引文格式:张唯,吉宸佳,李文凯,等.降水融合数据特征驱动下基于GAN的遥感降水产品空间超分辨率重建[J].武汉大学学报(信息科学版),2025,50(10):2035-2047.DOI:10.13203/j.whugis20230013

Citation: ZHANG Wei, JI Chenjia, LI Wenkai, et al. Spatial Super-Resolution Reconstruction of Remote Sensing Precipitation Products Using GAN by Multi-source Fused Precipitation Data Features[J].Geomatics and Information Science of Wuhan University,2025,50(10):2035-2047.DOI:10.13203/j.whugis20230013

降水融合数据特征驱动下基于GAN的遥感降水产品空间超分辨率重建

张 唯[1] 吉宸佳[2] 李文凯[3] 孙晓娜[1] 梁天欣[1]
韩松洁[1] 卫鸿飞[1]

1 中国地质大学(武汉)地理与信息工程学院,湖北 武汉,430074
2 陕西省建筑职工大学,陕西 西安,710043
3 中铁第四勘察设计院集团有限公司,湖北 武汉,430061

摘 要:高分辨率降水资料是研究区域尺度水和能量循环的重要基础数据之一,但现有降水产品的空间分辨率尚无法满足区域精细化研究的需要。充分利用中国智能网格实况分析产品中的降水融合数据(China's Meteorological Administration (CMA) multi-source precipitation analysis,CMPA)的高空间分辨率(5 km)特征,构建基于生成对抗网络(generative adversarial network,GAN)的两倍超分辨率重建模型,分别应用于全球降水观测计划多卫星集成反演产品(integrated multi-satellite retrievals for GPM,IMERG)日降水数据和第五代欧洲中期数值预报中心大气再分析数据集(The fifth generation ECMWF atmospheric reanalysis,ERA5)日降水数据的空间降尺度,得到高分辨率的IMERG日降水产品(0.05°)和ERA5日降水产品(0.125°),并使用实测气象站点数据对降尺度后的两种数据进行精度评价。结果表明:(1)基于CMPA的超分辨率重建模型能够用于其他降水产品的空间降尺度。重建后的IMERG和ERA5产品的各项精度指标均有不同程度的提升。(2)所构建的模型能够有效保留CMPA数据的基本数据特征。重建后的IMERG产品和ERA5产品在相似性、命中率和偏差方面都更贴合CMPA数据,整体准确性更高。(3)模型重建效果受原始数据的空间分辨率、精度以及与CMPA数据的相关性等因素的影响。原始空间分辨率和精度越高,与训练数据集的相关性越好,降尺度效果也更佳。因而,IMERG产品的应用效果明显优于EAR5产品。(4)所构建的GAN模型在IMERG日降水数据上的适用性优于多重分形模型和随机森林模型,细节重建效果更好。相较于原始的IMERG数据,GAN重建后的日降水在年、季、月3个时间维度上的精度统计结果大多有所改善,模型稳定性较强。

关键词:GAN;超分辨率重建;CMPA;IMERG;ERA5;降水空间降尺度

中图分类号:P237 **文献标识码**:A **收稿日期**:2023-05-23

DOI:10.13203/j.whugis20230013 **文章编号**:1671-8860(2025)10-2035-13

Spatial Super-Resolution Reconstruction of Remote Sensing Precipitation Products Using GAN by Multi-source Fused Precipitation Data Features

ZHANG Wei[1] *JI Chenjia*[2] *LI Wenkai*[3] *SUN Xiaona*[1] *LIANG Tianxin*[1]
HAN Songjie[1] *WEI Hongfei*[1]

1 School of Geography and Information Engineering, China University of Geosciences (Wuhan), Wuhan 430074, China
2 SCEGC Architecture Labor University, Xi'an 710043, China
3 China Railway Siyuan Survey and Design Group Co. Ltd, Wuhan 430061, China

Abstract: **Objectives**: High-resolution precipitation data is one of the important basic data for studying water and energy cycles at regional scales, but the spatial resolution of existing precipitation products cannot yet meet the needs of regional refinement studies. **Methods**: We make full use of the high spatial resolution (5 km) feature of the fused precipitation data (China's Meteorological Administration (CMA) multi-source

基金项目:湖北省自然科学基金联合基金(2025AFD425);国家自然科学基金(41501584);湖北省重点研发计划(2022BCA080)。

第一作者:张唯,博士,副教授,主要从事地理信息科学相关研究。weizhang@cug.edu.cn

precipitation analysis, CMPA) from the national intelligent grid live analysis product to construct a two sets of super-resolution reconstruction model based on generative adversarial networks (GAN), which is applied to the spatial downscaling of the integrated multi-satellite retrievals for GPM (IMERG) daily precipitation data and the fifth generation ECMWF atmospheric reanalysis (ERA5) daily precipitation reanalysis data, respectively. The final high-resolution IMERG daily precipitation product (0.05°) and ERA5 daily precipitation product (0.125°) were obtained and the accuracy of both downscaled data was evaluated using measured meteorological station data. **Results:** The results show that: (1) The super-resolution reconstruction model based on CMPA data can be used for spatial downscaling of other precipitation products. The reconstructed IMERG products and ERA5 products show different degrees of improvement in all accuracy indicators. (2) The model is able to effectively retain the essential data characteristics of the CMPA data. The reconstructed IMERG product and ERA5 product are more closely matched to the CMPA data in terms of the correlation coefficient, probability of detection and bias, with higher overall accuracy. (3) The effect of model reconstruction is influenced by the spatial resolution and accuracy of the original data, as well as the correlation between the original data and the CMPA data. The higher the spatial resolution and accuracy of the raw data, the better the correlation between the raw data and the training data set, and the better the downscaling effect. As a result, the application on the IMERG product is significantly better than the EAR5 product. (4) The constructed GAN model outperforms the multifractal and random forest models in terms of applicability to IMERG daily precipitation data, and its detailed reconstruction is better. Compared with the original IMERG data, the accuracy statistics of the GAN reconstructed daily precipitation in the three time dimensions of year, season and month are mostly improved, which indicates better model stability. **Conclusions:** Therefore, the constructed GAN model is able to provide a more refined precipitation distribution and is of some research value.

Key words: GAN; super-resolution reconstruction; CMPA; IMERG; ERA5; precipitation spatial downscaling

降水是研究气候变化和水文模拟的重要气象要素之一,具有复杂的时空异质性,已经广泛应用于农业、海洋、生态、地质、大气环境等各行各业[1-2]。高精度的降水空间分布是准确理解区域水文空间模拟和气候变化过程的关键因素,对于促进高质量的全球气象分析如洪水预报、极端气候预警等具有重要意义[3]。因此,提高降水空间分布的分辨率已经成为气象、测绘、水文等领域的重点研究问题。目前,降雨量的空间估算方式主要包括地面站(如雨量计)插值、气象雷达、卫星遥感、大气再分析以及多源数据融合等[4-6]。传统的降雨量估计(地面站插值和气象雷达)虽能较为真实地记录区域降水情况,但测量方式受观测密度、间距、观测环境以及仪器设备等诸多因素限制,无法保证一些偏远山区能够获得准确的降水数据,在实际应用中面临着一些问题与挑战[7]。随着遥感卫星技术的发展,利用遥感卫星获取降水产品已成为降水观测的主要方式,其能够提供覆盖面更广、普适性更强的气象降水资料[8]。目前,已有许多卫星获取的降水资料广泛应用于气象遥感分析研究,如气候预测中心变形技术[9]、热带降雨测量任务(tropical rainfall measuring mission, TRMM)[10]、全球降水气候计划[11]、全球降水测量(global precipitation measurement, GPM)[12]。然而,单一卫星反演降雨数据的误差特征和影响因素极其复杂,难以直接应用于气象和水文实践[7]。因此,对历史观测资料、数值模式预报输出、卫星、地面等观测产品进行同化分析处理的再分析降水资料应运而生[13]。但目前的区域水文分析已经对降水产品的精度提出了更高的要求。因此,为了准确获取区域局部细节降水特征,提高降水资料的空间分辨率就显得十分必要。

国内外已有诸多学者开展降水降尺度方面的研究。降尺度的目标是提高原始降水产品的空间分辨率,并借助大范围区域特征来丰富局部降水特征。目前的降尺度方式包括两大类:动力降尺度和统计降尺度[14]。动力降尺度能够保留物理过程,但繁杂的计算量限制了其发展进程[15]。然而,统计降尺度凭借着其计算量小、方法灵活、开发相对简单等特点已经广泛应用于遥感降水数据降尺度研究[14]。文献[16-17]考虑到

降水和环境变量之间的关系随位置变化而变化，将地理加权回归模型(geographically weighted regression，GWR)应用于小尺度卫星降水数据。结果表明，GWR提高了降水量空间估计的准确性。文献[18]基于数字高程模型和归一化植被指数(normalized difference vegetation index，NDVI)分别构建地理加权回归模型和多元线性回归模型，将GPM年、月降水数据的空间分辨率提高至1 km。随着机器学习的快速发展，随机森林(random forest，RF)、支持向量机等机器学习回归模型在降尺度方面也产生了良好的效果。文献[19]利用RF回归模型分别对IMERG和TRMM年降水数据进行降尺度，分析得到RF在IMERG降水降尺度方面效果更好。近年来，有学者研究发现降尺度技术和图像超分辨率存在相似之处[20]。文献[21]构建卷积神经网络模型，将ERA5再分析降水数据的空间分辨率提高至0.033 3°。随着计算机性能的大幅提高，基于深度学习的超分辨率重建方法在高分辨率降水推算方面显示出较好的应用前景[22]。

全国智能网格实况分析产品中的降水融合产品(China's Meteorological Administration (CMA) multi-source precipitation analysis，CMPA)于2017年进行发布推广[23]，空间分辨率高达5 km，数据质量较好。本文以中国大陆地区为研究区域，借助CMPA数据构建两倍的超分辨率重建模型，对IMERG日降水数据和第五代欧洲中期数值预报中心大气再分析数据集(The fifth generation ECMWF atmospheric reanalysis，ERA5)日降水再分析资料进行降尺度，将它们的空间分辨率分别提升至0.05°和0.125°，对比分析各模型结果的优劣，提供实用性更强的降水空间降尺度方法。

1　研究区域和数据

1.1　研究区域

中国大陆地区位于亚洲东部、太平洋西侧，如图1所示。中国大陆地貌多以山地、丘陵以及高原为主，盆地和平原占地面积较少，西部地势高，东部地势低，自西向东呈三级阶梯状分布。受地形地貌复杂多样、河流湖泊散状分布、幅员广阔等因素的影响，中国大陆的气候复杂多样，从南到北跨5个温度带。由于中国西北部地区距海遥远，全年降水量较少，整体降水量呈现从东南部向西北部递减的趋势。

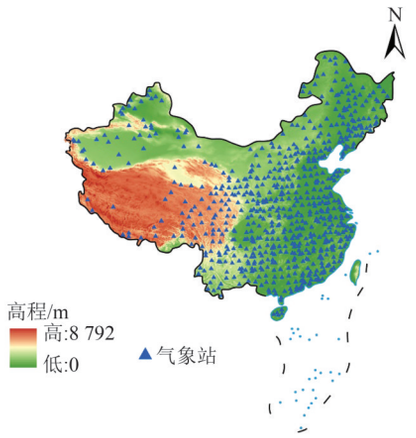


图1　研究区概况

Fig. 1　Overview of the Study Area

1.2　实验数据

1)站点数据。本文采用的气象站点观测数据来自中国气象数据网提供的全国地面气候数据日值数据集V3.0(数据集代码为SURF_CLI_CHN_MUL_DAY)。此数据集在中国大陆区域内共有699个观测站点，密集程度从东南向西北递减，分布如图1所示。本文对该数据集的两个12 h观测结果(20时至8时、8时至20时)进行合成以提供8时至8时的逐日气象站点观测数据，用于评价基于生成对抗网络(generative adversarial network，GAN)超分辨率重建后降水数据的数值统计精度。

2)全国智能网格实况分析产品。全国智能网格实况融合分析产品能够提供空间分辨率为5 km的产品，整点15 min内发布风速、降水等产品，主要用于为全国智能网格预报业务提供预报检验和滚动订正[23]。本文采用的CMPA是在原有的亚洲区域CMA陆面数据同化业务系统(CMA land data assimilation system，CLDAS)、中国区域融合降水分析系统(CMA multi-source precipitation analysis system，CMPAS)等成熟业务系统的基础上，通过多源数据融合方法制作而成的空间分辨率为5 km的降水产品。CMPA降水数据的空间覆盖范围为整个中国区域。本文选取2018—2019年的24 h CMPA降水数据作为超分辨率重建模型的高分辨率端。

3)GPM卫星降水数据。GPM作为TRMM的后续卫星降水观测计划，自2014年起可以提供时间分辨率为30 min的降水产品，其空间分辨率高达0.1°，覆盖范围增至南北极圈[24]。相较于TRMM，GPM的轨道倾角增大，能够获得覆盖面更广的降水数据，并且提高了GPM在微量降水和固态降水的观测上的敏感性，能够获取更加精

细化的降雨分布[25]。现阶段,GPM-IMERG数据已经被广泛应用于降水研究[26-27]。IMERG中包含Early Run、Late Run、Final Run 3类降水产品。本文选用Final Run降水产品,从GES DISC官方网站下载V06B日降水数据。

4)ERA5再分析数据集。ERA5数据集是欧洲中期数值预报中心(European Centre for Medium-Range Weather Forecasts, ECMWF)最新一代的再分析数据集,其空间分辨率为0.25°×0.25°,时间范围从1950年1月至今。ERA5降水再分析数据采用4VDAR系统,利用同化系统对降水数据进行同化分析处理,获得更加精准的降水再分析数据集[28]。本文采用ECMWF官网提供的日降水再分析数据集计算器,获得中国大陆的ERA5日降水数据。

5)NDVI。本文采用的NDVI数据来自LAADS DAAC官网提供的MOD13A3产品,在ArcGIS中去除异常值,乘以0.000 1后得出-1~1之间的NDVI值,并分别重采样为0.1°和0.05°,将其作为RF模型对IMERG日降水数据降尺度时的影响因子。

2 研究方法

2.1 图像超分辨率重建

2.1.1 GAN

图像超分辨率重建是计算机图像视觉领域的一个经典问题,旨在利用已有的低分辨率图像重建所需的高分辨率图像[29-30]。文献[31]首次将卷积神经网络应用到超分辨率重建中,提出超分辨率卷积神经网络模型,主要分为图像放大和图像质量修复两个部分。首先对输入端的低分辨率图像进行预处理,采用双三次插值的方法放大图像尺寸,再通过3层卷积层提取特征信息,训练得到低分辨率和高分辨图像之间的非线性映射关系,完成高分辨率图像的重建工作。

随着文献[32]提出GAN的概念,基于GAN的超分辨率重建就有了广泛的研究应用。GAN模型包括生成器和判别器两个部分,如图2和图3所示。生成器G用来重建高分辨率图像,并将生成的高分辨率图像输入判别器D中,与真实的高分辨率图像进行判别比较,不断迭代优化至判别器无法辨别真伪,即完成高分辨率图像的重建[33]。生成器主要由6个组织结构相同的残差单元组成,每个残差单元都包含卷积层(两个)、激活层(两个)和残差连接层(1个)。前3个残差单元采用ReLU函数作为激活函数,后3个残差单元采用LeakyReLU函数作为激活函数。此外,生成器还包含两个卷积层、一个全连接层、一个上采样层(双线性插值)和一个激活层(Sigmoid函数)。在本文中,所有卷积层的卷积核大小为3×3,步长为1,低分辨率端输入图像尺寸为64×64像素,高分辨率端输出图像尺寸为128×128像素。

判别器主要包括12个残差结构,每个残差单元都包含两个卷积层、两个激活层(激活函数为LeakyReLU函数)和一个残差连接层。除了全连接层外,判别器还包括全局平均池化层(global average pooling, GAP)和紧密连接层(dense block,DB)各两个。卷积层的卷积核尺寸(K)、步长(s)和特征图维数(n)等参数设置如图3所示。

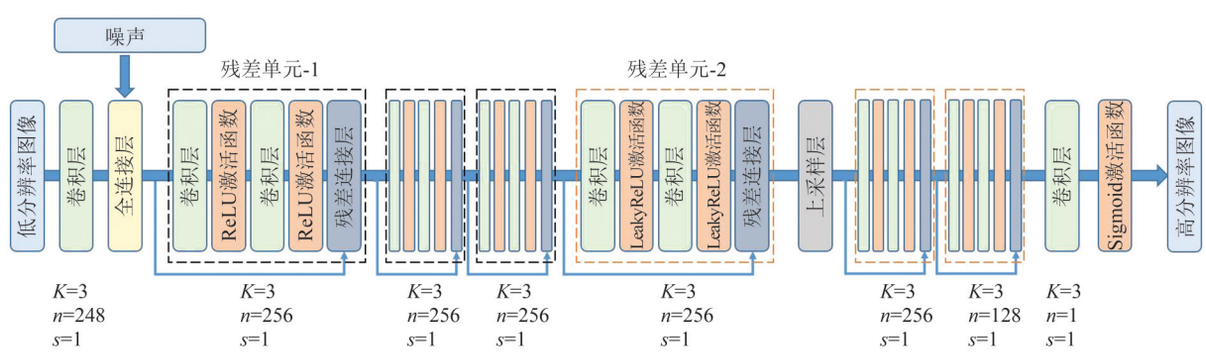


图2 生成器架构

Fig. 2 Architecture of the Generator

2.1.2 CMPA降水样本集的构建

本文基于中国大陆2018—2019年的CMPA日降水数据构建样本集,制作流程如下:

1)通过窗口遍历制作初始切片。在0.1°的低分辨率端选用64×64的窗口,步长设置为32;在0.05°的高分辨率端选用128×128的窗口,步长设置为64。低分辨率端和高分辨率端一一对应。

2)样本质量筛选。若样本的无降水数据过

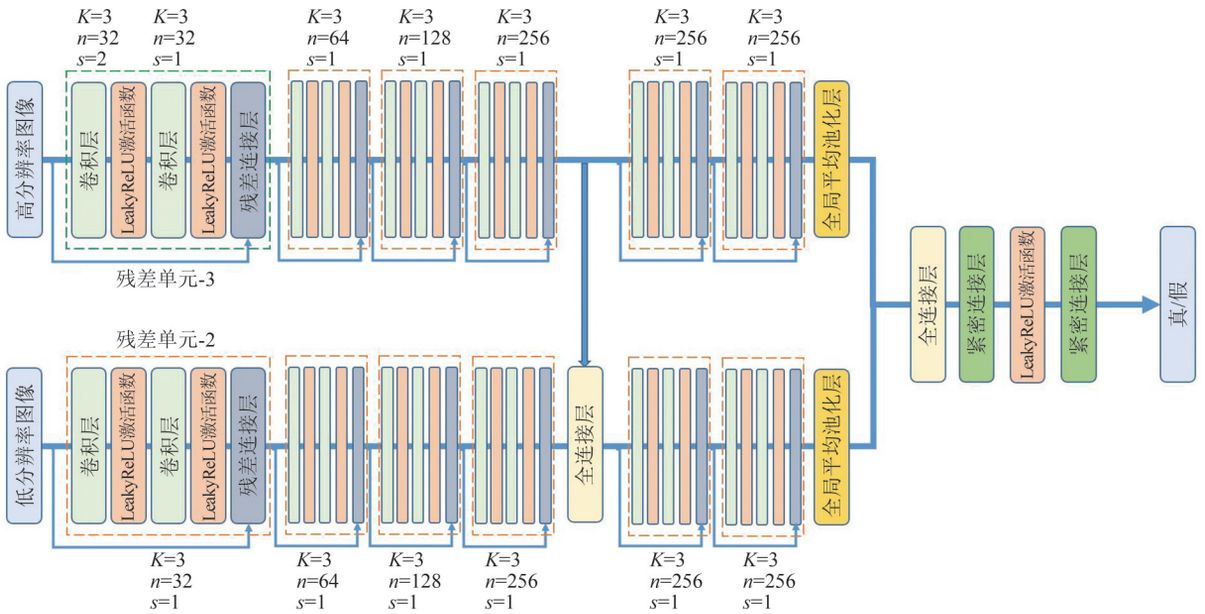


图 3　判别器架构

Fig. 3　Architecture of the Discriminator

多或存在异常值,会对网络模型的输出造成影响,因此需进行样本筛选,剔除无降水样本对以及降水异常值。当像元值小于0.1 mm/d时,即认定此像元为无降水像元;若样本对中无降水像元占比大于90%,即认定此样本对为无降水样本对,予以剔除。此外,一个样本对中可能会出现少许降水异常值,本文将像元值小于0 mm/d和大于1 000 mm/d的降水值以0值替代,完成样本对的制作。本文共选取47 814对CMPA降水样本用于训练模型,并分别制作76 495个IMERG测试切片和29个ERA5测试切片。

2.1.3　模型训练

本文采用轮流训练(生成器一个批次、判别器5个批次)的方式,每个批次的大小设置为8。训练起始阶段,本文采用自适应矩估计优化算法来更新模型参数,判别器和生成器的学习率为1×10^{-5};在训练接近结束(训练序列=36 800)时更换优化算法为随机梯度下降算法,并且将判别器和生成器的学习率设置为1×10^{-6}。训练一次需20 h。

2.2　多重分形(multifractal, MF)模型

MF分析是描述分形空间质量分布的定量工具。降雨由于其空间不连续性和空间变异质性等特点,可以表示为一个MF过程。MF模型将降水的空间变化特征描述为随机级联过程。均质化的MF降水公式为:

$$M_{(i,j)}^{l}=\begin{cases}M_{(\lceil i/2\rceil,\lceil j/2\rceil)}^{l-1}\times W_{(i,j)}, l>0\\ M_{(i,j)}^{0}, l=0\end{cases}\tag{1}$$

式中,i表示网格中的行索引,j表示网格中的列索引,并且$i,j=0,1,2,\cdots$;$M_{(i,j)}^{l}$表示在(i,j)位置第l次均质化MF的降水量;$W_{(i,j)}$表示在(i,j)位置以获得第l次均质化MF降水量的权重;$M_{(i,j)}^{0}$表示分辨率为0.1°时所代表的第一级均匀MF降水量。

随机级联权重$W_{(i,j)}$的计算公式为:

$$W_{(i,j)}=BY_{(i,j)}\tag{2}$$

$$\begin{cases}P(B=0)=1-b^{-\beta}\\ P(B=b^{\beta})=b^{-\beta}\end{cases}\tag{3}$$

$$Y_{(i,j)}=b^{-\delta^{2}\frac{\ln b}{2}+\delta X_{(i,j)}}\tag{4}$$

式中,β和δ分别是β模型和Lognormal模型的未知参数;b在本文中值为4;X是遵循标准正态分布的随机变量;B是β模型;P表示概率分布。

2.3　RF模型

RF算法是一种基于集成学习思想的机器学习算法,其基本构建单元是可用于分类和回归的决策树。RF回归是为连续随机变量构建的多元回归树,其最终预测结果是所有决策树输出的平均值。计算过程为:

$$g(x_t)=\frac{1}{S}\sum_{f=1}^{S}h_f(x_t)\tag{5}$$

式中,$g(x_t)$表示预测结果;x_t、$h_f(x_t)$和S分别表示要测试的样本数据、每棵回归树的预测结果以及树的数量。本文选取100棵树来构建随机森林回归模型。

2.4　评价指标

本文基于气象站点观测数据对IMERG日降

水数据、ERA5日降水数据及两种数据的超分辨率重建结果进行精度评价,选择相关系数E_{CC}、偏差E_{Bias}、命中率E_{POD}作为评价指标。相关系数表示栅格降水数据和站点数据之间的相关程度,越靠近1,则相关性越强;偏差表示栅格降水数据相对于站点数据的数值偏差,反映了栅格降水数据被高估或低估的程度,越靠近0,则偏差越小,大于0表示高估,小于0表示低估;命中率表示栅格降水产品正确检测到降水事件且误差小于5 mm的概率。计算公式为:

$$E_{CC}=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2\sum_{i=1}^{n}(y_i-\bar{y})^2}} \tag{6}$$

$$E_{Bias}=\frac{\sum_{i=1}^{n}(x_i-y_i)}{\sum_{i=1}^{n}y_i} \tag{7}$$

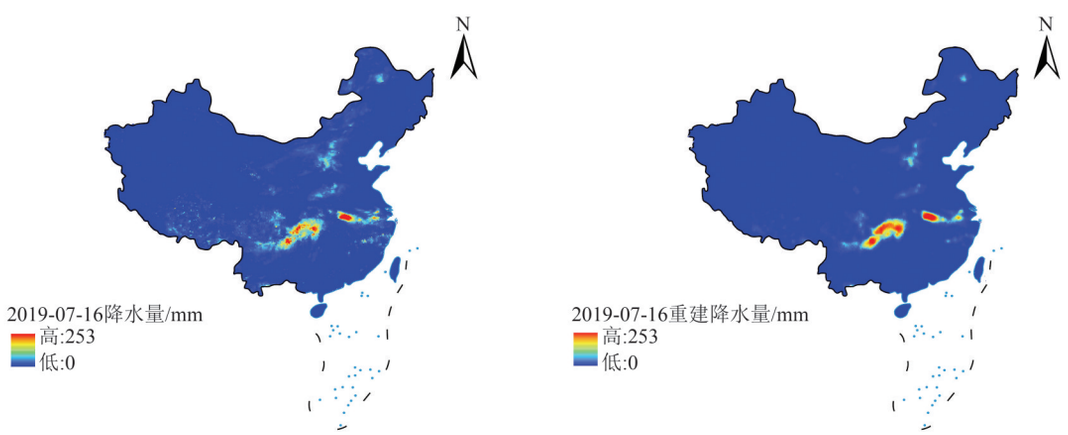
$$E_{POD}=\frac{H}{H+M} \tag{8}$$

式中,x表示栅格降水产品检测到的降水量;y表示地面站点观测到的降水量;n表示样本总数;H为栅格降水产品及地面站点均探测到有雨,并且二者之间的数值误差小于5 mm的天数;M为栅格降水产品及地面站点均探测到有雨,但二者之间的数值误差大于或等于5 mm的天数。

3 结果分析

3.1 IMERG和ERA5重建结果对比分析

本文以2019-07-16为例,对IMERG日降水数据和ERA5日降水数据进行超分辨率重建。原始的IMERG降水场和超分辨率重建后的降水场横向对比如图4所示。可以看出,重建后降水场整体的空间结构与原始降水场基本保持一致,降水高值区域也有较准确的对应,表明模型重建后基本还原了原始降水场的空间分布趋势。原始的ERA5降水场和超分辨率重建后的降水场横向对比如图5所示。可以看出,重建后降水场的空间分辨率有一定程度的提升,强降水区域的分布形式与原始ERA5降水场高度相似。此外,之前的研究显示ERA5再分析资料在横断山脉和青藏高原区域过高地估计了降水量[34],而重建后的降水场在这个两个区域出现较明显的降水高值减少现象,微量降水(蓝色区域)占比增加,这表明GAN模型可以获得更加精准的降水分布。

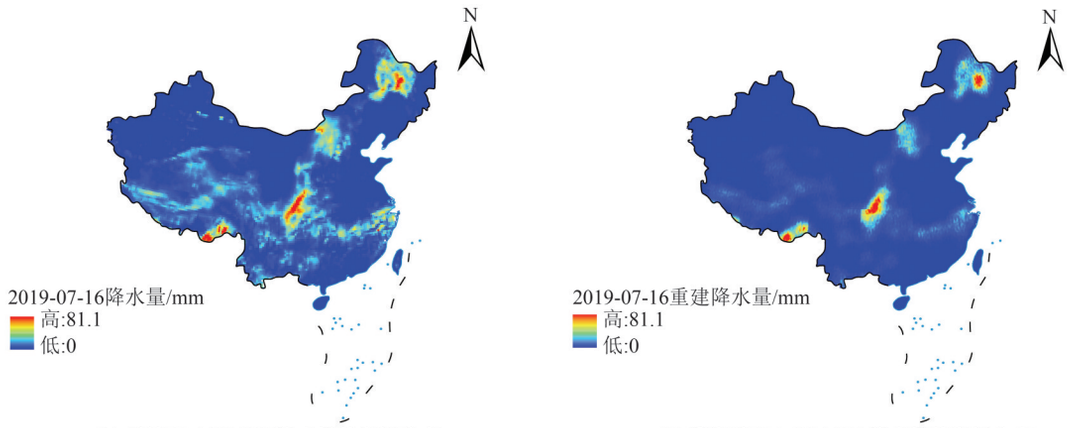


(a) 原始IMERG(0.1°)降水场的空间分布 (b) 重建后IMERG(0.05°)降水场的空间分布

图4 2019-07-16中国大陆原始IMERG(0.1°)和重建后IMERG(0.05°)降水场的空间分布图

Fig. 4 Spatial Distribution of Original IMERG (0.1°) and Downscaled IMERG (0.05°) Precipitation Fields in Chinese Mainland on 16 July 2019

本文采用中国大陆699个气象站点的观测数据,验证了CMPA数据、原始IMERG数据、原始ERA5数据和相应的GAN模型重建结果的数值精度,精度评估如表1所示。从评价结果来看,重建后IMERG降水场的数值精度大幅提升,相关系数和命中率分别增加了3.1%和9%,偏差减少了14.8%,甚至超过了CMPA数据,表明本文构建的GAN模型可以在提高空间分辨率的同时,获得数据准确性更高的IMERG日降水场;相比于原始的ERA5降水场,重建后ERA5降水场的各项精度指标均有明显的改善,其中相关系数和命中率分别增加了4%和11.2%,偏差降低了38%,表明重建后的降水场与观测站点数据之间的误差大幅减少,并且能够更加精准地检测到降水事件,在一定程度上增强了数据一致性和准确性。然而,受限于原始ERA5数据的空间分辨率



(a) 原始ERA5(0.25°)降水场的空间分布　(b) 重建后ERA5(0.125°)降水场的空间分布

图5　2019-07-16中国大陆原始ERA5(0.25°)和重建后ERA5(0.125°)降水场的空间分布图

Fig. 5　Spatial Distribution of Original ERA5 (0.25°) and Downscaled ERA5 (0.125°) Precipitation Fields in Chinese Mainland on 16 July 2019

表1　2019-07-16数值评价表

Table 1　Numerical Assessment for 16 July 2019

数据		相关系数	偏差	命中率
CMPA		0.702	0.063	0.668
IMERG	原始IMERG	0.685	0.153	0.592
	重建后IMERG	0.716	−0.005	0.682
ERA5	原始ERA5	0.480	0.464	0.531
	重建后ERA5	0.520	−0.084	0.643

和数值精度,重建后ERA5降水场的整体精度低于IMERG,且未能超过CMPA数据,其重建效果不如IMERG降水数据。

为进一步评估本文构建的GAN模型对两种降水数据的降尺度效果,以CMPA数据为基准,重新计算IMERG日降水数据和ERA5日降水数据的各项精度指标,如表2所示。结果表明,重建后IMERG降水数据和ERA5降水数据与CMPA的相关性均有升高,其中,IMERG数据的相关系数和命中率分别增加了4.2%和3.9%,偏差减少了2%;ERA5数据的相关系数和命中率分别增加了1.9%和13.5%,偏差减少了23.8%。从整体来看,相较于ERA5降水数据,重建后IMERG降水数据的数值精度更高,与CMPA数据的一致性更强,重建效果更好。实验结果也从另一方面证实,使用目前公认数据质量较好的CMPA数据进行其他降水产品校正的可行性。

考虑到训练数据对重建效果的影响,本文从这3种日降水数据中选取2019年1月至10月中国大陆699个站点所对应的数据,剔除缺测值和异常值,共保留198 383对降水数据,进一步计算CMPA数据与IMERG日降水数据和ERA5日降水数据长时间序列的相关性,结果如图6所示。

表2　2019-07-16 CMPA与IMERG和ERA5之间的数值评价表

Table 2　Numerical Assessment of CMPA with IMERG and ERA5 Respectively on 16 July 2019

数据		相关系数	偏差	命中率
IMERG-CMPA	原始IMERG	0.698	0.084	0.737
	重建后IMERG	0.740	−0.064	0.776
ERA5-CMPA	原始ERA5	0.435	0.377	0.586
	重建后ERA5	0.454	−0.139	0.721

可以看出,相较于IMERG降水数据,CMPA数据与ERA5降水数据之间的相关性仅为0.556,这也是导致重建效果不如IMERG降水数据的原因之一。综上,GAN模型重建后两种降水数据的空间分辨率和数值精度都有不同程度的提升,但在IMERG降水数据上的适用性更好,整体精度更高,分析原因是原始ERA5日降水数据的空间分辨率较低,相较于原始IMERG降水数据,ERA5降水数据精度较差,并且CMPA数据与ERA5降水数据之间的相关性不够高,这对重建后的结果产生了一定的影响。

3.2　基于GAN的IMERG降水产品的改进

3.2.1　多模型降尺度效果对比

为了进一步检验本文构建的GAN模型在IMERG日降水产品上的适用性,统计2020年中国大陆699个气象站点月平均降水量,选取降水量较大的月份(2020年7月)作为研究时段,在中旬和下旬各选取一个降水过程(2020-07-12—2020-07-15,2020-07-27—2020-07-31),对IMERG日降水数据进行GAN超分辨率重建,获得更加

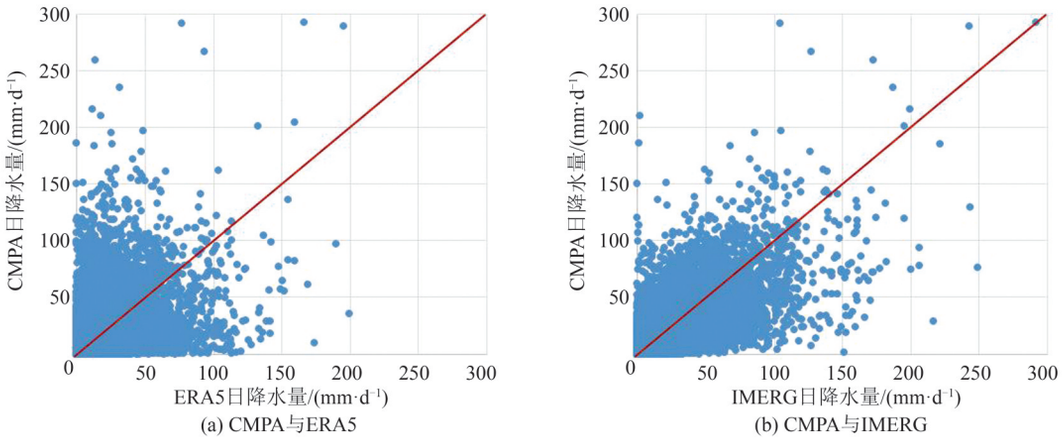


图6 CMPA与ERA5以及CMPA与IMERG日降水散点图

Fig. 6 Daily Precipitation Scatter Plots for CMPA with ERA5 and CMPA with IMERG

精确的IMERG（0.05°×0.05°）降水空间分布。同时，选择MF模型以及考虑NDVI和经纬度影响的RF模型进行降尺度效果对比，并利用站点数据进行精度评价。

以2020-07-29为例，将其原始的IMERG降水场与不同模型降尺度后的降水场横向对比，如图7所示。可以看出，3种降尺度后的降水场都基本保留了原始IMERG降水场的整体分布特征，较为准确地再现了强弱降水的空间分布。细节对比如图8所示，相较于原始的IMERG数据，GAN重建后的降水场细节特征最为丰富，降水信息的空间连续性更好；MF模型降尺度后出现了马赛克现象，细节特征不明显；RF模型降尺度后存在一定的块状效应[35]，无法反映更真实的降水细节。为深入分析GAN重建后的细节重建情况，本文从9 d的测试样本中再随机选取两对，将其原始IMERG和重建后的降水场进行横向对比，结果如图9所示，重建后的IMERG降水空间分布图像出现了一些斑状细节特征。综合来看，本文构建的GAN模型可以在提高降水空间分辨率的同时，在空间分布图像上获得更加细致的局部纹理特征。

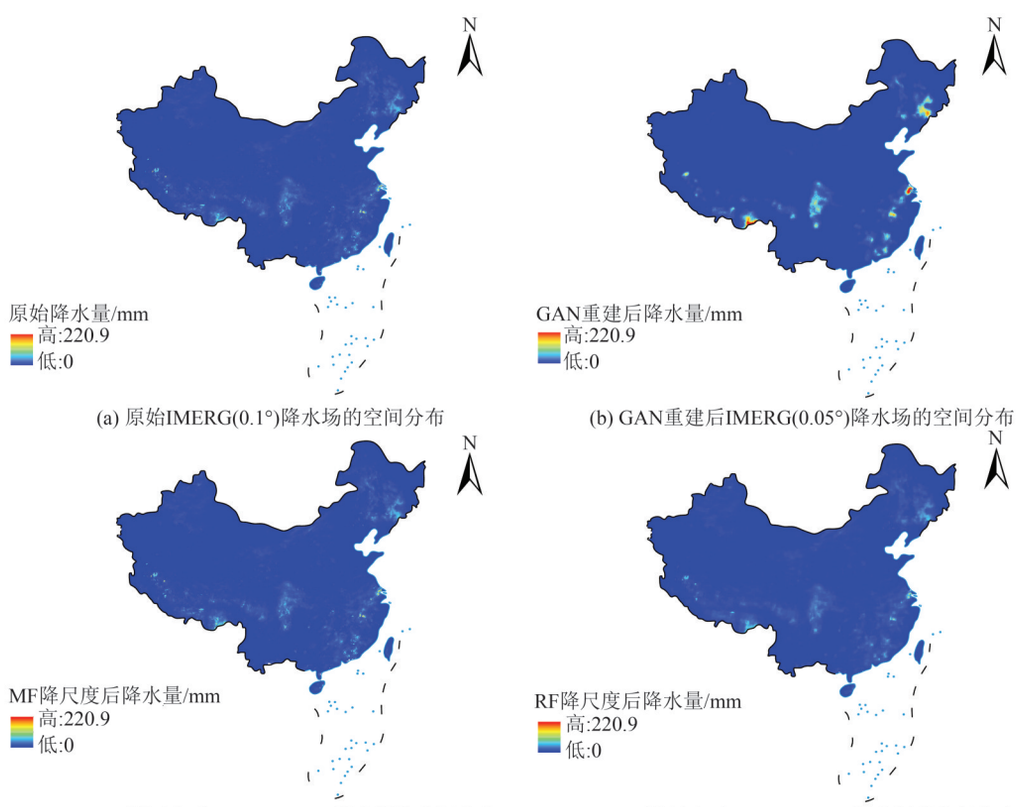


图7 2020－07－29中国大陆降水场的空间分布图

Fig. 7 Spatial Distribution of the Precipitation Field in Chinese Mainland on 29 July 2020

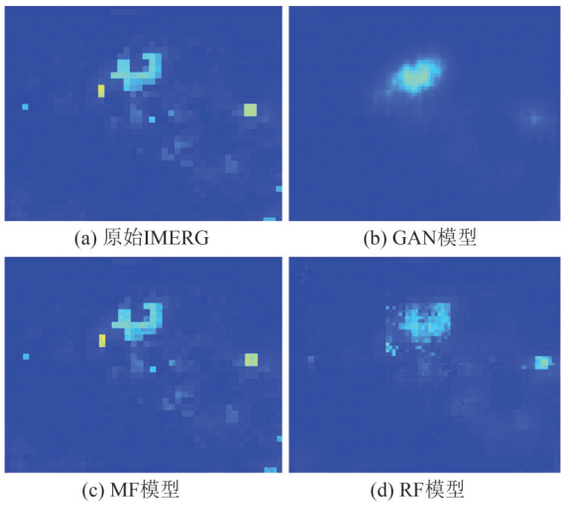


图 8　4种降水场的细节对比图

Fig. 8　Comparison of the Details of 4 Precipitation Fields

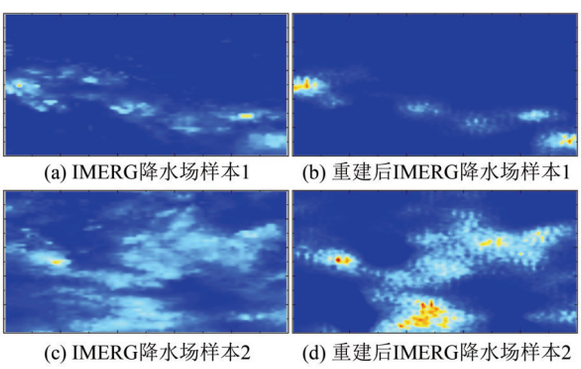


图 9　样本及其重建降水场的对比图

Fig. 9　Comparison of Samples and Their Reconstructed Precipitation Fields

本文利用中国大陆699个气象站点的观测数据,验证了原始IMERG数据和3种模型降尺度后结果的数值精度。9 d的平均数值评估如表3所示。从验证结果来看,GAN模型提高了原始IMERG降水数据的数值精度,偏差减少了10%,虽仍存在一定的高估现象,但相比原始IMERG数据,重建后的降水场与站点观测数据更加贴合,数据整体准确性更高;相关系数和命中率分别提高了0.2%和7.1%,其中有67%的天数相关系数效果变好,且在选取的两个降水事件中,重建后每一天的降水命中率都有一定改善,表明GAN模型能够在基本保持数据相关性的同时提升降水场的命中率。MF和RF模型降尺度效果不如GAN。MF模型降尺度后各项精度指标均低于原始降水数据,而RF模型虽在相关系数和*命中率*上有一定的提升效果,但偏差却高于原始IMERG数据。综合来看,本文构建的GAN模型能构建数据准确性更强的高分辨率降水场。

图10给出了原始IMERG数据及3种模型降尺度结果相对于站点数据的平均偏离程度。可以看出,GAN粉色圆点覆盖区域多于MF和RF模型,表明GAN重建后的降水场在减少或基本保持数据偏离情况方面表现最佳。此外,这两种模型在中国中西部地区和东北地区的蓝色圆点覆盖区域明显多于GAN模型,表明本文构建的GAN模型重建后数据准确性更好,更加贴合原始降水数据。

表 3　IMERG 的平均数值评价表

Table 3　Average Numerical Assessments of IMERG

方法	相关系数	偏差	命中率
原始 IMERG	0.574	0.227	0.547
GAN	0.576	0.127	0.618
MF	0.554	0.242	0.540
RF	0.582	0.253	0.574

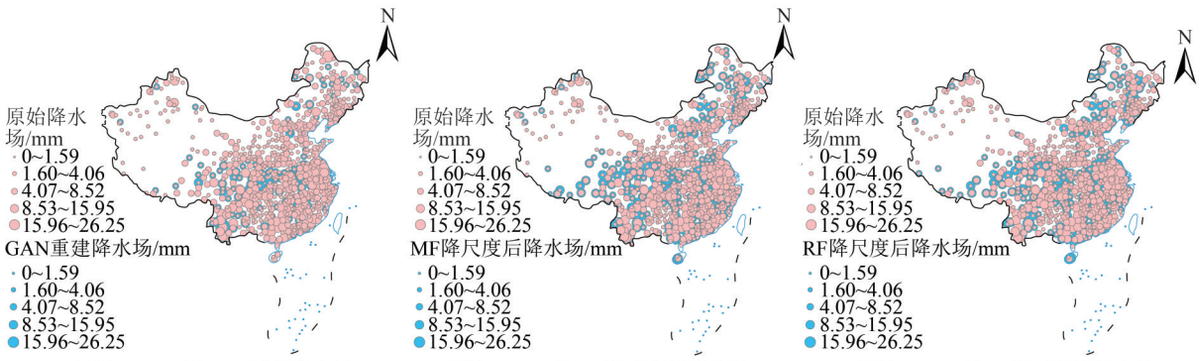
综合来看,相比于MF模型和考虑NDVI和经纬度影响的RF模型,本文构建的GAN模型可以在原始IMERG数据的基础上重建出更精细的降水场,将IMERG降水数据的空间分辨率从0.1°提升至0.05°,同时在提升原有IMERG精度的情况下,有效降低偏差,提升准确率,实用价值更高。

3.2.2　年、季、月尺度GAN降尺度效果的统计分析

为深入分析验证GAN模型的适用性,本文对2020年IMERG日降水数据进行GAN重建,生成中国大陆2020年逐日高分辨率降水数据集(0.05°×0.05°),并从年、季、月3个时间维度对降尺度效果进行精度评价。

年尺度的精度统计分析如图11所示,可以看出,相关系数整体变化趋势与原始降水数据基本保持一致,其中有超过73%的天数相关系数提升或与原始数据的相关系数差距在0.05以内,表明本文构建的GAN模型可以提升或基本维持原始降水数据的一致性;相较于原始降水数据,重建后偏差更加平稳,其中有76 d偏差的提升效果超过20%,表明模型能够有效减少降水数据的异常值,极大增强数据准确性;GAN重建后降水数据的命中率明显高于原始降水数据,其中有超过87%的天数命中率提升。

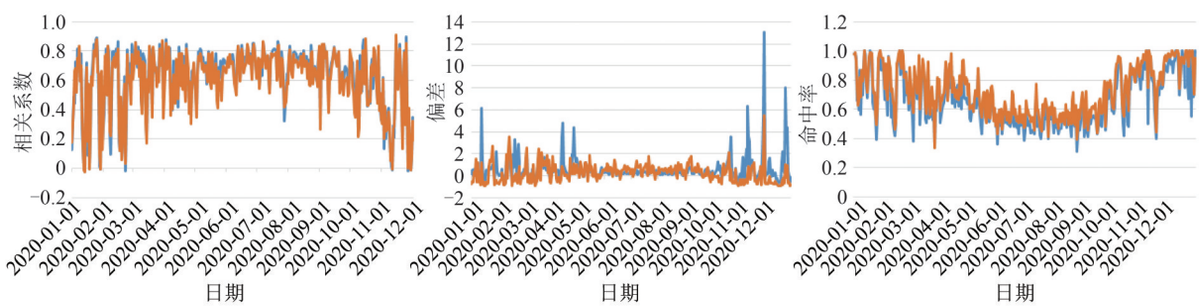
在季节尺度上,GAN重建后,每个季度均有超过65%的天数相关系数提升或与原始数据差距在0.05以内,其中第四季度有将近86%的天数



(a) 原始IMERG及GAN重建降水场相对于站点的平均偏离程度 (b) 原始IMERG及MF降尺度后降水场相对于站点的平均偏离程度 (c) 原始IMERG及RF降尺度后降水场相对于站点的平均偏离程度

图10 原始IMERG数据及3种降尺度结果相对于站点数据的平均偏离程度图

Fig. 10 Average Deviation of Original IMERG Data and 3 Downscaled Results Relative to the Site Data



(a) 原始IMERG数据及其重建结果的相关系数 (b) 原始IMERG数据及其重建结果的偏差 (c) 原始IMERG数据及其重建结果的命中率

图11 2020年原始IMERG数据及其重建结果的精度评价指标

Fig. 11 Accuracy Evaluation Indices of Original IMERG Data and Its Reconstruction Results in 2020

相关性增强,整体提升效果最佳;相较于原始降水数据,重建后每个季度的命中率都有明显的提升,均有超过79%的天数降水命中率增加。图12给出了每个季度原始IMERG数据及其重建结果相对于站点数据的平均偏离程度。可以看出,每个季度重建后降水场相对于站点数据的偏离程度均低于或基本维持原始数据的偏离度,其中第四季度偏离程度明显低于其他季度,数据准确性最高。

为检验分析GAN模型在不同月份的适用性,本文统计了2020年每个月的精度评价指标,如图13所示。相较于原始IMERG数据,GAN重建后降水场的相关性与原始数据基本保持一致,其中每个月均有超过50%的天数相关系数提升或与原始数据的差距小于0.05,且有8个月超过67%,表明本文构建的GAN模型在改善数据一致性方面表现较稳定;重建后的偏差大多低于原始降水数据,且更加平稳,尤其是在4月、10月、11月和12月改善较为明显,表明模型对枯水月偏差的改进效果更好;相较于原始降水数据,重建后降水数据的命中率明显增加,每个月都有超过60%的天数命中率提升,其中有8个月超过90%。综上所述,本文构建的GAN模型在IMERG日降水上的适用性较强,能够提供数据质量更好的高分辨率日降水数据,具有一定的研究意义。

4 结 语

本文以中国大陆为研究区域,基于全国智能网格实况分析降水产品,构建两倍的GAN模型,分别应用在IMERG(0.1°)和ERA5(0.25°)日降水数据中,实现两种数据的空间降尺度,分别将空间分辨率提升至0.05°和0.125°。同时,本文对比分析了GAN模型、MF模型和RF模型在IMERG日降水数据上的降尺度效果,并基于站点降水从年、季、月3个角度统计验证了GAN模型重建后降水数据的数值精度,生成2020年中国大陆高分辨率日降水数据集(0.05°×0.05°)。主要结论如下:

1)本文所构建的GAN模型能够很好地应用于其他降水产品的超分辨率重建,但在卫星降水数据上的适用性更好。重建后IMERG降水场的各项精度指标都有大幅提升,数据一致性和可靠

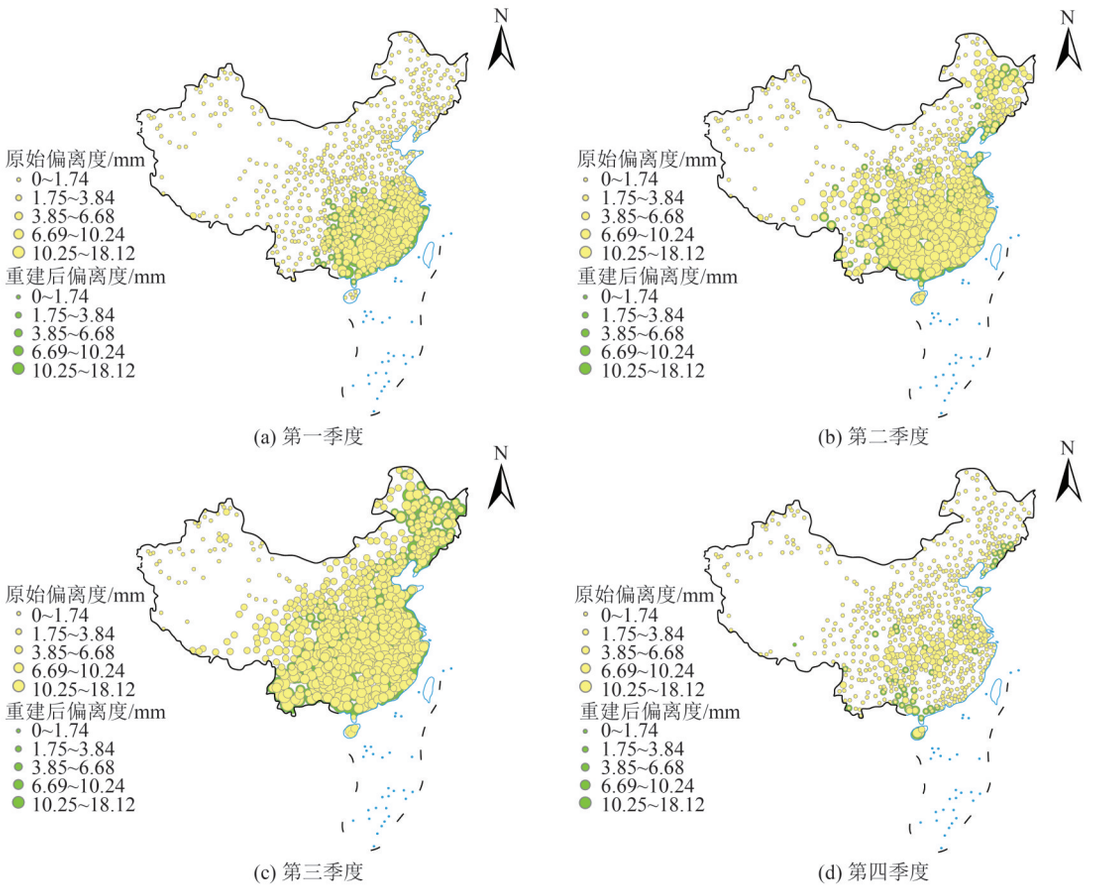


图 12　2020年每个季度原始IMERG数据及其重建结果相对于站点数据的平均偏离程度

Fig. 12　Average Deviation of Original IMERG Data and Its Reconstruction Results Relative to the Site Data for Each Quarter of 2020

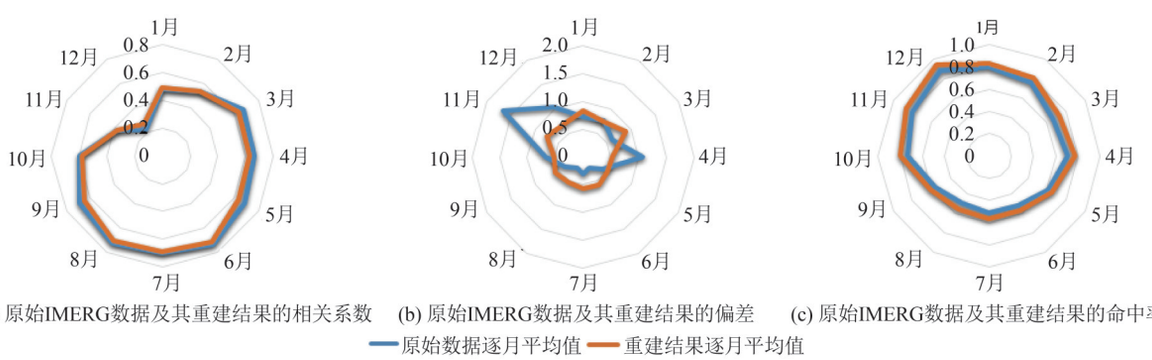


图 13　2020年逐月原始IMERG数据及其重建结果的精度评价指标

Fig. 13　Accuracy Evaluation Indices of Original IMERG Data and Its Reconstruction Results at Monthly Scale in 2020

性增强,并且在空间分布图像上出现了一些细节纹理特征,表明本文制作的CMPA降水样本集在卫星日降水数据降尺度中具有很好的实用价值。重建后的ERA5降水场在提高空间分辨率的同时,保留了降水的空间分布结构,全面提升了降水场的总体精度,偏差和命中率分别改善了38%和11.2%,但整体精度要低于IMERG降水数据。

2)模型可以保留CMPA产品的基本数据特征。重建后两种降水数据的各项精度指标都更贴近CMPA数据,而CMPA作为目前数据质量较好的高精度降水产品,即表明本文构建的GAN模型能够在一定程度上提升数据的精确性。

3)整体来看,不论是基于站点数据或是基于CMPA数据的精度检验,重建后ERA5降水数据的数值精度都低于IMERG降水数据,表明降尺度效果受原始数据的空间分辨率和精度的影响。此外,ERA5降水数据与CMPA数据之间的相关性不够高,致使GAN模型重建效果不佳。

4)本文构建的GAN模型在IMERG日降水数据上的降尺度效果比MF模型要好,也优于考

虑NDVI和经纬度影响的RF模型。在空间分布方面,GAN重建后的高分辨率降水场基本保留了原始降水的空间分布趋势,降水细节特征更明显,没有出现MF模型的马赛克现象和RF模型的块状效应;在数值评价方面,GAN重建后精度指标大多优于其他两种模型。RF模型虽在相关性方面有些许提升,但其偏差远不如GAN模型。

5)综合年、季、月3个维度的统计分析结果,本文构建的GAN模型在IMERG日降水数据上的适用性较强,模型表现较稳定,重建后的各项精度指标大多提升或保持原有数据的精度,表明本文生成的高分辨率日降水数据集具有一定的实用价值。

本文构建的GAN模型受限于CMPA数据的空间分辨率,空间降尺度效果无法达到最优,之后的研究可以考虑使用更高分辨率的数据作为模型高分辨率端的输入数据,以期获得更精细化的降水分布。

参 考 文 献

[1] 张寒博,徐勇,窦世卿,等. 基于GWR模型的长江流域TRMM数据降尺度[J]. 水土保持研究,2021,28(3):149-155.
ZHANG Hanbo, XU Yong, DOU Shiqing, et al. TRMM Downscaling Data of Yangtze Based on GWR Model[J]. *Research of Soil and Water Conservation*, 2021, 28(3): 149-155.

[2] 李国平,陈娇娜,郝丽萍. 基于GPS-PWV的不同云系降水个例的综合分析[J]. 武汉大学学报(信息科学版),2011,36(4):384-388.
LI Guoping, CHEN Jiaona, HAO Liping. Case Study of the Rainfall Processes in Different Cloud Systems Based on GPS-PWV Data in Chengdu Plain[J]. *Geomatics and Information Science of Wuhan University*, 2011, 36(4): 384-388.

[3] 刘舒雅,郑胜杰,张唯. 时序变差函数特征驱动下的月平均降水空间分布模拟[J]. 武汉大学学报(信息科学版),2022,47(7):1043-1051.
LIU Shuya, ZHENG Shengjie, ZHANG Wei. Simulation of Spatial Distribution of Monthly Average Precipitation Driven by Temporal Variation Function [J]. *Geomatics and Information Science of Wuhan University*, 2022, 47(7): 1043-1051.

[4] DUAN Z, LIU J Z, TUO Y, et al. Evaluation of Eight High Spatial Resolution Gridded Precipitation Products in Adige Basin (Italy) at Multiple Temporal and Spatial Scales[J]. *Science of the Total Environment*, 2016, 573: 1536-1553.

[5] LI Z, YANG D W, GAO B, et al. Multiscale Hydrologic Applications of the Latest Satellite Precipitation Products in the Yangtze River Basin Using a Distributed Hydrologic Model[J]. *Journal of Hydrometeorology*, 2015, 16(1): 407-426.

[6] 唐国强. 卫星遥感降水在全球及典型区域的检验、应用和改进[D]. 北京:清华大学,2019.
TANG Guoqiang. Validation, Application and Improvement of Satellite Remote Sensing Precipitation in the Globe and Typical Regions[D]. Beijing: Tsinghua University, 2019.

[7] HU Q F, LI Z, WANG L Z, et al. Rainfall Spatial Estimations: A Review from Spatial Interpolation to Multi-source Data Merging[J]. *Water*, 2019, 11(3): 579.

[8] 汤秋鸿,张学君,戚友存,等. 遥感陆地水循环的进展与展望[J]. 武汉大学学报(信息科学版),2018,43(12):1872-1884.
TANG Qiuhong, ZHANG Xuejun, QI Youcun, et al. Remote Sensing of the Terrestrial Water Cycle: Progress and Perspectives[J]. *Geomatics and Information Science of Wuhan University*, 2018, 43(12): 1872-1884.

[9] JOYCE R J, JANOWIAK J E, ARKIN P A, et al. CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution [J]. *Journal of Hydrometeorology*, 2004, 5(3): 487-503.

[10] MANTAS V M, LIU Z, CARO C, et al. Validation of TRMM Multi-satellite Precipitation Analysis (TMPA) Products in the Peruvian Andes[J]. *Atmospheric Research*, 2015, 163: 132-145.

[11] ADLER R F, HUFFMAN G J, CHANG A, et al. The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979-Present)[J]. *Journal of Hydrometeorology*, 2003, 4(6): 1147-1167.

[12] TANG G Q, WAN W, ZENG Z Y, et al. An Overview of the Global Precipitation Measurement (GPM)Mission and It's Latest Development[J]. *Remote Sensing Technology and Application*, 2015, 30(4):607-615.

[13] 王彧蓉,周建中,杨鑫,等. 再分析降水资料的适用性评估与偏差校正:以长江中下游地区为例[J]. 人民长江,2021,52(9):93-100.
WANG Yurong, ZHOU Jianzhong, YANG Xin, et al. Applicability Evaluation and Deviation Correction of Reanalysis Precipitation Data: Case of Middle and Lower Reaches of Changjiang River[J].

Yangtze River, 2021, 52(9): 93-100.

[14] 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 Lan, HE Qiquan, YANG Jiao, 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.

[15] WRIGHT D B, KNUTSON T R, SMITH J A. Regional Climate Model Projections of Rainfall from U. S. Landfalling Tropical Cyclones[J]. *Climate Dynamics*, 2015, 45(11): 3365-3379.

[16] CHEN F R, LIU Y, LIU Q, et al. Spatial Downscaling of TRMM 3B43 Precipitation Considering Spatial Heterogeneity[J]. *International Journal of Remote Sensing*, 2014, 35(9): 3074-3093.

[17] WANG M M, HE G J, ZHANG Z M, et al. Comparison of Spatial Interpolation and Regression Analysis Models for an Estimation of Monthly Near Surface Air Temperature in China[J]. *Remote Sensing*, 2017, 9(12): 1278.

[18] ZHAN C S, HAN J, HU S, et al. Spatial Downscaling of GPM Annual and Monthly Precipitation Using Regression-Based Algorithms in a Mountainous Area[J]. *Advances in Meteorology*, 2018, 2018(1): 1506017.

[19] MA Z Q, HE K, TAN X, et al. Comparisons of Spatially Downscaling TMPA and IMERG over the Tibetan Plateau[J]. *Remote Sensing*, 2018, 10(12): 1883.

[20] 程建新. 基于深度学习超分辨率的气象预报系统研究[D]. 武汉: 武汉大学, 2021.
CHENG Jianxin. Research on Meteorological Forecast System Based on Deep Learning Super Resolution[D]. Wuhan: Wuhan University, 2021.

[21] JIANG Y Z, YANG K, SHAO C K, et al. A Downscaling Approach for Constructing High-Resolution Precipitation Dataset over the Tibetan Plateau from ERA5 Reanalysis[J]. *Atmospheric Research*, 2021, 256: 105574.

[22] VANDAL T, KODRA E, GANGULY S, et al. Deepsd: Generating High Resolution Climate Change Projections Through Single Image Super-Resolution[C]//The 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, 2017.

[23] 师春香, 谷军霞, 韩帅, 等. 全国智能网格实况融合分析产品进展[C]//第35届中国气象学会年会, 中国合肥, 2018.
SHI Chunxiang, GU Junxia, HAN Shuai, et al. Advances in National Intelligent Grid Live Fusion Analysis Products [C]//The 35th Annual Meeting of the Chinese Meteorological Society, Hefei, China, 2018.

[24] 卢新玉. 融合多源遥感数据与地面观测的新疆地区降水量估算方法研究[D]. 南京: 南京信息工程大学, 2019.
LU Xinyu. Merging Method of Precipitation based on Multi-source Remote Sensing Data and Ground Observation in Xinjiang[D]. Nanjing: Nanjing University of Information Science and Technology, 2019.

[25] 许智铭, 王正涛. 联合GRACE和水文数据探测松花江流域地下水时空变化[J]. 武汉大学学报(信息科学版), 2023, 48(9): 1409-1415.
XU Zhiming, WANG Zhengtao. Detect Songhua River Basin Groundwater Spatiotemporal Variation Characteristics by GRACE and Multi-source Hydrological Data[J]. *Geomatics and Information Science of Wuhan University*, 2023, 48(9): 1409-1415.

[26] CHEN F R, GAO Y Q, WANG Y G, et al. Downscaling Satellite-Derived Daily Precipitation Products with an Integrated Framework[J]. *International Journal of Climatology*, 2019, 39(3): 1287-1304.

[27] GEMITZI A, KOUTSIAS N, LAKSHMI V. A Spatial Downscaling Methodology for GRACE Total Water Storage Anomalies Using GPM IMERG Precipitation Estimates[J]. *Remote Sensing*, 2021, 13(24): 5149.

[28] 叶梦姝. 中国大气再分析资料降水产品在天气和气候中的适用性研究[D]. 兰州: 兰州大学, 2018.
YE Mengshu. Reasearch on the Applicability of Precipitation Reanalysis Data in CRA-Interim of Climate and Weather Characters in China[D]. Lanzhou: Lanzhou University, 2018.

[29] 鹿璇, 汪鼎文, 石文轩. 利用在线字典学习实现图像超分辨率重建的算法[J]. 武汉大学学报(信息科学版), 2018, 43(5): 719-725.
LU Xuan, WANG Dingwen, SHI Wenxuan. Image Super-Resolution with On-Line Dictionary Learning [J]. *Geomatics and Information Science of Wuhan University*, 2018, 43(5): 719-725.

[30] 陈行, 罗斌. 利用动态上采样滤波深度网络进行多角度遥感影像超分辨率重建[J]. 武汉大学学报(信息科学版), 2021, 46(11): 1716-1726.
CHEN Hang, LUO Bin. Multi-angle Remote Sensing Images Super-Resolution Reconstruction

(下转第2085页)