

doi: 10. 6046/zrzyyg. 2022270

引用格式: 李新同,史岚,陈多妍. 基于深度学习的闽浙赣 GPM 降水产品降尺度方法[J]. 自然资源遥感,2023,35(4):105 – 113. (Li X T,Shi L,Chen D Y. A deep learning – based study on downscaling of GPM products in Fujian – Zhejiang – Jiangxi area[J]. Remote Sensing for Natural Resources,2023,35(4):105 – 113.)

基于深度学习的闽浙赣 GPM 降水产品降尺度方法

李新同,史 岚,陈多妍

(南京信息工程大学地理科学学院,南京 210044)

摘要: 及时准确评估降水的空间分布对国民经济发展有着重要的意义,目前遥感降水产品大多依靠多元回归模型和物理模型来提高降水监测的精度,很少涉及深度学习模型来改善降水精度。文章改进长短时记忆网络(long short – term memory neural network,LSTM)深度学习模型,得到优化后的 LSTM 深度学习模型;引入植被、坡向、坡度等多个降水主导因子,以闽浙赣为研究区域,基于 2015—2019 年 69 个气象站点的逐日降水数据,首先对 Integrated Multi – satellite Retrievals for Global Precipitation Measurement(GPM IMERG)逐日降水产品进行降尺度,继而分别从加密站验证和个例年验证 2 个角度评估模型的可靠性。结果发现:降尺度结果与气象站降水的时空分布趋于一致,比 GPM IMERG 降水产品更能体现出闽浙赣地区的降水空间分布,GPM 降水产品对降水区域存在低估和高估的降水数据得到了校正;通过加密站验证,降尺度模型在 7 月和 10 月表现较好,相关系数不低于 0.9,4 月次之,1 月最低,相关系数是 0.7;通过个例年验证,发现 2020 年日降水降尺度结果与实测值的相关性超过了 0.8 以上,均方根误差为 5.23 mm,平均相对误差为 9.43%。可见,基于深度学习的降尺度模型无论在日尺度还是月尺度都取得了较高的精度,且在时间和空间具有一定的普适性。

关键词: 闽浙赣; 鄱阳湖流域; 降水; 降尺度; 深度学习

中图法分类号: TP 79 文献标志码: A 文章编号: 2097 – 034X(2023)04 – 0105 – 09

0 引言

在地球水循环的环节中,水循环通过降水连接陆地、大气和海洋的水分交换过程。降水对气候、水文、植被等方面有着重要的意义,降水量影响大气中的水循环过程、局地气候和全球粮食作物的收成^[1–4]。研究降水的空间分布,有利于监测洪涝灾害,做到提前预警,同时也能提前应对干旱地区的缺水情况。虽然获得降水数据依靠地面气象站点监测,但是受到地形的限制和气象仪器设备缺少的影响,有些区域甚至没有完整连续的多年降水数据,因此很难获取某些地区的实测降水数据^[5]。在海洋上,由于没有地面气象站点,所以对研究降水时空分布造成了很大的困扰,而卫星遥感则没有这些限制,它为获得全球降水数据提供了便利。

全球降水测量计划(Global Precipitation Measurement,GPM)卫星于 2014 年由美国国家航空航天局地球科学办公室和日本空间发展署发射,目的是

为星载定量降水估计建立新的标准,提供下一代新的降水产品^[6–7]。GPM 产品与上一代的卫星降水产品热带降雨测量(Tropical Rainfall Measuring Mission,TRMM)比较,观测范围更大,由原来的 50°N~50°S 扩大成 90°N~90°S,也更加适用于研究中国区域降水。

GPM 的空间分辨率(0.1°×0.1°)虽然比 TRMM 高,但在地形复杂的地区,其精度仍然不高,需要用降尺度的方法提升 GPM 降水产品的精度。目前已有国内外学者基于降水主导因子对 GPM 降水卫星产品建立降尺度模型,如 Lu 等^[8]以中国天山山脉为研究区域,基于地形因子和植被指数构建地理加权回归降尺度模型,与逐步回归降尺度模型进行对比;史岚等^[9]引入水汽数据,结合地形因子和植被因子,构建地理加权回归降尺度模型;胡实等^[10]研究将 GPM 降水产品与植被指数和高程结合,构建时滞地理加权回归模型;Brocca 等^[11]结合 GPM 早期降水产品、雨量计和欧洲中期降水中心天气预报再分析 3 种降水产品,提出 3 种不同的降水

收稿日期: 2022 – 07 – 04; 修订日期: 2022 – 11 – 21

基金项目: 江苏省研究生科研与实践创新计划项目“鄱阳湖流域卫星降水产品降尺度研究与径流模拟”(编号: KYCX22 – 1130)资助。

第一作者: 李新同(1995 –),女,硕士研究生,主要从事气象 GIS 与应用。Email: AlZnCu1995@outlook.com。

通信作者: 史 岚(1978 –),女,博士,副教授,主要从事 3S 技术与气象应用研究。Email: sl_nim@163.com。

径流模型。虽然 GPM 降水产品基于降水主导因子的研究众多,但是运用深度学习研究方法研究降水降尺度的相关研究较少。

深度学习由 Hinton 等^[12]在 2006 年提出,它由机器学习演变而来,可以一次性学习所有特征,具有更强大的学习能力。深度学习有多种算法,如前馈神经网络、卷积神经网络、循环神经网络(recurrent neural network, RNN)等。深度学习在其他领域获得了不小的成果,使得更多的学者将深度学习应用在降尺度方面。如 Afshin 等^[13]选取了卡隆盆地的降水量、南方涛动指数、海平面气压和海表面温度等数据,将人工神经网络、模糊逻辑和小波函数结合,海平面气压和海表温度达到了最佳的结果。国内外学者不仅将深度学习方法应用在海温降尺度方面,而且更偏向于利用深度学习方法预报降尺度的研究。Bonnet 等^[14]运用视觉预测深度学习算法预测巴西圣保罗的降水量,指出视觉预测深度学习算法作为辅助临近预报具有很大的潜力;王慧媛^[15]基于雷达回波数据和气象站数据,运用多时间尺度支持向量机估测华东地区短时动态降水;周康辉等^[16]结合多源观测数据和高分辨率数值天气预报,运用深度学习算法,可视化强对流天气并解释预测过程,发现该算法得出的结果比单一数据的预报更加可靠;郭瀚阳等^[17]利用卷积循环网络模型预报高分辨率的强对流天气;徐海龙等^[18]基于长短时记忆网络(long short-term memory neural network, LSTM)算法,结合最小二乘法,提高了日长变化预报的精确性;袁建刚等^[19]根据递归神经网络算法建立全球电离层模型,成功模拟出电离层的物理变化;Xiang 等^[20]提出了新的基于参考和梯度引导的降尺度深度学习模型。深度学习的算法多应用于图像识别、无人机和预测等领域^[21-24],在降水降尺度研究领域应用较少。

目前,降水降尺度研究大多基于年尺度、季尺度和月尺度,很少研究日尺度,同样也缺乏构建深度学习降尺度模型。本文选择闽浙赣作为研究区,基于 2015—2019 年日降水数据,对 GPM 卫星遥感降水产品运用深度学习方法构建降尺度模型,获得研究区高精度的降水数据,并对模型结果进行评估。

1 研究区概况及数据源

1.1 研究区概况

闽浙赣地区(如图 1)位于中国的东南部,处于中国地形第三级阶梯,东濒东海,西临湖南和湖北,南近广东,北与江苏和安徽接壤。该区域经纬度范

围为 $E113.5^{\circ} \sim 122^{\circ}$, $N23.5^{\circ} \sim 31^{\circ}$ 。研究区包含福建、浙江和江西三省行政区域,有杭州市、南昌市、福州市等 31 个地级城市,三省总面积大约是 39 万 km^2 。海拔由西向东升高,地势起伏较大,地形多为高山、山地、丘陵,同时也有平原。该研究区域的气候类型属于亚热带季风气候和亚热带季风性湿润气候,具有雨热同期的特点,夏季高温多雨,盛行东南季风,冬季低温少雨,盛行西北风。研究区域的年降水量可以达到 1 000 ~ 2 000 mm 左右。从春末到夏季都是多雨时节,容易发生滑坡、泥石流等地质灾害和洪涝灾害,同时也容易出现风雹、台风等气象灾害。因此,研究闽浙赣地区降水的时间和空间分布具有十分重大的意义,可以为监测、预测自然灾害提供技术支撑。

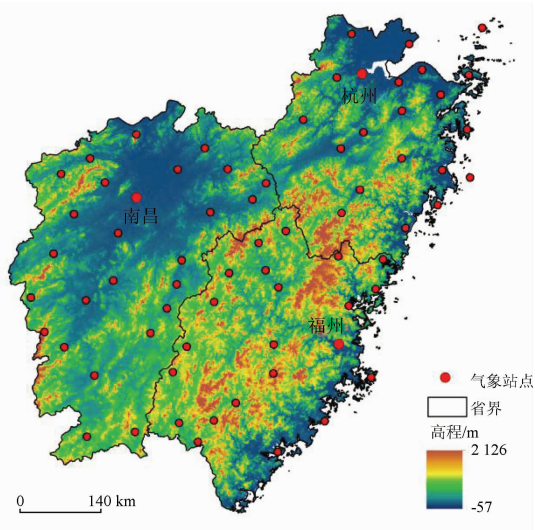


图 1 闽浙赣地区的地形和气象站点分布图

(该图基于自然资源部标准地图服务网站下载的审图号为 GS(2016)2923 号的标准地图制作,底图无修改。下文同。)

Fig. 1 Distribution map of topographic and meteorological stations in Fujian-Zhejiang-Jiangxi regions

1.2 数据源及其预处理

1.2.1 GPM 数据

本文采用的 GPM 数据来自美国国家航空航天局(<http://disc.gsfc.nasa.gov/datasets?keywords=i-merg&page=1>)。该降水产品的数据范围是 $180^{\circ}W \sim 180^{\circ}E$, 产品格式是 nc4。研究时间为 2015 年 1 月—2020 年 12 月,时间分辨率为 1 d,空间分辨率为 $0.1^{\circ} \times 0.1^{\circ}$ 。在 ArcGIS 软件平台上对 GPM 降水数据进行预处理操作。

1.2.2 气象数据

气象站点数据来自于中国气象科学共享服务网(<http://cdc.nmic.cn/>)的中国地面日值数据集(V3.0),时间为 2015 年 1 月—2020 年 12 月。数据文件有站点编号、高程、逐日降水量和站点的经纬度

等信息。常规气象站点一共有 78 个,鄱阳湖流域加密站 84 个,经质量控制之后得到了 153 个站点,根据气象观测数据与 GPM IMERG 降水产品的观测时间不一致,将观测数据的北京时间订正为世界时(UTC)时间。其中将常规气象站点 2015—2019 年数据集作为训练集,2020 年降水数据集作为个例年验证数据集,鄱阳湖流域 84 个加密站作为加密站验证数据集。

1.2.3 基础地理数据

数字高程模型(digital elevation model, DEM)来自国家基础地理信息数据库(<http://www.g-scloud.cn>),分辨率为 1 km × 1 km,比例为 1:100 万,采取正轴等积割圆锥投影(Albers equal area project),在 ArcGIS 软件平台上投影,提取 DEM 的坡度、坡向等地形信息。

1.2.4 植被指数数据

归一化植被指数(normalized difference vegetation index,NDVI)数据为 MOD13A2 中国区域 NDVI 合成产品,来自于美国国家航空航天局的官方网站(<http://ladsweb.modaps.eosdis.nasa.gov/>),空间分辨率是 1 km × 1 km。NDVI 产品从 MODIS 的 TERRA 卫星上获取,产品格式是 HDF。数据前期处理包括:校正、拼接、裁剪、投影转换、单位换算等。

2 研究方法

2.1 降尺度模型构建

2.1.1 模型构建

根据前人的研究^[25-26],地理因子和植被指数均对 Integrated Multi-satellite Retrievals for GPM(GPM IMERG)降水精度有影响。本文引入高程(H)、坡向(aspect)、坡度(slope)、经度(lon)、纬度(lat)和植被指数(NDVI)作为自变量的因子,模型运行成功后的 GPM 降水数据作为降水主导因子,建立降尺度模型,公式为:

$$GPM=f(Lon, Lat, Slope, Aspect, H, NDVI)。(1)$$

将降水主导因子输入深度学习降尺度模型中,构建空间分辨率 1 km × 1 km 的 GPM 降尺度模型。

2.1.2 基于 LSTM 的降尺度模拟

RNN 与普通神经网络相比没有太多的限制因素,应用于序列数据集更优,但是在训练过程中容易出现梯度消失或者梯度爆炸的问题,尤其是长序列的时候,问题更加明显。在此基础上,Hochreiter 等^[27]对 RNN 进行了算法提升,提出了 LSTM,LSTM 克服了 RNN 的不足之处^[28]。如图 2 所示,LSTM 由

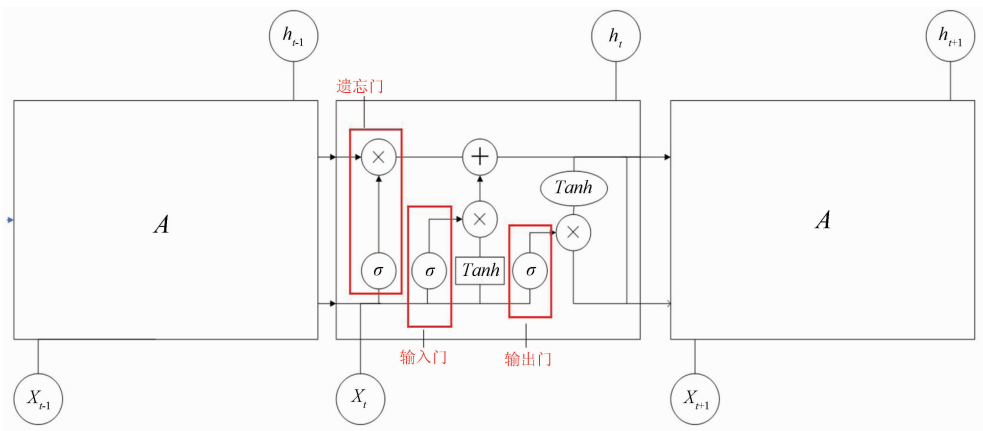


图 2 LSTM 结构图
Fig.2 LSTM structure diagram

单元状态(cell state)和门控(gates)构成,在每个序列索引位置 t 的门是由输入门、遗忘门和输出门构成,输入门控制信息流入,遗忘门控制状态更新,输出门控制信息输出。图 2 中, X_t 和 h_t 分别为 t 时刻的输入和输出, σ 和 $Tanh$ 分别为 sigmoid 和 Tanh 激活函数, $\otimes$ 表示乘法,输出结果的值域为 $[0,1]$,表示允许输出多少数据,值越小,表示输出数据越少,反之,输出数据越多, $\oplus$ 为加法, A 为其他时刻的 LSTM 结构。

一般 LSTM 可解决长时间序列的问题,随之而

来的局限性是当前时间 t 的目标值不仅与 $t-1$ 时刻的变量有关,还和当前时间 t 的变量有关,事实上,当前时间 t 的变量并不存在,因为 t 时间的变量与输出结果共同存在。在此基础上,本文考虑将第一个数据集输入 LSTM 模型中,输出的结果和下一个数据集再次输入到 LSTM 模型中,对 LSTM 进行优化,得到优化后的 LSTM (optimized LSTM, OLSTM) 算法,具体如下: ①调整算法网络,初始化权重值、迭代次数等参数; ②将 2015—2019 年的数据集分为训练集和测试集,输入第一个训练集主导因子,输出

的结果和下一个训练集主导因子输入模型中,得到 OLSTM 模型;③将数据的测试集输入降尺度模型中进行测试并计算精度指标,以检验降尺度模型。

2.2 误差验证

本文从加密站和个例年 2 方面进行验证,以说明模型在空间和时间上的普适性。采用相关系数 R 、平均相对误差 (mean relative error, MRE) 以及均方根误差 (root mean square error, RMSE) 对估算降水量与实测降水量比较,分析两者之间的误差,评估降尺度模型结果精度。采用标准差 σ 、相关系数 R 和 RMSE,3 个指标组成一个极坐标图,评估降尺度模型在极坐标图中的监测精度。各指标计算公式分别为:

$$R = \frac{\sum_{i=1}^n (\alpha_i - \bar{\alpha})(\beta_i - \bar{\beta})}{\sqrt{\sum_{i=1}^n (\alpha_i - \bar{\alpha})^2(\beta_i - \bar{\beta})^2}}, \quad (2)$$

$$MRE = \frac{1}{n} \sum_{i=1}^n \frac{|\alpha_i - \beta_i|}{\alpha_i} \times 100\% \quad , \quad (3)$$

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

$$\sigma_{\alpha} = \sqrt{\frac{\sum_{i=1}^n (\alpha_i - \bar{\alpha})^2}{n}} \quad , \quad (5)$$

$$\sigma_{\beta} = \sqrt{\frac{\sum_{i=1}^n (\beta_i - \bar{\beta})^2}{n}} \quad , \quad (6)$$

式中: n 为气象站点的数量; α_i 为第 i 个气象站点位置的实测降水数据; β_i 为第 i 个气象站点位置的 GPM IMERG 估测降水数据; $\bar{\alpha}$ 和 $\bar{\beta}$ 分别为实测降水数据的平均值和 GPM IMERG 估测降水数据的平均值。

3 分析与讨论

3.1 模型调整参数

不断调整隐含层 L 、神经元 N 和迭代次数 $epochs$,得到降水数据的训练集和测试集的决定系数 R^2 、RMSE 和 MRE,根据降水数据集的长度,神经元数量 N 为 300,隐含层 L 和迭代次数的调整结果如表 1 所示。从表 1 可知,当 $L = 5, N = 300, epochs = 250$ 时,训练集的决定系数 R^2 为 0.89, RMSE 为 0.39 mm,此时决定系数最高, RMSE 最低,拟合出的降水精度理论上是最优的,但是此时的测试集 R^2 和

表 1 模型调整参数表
Tab.1 Model adjustment parameter table

L	N	$epochs$	训练集			测试集		
			R^2	RMSE/ mm	MRE/%	R^2	RMSE/ mm	MRE/%
4	300	150	0.88	1.17	14.76	0.94	1.34	14.31
4	300	200	0.89	0.76	71.37	0.87	0.88	62.82
4	300	250	0.87	0.56	89.25	0.66	0.60	83.96
5	300	150	0.87	1.38	18.67	0.88	1.96	16.52
5	300	200	0.89	0.89	30.28	0.90	0.93	34.54
5	300	250	0.89	0.39	33.36	0.75	0.11	32.63
...	...	...	...	...	...	...	...	...

MRE 不是最优解,过分追求评估指数最优,导致结果过于拟合,模型将会表现出极差的泛化能力,甚至得到的降水精度不如预期的理想,这是因为迭代次数过多,拟合了训练集数据的噪声和训练样例中没有代表性的特征。为了解决过拟合的问题,本文逐渐减少隐含层的层数和迭代次数,直到隐含层为 4,迭代次数为 150 时,损失值 (loss) 慢慢收敛稳定,故选取模型的参数为: $L = 4, N = 300, epochs = 150$,将参数和降水数据集输入模型中,得到 2015—2019 年降水数据集。

3.2 降尺度的结果分析与验证

图 3—6 分别为 2019 年四季中某一天的气象站降水、遥感卫星 GPM IMERG 和降尺度结果的对比图,可以看出降尺度结果降水的分布图比气象站降水分布图和遥感卫星 GPM IMERG 降水分布图具有更高的空间分辨率,还精细地刻画了研究区降水的空间分布差异。从整体上来看, GPM IMERG 降水产品对研究区进行了严重的低估,降尺度结果则与气象站降水量比较接近,但是在 7 月 12 日, GPM 降水产品对研究区局地进行了高估,比如江西宜春。降尺度结果的空间分布情况与气象站降水数据的空间格局也更为一致,但是空间差异性较气象站降水数据更为精细,在降水的最值区域三者有差异:在 1 月 17 日 (图 3),气象站降水最大值在浙江玉环,降水量是 42.1 mm, GPM 和降尺度结果的最大降水量分别在浙江上虞和平湖,数值较低,分别是 11.2 mm 和 25.1 mm;在 4 月 7 日 (图 4),浙江云和的气象站降水量最大,为 52.8 mm, GPM 和降尺度结果的最大降水量都位于福建福鼎;在 7 月 12 日 (图 5),福建福安的气象站最大降水量达到了 75.7 mm, GPM 和降尺度结果的最大降水量都在福建永定;在 10 月 16 日 (图 6),江西靖安的气象站降水量最大,为 14.5 mm, GPM 降水量在江西遂川最大,达 6 mm,降尺度结果的最大降水量位于江西靖安,为 13.5 mm,与气象站最大降水量所在位置和数值都更为一致。总体而言气象站降水的最值区域与 GPM 的最值降水

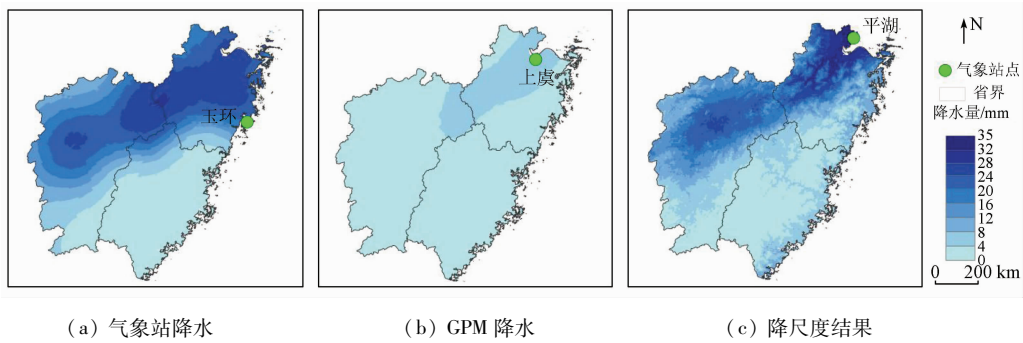


图 3 2019 年 1 月 17 日降水空间分布图

Fig. 3 Spatial distribution of precipitation on January 17, 2019

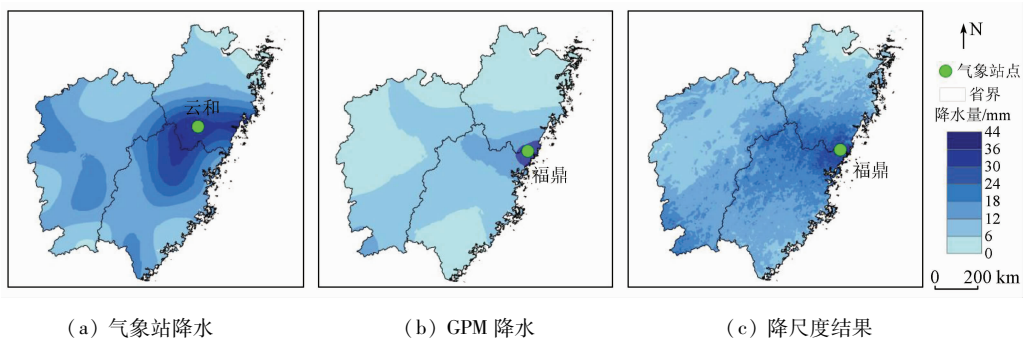


图 4 2019 年 4 月 7 日降水空间分布图

Fig. 4 Spatial distribution of precipitation on April 7, 2019

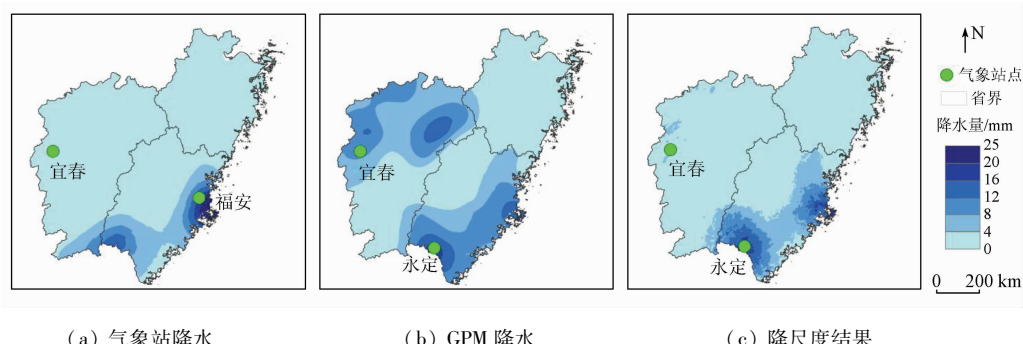


图 5 2019 年 7 月 12 日降水空间分布图

Fig. 5 Spatial distribution of precipitation on July 12, 2019

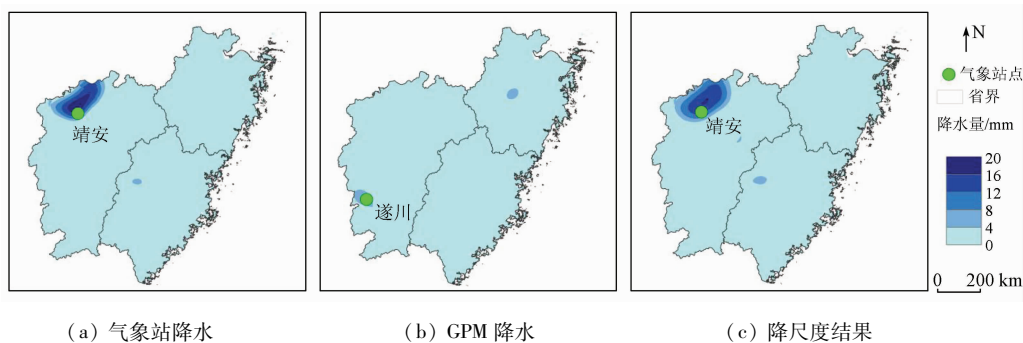


图 6 2019 年 10 月 16 日降水空间分布图

Fig. 6 Spatial distribution of precipitation on October 16, 2019

区域有区别,与降尺度结果的最值区域更为接近。

从降水时间分布来看,1 月 17 日,图 3(a) 和图 3(c) 可知,闽浙赣地区降水分布由西北向东南方向递减,降水集中在浙江北部,浙江东北沿海地区降水量最多,达到了大雨级别,而降尺度结果的级别仍是中雨,在福建地区降水最少; 4 月 7 日,从图 4(a) 和

图 4(c)可以看出,气象站降水集中在浙江和福建交界处、江西中部和西部,降尺度结果在浙江和福建交界处降水量较多,在其他地方是少量降水;7 月 12 日,由图 5(a)可以见到,研究区大部分地区是少雨,除了福建地区沿海和南部地区仍有降水,降尺度结果的空间分布格局与气象站一致;10 月 16 日,图 6(a)和图 6(c)可知,气象站降水集中在江西西北部,其他地方仍有少量降水,而降尺度结果也集中在江西西北部。运用深度学习算法订正降水数据的方法提高了原始 GPM IMERG 数据的精度,该方法在一定程度上是可靠的。本文降尺度结果虽然能够较好地反映出闽浙赣降水的时间和空间分布情况,但是在一些细节上仍有偏差。主要在 4 月 7 日这一天,降水集中在浙江南部和福建北部,处于暴雨强度,而降尺度结果并没有达到暴雨强度,即降尺度结果在 4 月 7 日订正效果并不好,未来将引入水汽含量、风向等其他降水因子,提升深度学习降尺度模型的效果。

3.3 模型验证

3.3.1 典型区域加密站验证

利用鄱阳湖流域 84 个加密气象站点 2015—2019 年平均月降水量对 OLSTM 降尺度模型进行验证。为更好地验证 OLSTM 深度学习的降尺度效果,本文将降尺度结果绘制成平均月降水的泰勒图。从图 7 可知,平均月降水量的相关系数 R 都超过了 0.7,其中,10 月的降尺度结果最优,比其他月份表现更佳, R 达到了 0.91, $RMSE$ 为 1.19 mm,标准差是 2.46,最接近标准差之比 2.5;7 月和 4 月次之,7 月的 R 是 0.9, $RMSE$ 是 3.01 mm,标准差是 5.23,4 月的相关系数是 0.8, $RMSE$ 是 1.66 mm,标准差是 1.66。但是 1 月误差偏差较大, R 是 0.7, $RMSE$ 是 1.9 mm,标准差是 0.78。这是因为 1 月降水量少,

降尺度模型存在高估的现象。

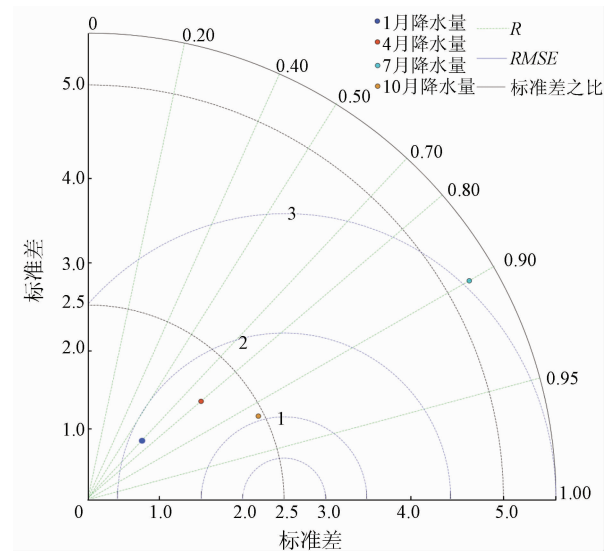


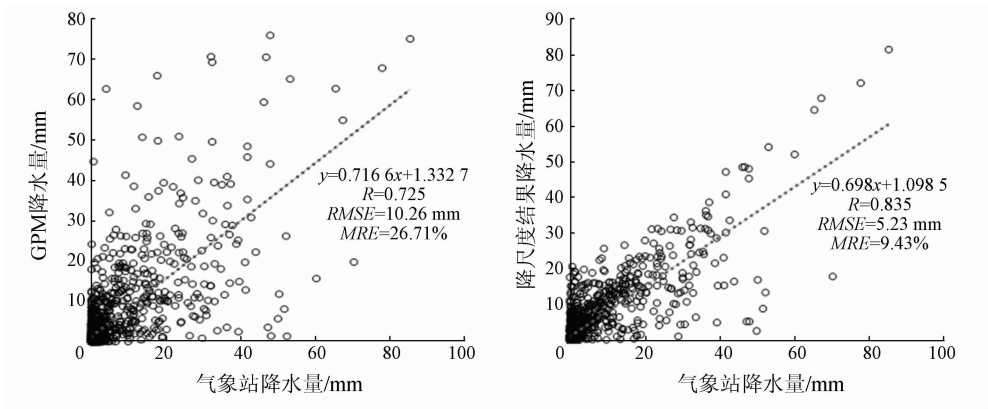
图 7 平均月降水的泰勒图

Fig. 7 Taylor chart of average monthly precipitation

3.3.2 个例年验证

为了验证深度学习降尺度模型在时间尺度的模拟水平,本文选取未参与模型训练的 2020 年降水数据,对研究区的空间降水进行降尺度,将降尺度后的降水数据与气象站点数据分别从日尺度和月尺度 2 个时间尺度进行精度评估,计算降尺度前后的精度指标 R , $RMSE$ 和 MRE 以评估模型的实用性。

1) 日尺度降水精度评估。为对比分析日尺度下 GPM 降水数据与降尺度结果的降水数据的精度,图 8 是气象站点降水数据分别与 GPM 降水数据、降尺度结果据的对比图。从图 8 可知,降尺度结果在 R 指标上相比较 GPM 有提升,提升了 0.115, MRE 和 $RMSE$ 均有降低, MRE 从 26.71% 降至 9.43%, $RMSE$ 从 10.26 mm 降到了 5.23 mm,通过深度学习降尺度模型,降尺度后的降水数据误差更小,相关系



(a) 站点降水量与 GPM 降水量散点图 (b) 站点降水量与降尺度结果散点图

图 8 日尺度下站点降水量分别与 GPM 降水量与降尺度结果的散点图

Fig. 8 Scatter plot of station precipitation, GPM precipitation and downscaling result precipitation at daily scale

数更大,说明降尺度结果的精度比 GPM IMERG 更高。

2) 月尺度降水精度评估。为对比分析月尺度下 GPM 降水数据与降尺度结果的精度,表 2 是 2020 年各月的 2 种降水数据的误差对比。从表 2 可以看出,降尺度结果的质量比降尺度前更优一点,每月的相关系数 R 都提升了, R 的平均值达到了 0.882,10 月的 R 甚至达到了 0.893,但 4 月的 R 比较低,只有 0.879,2 个月相差了 0.14,体现出降尺度结果与实测值有较好的一致性;4 个月的 MRE 在 5.71%~7.87% 之间,大大降低了 GPM 降水数据与实测数据的误差;每月的 $RMSE$ 在 0.324~4.123 之间,降低了 0.359~3.269,提升了 GPM 降水产品的精度。其中,降尺度结果最好的是 7 月和 10 月,1 月和 4 月降尺度结果稍微较差,这与以前学者研究的结果一致^[29]。GPM 由于自身的算法的限制,对微量降水和固态降水敏感,有可能出现少雨强高估的现象^[30],也有台风登陆,带来强降水,造成误差的原因。总的来说,利用深度学习降尺度模型对 GPM 降水产品进行降尺度具有一定的效果。

表 2 GPM 降水数据与降尺度结果精度指标
Tab.2 GPM precipitation data and precision
index of downscaling result

月份	R		$MRE/\%$		$RMSE/\text{mm}$	
	GPM	降尺度 结果	GPM	降尺度 结果	GPM	降尺度 结果
1 月	0.784	0.868	12.01	5.71	0.683	0.324
4 月	0.770	0.879	11.16	7.76	0.885	0.268
7 月	0.785	0.888	15.15	7.59	7.392	4.123
10 月	0.789	0.893	15.61	7.87	2.673	0.360
平均	0.782	0.882	13.50	7.20	2.908	1.269

4 结论

本文根据深度学习模型的特点,构建 OLSTM 深度学习降尺度模型,基于 2015—2019 年的逐日降水数据和植被数据以及高程数据,对 GPM 降水产品订正,得到研究区的降水空间分布。通过降水实测值和降尺度结果的误差验证,得出如下结论:

1) 从降水空间分布来看,降尺度结果与气象站降水数据的空间分布趋于一致,比 GPM IMERG 降水产品更能体现出闽浙赣地区的降水空间分布,气象站降水的最值区域与降尺度结果的最值区域相同;从降水的时间分布来看,降尺度结果在 1 月 17 日,7 月 12 日和 10 月 16 日与气象站降水具有一致性,1 月 17 日降水都集中在浙江北部,7 月 12 日降水全集中在福建地区沿海和南部地区,10 月 16 日

降水中心皆集中在江西西北部。

2) 加密站验证表明,OLSTM 降尺度模型在 7 月和 10 月的降水量表现较好,1 月和 4 月相对较差,即在每月的降水量的相关系数大于 0.7,说明该模型较为可靠。

3) 个例年验证表明,OLSTM 降尺度模型的空间降水估算方法效果良好,能够在一定程度上提升 GPM 产品的降水精度和空间分辨率,2020 年降水数据经过降尺度后 R 在 0.8 以上,说明模型适用于闽浙赣区域。但是 2020 年降尺度结果有季节性区别,7 月和 10 月拟合效果较好,1 月和 4 月次之。主要原因是 4 月台风登陆南边沿海城市,带来的强降水使得 GPM 产品的捕捉降水能力减弱。

综上,本文提出的优化深度学习降尺度模型无论在日尺度还是月尺度都有较高的精度,且在时间和空间具有一定的普适性,后续将引入水汽含量、风向等其他降水因子,以期进一步提升深度学习降尺度模型的效果。

参考文献 (References):

[1] Wang Q X, Wang M B, Fan X H, et al. Trends of temperature and precipitation extremes in the Loess Plateau region of China, 1961—2010[J]. Theoretical and Applied Climatology, 2017, 129(3): 949—963.

[2] Xu C, Luo Y, Xu Y. Projected changes of precipitation extremes in river basins over China[J]. Quaternary International, 2011, 244(2): 149—158.

[3] Yu S, Xia J, Yan Z, et al. Changing spring phenology dates in the Three—Rivers Headwater region of the Tibetan Plateau during 1960—2013[J]. Advances in Atmospheric Sciences, 2018, 35(1): 116—126.

[4] Huang L, Liu J, Shao Q, et al. Changing inland lakes responding to climate warming in Northeastern Tibetan Plateau[J]. Climatic Change, 2011, 109(3): 479—502.

[5] 张华龙,肖柳斯,陈生,等. 基于 GPM 卫星的广东汛期降水日变化特征与评估[J]. 热带气象学报, 2020, 36(3): 335—346.

Zhang H L, Xiao L S, Chen S, et al. Characteristics and evaluation of diurnal rainfall variation in rainy seasons in Guangdong based on GPM satellite[J]. Journal of Tropical Meteorology, 2020, 36(3): 335—346.

[6] Hou A Y, Kakar R K, Neeck S, et al. The global precipitation measurement mission[J]. Bulletin of the American Meteorological Society, 2014, 95(5): 701—722.

[7] Sodunke M A, Ojo J S, Adedayo K D, et al. Performance evaluation of metric measures for converting 30—min GPM rain data to 1—min for microwave applications in Tropical region of Nigeria: A multivariate approach[J]. Advances in Space Research, 2022, 69(8): 3117—3129.

[8] Lu X, Tang G, Wang X, et al. Correcting GPM IMERG precipitation data over the Tianshan Mountains in China[J]. Journal of Hydrology, 2019, 575: 1239—1252.

- [9] 史 岚,何其全,杨 娇,等. 闽浙赣地区 GPM IMERG 降水产品降尺度建模与比较分析[J]. 地球信息科学学报,2019,21(10):1642–1652.
- Shi L, He Q Q, Yang J, et al. Downscaling modeling of the GPM IMERG precipitation product and comparative analysis in the Fujian – Zhejiang – Jiangxi region[J]. Journal of Geo – Information Science, 2019, 21(10):1642–1652.
- [10] 胡 实,韩 建,占车生,等. 太行山区遥感卫星反演降雨产品降尺度研究[J]. 地理研究,2020,39(7):1680–1690.
- Hu S, Han J, Zhan C S, et al. Spatial downscaling of remotely sensed precipitation in Taihang Mountains[J]. Geographical Research, 2020, 39(7):1680–1690.
- [11] Brocca L, Massari C, Pellarin T, et al. River flow prediction in data scarce regions; Soil moisture integrated satellite rainfall products outperform rain gauge observations in West Africa[J]. Scientific Reports, 2020, 10(1):12517.
- [12] Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7):1527–1554.
- [13] Afshin S, Fahmi H, Alizadeh A, et al. Long term rainfall forecasting by integrated artificial neural network – fuzzy logic – wavelet model in Karoon basin[J]. Scientific Research and Essays, 2011, 6(6):1200–1208.
- [14] Bonnet S M, Evsukoff A, Morales R C A. Precipitation nowcasting with weather Radar images and deep learning in São Paulo, Brasil[J]. Atmosphere, 2020, 11(11):1157.
- [15] 王慧媛. 基于深度学习的短时定量降水预测研究[D]. 金华: 浙江师范大学, 2020.
- Wang H Y. Short – term quantitative precipitation prediction based on deep learning [D]. Jinhua: Zhejiang Normal University, 2020.
- [16] 周康辉,郑永光,韩 雷,等. 机器学习在强对流监测预报中的应用进展[J]. 气象,2021,47(3):274–289.
- Zhou K H, Zheng Y G, Han L, et al. Advances in application of machine learning to severe convective weather monitoring and forecasting[J]. Meteorological Monthly, 2021, 47(3):274–289.
- [17] 郭瀚阳,陈明轩,韩 雷,等. 基于深度学习的强对流高分辨率临近预报试验[J]. 气象学报,2019,77(4):715–727.
- Guo H Y, Chen M X, Han L, et al. High resolution nowcasting experiment of severe convections based on deep learning[J]. Acta Meteorologica Sinica, 2019, 77(4):715–727.
- [18] 徐海龙,乔书波,林家乐. 利用长短时记忆网络的日长变化预报[J]. 测绘科学技术学报,2020,37(5):474–478.
- Xu H L, Qiao S B, Lin J L. Prediction of length – of – day variations by long short – term memory network [J]. Journal of Geomatics Science and Technology, 2020, 37(5):474–478.
- [19] 袁建刚,李 旺,刘胜男. 基于深度学习构建的全球电离层 NmF2 模型[J]. 测绘科学技术学报,2020,37(1):15–20.
- Yuan J G, Li W, Liu S N. A global ionospheric NmF2 model developed by deep learning [J]. Journal of Geomatics Science and Technology, 2020, 37(1):15–20.
- [20] Xiang L, Xiang J, Guan J, et al. A novel reference – based and gradient – guided deep learning model for daily precipitation downscaling[J]. Atmosphere, 2022, 13(4):511.
- [21] 吴海平,黄世存. 基于深度学习的新增建设用地信息提取试验研究——全国土地利用遥感监测工程创新探索[J]. 国土资源遥感,2019,31(4):159–166. doi:10.6046/gtzyyg. 2019.04.21.
- Wu H P, Huang S C. Research on new construction land information extraction based on deep learning; Innovation exploration of the national project of land use monitoring via remote sensing[J]. Remote Sensing for Land and Resources, 2019, 31(4):159–166. doi:10.6046/gtzyyg. 2019.04.21.
- [22] 徐 佳,袁春琦,程圆娥,等. 基于主动深度学习的极化 SAR 图像分类[J]. 国土资源遥感,2018,30(1):72–77. doi:10.6046/gtzyyg. 2018.01.10.
- Xu J, Yuan C Q, Cheng Y E, et al. Active deep learning based polarimetric SAR image classification [J]. Remote Sensing for Land and Resources, 2018, 30(1):72–77. doi:10.6046/gtzyyg. 2018.01.10.
- [23] 王永全,李清泉,汪驰升,等. 基于系留无人机的应急测绘技术应用[J]. 国土资源遥感,2020,32(1):1–6. doi:10.6046/gtzyyg. 2020.01.01.
- Wang Y Q, Li Q Q, Wang C S, et al. Tethered UAVs – based applications in emergency surveying and mapping[J]. Remote Sensing for Land and Resources, 2020, 32(1):1–6. doi:10.6046/gtzyyg. 2020.01.01.
- [24] 蔡 祥,李 琦,罗 言,等. 面向对象结合深度学习方法的矿区地物提取[J]. 国土资源遥感,2021,33(1):63–71. doi:10.6046/gtzyyg. 2020111.
- Cai X, Li Q, Luo Y, et al. Surface features extraction of mining area image based on object – oriented and deep – learning method[J]. Remote Sensing for Land and Resources, 2021, 33(1):63–71. doi:10.6046/gtzyyg. 2020111.
- [25] 崔林丽,史 军,杨引明,等. 中国东部植被 NDVI 对气温和降水的旬响应特征[J]. 地理学报,2009,64(7):850–860.
- Cui L L, Shi J, Yang Y M, et al. Ten – day response of vegetation NDVI to the variations of temperature and precipitation in Eastern China[J]. Acta Geographica Sinica, 2009, 64(7):850–860.
- [26] 舒守娟,王 元,熊安元. 中国区域地理、地形因子对降水分布影响的估算和分析[J]. 地球物理学报,2007,50(6):1703–1712.
- Shu S J, Wang Y, Xiong A Y. Estimation and analysis for geographic and orographic influences on precipitation distribution in China [J]. Chinese Journal of Geophysics, 2007, 50(6):1703–1712.
- [27] Hochreiter S, Schmidhuber J. Long short – term memory[J]. Neural Computation, 1997, 9(8):1735–1780.
- [28] Yu Y, Si X, Hu C, et al. A review of recurrent neural networks: LSTM cells and network architectures [J]. Neural Computation, 2019, 31(7):1235–1270.
- [29] 金晓龙,邵 华,张 弛,等. GPM 卫星降水数据在天山山区的适用性分析[J]. 自然资源学报,2016,31(12):2074–2085.
- Jin X L, Shao H, Zhang C, et al. The applicability evaluation of three satellite products in Tianshan Mountains[J]. Journal of Natural Resources, 2016, 31(12):2074–2085.
- [30] Tan J, Petersen W A, Tokay A. A novel approach to identify sources of errors in IMERG for GPM ground validation[J]. Journal of Hydrometeorology, 2016, 17(9):2477–2491.

A deep learning – based study on downscaling of GPM products
in Fujian – Zhejiang – Jiangxi area

LI Xintong, SHI Lan, CHEN Duoyan

(School of Geographical Sciences, Nanjing University of Information Science & Technology, Nanjing 210044, China)

Abstract: A timely and accurate assessment of the spatial precipitation distribution holds great significance to the development of the national economy. At present, most remote sensing – based precipitation products improve their accuracy using multiple regression models and physical models rather than deep learning models. This study improved a long short – term memory neural network (LSTM) deep learning model, yielding an optimized LSTM deep learning model. With the Fujian – Zhejiang – Jiangxi area as the study area, this study conducted downscaling for an integrated multi – satellite retrievals for global precipitation measurement (IMERG) product based on the daily precipitation data of 69 meteorological stations from 2015 to 2019 by introducing multiple factors controlling precipitation such as vegetation, slope aspect, slope gradient. Finally, this study assessed the reliability of the optimized model through verifications based on high – density meteorological stations and individual years. The results show that the downscaling results are consistent with the spatio – temporal distribution of precipitation measured at meteorological stations and, thus, can better reflect the spatial distribution of precipitation in the study area than the original IMERG. Furthermore, underestimated and overestimated precipitation data of the study area from the GPM product were corrected. As indicated by the verification based on high – density meteorological stations, the downscaled model yielded correlation coefficients of 0.9 or above for July and October, which were followed by April. The correlation coefficient was the lowest of 0.7 in January. As shown by the verification based on individual year data, the correlation coefficient between the daily precipitation downscaling results and the measurement results in 2020 was above 0.8, with a root mean square error of 5.23 mm and an average relative error of 9.43%. Therefore, the deep learning – based downscaling model enjoys high accuracy on both daily and monthly scales and can be widely applied in the assessment of both spatial and temporal precipitation distributions.

Keywords: Fujian – Zhejiang – Jiangxi; Poyang Lake basin; precipitation; downscaling; deep learning

(责任编辑: 张 仙)