

基于深度学习的降水降尺度方法构建及优化

牛子贤¹, 陈杰¹, 熊立华¹, 李爽¹, 柏雄风¹

1 水资源工程与调度全国重点实验室 武汉大学

摘要: 为了提高深度学习对全球气候模式(GCMs)日降水的降尺度效果,以长江流域为研究对象,基于20种GCMs输出的日降水数据,构建了4种深度学习降尺度模型,并与日偏差校正方法(DBC)结合,提出一种混合降尺度方法(DL-DBC)。4种深度学习对GCMs日降水的降尺度表现接近;与DBC方法相比,其降尺度后的多年平均日降水的平均绝对相对误差(MARE)更低,但多年平均月降水和多年平均年降水的MARE略高,与深度学习相比,DL-DBC得到的多年平均年降水的MARE降低了6.7%-11.3%,多年平均月降水的MARE降低6.3%-7.6%,且在降水量频率分析等方面同样表现更好。混合方法DL-DBC能提高深度学习模型对GCMs日降水数据的降尺度效果,进一步减小GCMs日尺度降水数据的偏差。

关键词: 降尺度、深度学习、降水、全球气候模式、长江流域

中图分类号: P339 收稿日期: 2025-02-17

Improvement of Deep Learning Method for daily Precipitation Downscaling

Zixian Niu¹, Jie Chen¹, Lihua Xiong¹, Shuang Li¹, Xiongfeng Bai¹

1 State Key Laboratory of Water Resources Engineering and Management, Wuhan University

基金项目: 湖北省自然科学基金创新群体项目(2025AFA023)

国家自然科学基金长江水科学研究联合基金(No. U2240201)。

作者简介: 牛子贤(2001-),男,硕士研究生,主要研究方向为气候变化的水文响应。E-mail: 2019302060100@whu.edu.cn, ORCID: 0009-0001-5809-5933

通讯作者: 陈杰, E-mail: jiechen@whu.edu.cn; 熊立华, E-mail: xionglh@whu.edu.cn

Abstract: To improve the downscaling effectiveness of deep learning methods for daily precipitation from Global Climate Models (GCMs). Targeting the Yangtze River Basin, we constructed four deep learning downscaling models based on historical daily precipitation outputs from 20 GCMs. A hybrid method (DL-DBC) was proposed by integrating these models with the Daily Bias Correction (DBC) method. The four deep learning models exhibited comparable performance in daily precipitation downscaling. Compared to DBC, they achieved lower Mean Absolute Relative Error (MARE) for multi-year average daily precipitation but slightly higher MARE for multi-year average monthly and annual precipitation. The DL-DBC method outperformed standalone deep learning models, reducing MARE for multi-year average annual precipitation by 6.7%–11.3% and monthly precipitation by 6.3%–7.6%, while also demonstrating superior performance in precipitation frequency analysis. The DL-DBC method enhances the downscaling effectiveness of deep learning models and further reduces biases in on daily precipitation data from GCMs.

Keywords: Downscaling, Deep learning, Yangtze River Basin, Precipitation, Global Climate Models

引言

全球变暖已成为学术界广泛关注的议题，全球变暖暴雨等极端水文事件更为频发（张仲石等, 2022）不仅影响自然生态系统，还对人类社会产生了深远的影响。由全球气候模式(GCMs)提供的降水等气候变量涵盖过去和未来时期，已成为科学界探测未来气候变化信息的有效工具(Meehl et al., 2007; Taylor et al., 2012)。然而，GCMs 输出的气象数据往往分辨率较低，难以直接与流域水文模型相结合，且与同一网格尺度的观测气象要素相比，GCMs 模拟数据通常具有较大的偏差，不能直接应用于流域水文模型开展水文响应评估(Cannon, 2017; Mearns et al., 2012; Sillmann et al., 2013)，降尺度是解决气候模式输出与响应模型输入之间尺度不匹配以及偏差等问题的重要手段。

常用的降尺度方法可以分为动力降尺度、统计降尺度以及动力和统计相结合的降尺度方法(Chen et al., 2016)。动力降尺度，即区域气候模型(Regional Climate Models, RCMs)，是建立在一套大气动力学方程基础之上，基于 GCMs 提供的初始和侧边界条件，通过动力学方程之间的关系，推导高分辨率的气候变化情景(Dickinson et al., 1989; Yang et al., 2010)。统计降尺度是一种通过建立气候模式输出的历史时段网格尺度气候变量与站点或局地观测的气象要素之间的统计关系，

然后将这种关系应用于气候模式输出的未来时段网格数据,从而获得站点或区域尺度未来气候变化情景的方法。相比动力降尺度,统计降尺度计算简单,模型易于构建(刘永和 et al., 2011),灵活度更高,因此得到了更广泛的应用(Bennett et al., 2011; Eden et al., 2014; Eum et al., 2017; Guo et al., 2019; Monjo et al., 2016; Pierce et al., 2015; L. Wang et al., 2014)。目前,统计降尺度方法包括从简单的空间分解偏差校正(BCSD)(Wood et al., 2004)到分位数校正(如QM、QDM等)到机器学习方法,如:人工神经网络(Schoof et al., 2001; Wilby et al., 1998)、随机森林等(He et al., 2016; Legasa et al., 2022; Mei et al., 2020; Pour et al., 2016)。

目前,机器学习已经在水文的各个领域得到了运用(刘杰等, 2024),但以往的研究表明,与传统统计降尺度方法相比,机器学习方法所获得的改进甚微(He et al., 2016; Sachindra et al., 2019; Vandal et al., 2019; Yang et al., 2018)。作为机器学习的一种进化,深度学习能够构建更复杂的层次结构,在提取数据的高级特征上有更好的表现,有助于捕获复杂的非线性关系,处理高维数据,因此逐渐被应用于降尺度研究。尤其是具有卷积的深度学习神经网络(CNN)在气候模式输出(Y. Liu et al., 2020; Pan et al., 2021; Rodrigues et al., 2018; F. Wang et al., 2022)、再分析数据(Baño-Medina et al., 2020; Sun et al., 2021)和气象预报模型输出(Harris et al., 2022; Li et al., 2022)等的降尺度中成功应用。具体而言,基于深度学习的降尺度方法相比过去的统计降尺度方法能够更好的优化降水发生的时间结构。

然而,深度学习方法在针对GCMs日降水的降尺度中应用仍然较少,主要是由于GCMs模拟降水与观测值在时间上并不存在同步关系,因此具体到每日的降水偏差较大,使得模型输入的样本与实测数据无法一一对应。然而,现有部分研究采用深度学习方法直接针对GCMs日降水数据进行降尺度,如Sharma等(2024)在印度基于两种深度学习方法SRDRN和CNNBC对GCMs日降水开展降尺度研究,结果表明两种深度学习方法的网格协议指数IOA并不优于简单的分位数增量映射(QDM)方法,从理论上而言,这种降尺度忽略了GCMs数据与实测数据之间的非对应关系。

相比较来自GCMs的降水数据,再分析数据可能更适用于深度学习模型的训练,这是因为再分析数据集通过同化多元观测数据,降低了与实际地表观测数据的偏差,因此基于深度学习模型的降尺度方法在再分析数据中的降尺度表现往往

更好，例如 Sun and Lan (2021)对于 ECMWF 再分析降水数据的降尺度测试中，CNN 模型取得了比广义线性回归 GLM 更好的降水降尺度效果，对比训练期和测试期的降尺度效果，也展现出了更好的稳定性。然而，GCMs 数据与再分析数据之间的偏差往往较大，在将基于再分析数据训练的深度学习模型应用于 GCMs 数据时，需要对 GCMs 数据进行偏差校正。

基于以上背景，本研究选择 20 种 GCMs 输出的历史日降水数据，并以长江流域为对象，构建了 4 种深度学习（CNN、SRDRN、ConvNext、DenseNet）降尺度模型，评估了其在 GCMs 降水降尺度方面的表现，作为常用的降水偏差校正方法，日偏差校正方法（DBC 校正）能考虑干湿频率并基于分位数进行校正，具有较为稳定的效果，研究选择 DBC 与深度学习降尺度方法结合，并引入 ERA5 再分析日降水数据作为训练数据提出了一种混合降尺度方法（DL-DBC），与传统的线性缩放（LS）偏差校正和 DBC 方法进行对比验证，从降水的整体偏差、降水频率分布特征、降水时间序列和空间结构等方面对多种降尺度方法进行综合比较分析。

1. 数据与方法

1.1 研究流域与数据

本研究选择的区域为长江流域，范围从 90° E 到 122° E, 24° N 到 36° N，如图 1 所示。长江流域是世界第三大流域，源区地处高原气候区，中下游则主要是丘陵和盆地，降水量从上游到下游呈现出明显的梯度变化，其中四川盆地和长江中下游平原是降水最多的区域，部分地区超过 1500mm，上游地区降水量较少，一般在 500mm 以下。长江流域地处东亚季风区，降水量在时间上分配不均匀，降水主要集中在夏季，占全年降水量的 50%以上，冬季降水量仅占全年的 10%左右，且长江流域是夏季大范围旱涝灾害发生频繁且严重的地区(张琼 et al., 2013)。

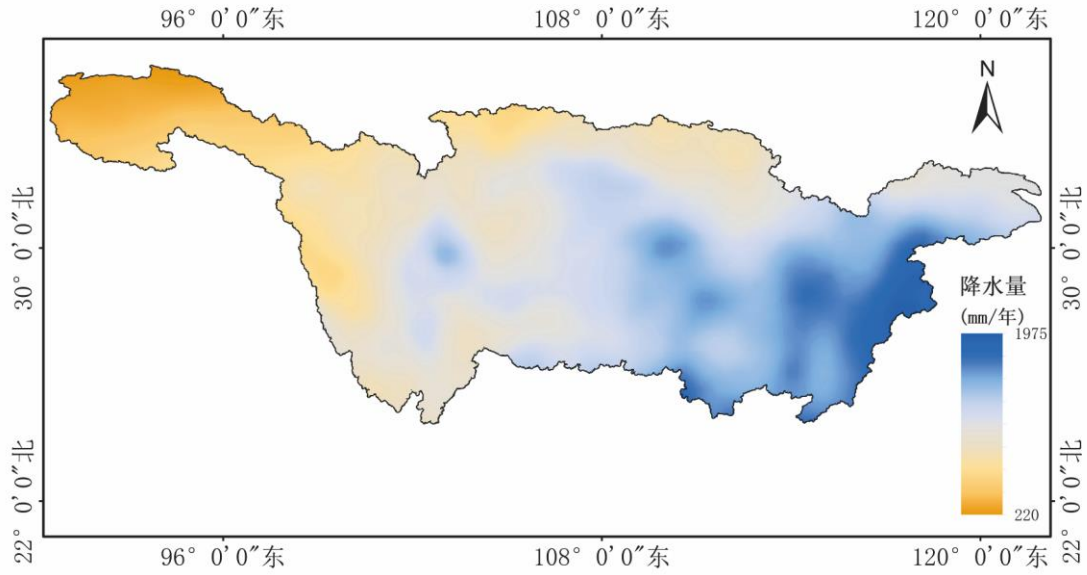


图 1 长江流域 1979-2014 年间多年平均年降水量（数据来源于 CN05.1 数据集）

Fig.1 The multi-year average annual precipitation in the Yangtze River Basin from 1961 to 2014 (data sourced from the CN05.1 dataset).

本研究同时使用来源于国际模式第六次比较计划（CMIP6）的 20 种 GCMs 模拟的降水数据以及 ECMWF 发布的第五代大气再分析数据集 ERA5 降水数据，20 种 GCMs 的模式、分辨率及开发机构如表 1。大多数 GCMs 包含 1850-2014 年历史条件驱动下的降水序列。由于 ERA5 降水数据在 1979 年之前的质量较差，因此为了将数据的时间范围保持一致，研究的时间范围选择 1979-2014 年。

本研究采用 CN05.1 的网格降水数据作为观测数据。CN05.1 是一个使用基于中国 2416 个气象站观测数据的插值建立的格网数据集(Wu et al., 2013)，其分辨率为 $0.25^{\circ} \times 0.25^{\circ}$ ，在本研究区域内共有 48×128 个栅格。CN05.1 包含中国境内 1961-2014 年的多种气象变量，为了与模拟降水数据保持一致，本研究仅使用 1979-2014 年的日降水数据序列。

表 1 GCMs 降水数据的初始分辨率及来源

Table 1 The initial resolution and sources of precipitation data from GCMs.

GCMs	分辨率 (纬度° × 经度°)	国家/机构
GFDL-ESM4	$1^{\circ} \times 1.25^{\circ}$	NOAA.USA
ACCESS-CM2	$1^{\circ} \times 1^{\circ}$	Australia
MPI-ESM1	$1^{\circ} \times 1^{\circ}$	Germany

INM-CM4	1.5°×2°	Russia
CSM2-MR	1°×1°	NCAR.USA
CMCC-CM2-SR5	1.125°×1.112°	Italy
CNRM-CM6	1.41°×1.41°	France
FGOALS-f3	1°×1.25°	China
IPSL-CM6A	1.26°×2.5°	France
MIROC6	1.41°×1.41°	Japan
CESM2-WACCM	0.94°×1.25°	NCAR.USA
TaiESM1	1.25°×0.9424°	Taiwan.rpc
NorESM2-MM	0.94°×1.25°	Norway
INM-CM5	1.5°×2°	Russia
CESM2	0.94°×1.25°	NCAR.USA
EC-Earth3	0.70°×0.70°	EC-Earth-Consortium
UKESM1-0-LL	1.25°×1.88°	UK
CAMS-CSM1-0	1.11°×1.13°	China
BCC-CSM2-MR	1.11°×1.13°	China
AWI-ESM-1-1-LR	1.85°×1.88°	Germany

1.2 研究方法

1.2.1 偏差校正方法

本研究选择了两种偏差校正方法作为基准方法，分别是基于均值的线性缩放法 LS 和基于分布的日偏差校正 DBC(Chen et al., 2013)。

LS 校正(Linear scaling bias correction)是一种基于均值的偏差校正方法，通过将 GCMs 的历史模拟数据与观测数据在各月份的差异确定为各月的修正因子，用于修正测试期 GCM 各月的日降水量，公式如下：

$$P_{cor,cali,m} = \frac{P_{obs,m}}{P_{cali,m}} * P_{cali,m} \quad (1)$$

$$P_{cor,vali,m} = \frac{P_{obs,m}}{P_{cali,m}} * P_{vali,m} \quad (2)$$

式中，下标 obs 、 $cali$ 和 $vali$ 分别表示训练期观测数据、训练期待校准数据和

验证期待校准数据; m 表示月份, cor 表示校正后的数据。

DBC (Daily bias correction)是一种基于分布的偏差校正方法,它能够考虑逐日降水频率分布的不同,并修正降水的雨日频率。DBC 方法的步骤是基于历史时期的观测数据对相应时期的 GCMs 模拟数据进行雨日频率校正和分位数校正,然后基于历史时期的校正因子对验证期或未来 GCMs 的降水分别进行雨日频率和分位数校正。DBC 校正的公式如下所示:

$$P_{cor,cali,m,p} = \frac{P_{obs,m,p}}{P_{cali,m,p}} * P_{cali,m,p} \quad (3)$$

$$P_{cor,valid,m,p} = \frac{P_{obs,m,p}}{P_{cali,m,p}} * P_{valid,m,p} \quad (4)$$

式中,下标 obs 、 $cali$ 和 $valid$ 分别表示训练期观测数据、训练期待校准数据和验证期待校准数据; m 表示月份, cor 表示修正后的数据, p 为校正因子对应分位数。

1.2.2 深度学习降尺度方法

本研究构建了 4 种基于监督学习的深度学习降尺度方法,均由深度学习主干部分和上采样部分构成。模型需要一组配对的高分辨率观测数据和低分辨率模拟数据作为输入,并划分为训练期与验证期,利用主干部分的卷积层提取训练期低分辨率数据的特征,训练模型参数,并在上采样部分生成高分辨率数据,通过计算损失函数衡量模型在验证期输出的高分辨率数据与观测数据的接近程度,并不断调整参数权重,逐渐优化模型的模拟效果。

研究选择的 4 种模型主干如图 2(A)所示,该部分用于提取低分辨率数据的特征。CNN(a)是基础的卷积神经网络,由标准卷积块构成,本研究构建的 CNN 由 4 个标准卷积块组成,每个卷积块包括两个卷积层 (3x3 的卷积核)和批归一化层 (Batch Normalization); SRDRN(b)通过引入残差块,能够避免普通 CNN 网络受梯度消失/爆炸和退化等问题,在更深的情况下也无法减少训练损失的缺点(Pan et al., 2019),主要包括 16 个残差块以及其余卷积层(F. Wang et al., 2021); ConvNext(c)由 Z. Liu et al. (2022)提出,借鉴了 Vision Transformer (ViT) 的设计理念,旨在提高其在计算机视觉任务中的表现,同时保持 CNN 本身良好的局部特征学习能力和高效的推理速度,包含 8 个 ConvNext 风格的卷积块; DenseNet(d)通过建立密集连接在残差网络的基础上进一步解决网络退化问题,并增加对每层特征の利用(Huang et al., 2017),包含 8 个密集连接块。本研究构建的 CNN、SRDRN、

DenseNet 均选择修正线性单元 ReLU (Rectified Linear Unit) 作为激活函数, 归一化方法均为批归一化, ConvNext(c)采用高斯误差线性单元 GeLU 作为激活函数, 并选择了更大的卷积核尺寸 (7x7) 及层归一化方法 (Layer Normalization)。

研究选择的上采样方法为次像素卷积法 SPC (Sub-pixel Convolution), 该部分模型由次像素卷积块组成(Lim et al., 2017), 如图 2 (B)所示, 每个上采样块由一个卷积层和一个上采样层组成, 后面是参数化 ReLU 激活函数。上采样块接收到模型主干输出的低分辨率特征图, 并通过重排通道的方式生成高分辨率的输出, 其中每个上采样块负责 2 倍的分辨率放大倍数, 本研究中分辨率放大倍数为 4 倍, 因此该部分由两个上采样块组成。

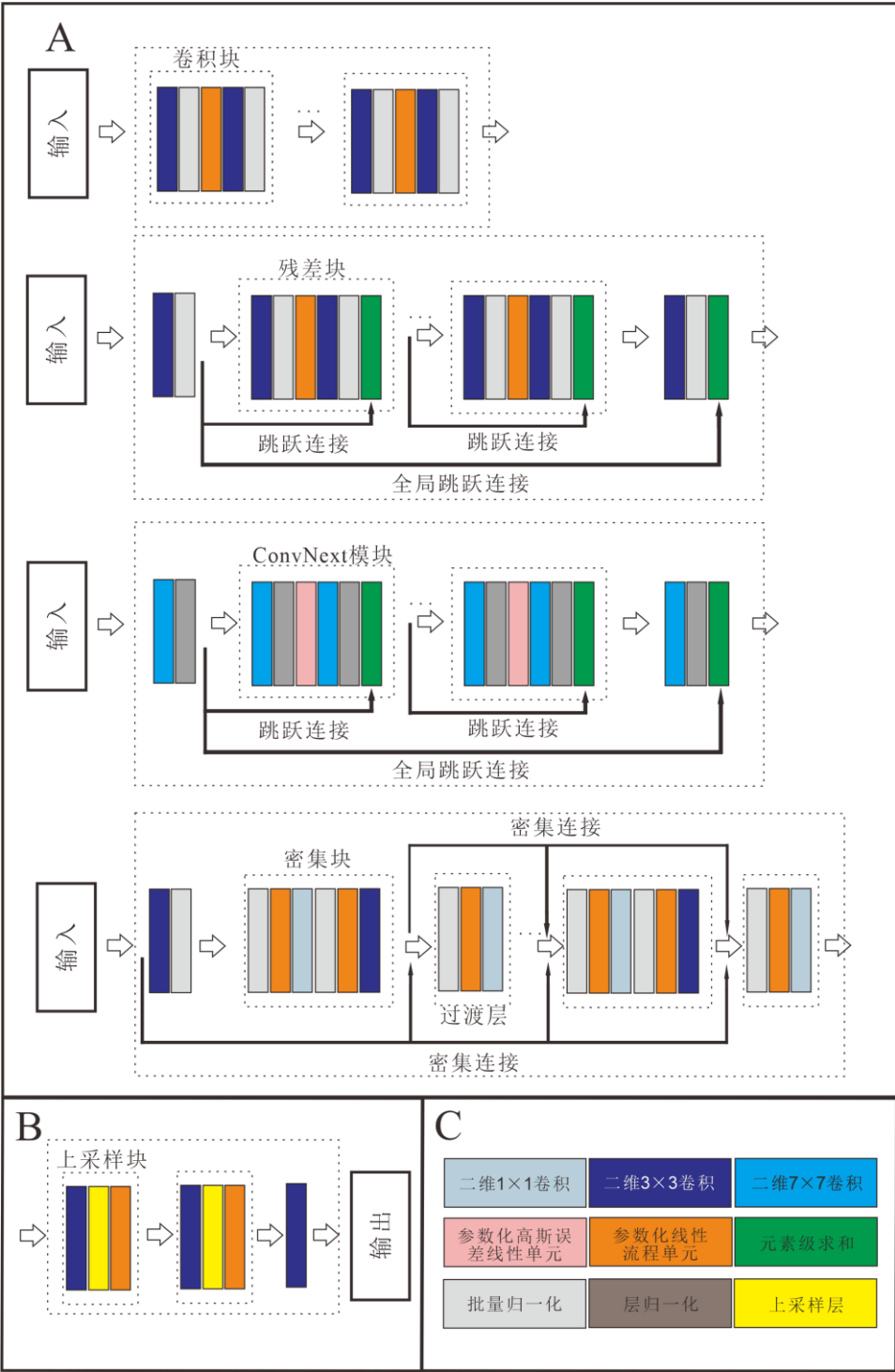


图 2 深度学习模型架构，A 为 4 种模型主干架构，B 为 SPC 上采样法的架构，C 介绍了深度学习架构中所用到的具体结构

Fig. 2 Deep learning model architecture, Part A shows the main architectures of the four models, Part B illustrates the architecture of the SPC upsampling method, and Part C introduces the specific structures used in the deep learning architecture.

研究中所有深度学习模型的训练损失函数为平均绝对误差 MAE, 使用 Adam 优化算法(Diederik, 2014)对网络进行训练, 学习率初始值设置为 0.0005, 随着轮数的增加, 学习率会按一定比例减少, 最小值为 0.0001。所有深度学习模型的 epoch 数量均设置为 200, 这意味着在训练时所有的样本均会通过模型 200 次, 其中每个 epoch 中的 Batch size 被设置为 16, 即每批进入模型的样本数量为 16, 每处理完这一批样本, 模型参数就会进行一次更新。

在使用深度学习模型进行降尺度前, 研究对 GCM 降水数据进行了简单的数据增强, 即通过减去平均值 (μ)并除以标准差 (σ) 的方法对降水数据集进行归一化, 并通过旋转和水平/垂直翻转来提高样本的多样性, 以增强模型在面对未知时期的适应能力。为了减小噪声, 降低模型训练的难度, 在训练模型时将 GCMs 降水数据和观测降水进行 3 日滚动平均, 而在对测试期降水进行模拟时, 依旧采用原始的 GCMs 降水数据作为输入数据。

1.2.3 多源数据混合降尺度方法 (DL-DBC)

DL-DBC 方法首先采用 ERA5 降水数据作为深度学习模型的训练数据, 然后根据训练期的 ERA5 和 GCMs 降水数据的差异, 采用 DBC 方法对测试期的 GCMs 数据进行偏差校正, 并作为深度学习降尺度模型测试期的输入。该方法考虑了 ERA5 与 GCMs 降水数据之间的差异, 在使用 GCMs 数据之前, 先以 ERA5 降水数据为基本, 采用 DBC 对其进行偏差校正, 同时, 通过 ERA5 数据提高模型的训练效果, 并可用于未来时期的 GCMs 模拟降水数据的降尺度。

对于不同降尺度和偏差校正方法, 采用表 2 中的编号表示。A1 和 A2 分别是 LS 方法与 DBC 方法, B1-B4 为直接使用 GCMs 训练的几种深度学习模型降尺度方法, B1A2-B4A2 分别为基于四种深度学习模型的混合方法 DL-DBC。

1.3 降尺度实验设计

在进行偏差校正和降尺度之前, 首先使用双线性插值法将所有 GCMs 降水数据重新网格化为统一的 $1^\circ \times 1^\circ$, 目标分辨率是与观测数据集相同的 0.25° 纬度 $\times 0.25^\circ$ 经度, 同时在 DL-DBC 中, 为了与模拟用的 GCMs 数据的分辨率一致, 训练用的 ERA5 数据被重采样到 $1^\circ \times 1^\circ$ 。针对少数缺失值, 采用反距离加权法进行插值, 为了避免极小数据影响训练效果, 在训练期设置的阈值为 0.1mm, 低于阈值的数据被视为无降水; 在使用基于深度学习算法的降尺度方法时, 为了尽可

能提高测试期的长度，同时保证训练数据足够，基于 GCMs 直接训练的深度学习降尺度方法 (B1-B4) 和引入 ERA5 数据训练的 DL-DBC (B1A2-B4A2) 将 1979-1989 年作为模型监督学习的训练期，而 1990-2000 年作为监督学习的验证期，测试期选择 1997-2014 年，经过分析，以上三个时期均具有较好的代表性，不存在降水量非常极端的年份；在进行 LS 和 DBC 偏差校正时，选择 1979-2000 作为训练期，2001-2014 作为测试期，从而与两种偏差校正方法的测试期一致。

本研究基于测试期的降尺度结果，从降水的偏差、降水量的频率分布、流域多年平均月降水序列、降水空间偏倚结构等维度衡量了几种降尺度方法。

在偏差和降水量频率分析中，分别采用了 MARE 和 MAE 作为评估指标，其中 MAE 计算公式如下：

$$MAE = \frac{1}{X \times Y \times T} \sum_{x=1}^X \sum_{y=1}^Y \sum_{t=1}^T |p_i - \hat{p}_i| \quad (5)$$

MARE 计算公式如下：

$$MARE = \frac{1}{X \times Y \times T} \sum_{x=1}^X \sum_{y=1}^Y \sum_{t=1}^T \left| \frac{p_i - \hat{p}_i}{p_i} \right| \quad (6)$$

式中， p_i 为观测降水， $\hat{p}_i$ 为模拟结果， X 、 Y 为空间维度长度， T 为时间维度长度。

在分析模拟降水偏差时，采用一致性指数 (Index of Agreement, IOA) 作为分析指标，以衡量模拟降水与观测值一致性 (Sharma et al., 2024)。IOA 计算公式如下：

$$IOA = 1 - \frac{\sum_{i=1}^n |\hat{y}_i - y_i|}{\sum_{i=1}^n (|\hat{y}_i - \bar{y}| + |y_i - \bar{y}|)} \quad (8)$$

式中， $\hat{y}_i$ 和 y_i 分别是模拟和观测降水序列， $\bar{y}$ 是观测降水均值。

在流域多年平均月降水序列分析中，分别使用了均方根误差 RMSE 和皮尔逊相关系数 (Pearson correlation coefficient, r) 作为评估指标，均方根误差 RMSE 如下：

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (9)$$

式中， y_i 为观测降水， $\hat{y}_i$ 为模拟结果， n 为降水数据的总数。

皮尔逊相关系数的计算公式如下：

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

式中， x_i 和 y_i 分别是模拟和观测降水， $\bar{x}$ 和 $\bar{y}$ 分别是模拟和观测降水的均值。

研究通过比较观测降水和模拟降水的空间半变函数图，评估降尺度方法还原降水空间偏倚关系的效果。空间半变函数曲线 γ_s 可计算如下：

$$\gamma_s(h_s) = \frac{1}{2} [I(j + h_s) - I(j)]^2 \tag{9}$$

其中， $I(j)$ 代表位置 j 处的数据， h_s 是一个位移向量（即四个方向网格单元之间的间距）。

表 2 偏差校正与降尺度方法的名称

Table 2 Titles of various bias correction and downscaling methods.

编号	校正方法
A1	LS 校正 (Linear scaling bias correction)
A2	DBC 校正 (Daily Bias Correction Method)
B1	卷积神经网络 CNN (Convolutional Neural Network)
B2	超分辨率深度残差网络 SRDRN (Super resolution deep residual network)
B3	ConvNext 网络
B4	密集连接卷积网络 DenseNet (Densely Connected Convolutional Network)
B1A2	基于 CNN 与 DBC 校正的混合方法
B1A2	基于 SRDRN 与 DBC 校正的混合方法
B1A2	基于 ConvNext 与 DBC 校正的混合方法
B1A2	基于 DenseNet 与 DBC 校正的混合方法

2. 研究结果

2.1 整体偏差分析

研究基于几种降尺度方法得到的测试期模拟降水及 GCMs 直接插值到对应分辨率的结果，以格点为单位，分别计算了多年平均降水、多年平均月降水和多年平均日降水的 MARE（图 3）。

根据图 3 可知,所有降尺度方法得到的降水 MARE 均优于直接对 GCMs 进行双线性插值(BL)的结果,4 种深度学习方法的整体表现接近。在多年平均年降水的分析中,LS 方法得到的 20 种 GCMs 模拟降水的 MARE 均值为 19.4%,DBC 方法得到的结果为 8.9%,直接对 GCMs 进行降尺度的 4 种深度学习降尺度(B1-B4)的结果为 14.0%-18.7%,4 种 DL-DBC 的多年平均降水的 MARE 为 7.1%-7.5%。而在多年平均月降水的分析中,LS 方法的结果为 60.1%,DBC 方法为 27.3%,直接对 GCMs 进行降尺度的 4 种深度学习降尺度的结果为 33.1%-34.3%,4 种 DL-DBC 的结果为 26.1%-26.5%,表现优于其他方法。而在多年平均日降水的分析中,原始数据的 MARE 达到了 186%,直接对 GCMs 进行降尺度的深度学习得到的结果为 75%-81%,优于 LS 方法的 122%和 DBC 方法的 93%,DL-DBC 的结果为 71%-75%,同样优于其他方法。

偏差分析的结果表明,与 DBC 方法相比,LS 方法的结果整体偏差更大,而两种方法在均值上的校正效果本应该比较接近,这很可能是因为日降水数据分布高度偏态,而 LS 方法仅关注整体均值,而忽略极端值的局部偏差。直接使用 GCMs 训练的深度学习模型(B1-B4)得到的多年平均降水以及多年平均月降水 MARE 指标相比 DBC 方法同样略差,但多年平均日降水的 MARE 指标相比 DBC 方法略优,证明基于深度学习的降尺度方法较小的时间尺度下能更有效的缩小偏差,而基于分布的 DBC 方法在较大的时间尺度下表现更好,混合方法 DL-DBC 进一步提升了深度学习降尺度的效果,在不同时间尺度下均能更多的缩小模拟降水的偏差。

为了分析降水降尺度结果在不同季节的偏差分布情况,以 MPI-ESM1 的降尺度结果为例,绘制了降水偏差的季节空间分布图(图 4)。结果显示,在长江上游的高原区域,由于降水本身较少,所有降尺度方法的结果偏差均较小,除高原区域以外,LS 方法在各季节的表现均不稳定,4 种基于 GCMs 训练的深度学习降尺度表现优于 LS 方法,但仍略差于 DBC 方法,具体而言,在降水较高的季节,DBC 方案的结果在大多数地区的偏差均较小,对长江中下游区域的部分地区存在一定的高估,而基于 GCMs 训练的深度学习模型则在多数地区存在低估,下游部分区域略有高估,在冬季也展现出对中下游降水的略微低估;对比之下,4 种 DL-DBC 得到的降水在各个季节的偏差均更小,但雨季仍存在对长江中下游流域

降水的高估。

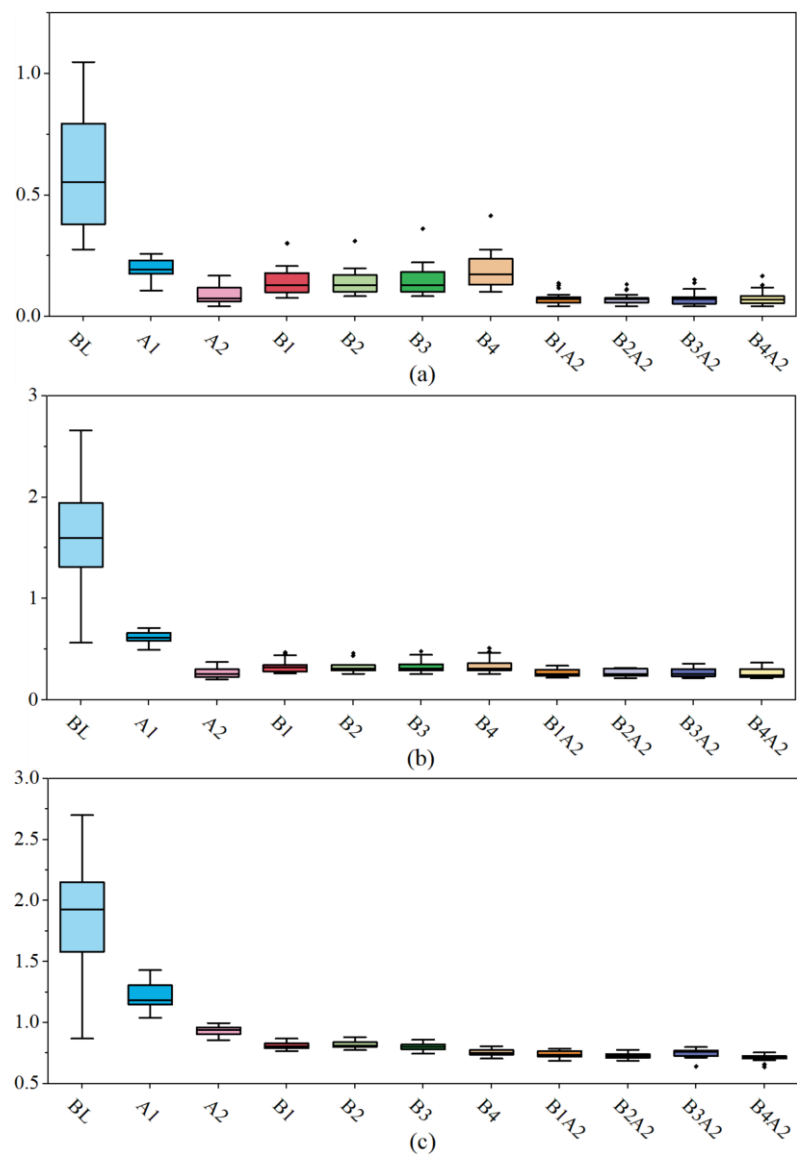


图 3.不同降尺度方法基于不同 CGMs 的降尺度降水与观测降水在测试期的 MARE，(a)为多年平均降水的计算结果，(b)为多年平均月降水计算结果，(c)为多年平均日降水计算结果

Fig. 3 MARE of downscaled precipitation based on different GCMs and observed precipitation during the testing period for various downscaling methods, (a) shows the results for multi-year average precipitation, (b) shows the results for multi-year average monthly precipitation, and (c) shows the results for multi-year average daily precipitation.

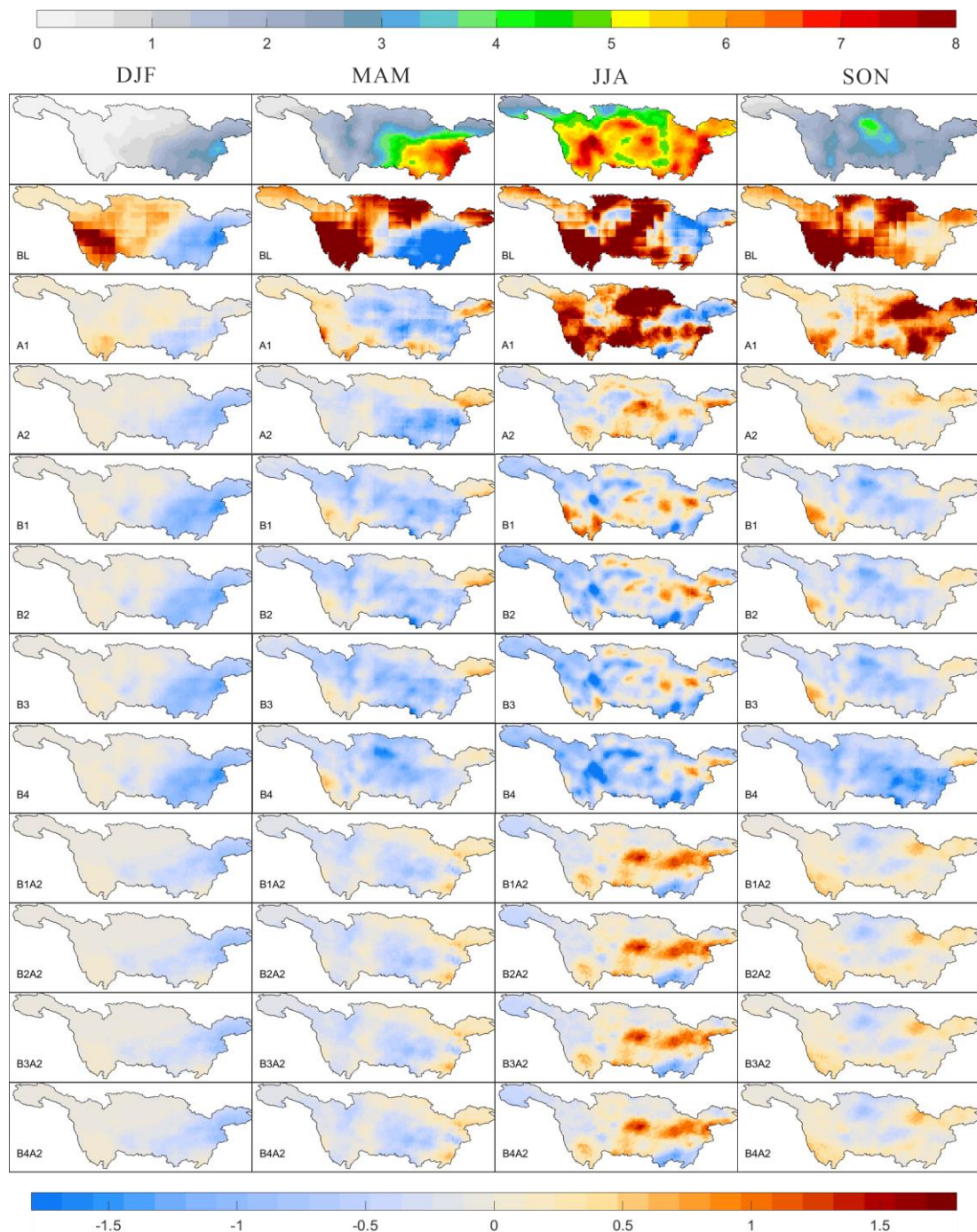


图 4 .任意一种 GCMs 降尺度结果的季节性平均降水量偏差 (毫米/日)的空间分布图以及偏差箱线图, 季节以 12-2 月, 3-5 月, 6-8 月和 9-11 月划分。

Fig. 4 Seasonal average precipitation bias (mm/day) spatial distribution and boxplot of bias for the downscaled results of an arbitrary GCM, with seasons defined as December-February, March-May, June-August, and September-November.

进一步, 对流域每个网格点的一致性指数 IOA 进行了评估。首先对测试期所有年份的降水序列进行每日平均, 得到流域每个格点的年平均日降水序列, 然后

分别计算各网格点的 IOA 指标。图 5(a)展示了 20 种 GCMs 的模拟结果在各网格点的 IOA 均值；图 5(b)统计了 20 种 GCMs 降尺度降水的流域平均 IOA。总体而言，4 种基于 GCMs 训练的深度学习降尺度降水的 IOA 均值为 0.775-0.784，高于 LS 方法的 0.689，与 DBC 方法的结果 0.778 接近，空间分布图可以看出 DBC 方法在降水更低的长江上游表现更好，而在长江中下游，基于 GCMs 训练的深度学习降尺度模型的效果更好。混合方法 DL-DBC 的整体表现更好一些，IOA 均值为 0.791-0.795。从图 5(a)可知，DL-DBC 相比较深度学习降尺度，在下游地区的 IOA 指标略有提升。考虑到长江流域下游的降水量高于中上游，DL-DBC 在降水量较高的区域可能会展现出更好的效果。

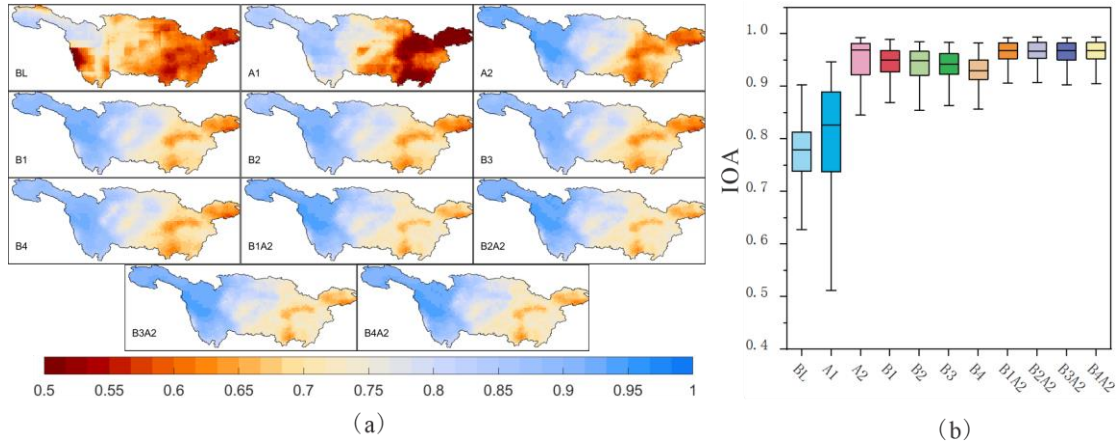


图 5. 模拟降水的 IOA 空间分布图(a)以及 20 种 GCMs 降尺度结果的流域 IOA 均值 (b)

Fig. 5 Spatial distribution of IOA for simulated precipitation (a) and the mean IOA of the watershed from the downscaling results of 20 GCMs (b).

2.2 降水量频率分析

为了分析降尺度降水与观测降水分布的同步性，绘制了分位数-分位数图 (Quantile-Quantile, Q-Q)(图 6)。将训练期观测数据的多年平均日降水大于 2mm 的格点作为高降水地区，低于 2mm 则为低降水区域，分别绘制了 QQ 图。如图所示，与 LS 和 DBC 方法相比，4 种基于 GCMs 训练的深度学习降尺度方法 (B1-B4) 对极端降水存在普遍的低估，无论是在降水偏高还是偏低的地区均展现出局限性，。而 4 种引入 ERA5 降水数据训练的 DL-DBC 与实测数据的 Q-Q 曲线更接近 1:1 线，这表明 DL-DBC 能改善深度学习降尺度难以模拟高降水的问题。

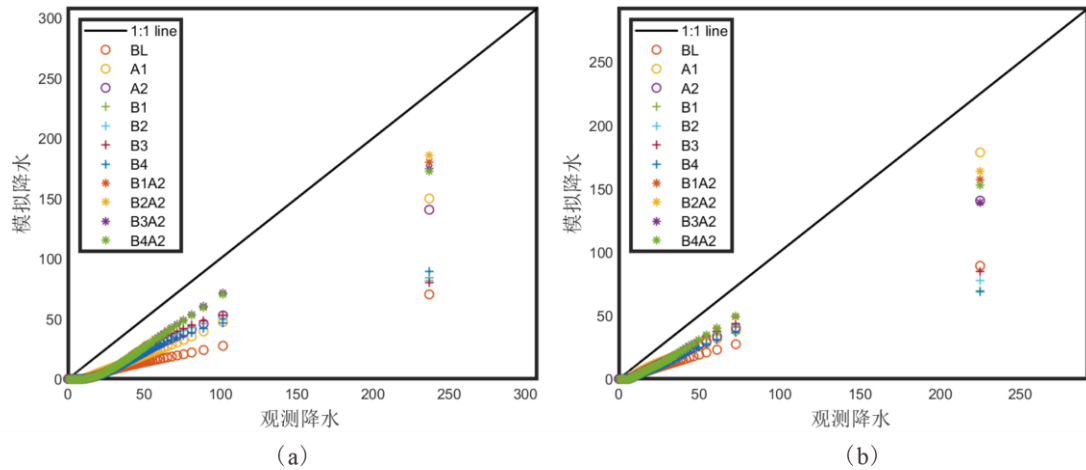


图 6. 任意 1 种 GCMs 模拟降水与观测降水的在高降水区域(a)和低降水区域(b)的流域日平均降水 Q-Q 图 (单位: mm)

Fig. 6. Quantile-Quantile (Q-Q) plots of basin daily average precipitation between simulated and observed precipitation for one selected GCM in (a) high-precipitation areas and (b) low-precipitation areas (unit: mm).

研究选择了四个指标对流域的降水特征进行分析, 包括年湿日 (流域日平均降水量 $>1\text{mm}$) 总数、年最长湿期、年高降水日 (流域日平均降水量 $>10\text{mm}$) 总数和年最高日降水量。逐格点计算了测试期降水的年平均日降水序列, 计算上述四种指标, 并采用平均绝对误差 (MAE) 评估了各格点模拟降水与观测降水的指标偏差, 将 20 种 GCMs 的结果取平均值展示 (见图 7)。图 7(a) 显示, 4 种基于 GCMs 训练的深度学习降尺度年湿日总数 MAE 为 37.9-38.6, 表现优于 LS 方法的 51.6, 但差于 DBC 方法的 33.4, 这可能是因为 DBC 方法会考虑流域的干湿频率, 因此表现更好, 而 4 种 DL-DBC 的年湿日总数 MAE 为 32.3-32.6, 表现与 DBC 方法接近。图 7(b) 中, 4 种深度学习降尺度的年最长湿期 MAE 达到了 24.5-27.7 左右, 明显高于 LS 方法的 18.4 和 DBC 方法的 13.8, 这可能是因为当模型训练效果不够好时, 深度学习模型的模拟曲线通常相对平缓, 很容易高估湿期的长度, 对比之下, DL-DBC 年最长湿期 MAE 均值为 13.7-14.9, 相比基于 GCMs 训练的结果有明显改进, 这证明 DL-DBC 的模拟效果明显更好。图 7(c) 则表明, 4 种深度学习降尺度的年高降水日总数 MAE 为 5.5-6.7, 优于 LS 方法的 11.99, 但略差于 DBC 方法的 4.54。结合图 6 分析, 这可能是因为深度学习降尺度对高降水存在较明显的低估, 而 4 种 DL-DBC 的年高降水日总数 MAE 均值为 3.64-

3.77，表现比 DBC 方法更好，这可能是因为 DL-DBC 利用 ERA5 数据实现了较好的训练效果，因此在模拟高降水方面表现比仅使用 GCMs 数据训练的结果好很多，图 7(d)同样可以验证该结论。

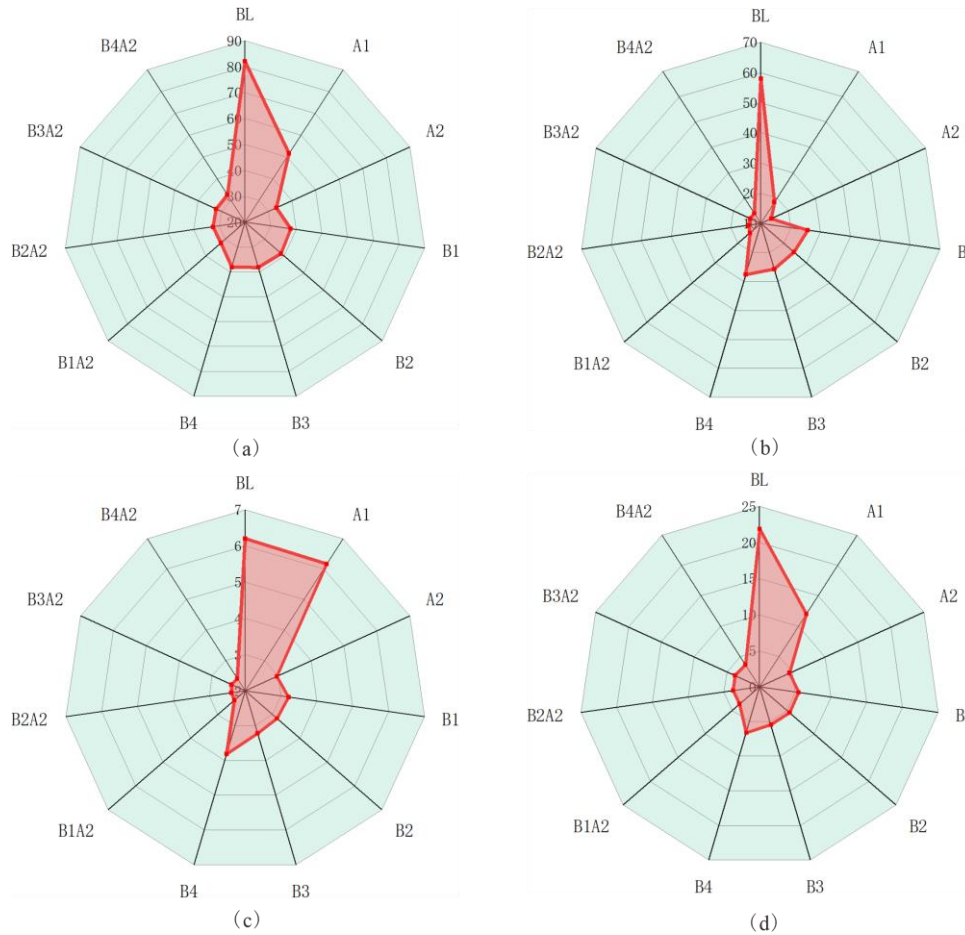


图 7.逐格点计算的模拟降水与观测降水的年湿日总数 (a)、年最长湿期 (b)，年高降水日总数 (c) 和年最高日降水量 (d)的 MAE 均值

Fig. 7 Mean Absolute Error (MAE) for grid cell-wise calculations of annual wet days count (a), annual longest wet spell (b), annual high precipitation days count (c), and annual maximum daily precipitation (d) between simulated and observed precipitation.

上述分析表明，在对 GCMs 日尺度降水进行降尺度时，直接利用 GCMs 数据和观测数据训练的 4 种深度学习降尺度方法在还原长江流域降水分布特征上的表现并不理想，问题主要表现在对湿期长度的高估以及高降水的低估，这是由于数据质量较差，且深度学习降尺度的模拟结果往往较为平缓导致的。混合方法 DL-DBC 有效的改善了深度学习降尺度的这些问题，并在部分指标上优于 DBC 方法，能够更有效的捕捉流域降水的分布特征。

2.3 时间偏差分析

研究计算了整个流域测试期模拟降水的多年平均月降水,并评估了模拟降水与观测降水的序列偏差(见图8)。结果表明,20种GCMs经DBC方法后的RMSE均值为11mm,MAE达到8.84mm,相比校正前偏差明显缩小,LS方法的RMSE和MAE表现均比DBC方法更差,但相比校正前仍有一定的改善,4种基于GCMs训练的深度学习的降尺度方法稍差于DBC方法,RMSE均值在13.5-18.2mm,MAE为11.3-15.9mm,混合方法DL-DBC的表现更好,RMSE稳定在10.5-10.8mm,MAE稳定在8.2-8.6mm。上述分析表明,相比仅使用GCMs训练的深度学习降尺度方法,DL-DBC降低流域平均降水的年均月降水偏差的效果明显提升,在对GCMs日降水数据进行降尺度时,能有效减少模拟降水与观测降水之间的偏差,提高模拟降水序列与观测值的相关性。

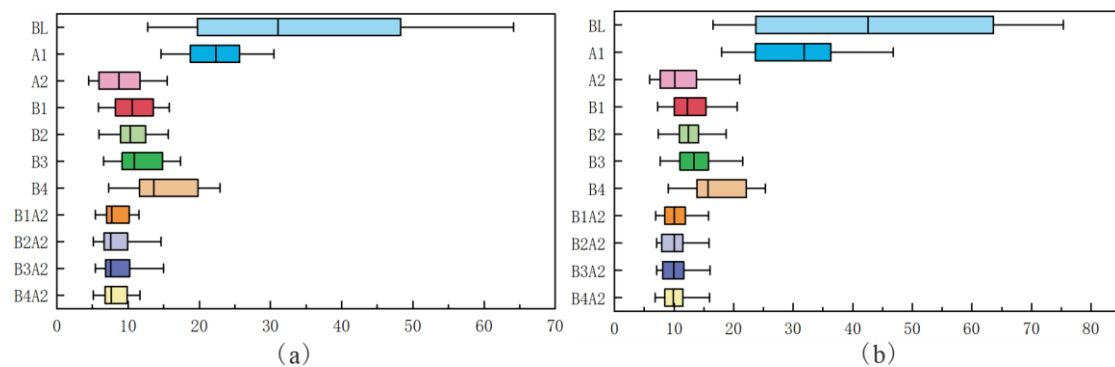


图8.所有GCMs模拟降水与观测降水的年均月降水序列的RMSE(a)和MAE(b)的统计结果

Fig. 8 Statistical results of Root Mean Square Error (RMSE) and MAE for the annual mean monthly precipitation sequences of observed and simulated precipitation from all GCMs (a, b).

2.4 空间偏差与偏倚结构

为了更好的分析降尺度方法在细节上模拟降水的准确性,研究分析了模拟降水与观测降水在空间偏倚结构上的差异。基于模拟降水与观测降水在流域各格点逐日计算的空间半边函数曲线,通过皮尔逊相关系数 r 与均方根误差(RMSE)衡量模拟降水与观测降水空间半边函数曲线的接近程度,如图9所示。

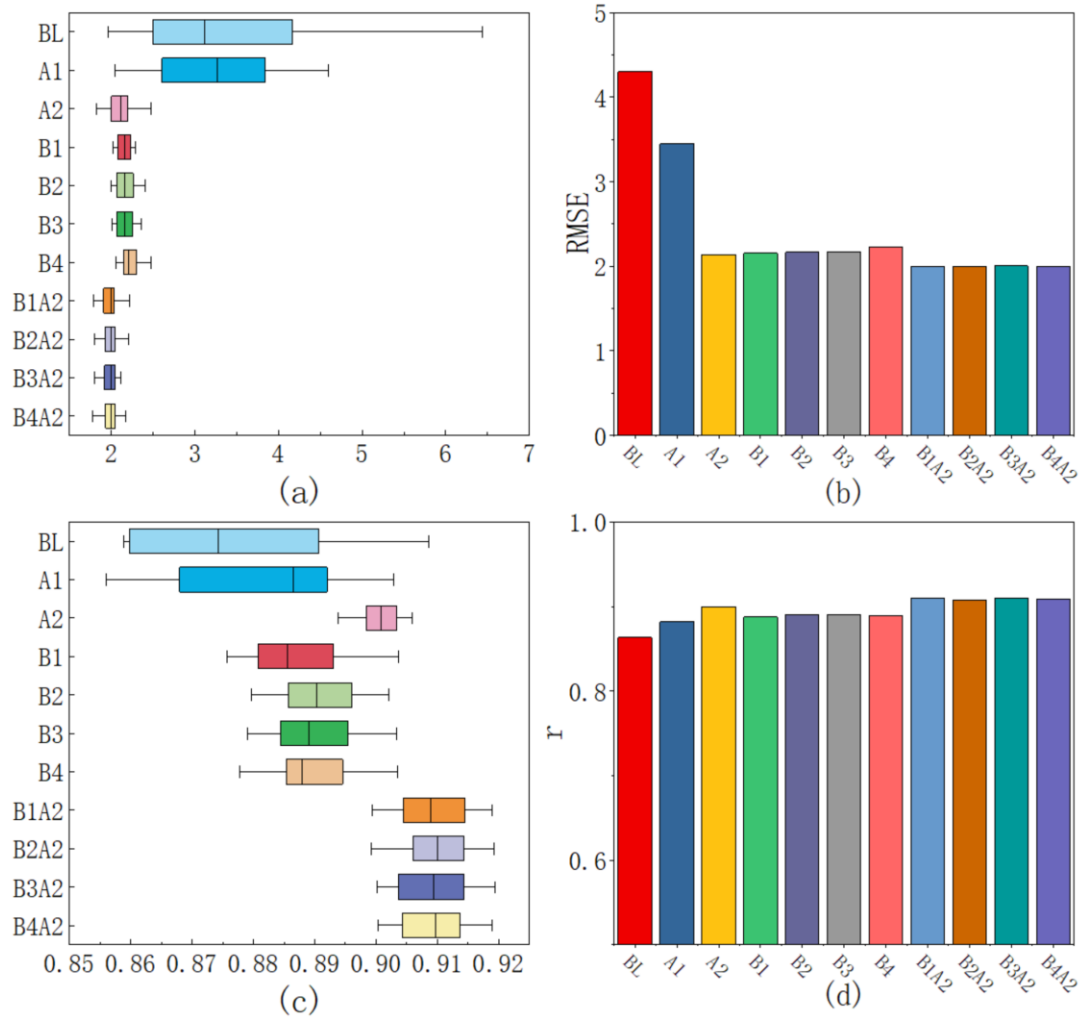


图 9. 20 种 GCMs 在测试期的模拟与观测年均日降水的空间半变函数图的 RMSE (a) 与 皮尔逊相关系数 r (c), (b)、(d) 为对 20 种 GCMs 测试的指标均值

Fig. 9 RMSE (a) and Pearson correlation coefficients r (c) of spatial semivariance for simulated and observed annual mean daily precipitation from 20 GCMs during the test period and mean values of these metrics across the 20 GCMs are presented in (b) and (d).

与 DBC 方法相比, 基于 GCMs 训练的 4 种深度学习降尺度方法的模拟降水的 RMSE 和皮尔逊相关系数略差, 但优于原始数据和 LS 方法后的结果。4 种 DL-DBC 方法表现比 DBC 方法更好且更加稳定, 在 RMSE 指标上, 四种 DL-DBC 的结果为 1.99-2.03, 相比四种基于 GCMs 训练的深度学习降尺度的结果下降了 0.16-0.22; 皮尔逊相关系数均值达到 0.91 左右, 相比基于 GCMs 训练的深度学习降尺度上升 0.02。整体来说, DL-DBC 在模拟降水的空间偏倚关系上表现优于 DBC 方法以及直接基于 GCMs 训练的深度学习降尺度模型。

表 3 是分不同季节展示了空间半边函数曲线的 RMSE 和皮尔逊相关系数。可以看到，相比深度学习降尺度方法，DL-DBC 对模拟降水的空间半边函数曲线优化最明显的是 6-8 月，两项指标均较好。对于降水较少的季节，如冬季和春季，DL-DBC 的皮尔逊相关系数 r 同样更高，但 RMSE 并不明显优于其他方法。考虑到长江流域降水主要集中在夏季，因此 DL-DBC 在模拟较高降水的空间偏倚结构的整体表现比 LS 方法、DBC 方法以及直接由 GCMs 训练的深度学习降尺度方法更好。

表 3 不同季节的空间半变函数图 RMSE 与皮尔逊相关系数 r

Table 3: RMSE and Pearson correlation coefficient r of spatial semivariance for different seasons.

	12-2 月		3-5 月		6-8 月		9-11 月	
	r	RMSE	r	RMSE	r	RMSE	r	RMSE
BL	0.836	0.842	0.877	5.238	0.852	8.133	0.889	3.191
A1	0.900	1.009	0.879	3.007	0.852	6.754	0.897	3.157
A2	0.918	0.669	0.912	2.233	0.863	4.022	0.905	1.709
B1	0.904	0.655	0.897	2.211	0.848	4.100	0.900	1.735
B2	0.908	0.661	0.903	2.253	0.847	4.092	0.904	1.735
B3	0.905	0.661	0.902	2.252	0.848	4.119	0.903	1.745
B4	0.905	0.650	0.903	2.178	0.848	4.322	0.900	1.823
B1A2	0.929	0.676	0.924	2.233	0.867	3.765	0.913	1.659
B2A2	0.930	0.674	0.925	2.214	0.867	3.785	0.916	1.664
B3A2	0.930	0.667	0.924	2.145	0.866	3.771	0.916	1.678
B4A2	0.930	0.670	0.924	2.237	0.866	3.789	0.917	1.676

3 结论

GCM 模拟降水的空间分辨率较低，且具有较大的偏差，在使用前通常需要进行降尺度。目前，基于深度学习模型的降尺度方法已成为该领域内一种新的方

向, 已有许多研究证明了基于深度学习的降尺度方法的可行性。然而, 对于带有较大时空偏差的粗分辨率数据 (GCMs 日尺度降水数据), 深度学习模型的降尺度效果仍有待研究。本研究以长江流域为研究对象, 基于 20 种 GCM 模拟的日降水数据, 评估了 4 种深度学习模型, 并与常用的 LS 方法和 DBC 方法进行了对比分析, 并提出了一种引入 ERA5 降水数据进行训练, 并由基于 ERA5 数据进行 DBC 方法后的 GCMs 数据进行模拟的混合方法 DL-DBC, 旨在结合 GCMs 和 ERA5 数据的优势, 提高模型效果。研究得到了以下几个结论:

1. 4 种基于 GCMs 训练的深度学习降尺度方法得到的降尺度结果的多年平均日降水 MARE 均值为 75%-81%, 表现优于 LS 方法与 DBC 方法, 多年平均降水及多年平均月降水的 MARE 均值分别是 7.1%-7.5% 和 33.1%-34.3%, 略差于 DBC 方法。混合方法 DL-DBC 提升了深度学习降尺度方法的模拟效果, 基于 4 种深度学习模型的 DL-DBC 的多年平均降水 MARE 相比深度学习降尺度降低 6.7%-11.3%, 多年平均月降水的 MARE 下降 6.3%-7.6%, 而多年平均日降水的 MARE 下降 4.3%-8.7%。

2. 4 种 DL-DBC 得到的流域平均 IOA 为 0.791-0.794, 略高于几种直接基于 GCMs 训练的深度学习方法以及 LS 方法和 DBC 方法, 主要改进表现在降水较高的长江下游地区, 这表明在降水较高的区域, DL-DBC 的模拟降水整体上更接近观测数据。

3. 基于 4 种深度学习降尺度的 DL-DBC 能够较好的捕捉降水频率分布特征, 有效改善了直接使用 GCMs 训练的深度学习降尺度方法对高降水的低估问题, 能够提高模拟降水量与观测降水量频率分布的一致性。

4. DL-DBC 的多年平均月降水的 RMSE 与 MAE 优于深度学习降尺度方法, 和 DBC 方法相比也更加稳定, 四种 DL-DBC 的 RMSE 均值相比 GCMs 训练的深度学习降尺度方法平均下降了 2.77-7.59mm, MAE 下降了 2.76-7.73mm。主要改善在于夏季, 这表明 DL-DBC 能够更稳定的模拟降水的年内分配情况。

5. 基于 GCMs 训练的深度学习降尺度方法的降水空间半边函数曲线的皮尔逊相关系数与 RMSE 略优于 LS 方法, 与 DBC 方法接近, DL-DBC 的效果更好, 皮尔逊相关系数达到 0.903-0.911, 而 RMSE 指标为 1.99-2.03, 与基于 GCMs 训练的深度学习降尺度方法相比, DL-DBC 的提升主要体现在降水较高的季节。

尽管过去的深度学习降尺度研究中，更高级、深层的网络往往有利于提取更高级的信息，但在本研究分析中，几种深度学习架构对长江流域 GCMs 日尺度的降水降尺度效果差异不大，这可能是因为 GCMs 日尺度降水数据本身巨大的时空偏差，在训练时模拟数据与观测数据无法进行有效的对应，因此使用更深的网络也难以更有效的提取数据特征，同样的，基于 4 种不同深度学习架构的 DL-DBC 之间也没有显著的效果差异，对于日尺度降水的降尺度，更复杂的网络模型是否能够展现出显著的优势仍有待进一步的测试。

References

- Baño-Medina, J., Manzananas, R., et al. (2020). Configuration and intercomparison of deep learning neural models for statistical downscaling. *13*(4), 2109-2124.
- Bennett, J., Grose, M., et al. (2011). *Performance of quantile-quantile bias-correction for use in hydroclimatological projections*. Paper presented at the MODSIM2011 International Congress on Modelling and Simulation, Perth.
- Cannon, A. J. (2017). Multivariate quantile mapping bias correction: an N-dimensional probability density function transform for climate model simulations of multiple variables. *Climate Dynamics*, *50*(1-2), 31-49. doi:10.1007/s00382-017-3580-6
- Chen, J., Brissette, F. P., et al. (2013). Performance and uncertainty evaluation of empirical downscaling methods in quantifying the climate change impacts on hydrology over two North American river basins. *479*, 200-214.
- Chen, J., Xu, C., et al. (2016). Progress and challenge in statistically downscaling climate model outputs. *5*(4), 299-313.
- Dickinson, R. E., Errico, R. M., et al. (1989). A regional climate model for the western United States. *15*, 383-422.
- Diederik, P. K. J. (2014). Adam: A method for stochastic optimization.
- Eden, J. M., & Widmann, M. J. J. o. C. (2014). Downscaling of GCM-simulated precipitation using model output statistics. *27*(1), 312-324.
- Eum, H. I., & Cannon, A. J. J. I. J. o. C. (2017). Intercomparison of projected changes in climate extremes for South Korea: application of trend preserving statistical downscaling methods to the CMIP5 ensemble. *37*(8), 3381-3397.
- Guo, Q., Chen, J., et al. (2019). A new two-stage multivariate quantile mapping method for bias correcting climate model outputs. *53*, 3603-3623.
- Harris, L., McRae, A. T., et al. (2022). A generative deep learning approach to stochastic downscaling of precipitation forecasts. *14*(10), e2022MS003120.
- He, X., Chaney, N. W., et al. (2016). Spatial downscaling of precipitation using adaptable random forests. *52*(10), 8217-8237.
- Huang, G., Liu, Z., et al. (2017). *Densely connected convolutional networks*. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
- Legasa, M., Manzananas, R., et al. (2022). A posteriori random forests for stochastic downscaling of precipitation by predicting probability distributions. *58*(4), e2021WR030272.
- Li, W., Pan, B., et al. (2022). Convolutional neural network-based statistical post-processing of ensemble precipitation forecasts. *605*, 127301.
- Liu, Y., Ganguly, A. R., et al. (2020). *Climate downscaling using YNet: A deep convolutional network with skip connections and fusion*. Paper presented at the Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining.
- Liu, Z., Mao, H., et al. (2022). *A convnet for the 2020s*. Paper presented at the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition.
- Mearns, L. O., Arritt, R., et al. (2012). The North American Regional Climate Change Assessment Program: Overview of Phase I Results. *Bulletin of the American Meteorological Society*, *93*(9), 1337-1362. doi:10.1175/bams-d-11-00223.1
- Meehl, G. A., Covey, C., et al. (2007). The WCRP CMIP3 multimodel dataset: A new era in climate

- change research. *88*(9), 1383-1394.
- Mei, Y., Maggioni, V., et al. (2020). A nonparametric statistical technique for spatial downscaling of precipitation over High Mountain Asia. *56*(11), e2020WR027472.
- Monjo, R., Gaitán, E., et al. (2016). Changes in extreme precipitation over Spain using statistical downscaling of CMIP5 projections. *36*(2), 757-769.
- Pan, B., Anderson, G. J., et al. (2021). Learning to correct climate projection biases. *13*(10), e2021MS002509.
- Pan, B., Hsu, K., et al. (2019). Improving precipitation estimation using convolutional neural network. *55*(3), 2301-2321.
- Pierce, D. W., Cayan, D. R., et al. (2015). Improved bias correction techniques for hydrological simulations of climate change. *16*(6), 2421-2442.
- Pour, S. H., Shahid, S., et al. (2016). A hybrid model for statistical downscaling of daily rainfall. *154*, 1424-1430.
- Rodrigues, E. R., Oliveira, I., et al. (2018). *DeepDownscale: A deep learning strategy for high-resolution weather forecast*. Paper presented at the 2018 IEEE 14th International Conference on e-Science (e-Science).
- Schoof, J. T., & Pryor, S. J. I. J. o. C. A. J. o. t. R. M. S. (2001). Downscaling temperature and precipitation: A comparison of regression-based methods and artificial neural networks. *21*(7), 773-790.
- Sharma, S. C. M., Kumar, B., et al. (2024). Deep learning-based bias correction of ISMR simulated by GCM. 107589.
- Sillmann, J., Kharin, V., et al. (2013). Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate. *118*(4), 1716-1733.
- Sun, L., & Lan, Y. J. I. J. o. C. (2021). Statistical downscaling of daily temperature and precipitation over China using deep learning neural models: Localization and comparison with other methods. *41*(2), 1128-1147.
- Taylor, K. E., Stouffer, R. J., et al. (2012). An overview of CMIP5 and the experiment design. *93*(4), 485-498.
- Wang, F., Tian, D., et al. (2021). Deep learning for daily precipitation and temperature downscaling. *57*(4), e2020WR029308.
- Wang, F., & Tian, D. J. C. d. (2022). On deep learning-based bias correction and downscaling of multiple climate models simulations. *59*(11), 3451-3468.
- Wang, L., & Chen, W. J. I. J. o. C. (2014). A CMIP5 multimodel projection of future temperature, precipitation, and climatological drought in China. *34*(6).
- Wilby, R. L., Wigley, T., et al. (1998). Statistical downscaling of general circulation model output: A comparison of methods. *34*(11), 2995-3008.
- Wood, A. W., Leung, L. R., et al. (2004). Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *62*, 189-216.
- Wu, J., & GAO, X.-J. J. C. J. o. G. (2013). A gridded daily observation dataset over China region and comparison with the other datasets. *56*(4), 1102-1111.
- Yang, W., Andréasson, J., et al. (2010). Distribution-based scaling to improve usability of regional climate model projections for hydrological climate change impacts studies. *41*(3-4), 211-229.
- 刘永和, 郭维栋, et al. (2011). 气象资料的统计降尺度方法综述. *26*(8), 837.

刘杰, 陈前, 许妍, 等. 长江流域洞庭湖区出入湖磷通量模拟及水质预测: 机器学习与传统水文模型耦合方法[J]. 地球科学, 2024, 49(11): 3995-4007. doi: 10.3799/dqkx.2024.061

张琼, & 气象学报, 吴. J. (2013). 长江流域大范围旱涝与南亚高压的关系. 59(5), 569-577.

张仲石, 李双林, 王会军, 等. 浅谈大气科学与地质学的学科交叉[J]. 地球科学, 2022, 47(10): 3569-3579. doi: 10.3799/dqkx.2022.350