et al. Climate effects of the Three Gorges Reservoir as simulated by a high resolution double nested regional climate model. *Quaternary International*, 2012, 282: 27-36.
- [9] Miller N, Jin J M, Tsang C F. Local climate sensitivity of the Three Gorges Dam. *Geophysical Research Letter*, 2005, 32, L16704.
- [10] Wu L G, Zhang Q, Jiang Z H. The Three Gorges Dam affects regional precipitation. *Geophysical Research Letter*, 2006, 33: L13806. doi: 10.1029/2006GL026780.
- [11] Li Bo, Tang Shihao. Local Precipitation changes induced by the Three Gorges Reservoir based on TRMM observations. *Resources and Environment in the Yangtze Basin*, 2014, 23(5): 617-625. [李博, 唐世浩. 基于 TRMM 卫星资料分析三

- 峡蓄水前后的局地降水变化. 长江流域资源与环境, 2014, 23(5): 617-625.]
- [12] Xiao C, Yu R C, Fu Y F. Precipitation characteristics in the Three Gorges Dam vicinity. *International Journal of Climatology*, 2010, 30: 2021-2024.
- [13] Ma Zhanshan, Zhang Qiang, Qin Yanyan. Numerical simulation and analysis of the effect of Three Gorges Reservoir Project on the regional climate change. *Resources and Environment in the Yangtze Basin*, 2010, 19(9): 1044-1052. [马占山, 张强, 秦琰琰. 三峡水库对区域气候影响的数值模拟分析. 长江流域资源与环境, 2010, 19(9): 1044-1052.]
- [14] Zhao F, Sheperd M. Precipitation changes near Three Gorges Dam, China (Part I): A spatiotemporal validation analysis. *Journal of Hydrometeorology*, 2012, 13: 735-745.
- [15] Tian, Y D, Peters-Lidard C D. A global map of uncertainties in satellite-based precipitation measurements. *Geophysical Research Letters*, 2010, 37: L24407.
- [16] Wood S N. Thin plate regression splines. *Journal of the Royal Statistical Society Series B (Statistical Methodology)*, 2003, 65(1): 95-114.
- [17] Parmentier B, McGill B, Wilson A M, et al. An assessment of methods and remote-sensing derived covariates for regional predictions of 1 km daily maximum air temperature. *Remote Sensing*, 2014, 6: 8639-8670.
- [18] Ouara T B M J, Charron C, Marpu P R. The generalized additive model for the assessment of the direct, diffuse, and global solar irradiances using SEVIRI images, with application to the UAE. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2016, 9(4): 1553-1566.
- [19] Zhang X, Liang S, Song Z, et al. Local adaptive calibration of the satellite derived surface incident shortwave radiation product using smoothing spline. *IEEE Transactions on Geoscience and Remote Sensing*, 2016. doi: 10.1109/TGRS.2015.2475615.
- [20] Park N W, Chi K H. Quantitative assessment of landslide susceptibility using high resolution remote sensing data and a generalized additive model. *International Journal of Remote Sensing*, 2008, 29(1): 247-264.
- [21] Package 'mgcv'. <https://cran.r-project.org/web/packages/mgcv/mgcv.pdf>.
- [22] Hamed K H. Trend detection in hydrologic data: The Mann-Kendall trend test under the scaling hypothesis. *Journal of Hydrology*, 2008, 349(3/4): 350-363.
- [23] Wang Yuanyuan, Li Guicai, Zhang Yan. Regional representativeness analysis of national reference climatological stations based on MODIS/LST product. *Journal of Applied Meteorological Science*, 2011, 22(2): 214-220. [王圆圆, 李贵才, 张艳. 利用MODIS/LST产品分析基准气候站环境代表性. 应用气象学报, 2011, 22(2): 214-220.]
- [24] Hu Ting, Zhou Jiangxing, Dai Kan. Application of USCRN Station Density Strategy to China Climate Reference Network. *Journal of Applied Meteorological Science*, 2012, 23(1): 40-46. [胡婷, 周江兴, 代刊. USCRN气候基准站网布局理论在我国的应用. 应用气象学报, 2012, 23(1): 40-46.]

Precipitation estimation and analysis of the Three Gorges Dam region (1979-2014) by combining gauge measurements and MSWEP with generalized additive model

WANG Yuanyaun, GUO Zheng, LI Guicai, GUO Zhaodi

(National Satellite Meteorological Centre, Beijing 100081, China)

Abstract: Long time-series, spatially-contiguous and accurate precipitation products are one of the most important inputs for various studies, including climate change detection, hydrological modelling, drought monitoring, etc. However, due to its high spatio-temporal variability, precipitation is one of the meteorological elements that are most difficult to estimate. Recently, a new global gridded precipitation dataset that merges gauge measurements, satellite products, and reanalysis data, has been produced. Owing to its high spatial resolution, long time span, and comprehensive combination of different precipitation data sources, MSWEP data have received wide attention since its release. In this study, we first evaluated accuracy of MSWEP monthly precipitation using local gauge measurements in the Three Gorges Dam region, then produced more accurate precipitation data by combining MSWEP and gauge measurements with the GAM (Generalized Additive Model) method, and finally analysed precipitation changes before and after the dam water level rose to 135 m in June 2003. Main conclusions are drawn as follows. (1) No matter what methods are used, estimation accuracy of precipitation shows strong seasonality: more accurate in cold-dry season (spring and winter), while less accurate in hot-wet season (summer and autumn). Gauge measurement and MSWEP are complementary, with the former being more accurate in cold-dry season and the latter being more accurate in hot-wet season. (2) GAM can take advantages of both gauge measurements and MSWEP with flexibility and achieve more accurate precipitation estimation (rmse decreases by 17%-50%, and r increases by 10%-30%). There are still great seasonal variations in accuracy, with rmse being 8-20 mm in spring and winter and 20-50 mm in summer and autumn. (3) Based on the precipitation estimation results obtained in step 2, we found the following phenomena after the water level rose to 135 m in 2003: 1) dam regions south of the Yangtze River show a precipitation reduction over May-October; 2) dam regions in the western part show a precipitation increase over November-April; 3) The northwestern part outside core dam regions shows a precipitation increase over May-October, which is consistent with the results of other studies; 4) precipitation shows an increase in spatial heterogeneity, but a slight decrease in seasonality conveyed by an increase in proportion of precipitation in dry season.

Keywords: precipitation; Three Gorges Dam region; MSWEP; GAM