



# 基于 CNN-Attention-BP 的降水发生预测研究

## 摘要

在综合分析降水统计预测模型特点的基础上,提出一种基于 Attention 机制、卷积神经网络(CNN)和 BP 神经网络的 CNN-Attention-BP 组合模型,并对 1961—2020 年不同气候类型的长春站、白城站、延吉站夏季降水进行实证分析。首先,运用卷积神经网络对 6—8 月 20—次日 20 时降水量、平均气压、平均风速、平均气温和平均相对湿度进行特征学习,利用 Attention 机制来确定气象影响因素对降水预测的权重;然后,使用 BP 神经网络进行降水发生预测,选用准确率、交叉熵损失函数和 F1-score 来综合评价 CNN-Attention-BP 组合模型的性能。最后,将单一的支持向量机、多层感知机和卷积神经网络模型与组合模型进行比较分析。结果表明,CNN-Attention-BP 组合模型具有自主学习和关注更重要信息的特征,能够有效提高吉林省夏季降水发生模型的预测能力,在样本越均衡、降水频率越接近于 0.5 的站点,预测精度越高,准确率最高可达 88.4%。CNN-Attention-BP 组合模型的准确率相较于其他单一模型最高可以提高近 17 个百分点。

## 关键词

降水预测;卷积神经网络;Attention 机制;BP 神经网络;交叉熵损失函数

中图分类号 P457.6

文献标志码 A

收稿日期 2021-05-31

资助项目 国家重点研发计划(2018YFC1507905);国家自然科学基金(42075068,41975087);2020 年江苏高校“大学素质教育与数字化课程建设”专项课题(2020JKT032);南京信息工程大学 2019 年教改研究课题(共建共享的概率论与数理统计“金课”的探索与实践);南京信息工程大学数统学院本科专业建设项目

## 作者简介

吴香华,女,博士,副教授,研究方向为气候预报和气候统计学。wuxianghua@nuist.edu.cn

- 1 南京信息工程大学 数学与统计学院,南京,210044
- 2 南京气象科技创新研究院/中国气象局交通气象重点开放实验室,南京,210008
- 3 南京信息工程大学 大气与实验环境教学中心,南京,210044

## 0 引言

吉林省位于中国东北地区中部,大部分地区属于温带季风气候,降水、气温和气象灾害的时空分布差异较大。吉林省是我国重要的粮食和农产品生产地,稳定的农业生产环境是保证粮食产量的根本因素<sup>[1]</sup>。研究降水预测,对预防气象灾害、稳定农业生产和经济发展具有重要的战略意义<sup>[2]</sup>。

近年来,基于现代统计方法的人工智能理论和应用迅速发展,相关技术被广泛应用于降水预测中。运用现代统计模型和人工智能方法改进传统的降水发生模型,提高预测精度是科研和业务中提升管理水资源决策可靠性的迫切需求<sup>[3-4]</sup>。在运用 Logistic 模型对全国范围内强降水的研究中,引进主成分分析和 Bootstrap 抽样技术能进一步改善模型过拟合的问题<sup>[5]</sup>。将随机森林应用到降水发生的预测,并与 C4.5 决策树和 SQL 决策树进行比较,发现随机森林的准确度高于其他类型的决策树,但建模所需时间远大于其他方法<sup>[6]</sup>。Ortiz-Garcia 等<sup>[7]</sup>使用支持向量机(Support Vector Machine, SVM)预测降水量,与决策树、极限学习机等方法比较,并评估不同资料的可靠性,发现随着样本量的增加,SVM 的核函数映射维度提高、计算量增加从而影响模型的时效性。

BP 神经网络具有较好的非线性映射能力,比传统统计方法更加适合非线性的降水研究。在对不同环流型下的降水观测资料进行降尺度预报的研究中,发现 BP 神经网络对暴雨预测的相对偏差指数最小,预测效果最优<sup>[8]</sup>。在预报中心地区和周围地区运用多层感知机(Multilayer Perceptron, MLP)建立不同地区与降水之间的关系以预测 3 h 降水,比日本气象厅数值模式的均方根误差降低了 1.5 mm<sup>[9]</sup>。针对较长时间尺度的历史数据,神经网络方法能够快速学习到数据的主要特征。李增<sup>[10]</sup>基于长达 56 年的东北地区气象观测数据进行干旱预测,结果发现,深度学习的神经网络模型比 MLP 的预测性能提高 24.19%,且明显优于传统的统计预测模型。

卷积神经网络能快速捕捉降水数据的非线性关系并缓解过拟合问题,能够显著改善降水预报模型的效果<sup>[11]</sup>。当输入数据的向量维度过大,易出现维数灾难问题而降低模型的学习效率。Attention 机制是一种有效获得良好结果的机制,其关键是对重要信息提高注意力,利用注意力机制确定气象因素对降水预测的影响权重。在卷积神经网络

中加入 Attention 机制,通过对中间层神经元分配不同的概率权重,使得神经网络模型更加关注对降水发生影响较大的信息,减少甚至忽略对降水预测影响较小信息的关注<sup>[12]</sup>.

本文提出将 Attention 应用于卷积神经网络(Convolutional Neural Networks, CNN)中,加入 BP 神经网络(Back Propagation, BP)对整个组合模型进行调参.将预处理后的降水影响因素作为输入向量,进行神经网络降维.在 CNN 隐藏层中,引入 Attention 机制来捕捉气压、风速、气温和相对湿度的特征对降水发生影响的重要程度.将经过 CNN 层和 Attention 层降维后的特征向量输入 BP 神经网络进行降水发生预测,并对模型性能进行评估分析.

## 1 研究数据与预处理

### 1.1 数据来源

本文的数据来源于国家气象科学数据中心(<http://data.cma.cn>)的中国地面气候资料日值数据集(V3.0).根据不同的气候类型,选取吉林省的长春站、白城站、延吉站为研究站点,各站点信息见表 1.选用 1961—2020 年夏季 6—8 月的 20—次日 20 时本站平均气压(hPa)、平均风速(m/s)、平均气温(°C)和平均相对湿度(%)为物理协变量,降水发生与否为研究的目标变量.

## 1.2 数据预处理

对研究数据进行质量控制.为了加快模型的收敛速度,减少离群点的影响,对数据记录中的物理协变量进行归一化处理<sup>[13]</sup>,公式如下:

$$x_i = \frac{x_i^{(0)} - x_{\min}}{x_{\max} - x_{\min}}, \quad (1)$$

其中,  $x_i$  为归一化处理后的序列,  $x_i^{(0)}$  为输入的原始气象因子序列,  $x_{\min}$  和  $x_{\max}$  分别为影响因子序列的最小值和最大值.经过归一化处理,数据的值都在  $[0, 1]$  内.最后,对日降水量数据进行类标号:将大于 0 的值标记为 1,表示样本为正;剩余样本标记为 0,表示负样本.

## 2 方法介绍

### 2.1 卷积神经网络原理

卷积神经网络是一种前馈式神经网络,由输入层、卷积层、池化层、全连接层及输出层构成,结构如图 1 所示.卷积层和池化层专门用来数据预处理,任务是过滤输入数据、提取有用的信息用作输入.卷积层在原始输入数据和生成新特征值的卷积核之间进行卷积运算.这种技术最初用于从图像数据集中提取特征<sup>[14]</sup>,输入数据必须是结构化的矩阵形式.

相对于输入矩阵,卷积核(滤波器)被认为是一个小窗口,包含矩阵形式的系数值.这个窗口滑动整

表 1 研究站点的地理特征和降水频率

Table 1 Geographic information and precipitation frequencies in the studied stations

区站号	站点名	经度/(°E)	纬度/(°N)	海拔高度/m	气候类型	降水频率
54161	长春	125.13	43.54	236.8	温带季风气候	0.459
50936	白城	122.50	45.38	155.3	温带大陆性气候	0.360
54292	延吉	129.30	42.52	257.3	中温带半湿润气候	0.481

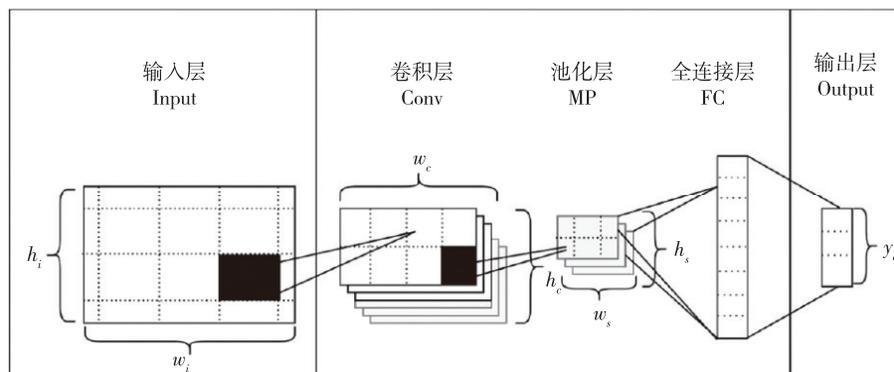


图 1 卷积神经网络的结构

Fig. 1 Architecture of convolutional neural network

个输入矩阵,对输入矩阵中的指定窗口遇到每个子区域进行卷积运算.所有数据经过操作后的结果称为卷积矩阵,表示由系数值和应用滤波器的尺寸大小所指定的特征值<sup>[15]</sup>.通过对输入数据应用不同的卷积核,可以生成多个卷积特征,这些特征通常比输入数据的原始特征更重要,从而可以提高模型的性能.池化层从卷积特征中二次采样,提取特征值并产生低维矩阵,以减少网络的参数,提高模型的鲁棒性.通过上述类似的过程,与卷积层上执行的操作一样,池化层利用最小滑动窗口,将卷积特征的每个子区域的值作为输入并输出一个新的值.

**2.2 注意力机制原理**

将 Attention 机制(图 2)加入神经网络中,考虑过去状态不同的影响因子,捕捉与日降水量相关的最重要部分.为了充分利用过去的状态信息,本文提出的 CNN-Attention-BP 组合模型在 CNN 层上增加一个注意力层,目的是降低甚至忽略与降水发生无关的信息,改变对重要信息的注意力.关键是对神经元分配不同的概率权重,使得模型更加注意到影响分类预测较大的数据信息以提高预测准确率<sup>[16]</sup>.

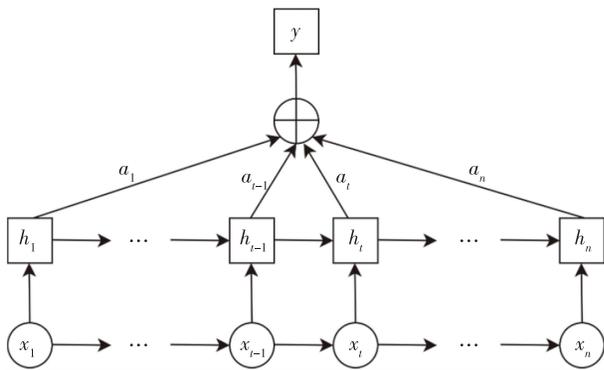


图 2 注意力机制的结构

Fig. 2 Architecture of attention mechanism

**2.3 BP 神经网络原理**

反向传播神经算法 (Back Propagation Neural Network) 简称 BP 神经网络,是一种基于误差反向传播的多层前馈网络.BP 网络包括输入层、隐藏层、输出层,训练过程由网络信号的正向传递和误差的反向传递组成<sup>[17]</sup>,结构如图 3 所示.前者为信号的前向传递,传递正方向计算输出值;后者采用梯度下降法调整权值和阈值来计算目标函数的最优值<sup>[18]</sup>.

**2.4 CNN-Attention-BP 组合模型**

本文提出的组合模型结构主要分为 Input 层、

CNN 层、Attention 层、BP 层和 Output 层,流程如图 4 所示.

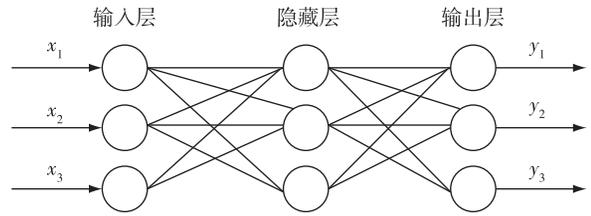


图 3 BP 神经网络结构

Fig. 3 Architecture of BP neural network

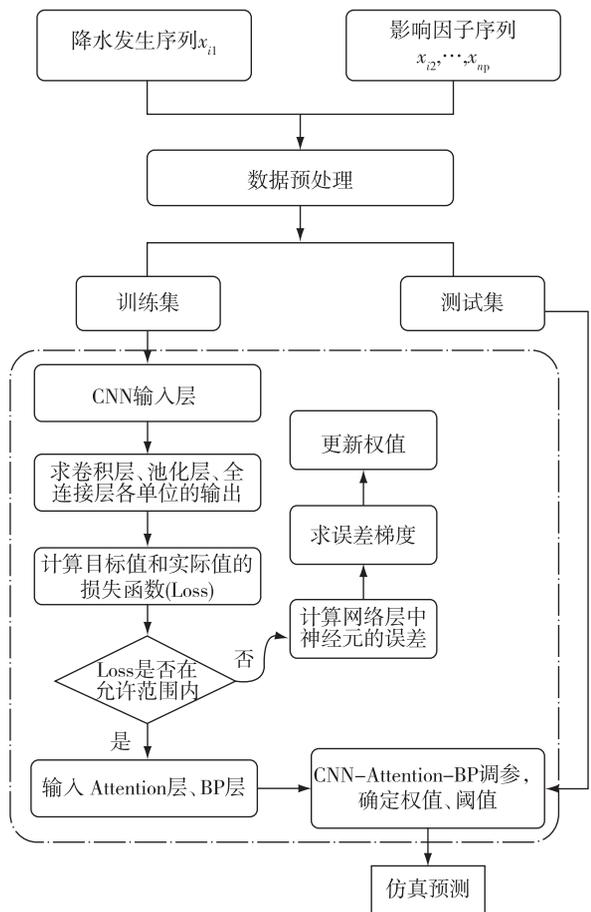


图 4 CNN-Attention-BP 组合模型流程

Fig. 4 Flow chart of the CNN-Attention-BP model

将降水相关的历史数据

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

作为组合模型的输入, $x_{i1}$  表示降水发生的标签值, $i=1, \dots, n; x_{i2}, \dots, x_{ip}$  表示第  $p$  个影响因子归一化后的序列.首先经过卷积层连接一个非线性激活函数

ReLU,然后选取常用的最大值池化进行运算,即提取池化层中元素的最大值.经过卷积层和池化层后原始数据被映射到隐层特征空间,再搭建选用 Sigmoid 激活函数的全连接层将其转换输出以解决降水分类问题,最后得到输出的特征向量<sup>[19]</sup>.

首先,将 CNN 层的输出特征向量  $Y_c$  表示为

$$Z_1 = f(X \otimes W_1 + b_1) = \text{ReLU}(X \otimes W_1 + b_1), \quad (2)$$

$$F_1 = \max(C_1) + b_2, \quad (3)$$

$$Z_2 = f(F_1 \otimes W_2 + b_3) = \text{ReLU}(F_1 \otimes W_2 + b_3), \quad (4)$$

$$F_2 = \max(Z_2) + b_4, \quad (5)$$

$$Y_c = f(F_2 \times W_3 + b_5) = \text{Sigmoid}(Z_2 \times W_3 + b_5), \quad (6)$$

式中:  $Z_1, Z_2$  分别为第 1 层和第 2 层卷积的输出;  $F_1, F_2$  分别为池化 1 层和 2 层的输出;  $W_1, W_2, W_3$  为权重矩阵;  $b_1, b_2, b_3, b_4, b_5$  为偏差;  $\otimes$  和  $\max()$  分别为卷积运算和最大值函数; CNN 层输出的长度为  $t$ ;  $Y_c = [Y_{c1}, Y_{c2}, \dots, Y_{ct}]^T$ .

然后,将经过 CNN 隐藏层激活处理后的向量  $Y$  作为 Attention 层的输入,经非线性转化成  $e_i$  表示第  $t$  个样本经过 CNN 层输出向量  $h_i$  的特征状态矩阵,如式(7):

$$e_i = \text{utanh}(w_i h_i + b), \quad (7)$$

$u$  和  $w_i$  为权重矩阵,  $b$  为偏置矩阵.其次,对每一个变量赋予初始权值,再通过 softmax 层进行归一化得到注意力权重,计算方式如下:

$$\theta_i = \frac{\exp(e_i)}{\sum_{j=1}^l \exp(e_j)}, \quad (8)$$

$$s_i = \sum_{t=1}^l \theta_i h_i, \quad (9)$$

其中,  $\theta_i$  为注意力向量,  $s_i$  为 Attention 层在第  $t$  时刻的输出.

将经过 CNN、Attention 层后的输出向量记为  $X_c = [x_{c1}, x_{c2}, \dots, x_{cl}]^T$ , 表示  $l$  维的 BP 神经网络输入向量,  $Y = [y_1, y_2, \dots, y_m]^T$  表示预测值,初始化权重值  $W_{pq}, W_{qk}$  以及隐层阈值  $a_q$  和输出层阈值  $b_k$  后,进行前向传播与后向传播.本文在 BP 各隐层节点使用 Sigmoid 激活函数,则神经节点  $l$  输入表示为

$$H_q = f\left(\sum_{l=1}^h W_{pq} x_l - a_q\right), l = 1, 2, \dots, h; q = 1, 2, \dots, s. \quad (10)$$

式中  $l$  为隐藏层节点数,  $f$  为隐藏层激励函数.经过隐藏层得到预测模型的输出:

$$O_k = \sum_{q=1}^s H_q W_{qk} - b_k, k = 1, 2, \dots, r. \quad (11)$$

当  $Y_k$  为实际值,有误差函数:

$$e_k = Y_k - O_k. \quad (12)$$

最后,计算输出目标向量和实际值的误差.若全局误差比设定的误差值小,固定当前的权重和阈值各参数值,学习结束,输出预测向量  $Y = [y_1, y_2, \dots, y_f]^T$ ; 若不在误差范围内,误差值进行反向传递以更新权值和阈值.

### 3 实例分析

#### 3.1 组合预测模型的步骤

基于 CNN-Attention-BP 组合模型进行降水发生预测的过程可划分为两个阶段.第一阶段为 1960—2000 年 6—8 月的训练阶段,第二阶段为 2001—2020 年 6—8 月的预测阶段.以长春站为例,改进后的卷积神经网络模型步骤如下:

Step1.将归一化后的长春站降水数据划分为训练集和测试集.

Step2.将已清洗的与降水相关的气象数据作为输入变量,实际降水标签值为输出变量,使用卷积神经网络方法进行训练.将训练周期(Epoch)设定为 1,初始化的学习率为 0.1, Batch\_Size 表示训练一次选取的样本数,初始设定为 32,捕捉降水相关序列的特征.

Step3.在 CNN 隐藏层引入 Attention 进行训练,提取与降水发生相关性较强的气象数据.

Step4.再将 Attention 层的输出结果输入到 BP 神经网络中,实现对整个神经网络进行调参,输出预测向量和实际值的误差;若不在误差范围内,返回 Step2 再次计算误差梯度,更新权值,直至在误差范围内,训练结束,固定各参数值.随着 Epoch 的增大,准确率开始下降,所以选取准确率最高时的 Epoch 为 9.由于批次过大会产生过拟合现象,故本文设定 Batch 为 32、64、128 和 256.经过前期训练,当 Batch 为 128 时训练误差最小.更新学习率为 0.000 1.具体见图 5、图 6.此时, BP 的训练次数为 10,学习率为 0.002,批次大小为 50.

Step5.利用参数优化后的模型对测试集数据进行建模,完成对降水发生模型的预测.改进后的神经网络方法的参数设置如表 2 所示.

#### 3.2 模型评价指标

本文在组合的预测模型训练过程中,选取 Adam (Adaptive moment estimation)<sup>[20]</sup> 算法最小化目标函

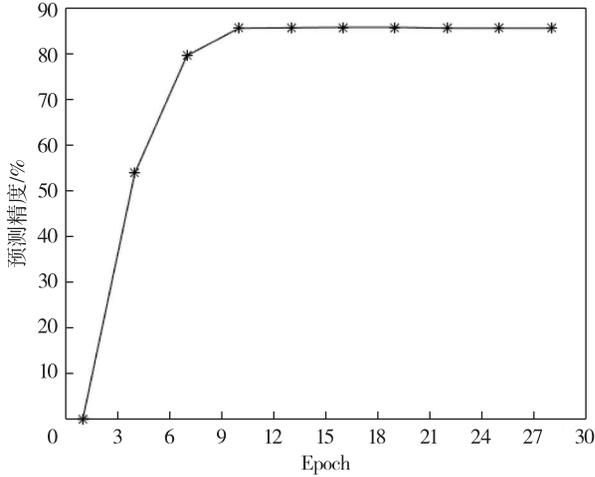


图5 模型在不同训练周期下的预测精度

Fig. 5 Forecast accuracy of the model with different epochs

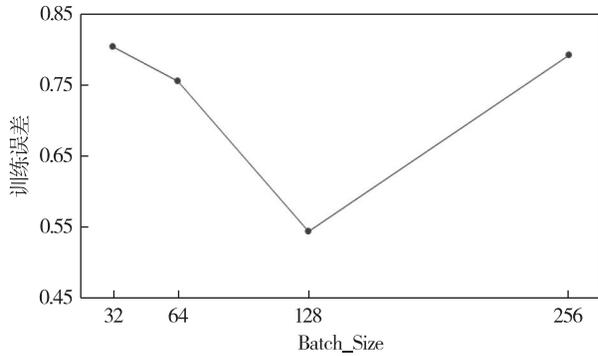


图6 模型在不同批次下的训练误差

Fig. 6 Training loss of the model with different Batch\_Size

表2 改进的神经网络方法的参数

Table 2 Parameters of the improved neural network

训练参数	方式/数值
注意力函数	点积模型
卷积层激活函数	ReLU
全连接层激活函数	Sigmoid
Droup 比例/%	30
损失函数	交叉熵损失函数
优化算法	Adam
初始学习率	0.000 1
Batch_Size	128
训练周期 (Epoch)	9

数,对神经网络参数进行优化.Adam 算法能够基于训练数据迭代更新神经网络的权重,使损失函数输出值达到最优.同时,使用交叉熵损失函数 (Cross-Entropy Loss Function) 作为评判指标,即:

$$f_{\text{loss}} = \frac{1}{N} \sum_{i=1}^f L_i = \frac{1}{N} \sum_{i=1}^f -y_i \cdot \log(p_i) +$$

$$(1 - y_i) \cdot \log(1 - p_i), i = 1, \dots, f, \quad (10)$$

式中,  $f_{\text{loss}}$  表示交叉熵损失函数,  $y_i$  为样本的期望输出,  $p_i$  为样本的实际输出.交叉熵其实是实际输出概率与期望输出概率的距离,即交叉熵的值越小,两个概率分布就越接近,损失就更小.

仅使用交叉熵损失函数不能充分评估模型的性能,故选取可表明模型稳定性的准确率 ( $A$ ) 以及评估模型泛化能力的 F1-score ( $F_1$ ):

$$A = \frac{TP + TN}{TP + TN + FP + FN}, \quad (11)$$

其中,TP 表示正确预测为有降水的样本数,TN 表示正确预测为无降水的样本数,FP 表示错误预测为有降水的样本数,FN 表示错误预测为无降水的样本数.

$$F_1 = \frac{2PR}{P + R} = \sum_{k=0}^n \frac{2T_k^k}{\sum_{i=1}^n P_i^k + \sum_{j=k}^n P_j^k} \times 100\%, \quad (12)$$

其中:

$$P = \sum_{k=0}^n \frac{T_k^k}{\sum_{i=1}^n P_i^k} \times 100\%,$$

$$R = \sum_{k=0}^n \frac{T_k^k}{\sum_{j=k}^n P_j^k} \times 100\%, k = 0, \dots, n.$$

$T_k^k$  表示样本实际为正且预测为正的样本个数,  $P_i^k$  表示模型预测为正的样本个数,  $P_j^k$  表示模型预测为负的样本个数.

### 3.3 预测结果与分析

本文选用吉林省 1961—2020 年夏季 6—8 月的 20—次日 20 时降水量、平均本站气压、平均风速、平均气温和平均相对湿度数据来进行组合模型的训练和测试.根据不同的气候类型选择具有代表性的长春站、白城站、延吉站为研究站点,使用 SVM、MLP、单一的 CNN 和组合的 CNN-Attention-BP 方法分别对降水发生模型进行预测.其中,1961—2000 年 6—8 月的气象数据作为训练样本,2001—2020 年 6—8 月观测资料用于预报效果的检验.

将上文提到的 SVM、MLP、单一的 CNN 和 CNN-Attention-BP 方法的评估结果进行对比.从图 7 可见:改进后的 CNN 组合模型相较于其他单一的模型具有更高的预测精度,准确率均大于 80%;在损失函数直方图中,CNN-Attention-BP 具有更小的损失,尤其在可表明模型综合性能的 F1-score 指标中,本文提出的方法具有更高的神经网络模型的质量.在不同

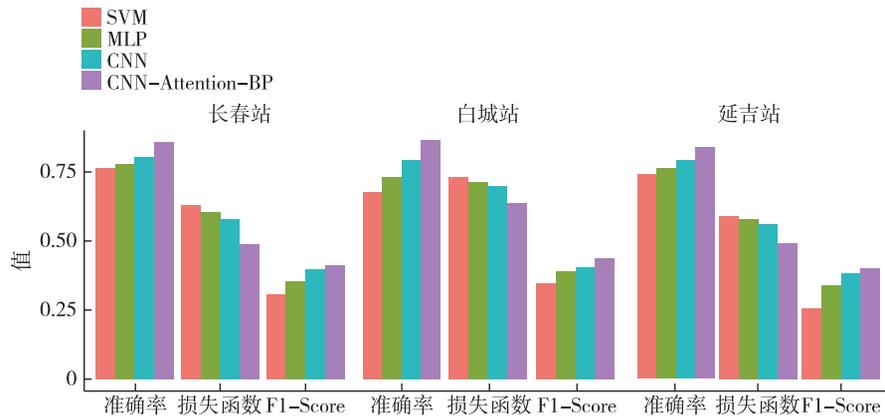


图7 代表性站点不同模型的性能

Fig. 7 Performance comparison of different models for the studied stations

站点,新的组合模型准确率更高,损失更小,效果明显优于单一模型。

根据定义可知,准确率越高表明模型具有更好的预测精度。从表3的预测结果中,延吉站的 CNN-Attention-BP 组合模型的准确率最高达到 88.4%,相对于其他单一模型至少提高 4.9 个百分点。交叉熵损失函数的值越小,损失就更小,表明实际输出概率与期望输出概率的距离越近。延吉站组合模型的交叉熵损失函数最小为 52%,相较于其他方法最多降低 10.5 个百分点;F1-score 综合考虑模型的精准率与查全率的计算结果,在 SVM 中 F1-score 最低达到 27.4%,MLP、CNN 以及 CNN-Attention-BP 对应的 F1-score 分别为 35.9%、40.1%、42.4%。CNN-Attention-BP 组合模型的 F1-score 最高,表明优化后的 CNN 质量更高。

表3 代表性站点不同模型的预测性能

Table 3 Performances of different forecast models for the studied stations

站点	模型	准确率	损失函数	F1-score
长春	SVM	0.761	0.632	0.307
	MLP	0.776	0.609	0.351
	CNN	0.805	0.578	0.396
	CNN-Attention-BP	0.859	0.544	0.411
白城	SVM	0.631	0.677	0.319
	MLP	0.679	0.665	0.363
	CNN	0.737	0.649	0.378
	CNN-Attention-BP	0.801	0.593	0.405
延吉	SVM	0.782	0.625	0.274
	MLP	0.803	0.609	0.359
	CNN	0.835	0.587	0.401
	CNN-Attention-BP	0.884	0.520	0.424

对于长春站的夏季降水发生模型, CNN-Attention-BP 组合模型的准确率最高可达 85.9%,相较于 SVM 方法提高 9.8 个百分点,相较于 MLP 方法提高 8.3 个百分点,相较于 CNN 方法提高 5.4 个百分点;交叉熵损失函数最高可降低 8.8 个百分点;在 F1-score 方面,组合的 CNN-Attention-BP 模型最稳定为 41.1%,SVM 最低为 30.7%。

在白城站, CNN-Attention-BP 的准确率最高达到 80.1%,相对于单一的 CNN 模型提高 6.4 个百分点。此站点的准确率比其他站点低,是因为白城属温带大陆性气候,降水整体偏少,降水频率只有 0.36,表明无降水与有降水的总天数存在不平衡,即正负样本数量的不均衡,导致传统的统计模型不能较好地学习其降水特征,而优化后的卷积神经网络可以快速地学习降水数据特征并找到重要信息进行训练。交叉熵损失函数相对于其他方法至少降低 5.6 个百分点;F1-score 在 SVM 中最低为 31.9%,MLP、CNN 以及 CNN-Attention-BP 对应的 F1-score 分别为 36.3%、37.8%、40.5%。综合分析,改进的卷积神经网络方法的预测准确率较为稳定,在交叉熵损失函数和 F1-score 指标上都有明显的改进,表明组合的降水发生预测模型具有更好的性能。

#### 4 结论与讨论

由于降水发生事件的不确定性,本文提出一种基于 CNN-Attention-BP 的组合模型,将复杂的降水发生预测任务进行分解。先采用卷积神经网络有效地学习与降水有关的气象信息进行特征提取,进而利用 Attention 机制对重要的信息进行权重分配,最后,使用 BP 对整个网络进行参数调整以达到最优的

预测结果.对吉林省代表性站点的夏季降水发生进行预测,结果表明:基于不同气候类型的夏季降水发生相关序列,在温带大陆性气候、降水频率为 0.36 的白城站,本文提出的 CNN-Attention-BP 组合模型在准确性方面,准确率最高可提高 17 个百分点;在衡量降水发生和不发生概率分布距离的损失函数指标中,最高降低 8.4 个百分点;代表模型泛化能力的综合性指标, F1-score 至少降低 2.7 个百分点.在降水频率近 0.5 的长春站、延吉站,准确率最高可达 88.4%,损失函数最低为 52%, F1-score 最高为 42.4%.比较代表性站点的性能评价指标,新的组合模型在降水发生预测方面稳定性和准确性更好,具有较好的适用性.

## 参考文献

### References

- [ 1 ] 王晓东,董炜华,褚丽娟,等.吉林西部旱涝灾害与降水对应关系的分析[J].灾害学,2015,30(4):53-58  
WANG Xiaodong, DONG Weihua, CHU Lijuan, et al. Relationship analysis of drought-flood disasters and precipitation in western Jilin province [J]. Journal of Catastrophology, 2015, 30(4):53-58
- [ 2 ] Ramana R V, Krishna B, Kumar S R, et al. Monthly rainfall prediction using wavelet neural network analysis [J]. Water Resources Management, 2013, 27 ( 10 ): 3697-3711
- [ 3 ] Radhakrishnan P, Dinesh S. An alternative approach to characterize time series data; case study on Malaysian rainfall data [J]. Chaos, Solitons & Fractals, 2006, 27 ( 2 ): 511-518
- [ 4 ] Barua S, Muttill N, Ng A W M, et al. Rainfall trend and its implications for water resource management within the Yarra River catchment, Australia [J]. Hydrological Processes, 2013, 27( 12 ): 1727-1738
- [ 5 ] 张芳华,曹勇,徐珺,等. Logistic 判别模型在强降水预报中的应用[J].气象,2016,42(4):398-405  
ZHANG Fanghua, CAO Yong, XU Jun, et al. Application of the logistic discriminant model in heavy rain forecasting [J]. Meteorological Monthly, 2016, 42 ( 4 ): 398-405
- [ 6 ] 张佳华,姚宜斌,曹娜.基于决策树对有无降水进行预测[J].测绘地理信息,2017,42(5):107-109  
ZHANG Jiahua, YAO Yibin, CAO Na. Prediction of whether precipitation based on decision tree [J]. Journal of Geomatics, 2017, 42(5):107-109
- [ 7 ] Ortiz-Garcia E G, Salcedo-Sanz S, Casanova-Mateo C. Accurate precipitation prediction with support vector classifiers; a study including novel predictive variables and observational data [J]. Atmospheric Research, 2014, 139: 128-136
- [ 8 ] 黎玥君,郭品文.基于 BP 神经网络的浙北夏季降尺度降水预报方法的应用[J].大气科学学报,2017,40 ( 3 ): 425-432  
LI Yuejun, GUO Pinwen. Application of downscaling forecast for the north of Zhejiang precipitation in summer based on the BP neural network model [J]. Transactions of Atmospheric Sciences, 2017, 40( 3 ): 425-432
- [ 9 ] 张鹏程,贾旻旻.一种基于多层感知器的动态区域联合短时降水预报方法[J].计算机应用与软件,2018, 35(11):153-158,183  
ZHANG Pengcheng, JIA Yangyang. A dynamic regional combined short-term precipitation forecasting approach based on multilayer perceptron [J]. Computer Applications and Software, 2018, 35(11):153-158,183
- [ 10 ] 李增.东北地区干旱特征及预测模型研究[D].沈阳:沈阳农业大学,2021  
LI Zeng. Study on drought characteristics and prediction model in Northeast China [J]. Shenyang: Shenyang Agricultural University, 2021
- [ 11 ] Pan B X, Hsu K, AghaKouchak A, et al. Improving precipitation estimation using convolutional neural network [J]. Water Resources Research, 2019, 55( 3 ): 2301-2321
- [ 12 ] 刘嘉琛,秦小麟,朱润泽.基于 LSTM-Attention 的 RFID 移动对象位置预测[J].计算机科学,2021,48(3):188-195  
LIU Jiachen, QIN Xiaolin, ZHU Runze. Prediction of RFID mobile object location based on LSTM-Attention [J]. Computer Science, 2021, 48( 3 ): 188-195
- [ 13 ] Junninen H, Niska H, Tuppurainen K, et al. Methods for imputation of missing values in air quality data sets [J]. Atmospheric Environment, 2004, 38( 18 ): 2895-2907
- [ 14 ] Rawat W, Wang Z H. Deep convolutional neural networks for image classification; a comprehensive review [J]. Neural Computation, 2017, 29( 9 ): 2352-2449
- [ 15 ] Zang H X, Liu L, Sun L, et al. Short-term global horizontal irradiance forecasting based on a hybrid CNN-LSTM model with spatiotemporal correlations [J]. Renewable Energy, 2020, 160: 26-41
- [ 16 ] Liang Z Y, Du J P, Li C Y. Abstractive social media text summarization using selective reinforced Seq2Seq attention model [J]. Neurocomputing, 2020, 410: 432-440
- [ 17 ] 谢浩.基于 BP 神经网络及其优化算法的汽车车速预测[D].重庆:重庆大学,2014  
XIE Hao. Prediction of driving condition for plug-in hybrid electric vehicles [D]. Chongqing: Chongqing University, 2014
- [ 18 ] 李静,徐路路.基于机器学习算法的研究热点趋势预测模型对比与分析: BP 神经网络、支持向量机与 LSTM 模型 [J]. 现代情报, 2019, 39( 4 ): 23-33  
LI Jing, XU Lulu. Comparison and analysis of research trend prediction models based on machine learning algorithm: BP neural network, support vector machine and LSTM model [J]. Modern Information, 2019, 39 ( 4 ): 23-33
- [ 19 ] 张瀚超.基于神经网络组合模型的风电场短期功率预测方法研究[D].株洲:湖南工业大学,2019  
ZHANG Hanchao. Research on short-term prediction of wind power based on optimized neural network [D]. Zhuzhou, China: Hunan University of Technology, 2019
- [ 20 ] Chang Z H, Zhang Y, Chen W B. Electricity price predic-

tion based on hybrid model of Adam optimized LSTM                      187:115804  
neural network and wavelet transform[J].Energy,2019,

## Application of CNN-Attention-BP to precipitation forecast

WU Xianghua<sup>1</sup> HUA Yajie<sup>1</sup> GUAN Yuanhong<sup>1</sup> WANG Weiwei<sup>3</sup> LIU Duanyang<sup>2</sup>

1 School of Mathematics and Statistics,Nanjing University of Information Science & Technology,Nanjing 210044

2 Nanjing Joint Institute for Atmospheric Sciences/Key Laboratory of Transportation Meteorology of CMA,Nanjing 210008

3 Experimental Teaching Center for Meteorology and Environment,  
Nanjing University of Information Science & Technology,Nanjing 210044

**Abstract** Here,a CNN-Attention-BP model is proposed for precipitation forecast based on analysis of the characteristics of precipitation statistical prediction models,then empirical analyses are made on the summer rainfall in Changchun,Baicheng and Yanji stations of Jilin province for the period of 1961–2020.First,a Convolution Neural Network (CNN) is used to study the characteristics of precipitation,air pressure,wind speed,air temperature and relative humidity.Second,the attention mechanism is used to determine the weight of meteorological factors for precipitation forecast.Then a BP neural network is applied to predict the precipitation probability.And the performance of the proposed CNN-Attention-BP model is evaluated by accuracy,cross-entropy loss function and F1-score,which is compared with that of the support vector machine,multi-layer perceptron and convolution neural network model.The results show that the CNN-Attention-BP model is characterized by autonomous learning and paying more attention to significant information,as well as the improved forecasting performance in summer precipitation occurrence for Jilin province.Meanwhile,the proposed model would perform better with more balanced sample and precipitation frequency closer to 0.5,when accuracy would reach up to 88.4%.Compared with single models,the CNN-Attention-BP improves the forecast accuracy by 17 percentage points.

**Key words** forecast of precipitation;convolutional neural network;attention mechanism;BP neural network;cross entropy loss function