DOI：10. 11714/acta. snus. ZR20260082

粤港澳大湾区短时降水CNN概率降尺度*

蓝玉峰[1]，郑浩阳[1]，张生奇[1]，程建培[1]，孙磊[1]，嵇萍[2]，邱国强[3]

1. 珠海市公共气象服务中心，广东 珠海 519082
2. 珠海市气象局，广东 珠海 519082
3. 珠海市国家气候观象台，广东 珠海 519082

摘　要：精细尺度短时降水预报是粤港澳地区防灾减灾的重要支撑。本研究利用卷积神经网络（CNN，convolutional neural network）构建短时降水概率降尺度预报模型，该模型以GFS（global forecast system）模式气象背景场预报资料作为输入信息，获得区域内各站点3 h降水预报的概率分布参数，并评价不同试验的预报性能。结果表明：1）确定性预报结果中，各试验均不同程度地降低了GFS原始预报误差，CNN-PR试验进一步强化了空间特征，预报效果最好（RMSE=3.31 mm）；SINGLE试验（未引入全局信息）和CNN0试验（未进行空间特征提取）表现较差；2）概率预报结果中，各试验总体校准性能较好，CONTROL试验的综合评分最优（CRPS=0.69 mm）；除SINGLE和CNN0试验外，其余试验对降水发生事件均有一定的区分能力，其中纳入时序演变信息的TREND试验对强降水的区分能力最优；CNN-PR试验的区间条件校准性最优，但以牺牲预报锐度为代价；3）各试验对强降水的预报结果偏小，概率预报水平受“整体−区间校准”“校准−锐度”2组内在矛盾的限制。本研究可为粤港澳大湾区短时降水精细化概率预报的业务化应用提供方法参考。

关键词：卷积神经网络；降尺度；大湾区；短时降水

中图分类号：P456　　**文献标志码：**A　　**文章编号：**2097－0137（XXXX）XX－0001－11

Probabilistic downscaling of short-range precipitation forecasting using CNN in the Guangdong-Hong Kong-Macao Greater Bay Area

Lan Yufeng[1], *Zheng Haoyang*[1], *Zhang Shengqi*[1], *Cheng Jianpei*[1], *Sun Lei*[1], *Ji Ping*[2], *Qiu Guoqiang*[3]

1. Zhuhai Public Meteorological Service Center, Zhuhai 519082, China
2. Zhuhai Meteorological Bureau, Zhuhai 519082, China
3. Zhuhai National Climate Observatory, Zhuhai 519082, China

Abstract: High-resolution short-range precipitation forecasting is critical for regional disaster prevention and mitigation over the Guangdong-Hong Kong-Macao Greater Bay Area (GBA). This study developed a probabilistic downscaling model for short-range precipitation using a Convolutional Neural Network (CNN), which takes large-scale meteorological forecast information from GFS (global forecast system) as input and outputs the parameters of the probabilistic distribution for 3 h precipitation at multiple stations. The model skill was evaluated through a set of carefully designed experiments. The results show that: 1) All experiments effectively reduce forecasting errors compared with the raw GFS outputs. Among them, the CNN-PR experiment, which incorporates enhanced spatial feature extraction, achieves the best deterministic performance (RMSE=3.31 mm). In contrast,

* **收稿日期：**2026－04－03　　**录用日期：**2026－09－08　　**网络首发日期：**XXXX－XX－XX
基金项目：广东省气象局科技项目(GRMC2024M40)；广东省海洋气象科学数据中心项目(2024B1212070014)
作者简介：蓝玉峰（1990年生），女；**研究方向：**大气科学；E-mail：lyfnuist@163.com
通信作者：孙磊（1992年生），男；**研究方向：**大气科学；E-mail：sunleimodel@gmail.com

全文阅读
ZR20260082

the SINGLE experiment, which relies solely on local information, and the CNN0 experiment, which omits spatial feature extraction, show inferior performance. 2) In terms of probabilistic skill, all models are well calibrated overall, with the CONTROL experiment achieving the optimal skill (CRPS= 0.69 mm). All experiments except SINGLE and CNN0 show improved discrimination ability for precipitation occurrence, and the TREND experiment incorporating temporal evolution information is optimal for discriminating heavy rainfall events. The CNN-PR experiment is superior in conditional calibration, albeit at the cost of reduced sharpness. 3) The study also reveals that all experiments systematically underestimate heavy rainfall events. Furthermore, the probabilistic forecasting skill is constrained by two key trade-offs: one between calibration and sharpness, and the other between overall and conditional calibration. This study provides a methodological reference for operational high-resolution short-range probabilistic precipitation forecasting in the GBA.

Key words: convolutional neural networks (CNN); statistical downscaling; Guangdong-Hong Kong-Macao Greater Bay Area; short-range precipitation

粤港澳大湾区（简称“大湾区”）的人口和经济活动高度聚集（王垚等，2024；Wang et al.，2021）。受全球气候变暖及区域地理气候环境等因素影响，大湾区极端降雨频发，易引发严重的次生灾害，威胁区域生命财产安全，制约高质量发展，故对精细化短时降水预报产品具有较高需求（李卓凡等，2023；Deng et al.，2023）。全球数值模式是降水预报最主要的参考产品，其预报性能持续提升（张进等，2023），但受到自身空间分辨率限制，对强降水时空特征的预报能力较弱（Liang et al.，2012；王蕾等，2024；张博等，2024；张欣等，2026）。因此，基于全球模式进行动力或统计降尺度，是获取精细化降水预报产品的重要途径。

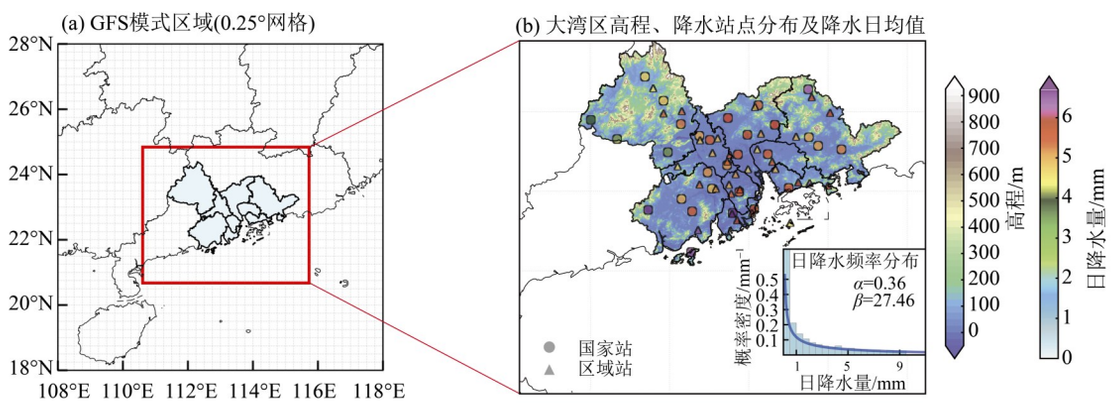
动力降尺度模式，如WRF（weather research and forecasting），在大尺度预报场的侧边界强迫作用下，通过完备的动力框架和尺度适配的物理参数化方案，输出高分辨率格点气象预报产品，有效提升各类气象要素的预报准确度（Xue et al.，2014；Liang et al.，2019；董晓云等，2020）。但动力降尺度对算力和存储要求较高，部署复杂、运行时间较长，限制了其在预报业务中的快速应用。统计降尺度通过构建预报因子与实际观测之间的统计关系实现高分辨率预报，常用方法包括模式输出统计（MOS，model output statistics method）和完全预报（PP，perfect prediction）（范丽军等，2005；Wilks，2011）。其中，MOS类方法主要利用模式输出建立统计订正关系，而PP类方法则以大尺度背景场作为预报因子，通过建立与实际观测之间的统计关系进行降尺度。统计降尺度计算开销小、运行速度快，已在预报业务中得到广泛应用，但该类方法对历史样本的代表性和预报因子的有效性要求较高，统计关系的稳定性也会影响降尺度效果。降水受多种气象环境要素非线性相互作用的影响，其时空分布呈现非正态特征（Vrac et al.，2007）。相较于确定性预报，概率预报能够描述降水预报的不确定性。在此方面，MOS类方法主要面向确定性预报，难以适配概率预报；而PP类方法可通过拟合概率分布参数开展概率预报，具有一定的技术优势（Sloughter et al.，2007；张亚萍等，2022）。

随着神经网络的兴起和发展，统计降尺度方法的拟合及推断能力进一步提升。尤其是深度学习中的卷积神经网络（CNN，convolutional neural network）具有层次化的特征提取结构，可有效提取和表达规则网格数据的空间关联特征，适用于统计预报、统计降尺度、模式订正等多个地球科学领域（Reichstein et al.，2019；Baño-Medina et al.，2020；孙健等，2021；杨绚等，2022；Sun et al.，2023）。已有研究采用CNN对我国日降水进行统计降尺度分析，并取得了较好的效果（Sun et al.，2021）；但该研究的分析对象为日尺度气象再分析资料，且未评估概率预报质量。鉴于此，本研究采用CNN构建了大湾区短时降水概率降尺度模型，模型利用卷积方法提取大尺度气象预报信息，以3 h站点降水实况为标签，通过优化定制的损失函数对模型进行训练，一次性即可得到各站点3 h降水预报概率分布的参数值。此外，本研究对比评估了模型的确定性预报和概率预报性能，可为其业务化应用提供一定的参考依据。

1　数据与方法

1.1　研究数据

本研究所选用的大尺度天气预报场数据来自美国国家环境预报中心（NCEP，National Centers for Environmental Prediction）开发的全球预报系统（GFS，global forecast system）。该数据包含 5 个气压层（925、850、700、500 和 200 hPa）的水平纬向风 U（m/s）、水平经向风 V（m/s）、位势高度 Z（gpm）、温度 T（K）和相对湿度 RH（%）。预报场的空间区域为覆盖大湾区的 41×41 个水平网格点（分辨率为 0.25°×0.25°）（图 1a），预报时效为 2 d（每日 20:00 起报），预报时间分辨率为 3 h，采用时间段为 2016—2018 年和 2021—2023 年。小时降水站点实况来自大湾区的 61 个站点，包括 29 个国家基本气象站和 32 个区域自动站（图 1b），时间段与 GFS 一致。所有站点的观测均作为标签数据用于模型训练，模型评估基于国家基本气象站。此外，本研究还利用 GEBCO（general bathymetric chart of the oceans）地形数据（线性插值到模式网格）探究了地形因素的影响。



a：实线为模式网格（0.25°×0.25°，41×41 个）；b：α 和 β 分别为日降水 Gamma 拟合分布的形状和尺度参数，蓝色实线为拟合曲线；审图号为 GS（2024）0650 号，底图边界无修改。

图 1　GFS 模式区域、地形、气象站点及降水统计

Fig. 1　GFS model domain, topographic elevation and used meteorological stations as well as precipitation statistics

1.2　模型概述

本研究使用 Bernoulli-Gamma 分布（Cannon，2008；Sun et al.，2021）以表征降水离散-连续的统计特性，其表达式如下：

$$f(y;\ p,\alpha,\beta)=\begin{cases}1-p, & y=0,\\ \dfrac{p\left(y/\beta\right)^{\alpha-1}\exp\left(-y/\beta\right)}{\beta\Gamma(\alpha)}, & y>0,\end{cases}\tag{1}$$

其中 y 为 3 h 降水量，p 为降水发生概率，α 和 β 分别为 Gamma 分布的形状和尺度参数，Γ 为 Gamma 函数。确定性降水预报则为分布的数学期望 $E(y)=p\times\alpha\times\beta$。

构建的深度学习降尺度模型 f 的表达式如下：

$$\boldsymbol{Y}_T=f\left(\boldsymbol{X}_{T-3}\right)=g(\boldsymbol{F}),\tag{2}$$

$$\boldsymbol{F}=\sigma\left(\boldsymbol{K}*\boldsymbol{X}_{T-3}+\boldsymbol{b}\right),\tag{3}$$

$$L=N_{\text{Bernoulli}}+N_{\text{Gamma}},\tag{4}$$

其中 $\boldsymbol{Y}_T=\left(p_{s,T},\ \alpha_{s,T},\ \beta_{s,T}\right)_{s=1}^{61}$ 为模型在预报时刻 T 输出的向量，$s=1,2,\cdots,61$ 为站点索引，$\boldsymbol{X}_{T-3}\in\mathbb{R}^{41\times41\times5\times5}$ 为预报时刻 T-3 输入的大尺度张量，维度分别为 41×41 个格点、5 个气压层（925、850、700、500 和 200 hPa）和 5 个变量（U、V、Z、T 和 RH）。$\boldsymbol{F}$ 为空间特征，g 为空间特征到 61 个站点分布参数的全连接映射，$*$ 为二维卷积操作，$\boldsymbol{K}$ 为卷积核，$\boldsymbol{b}$ 为偏置项，σ 为非线性激活函数。最后，N 为参数分布的负对数似然，模型损失函数 L 由 Bernoulli 分布的负对数似然项（$N_{\text{Bernoulli}}$）和 Gamma 分布的负对数似然项（N_{Gamma}）组成。

1.3　试验设计

为了评估模型的降尺度能力，本研究基于控制变量的思路设置了 6 组试验（表 1）。模型基于 PyTorch 框架建立，使用 Adam 作为优化器，学习率取 1×10^{-4}，设置 Dropout 失活率为 0.2。2016—2018 年的数据按照 8:2 的比例划分为训练数据和验证数据，进一步根据训练数据的统计量将大尺度背景逐格点标准化后作为模型的输入。6 组试验的最多训练次数为 500 轮，模型在验证集上的评价连续 50 轮无改进则停止。最后，训练好的模型利用 2021—2023 年数据做测试，测试结果作为模型性能评估的依据。

表1 试验设计概况

Table 1 Overview of the designed experiments

试验名称	核心架构	输入层	输出层	说明
CONTROL	1）卷积层：3×3卷积核，64个通道 2）最大池化层：2×2窗口 3）卷积层：3×3卷积核，128个通道 4）展平+时效特征融合	1）*T*-3时刻的25个特征：41×41个格点，5层5个气象要素 2）预报时效*T*	全连接61个站预报时刻*T*的3 h降水Bernoulli-Gamma分布的3个参数	控制试验，提供基准参照
TOPO	同CONTROL	1）*T*-3时刻的25个特征：41×41个格点，5层5个气象要素 2）地形静态特征：41×41个格点 3）预报时效*T*	同CONTROL	测试地形因子影响
TREND	同CONTROL	1）*T*-6，*T*-3和*T*共3个刻的75个特征：41×41个格点，5层5个气象要素 2）预报时效*T*	同CONTROL	测试时间趋势影响
CNN0	1）输入展平 2）全连接层：50个神经元 3）全连接层：50个神经元 4）时效特征融合	同CONTROL	同CONTROL	测试卷积操作的必要性
CNN-PR	1）卷积层：3×3卷积核，10个通道 2）卷积层：3×3卷积核，25个通道 3）卷积层：3×3卷积核，50个通道 4）展平+时效特征融合	同CONTROL	同CONTROL	测试强化空间特征提取能力的影响
SINGLE	1）1D卷积层：1×1卷积核，64个通道 2）1D卷积层：1×1卷积核，128个通道 3）展平+时效特征融合	1）*T*-3时刻的25个特征：站点插值，5层5个气象要素 2）预报时效*T*	全连接层输出单站预报时刻*T*的3 h降水Bernoulli-Gamma分布的3个参数	测试考虑全局空间信息的必要性

1.4 模型评价

1.4.1 确定性预报评价指标 本研究使用平均绝对误差（MAE）、均方根误差（RMSE）和相关系数（*R*）对模型的确定性预报能力进行评估，各指标表达式为：

$$\mathrm{MAE}=\frac{1}{N}\sum_{i=1}^{N}\left|f_i-o_i\right|, \tag{5}$$

$$\mathrm{RMSE}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(f_i-o_i\right)^2}, \tag{6}$$

$$R=\frac{\sum_{i=1}^{N}\left(f_i-\bar{f}\right)\left(o_i-\bar{o}\right)}{\sqrt{\sum_{i=1}^{N}\left(f_i-\bar{f}\right)^2}\sqrt{\sum_{i=1}^{N}\left(o_i-\bar{o}\right)^2}}, \tag{7}$$

其中N为总样本数，o为站点观测值，f为确定性预报值（或概率预报的数学期望）。此外，降水事件的预报技巧则利用ETS评分进行评估，其公式为：

$$\mathrm{ETS}=\frac{a-a_{\mathrm{ref}}}{a+b+c-a_{\mathrm{ref}}}, \tag{8}$$

其中a、b、c分别为命中样本数、空报样本数和漏报

样本数，$a_{\mathrm{ref}}=(a+b)(a+c)/N$ 为随机期望命中数。

1.4.2　概率预报评价指标　本研究使用连续分级概率评分（CRPS，continuous ranked probability score）对概率预报的综合性能进行评估，其值越小表示整体概率预报性能越优（理论完最优值为 0），表达式为：

$$\mathrm{CRPS}=\frac{1}{N}\sum_{i=1}^{N}\int_{-\infty}^{+\infty}\left[F_i(x)-\mathbb{I}(x\geqslant o_i)\right]^2\mathrm{d}x,\quad(9)$$

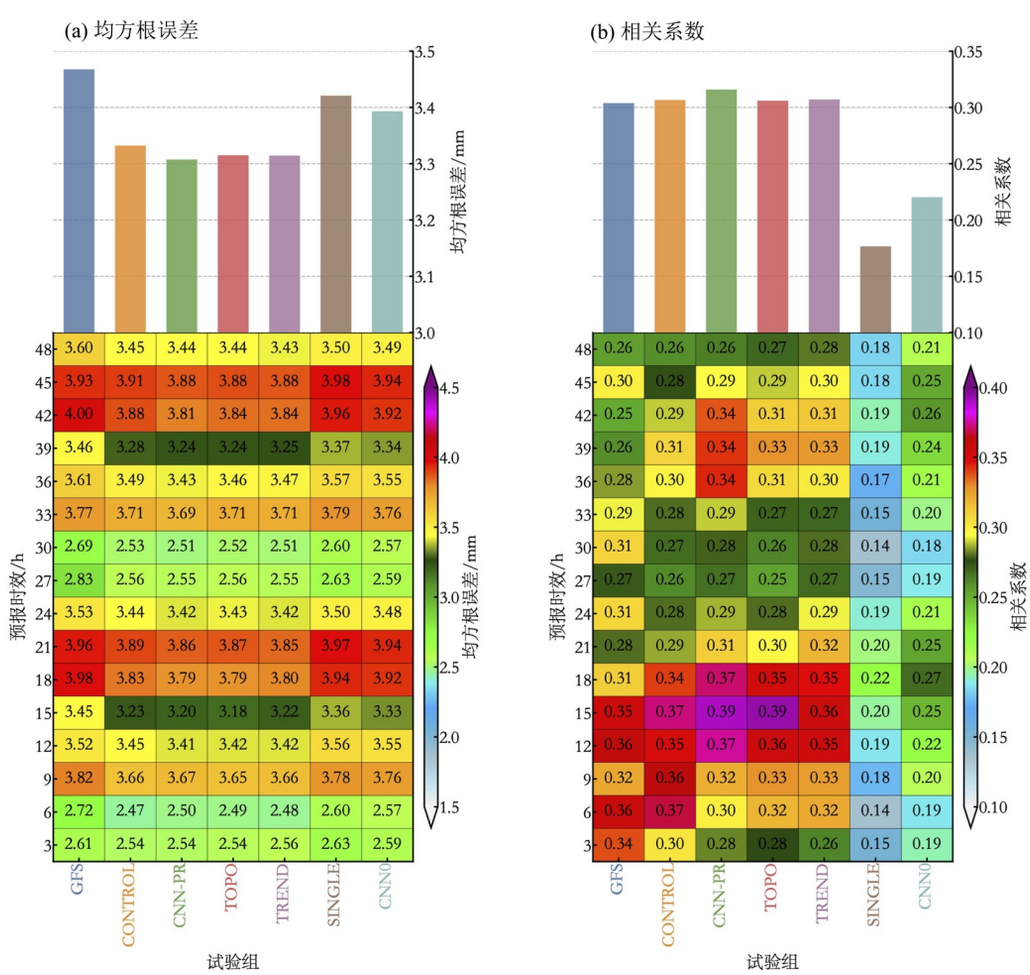
其中 F 为累积分布函数，$\mathbb{I}$ 为指示函数。

概率预报分布的整体校准度用概率积分变换（PIT，probability integral transform）直方图来表征并通过 Kolmogorov-Smirnov（KS）距离度量，PIT 越接近均匀分布，KS 距离越接近 0，表明校准度越高。此外，还使用 ROC 曲线来度量模型对特定降水事件的分辨能力，其曲线下面积（AUC）越接近 1，表明分辨能力越强；同时采用可靠性曲线评估概率预报的校准度，曲线越贴近对角线，表明预报概率与实际频率的一致性越好。最后，利用 80% 预测区间（10%~90% 的分位数）的观测覆盖率及其平均宽度评估区间校准度及其锐度；在区间覆盖率接近目标（即 80%）的情况下，区间宽度越窄，表明预报的锐度越好。

2　确定性预报能力评估

2.1　整体误差与相关性

图 2 比较了不同预报时效下 GFS 和各试验的 3 h 降水确定性预报的 RMSE 和 R。从观测来看，大湾区降水日变化呈现明显的双峰特征，主峰在午后至傍晚，次峰在凌晨至上午（Yu et al.，2007；李丽平等，2010；Li et al.，2021；Dong et al.，2023）。受预报时效和昼夜变化的影响，GFS 和各试验的降水预报偏差并非随着预报时效的增加而单调增大，其变化具有与降水观测昼夜变化相似的波动特征。与 GFS 相比（RMSE=3.47 mm，R=0.30），各试验均减小了预



热力图为各指标站点平均值，柱状图为各指标站点和预报时效平均值。

图 2　GFS 和各试验的均方根误差及相关系数对比

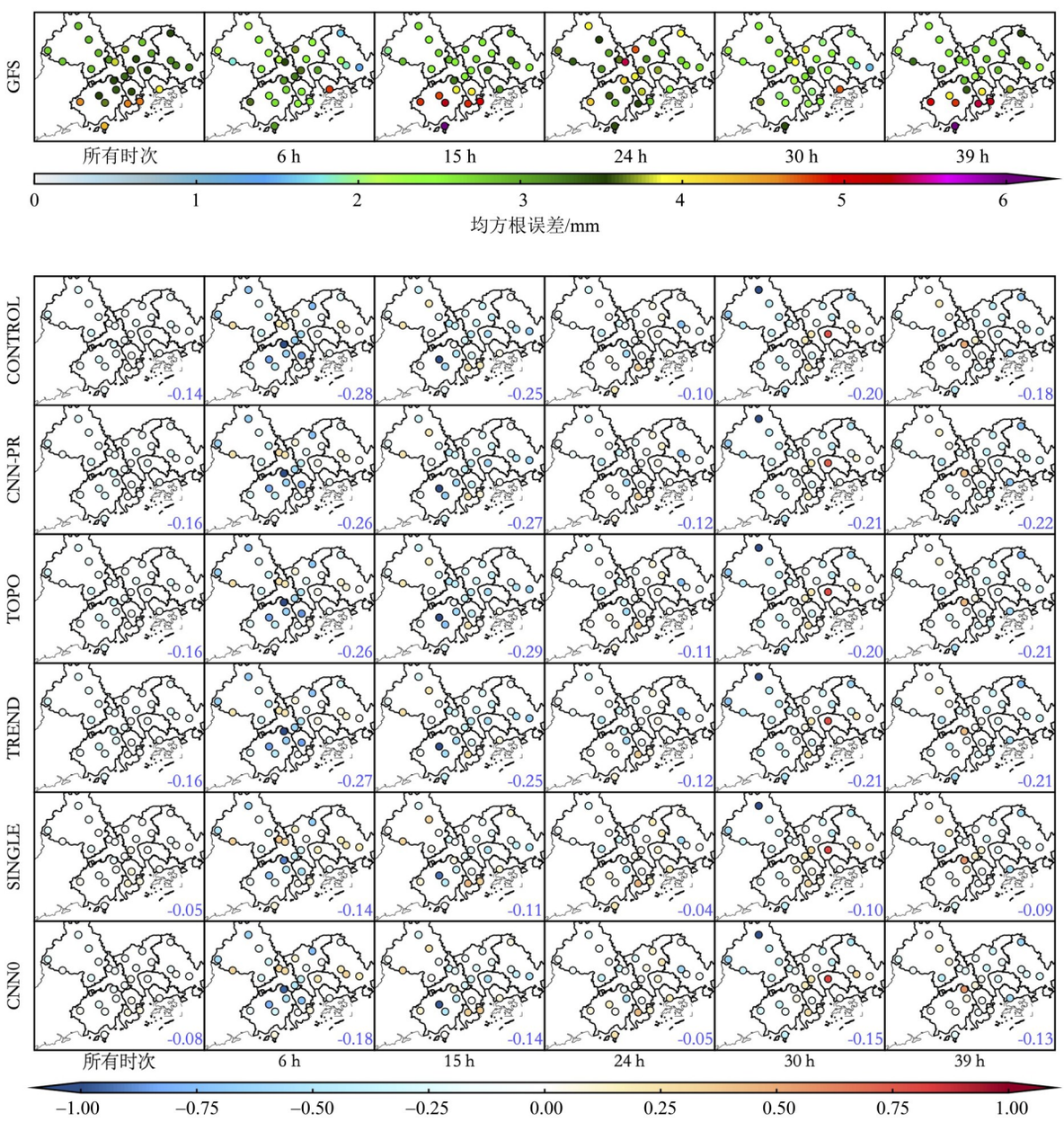
Fig. 2　Comparison of RMSE and R between GFS model and each experiment

报误差。各试验对相关性的改善较为有限，前 2 个预报时刻的相关性均低于 GFS，其中 CNN-PR 和 TREND 在后续多个预报时刻的相关性略高于 GFS，而 SINGLE 和 CNN0 的表现则存在明显局限性。

对比各试验后发现，CNN-PR 的预报性能最好，整体预报误差最小（RMSE=3.31 mm），相关系数也最大（R=0.32），说明渐进式卷积结构有利于提取更多的空间特征信息。TOPO 试验整体预报性能与 CONTROL 相当，表明增加静态地形特征信息对整个模型的预报性能无明显影响。此外，SINGLE 和 CNN0 与其他试验的差距较大，前者仅使用局地信息，后者不包含卷积结构。上述结果与已有研究结果一致，表明利用卷积提取全局空间特征有利于改善降尺度性能（Sun et al.，2021，2023）。

2.2 站点差异特征

不同预报时刻 GFS 预报的误差空间模态与降水实况均值分布相似，间隔 24 h 的 2 个预报时效的结果相近（图 3）。这表明，在降水量级及其自身时间变率的主导下，降水预报误差的时空分布特征具



蓝色数值为差异站点的平均值，审图号为 GS（2024）0650 号，底图边界无修改。

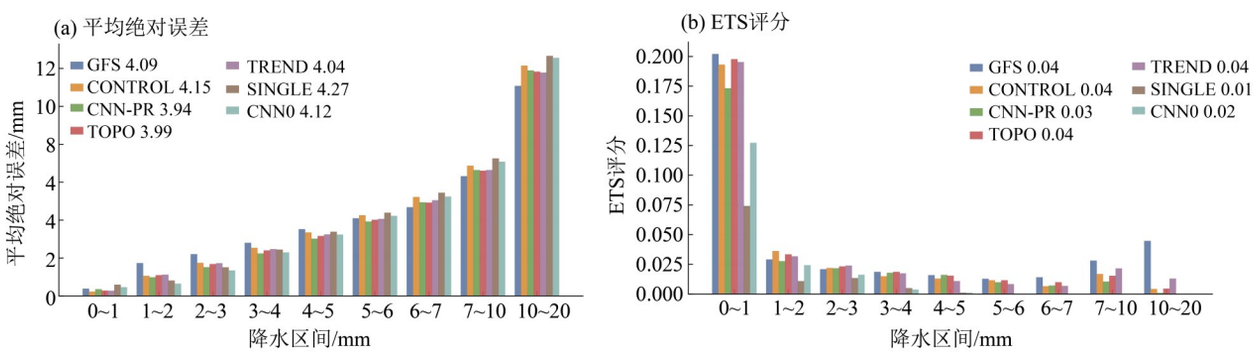
图 3 GFS 均方根误差的空间分布及其与各试验的差异

Fig. 3 Spatial distribution of GFS RMSE and its differences from each experiment

备固有的稳定性，并受预报时效影响。CNN-PR 效果仍为最优，相较于 GFS 提升了大部分站点的预报技巧；但受其自身泛化能力的制约，以及大湾区降水极端事件多、空间差异性大（Wang et al.，2021；李广宗等，2025）的影响，模型预报性能存在较显著的时空差异。CONTROL、TREND 和 TOPO 试验结果相近，但与 CNN-PR 有一定差距。SINGLE 和 CNN0 表现最差，订正能力有限。各试验的配置和性能虽不同，但预报误差的空间分布特征几乎一致，这主要归因于其受到统一的外部大尺度背景和内部损失函数制约。

2.3　区间差异特征

图 4 展示了不同降水区间内 GFS 和各试验 3 h 降水确定性预报的 MAE、ETS 值。在 1~5 mm 各降水区间内，各试验组的 GFS 降水预报误差均下降；在 5~6 mm 降水区间内，各试验组的平均预报性能与 GFS 相近；而在 6 mm 以上降水区间，各试验组的预报误差高于 GFS，且 GFS 的 ETS 评分也高于各试验组。这主要因为较高降水区间的样本占比较低，而降尺度模型以最小化整体误差为优化目标。因此，各试验均会出现对强降水低估的特点，也揭示了物理数值模式在强降水预报方面的优势（孙磊等，2024；Sun et al.，2024）。不同试验对比表明，CNN-PR 仍整体表现最佳，其 ETS 评分虽略低于 CONTROL 等试验，但在 2 mm 以上多个区间均取得最小预报误差。



柱条为各指标站点和预报时效的平均值，图例中数字为整体均值。

图 4　GFS 和各试验的平均绝对误差和 ETS 评分在不同观测降水区间的对比

Fig. 4　Comparison of MAE and ETS between the GFS model and each experiment across different observed precipitation intervals

3　概率预报能力评估

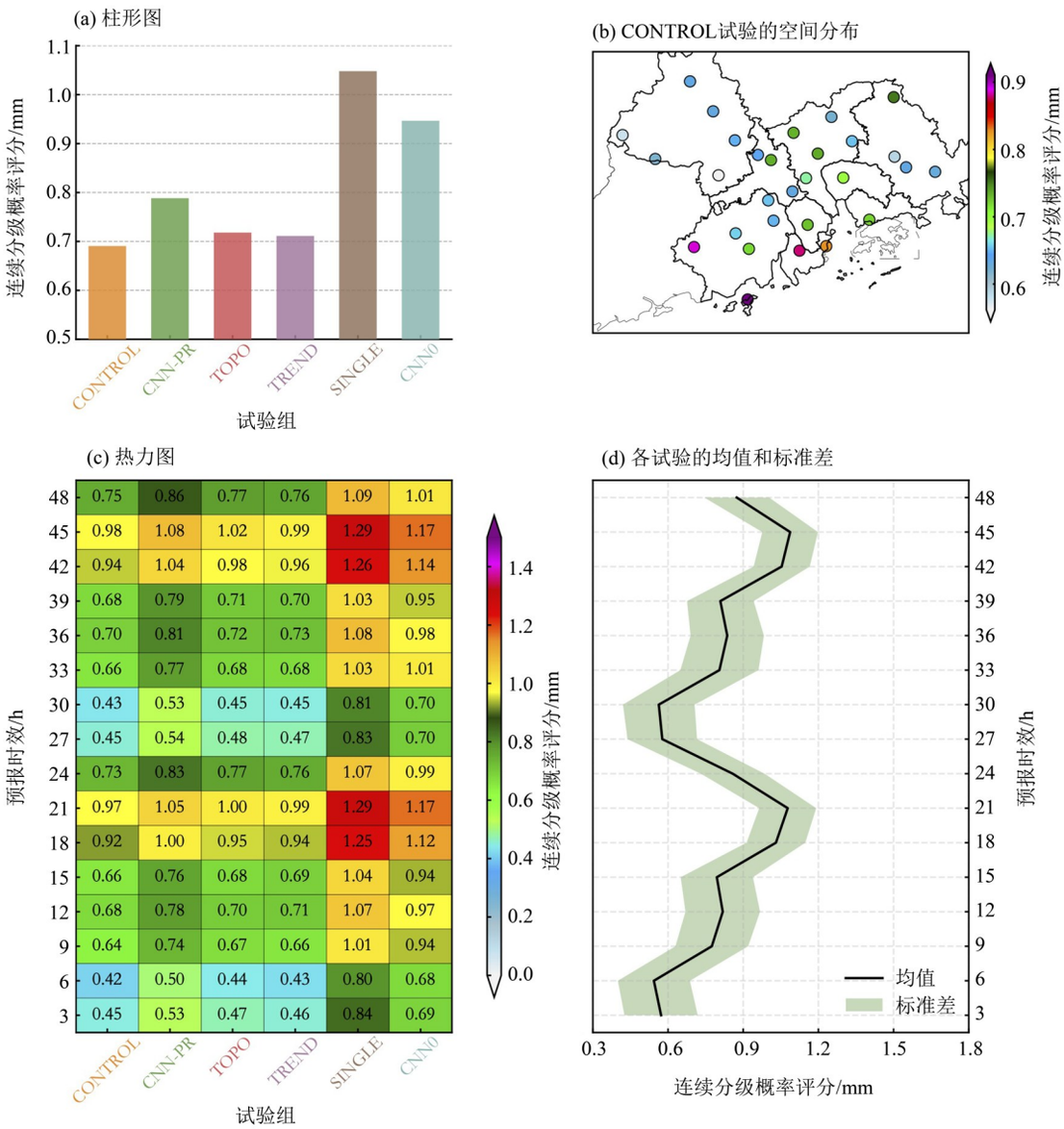
3.1　整体 CRPS 与校准度

概率预报评估中最为关键的是在满足良好校准性的前提下，尽量增大概率分布的锐度（Gneiting et al.，2007）。CRPS 是概率预报评估的核心指标之一，可以同时反映概率预报的校准性和锐度。比较可知，CONTROL 试验整体 CRPS 最优（0.69 mm），TREND（0.71 mm）和 TOPO（0.72 mm）次之（图 5）。CNN-PR 试验虽然在确定性预报评估中效果最好，但 CRPS（0.79 mm）结果却不如 CONTROL 试验。SINGLE（1.05 mm）和 CNN0（0.95 mm）试验的 CRPS 结果较差，再次说明利用卷积提取全局空间特征的重要性。与此类似的是，各试验不同预报时效的 CRPS 结果均表现出同样的变化特点，其空间分布模式也与降水量均值分布相似，这说明站点降水的时空统计特征主导着模型预报性能的时空差异。需要特别指出的是，除 SINGLE 以外的各组试验均为单模型配置，即模型使用所有预报时效数据进行训练和测试（预报时效是辅助特征），一次产生所有站点的 3 h 预报降水。这种设计会进一步强化模型学习降水整体时空统计分布特征的能力，但是也会削弱自身的泛化性能。

概率积分变换分析（图 6a、b）是研究概率预报整体校准性的重要方法，能够直观显示预报概率与实际观测的统计一致性。总体来看，各试验的整体校准水平均较好。以 CONTROL 和 CNN-PR 试验为例，二者的大多数柱高均接近于 1 且 KS 距离均较小；其中 CONTROL 试验在概率积分变换值接近 1 时柱高大于 1，表明其对强降水概率存在低估；而 CNN-PR 试验在概率积分变换值约为 0.9 时柱高略低于 1，反映出其对中等强度降水略有高估。

3.2　事件差异特征

本研究进一步利用 ROC 曲线和可靠性曲线，评估各试验对不同降水事件的区分性能和校准能力（图 6c~f）。对于降水发生事件（降水量>0.1 mm），CNN0 试验（AUC=0.811）具有一定的降水事件区分



a：指标站点和预报时效平均；b：地图中标记为指标预报时效平均；c：指标站点平均；
d：中均值和标准差基于各试验指标站点平均结果计算；审图号为GS（2024）0650号，底图边界无修改。

图5 连续分级概率评分对比

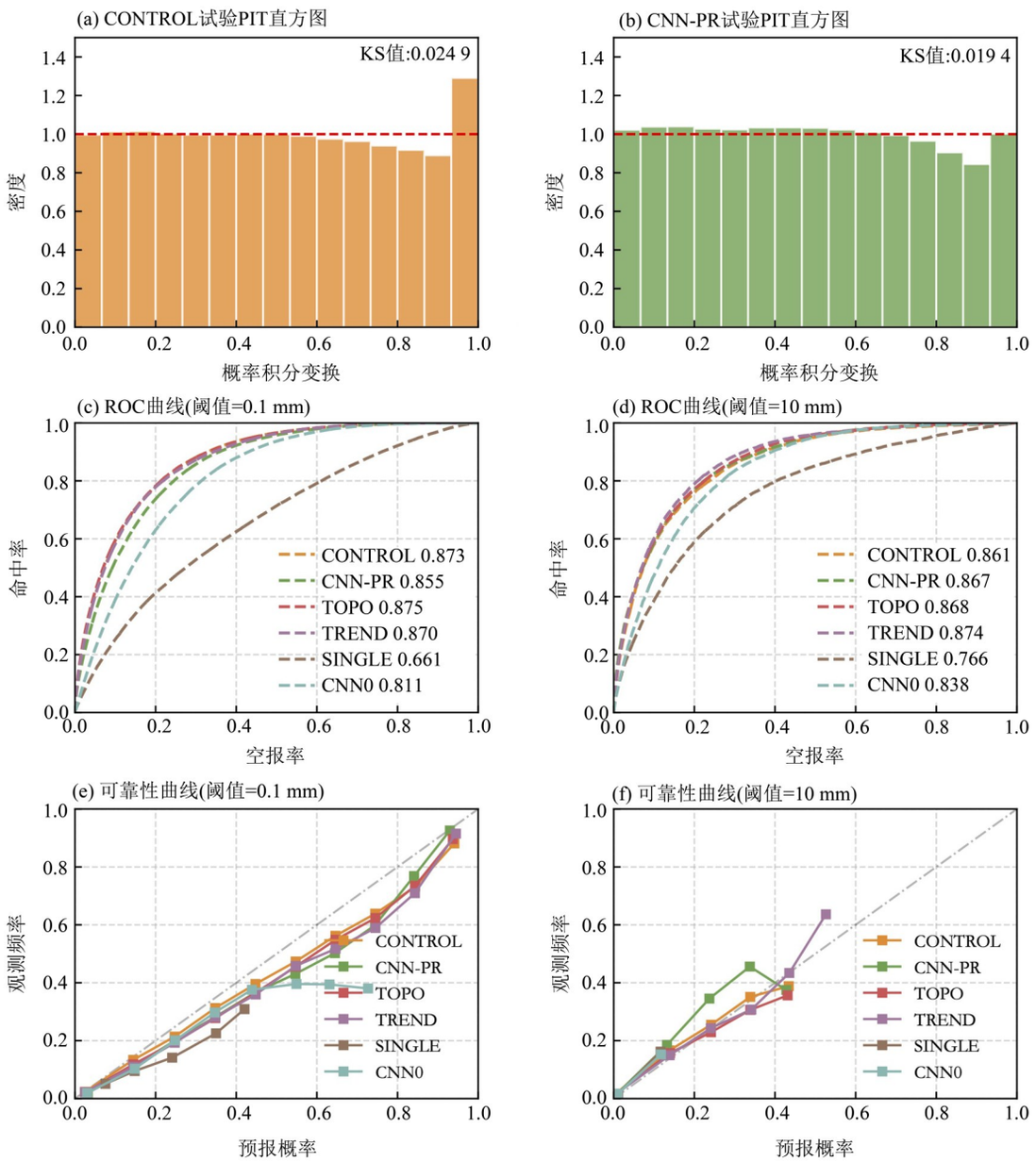
Fig. 5 Comparison of CRPS results

能力，但效果不如使用卷积结构的各试验（AUC>0.85），SINGLE试验（AUC=0.661）的技巧则最差。这表明，利用空间全局信息可以提高对降水事件的分辨能力，而空间上的卷积结构则有助于进一步强化这一效果。另外，大部分试验对事件的区分能力较好，但可靠性曲线在对角线下方，略有弯曲，反映出对降水发生的概率存在高估。CNN0和SINGLE试验的可靠性曲线在中高概率区间中断，反映出模型预报结果过于保守，对降水事件的预报技巧一般。

对于强降水事件（降水量>10 mm），TREND试验的区分能力最佳（AUC=0.874），CONTROL、CNN-PR和TOPO等3组试验的区分能力亦较好（AUC>0.86），但这4组试验的可靠性曲线在中等概率以上的区间均缺失。这表明，卷积模型预测强降水事件时的可靠性较差，存在较大的漏报风险，这与确定性预报结果一致。而CNN0和SINGLE试验的可靠性曲线缺失更为严重，说明其对强降水事件的校准能力进一步下降，无法给出强降水事件中等概率的预报结果。

3.3 校准与锐度分析

图7展示了不同降水区间下，各组试验的预报校准度与锐度的评估结果。在0~1 mm区间，各试



a~b:红色虚线为理想均匀分布参考线，数字为KS距离；c~d:图例中数值为AUC均值；
e~f:点划线为理想校准线。

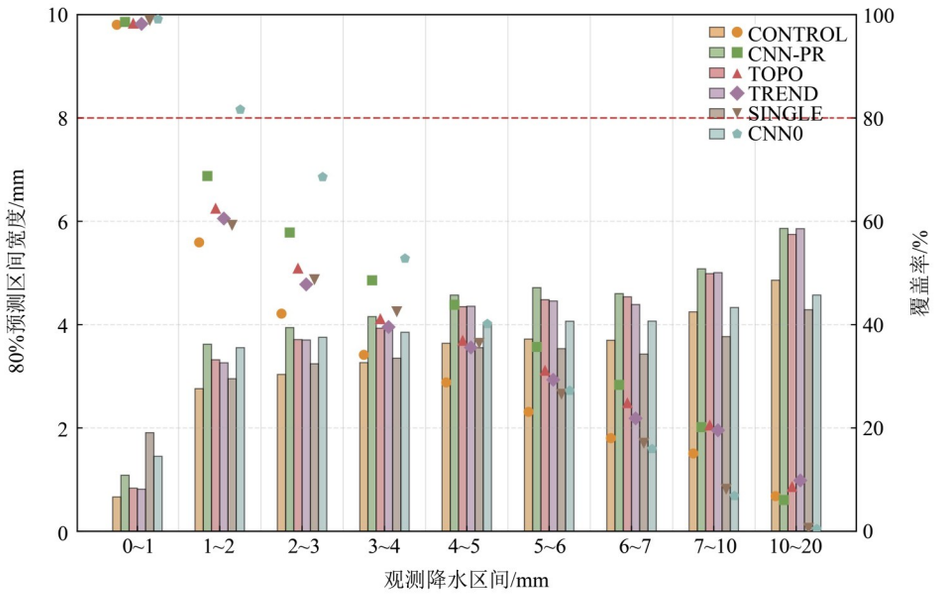
图 6　降水概率积分变换分析及降水事件 ROC 曲线、可靠性曲线对比

Fig. 6　Probability integral transform analysis and comparison of ROC curves and reliability diagrams for precipitation events

验预报覆盖率均接近 100%，存在过度覆盖现象；这表明模型对于干湿过渡区间的校准存在局限性。在 1~2 mm 区间，CNN0 试验的覆盖率在 80% 理想线附近，CNN-PR 表现次之，表明两者具有良好的校准度；其他试验覆盖率快速下降至 60% 附近，说明对该降水区间存在一定低估。随着降水区间的逐步增大，各试验覆盖率持续降低。这表明降水强度越强，各模型对降水的低估越明显，区间校准能力越差。总体而言，CNN-PR 试验覆盖率下降速度最慢，揭示出增强的空间特征提取结构能够有效提升降水预报的区间校准度。但在多个降水区间，CNN-PR 试验的区间宽度最大，表明校准度的提升以牺牲预报锐度为代价。

4　结论与展望

本研究采用 CNN 构建了大湾区短时降水概率降尺度模型，输入 GFS 模式的多层气象预报，输出区域多站点 3 h 降水概率降尺度分布参数，设计了不同敏感性试验组合检验降尺度模型的性能，主要结论如下。1）确定性预报方面，与 GFS 预报相比



标记为覆盖率结果，柱形图为80%预测区间宽度结果，红色虚线为理想校准线。

图7 各试验80%预测区间的观测覆盖率及其平均宽度在不同观测降水区间的对比

Fig. 7 Comparison of observation coverage and corresponding widths of 80% prediction intervals from different experiments under various observed precipitation intervals

（RMSE=3.47 mm），各试验均减小了预报误差。其中，CNN-PR能够更好地提取空间特征，整体性能最佳（RMSE=3.31 mm）；SINGLE试验（仅提取单点信息）和CNN0试验（未进行空间卷积）的性能低于其他试验。2）概率预报方面：各试验整体校准度较好，CONTROL试验校准度最佳（CRPS＝0.69 mm）。对于降水发生事件，采用卷积模型的各试验均具备较好的判别能力（AUC>0.85）；对于强降水发生事件，考虑时序演变信息的TREND试验分辨能力最佳（AUC=0.874）；但各模型预报保守，可靠性不足。此外，各试验区间校准度不佳，CNN-PR整体最佳，但牺牲了预报锐度。3）本研究构建的降尺度模型对弱降水区间的降水预报性能较好，降水强度增加后，模型概率预报过于保守，确定性预报技巧不如GFS。这是因为模型为最小化整体预报误差，对高频弱降水拟合程度较好，但对强降水表达能力较差。

综合以上结果，对于深度学习降尺度模型的发展和实际应用，可得到以下启示：1）在相同损失函数约束的情况下，试验之间的预报性能差异不大，时空特征较为相似。后续工作可以通过增加大值样本权重、融合物理机制构造非对称损失函数等方式，探索模型的改进效果。2）单一模型的设计能够使模型更好地捕捉到降水的整体规律，但难以适应降水昼夜、季节以及空间变化。随着数据积累，后续可以尝试设计多个模型去适配不同情景。3）降尺度概率预报的水平受“整体-区间校准”“校准度-锐度”两组内在矛盾的制约。因此，模型的设计与优化应结合具体应用场景，保证预报性能与实际需求相匹配。

参考文献：

董晓云，余锦华，梁信忠，等，2020. CWRF模式极端降水模拟误差订正[J]. 应用气象学报，31(4)：504-512.

范丽军，符淙斌，陈德亮，2005. 统计降尺度法对未来区域气候变化情景预估的研究进展[J]. 地球科学进展，3：320-329.

李广宗，孙赫敏，苏布达，等，2025. 珠江流域不同历时极端降水雨型时空分布特征研究[J]. 水利水电技术（中英文），56(10)：72-83.

李丽平，章开美，王超，等，2010. 近40年华南前汛期极端降水时空演变特征[J]. 气候与环境研究，15(4)：443-450.

李卓凡，王东海，曾智琳，2023. 粤港澳大湾区一次局地极端强降水过程的多尺度观测特征与演变分析[J]. 热带气象学报，39(6)：915-928.

孙健，曹卓，李恒，等，2021. 人工智能技术在数值天气预报中的应用[J]. 应用气象学报，32(1)：1-11.

孙磊，蓝玉峰，梁秀姬，等，2024. 基于图卷积记忆网络对珠海臭氧时空预测[J]. 中山大学学报(自然科学版)(中英文)，63(3)：48−59.

王蕾，陈起英，徐国强，等，2024. CMA−GFS 对一次强降水过程预报评估及诊断改进[J]. 中山大学学报(自然科学版)(中英文)，63(2)：46−58.

王垚，史海匀，2024. 粤港澳大湾区极端降雨的时空变化及其驱动因素探究[J]. 水资源保护，40(6)：28−37，77.

杨绚，代刊，朱跃建，2022. 深度学习技术在智能网格天气预报中的应用进展与挑战[J]. 气象学报，(5)：649−667.

张博，张芳华，李晓兰，等，2024. “23·7”华北特大暴雨数值预报检验评估[J]. 应用气象学报，35(1)：17−32.

张进，孙健，沈学顺，等，2023. CMA−GFS $V_{4.0}$ 模式关键技术研发和业务化[J]. 应用气象学报，34(5)：513−526.

张欣，刘凑华，戴建华，等，2026. 长三角地区不同环流型下强降水模式预报性能评估[J]. 气象，52(5)：580−594.

张亚萍，张焱，翟丹华，等，2022. 深度学习方法在强降水临近预报技术中的应用及其改进思考[J]. 暴雨 灾害，41(5)：506−514.

Baño−Medina J，Manzanas R，Gutiérrez J M，2020. Configuration and intercomparison of deep learning neural models for statistical downscaling [J]. Geosci Model Dev，13(4)：2109−2124.

Cannon A J，2008. Probabilistic multisite precipitation downscaling by an expanded bernoulli−gamma density network [J]. J Hydrometeorol，9(6)：1284−1300.

Deng Z，Wang Z，Wu X，et al，2023. Effect difference of climate change and urbanization on extreme precipitation over the Guangdong−Hong Kong−Macao Greater Bay Area[J]. Atmos Res，282：106514.

Dong F，Zhi X，Zhu S，et al，2023. Principal modes of diurnal cycle of rainfall over South China during the presummer rainy season[J]. J Clim，36(8)：2457−2470.

Gneiting T，Balabdaoui F，Raftery A E，2007. Probabilistic forecasts，calibration and sharpness[J]. J R Stat Soc Ser B Stat Methodol，69(2)：243−268.

Li Z，Chan J C L，Zhao K，et al，2021. Impacts of urban expansion on the diurnal variations of summer monsoon precipitation over the South China coast[J]. J Geophys Res Atmos，126(22)：e2021JD035318.

Liang X Z，Sun C，Zheng X，et al，2019. CWRF performance at downscaling China climate characteristics [J]. Clim Dyn，52(3)：2159−2184.

Liang X Z，Xu M，Yuan X，et al，2012. Regional climate−weather research and forecasting model [J]. Bull Am Meteor Soc，93(9)：1363−1387.

Reichstein M，Camps−Valls G，Stevens B，et al，2019. Deep learning and process understanding for data−driven Earth system science[J]. Nature，566(7743)：195−204.

Sloughter J M L，Raftery A E，Gneiting T，et al，2007. Probabilistic quantitative precipitation forecasting using Bayesian model averaging[J]. Mon Weather Rev，135(9)：3209−3220.

Sun L，Lan Y，2021. Statistical downscaling of daily temperature and precipitation over China using deep learning neural models：Localization and comparison with other methods[J]. Int J Climatol，41(2)：1128−1147.

Sun L，Lan Y，Jiang R，2023. Using CNN framework to improve multi−GCM ensemble predictions of monthly precipitation at local areas：An application over China and comparison with other methods[J]. J Hydrol，623：129866.

Sun L，Lan Y，Sun X，et al，2024. Deterministic forecasting and probabilistic post−processing of short−term wind speed using statistical methods[J]. J Geophys Res Atmos，129(7)：e2023JD040134.

Vrac M，Naveau P，2007. Stochastic downscaling of precipitation：From dry events to heavy rainfalls[J]. Water Resour Res，43(7)：2006WR005308.

Wang Y，Han Z，Gao R，2021. Changes of extreme high temperature and heavy precipitation in the Guangdong−Hong Kong−Macao Greater Bay Area [J]. Geomatics，Natural Hazards and Risk，12(1)：1101−1126.

Wilks D S，2011. Statistical methods in the atmospheric sciences [M]. 3rd ed. Oxford：Academic press：257−267.

Xue Y，Janjic Z，Dudhia J，et al，2014. A review on regional dynamical downscaling in intraseasonal to seasonal simulation/prediction and major factors that affect downscaling ability[J]. Atmos Res，147/148：68−85.

Yu R，Zhou T，Xiong A，et al，2007. Diurnal variations of summer precipitation over contiguous China[J]. Geophys Res Lett，34(1)：L01704.

（责任编辑　江　睿）