

基于PCA-XGBoost算法的洪水预测模型研究

蔡洽羽

汕头大学工学院, 广东 汕头

收稿日期: 2024年8月12日; 录用日期: 2024年9月5日; 发布日期: 2024年9月13日

摘要

洪水作为全球范围内频发的自然灾害, 其发生频率和严重程度受多种因素影响。本研究基于包含超过百万洪水事件数据的大规模数据集, 涉及20个关键指标, 通过斯皮尔曼相关性分析及随机森林分类模型, 精确识别出对洪水概率影响显著的因素。为克服多重共线性及过拟合问题, 本文创新性的提出一种结合主成分分析(PCA)与梯度提升算法(XGBoost)的复合预测模型——PCA-XGBoost洪水预测模型。此模型通过降维处理显著减少了输入变量之间的相互依赖性, 同时优化了XGBoost的超参数以增强其预测性能和泛化能力。通过严格的交叉验证和参数调优, PCA-XGBoost模型在洪水预测精度和操作效率上展现了卓越性能, 为全球多地区洪水预警系统提供了科学依据和有效工具。此研究不仅推动了洪水风险评估模型的创新, 也为相关政策制定和灾害管理提供了理论支持。

关键词

洪水预测, 主成分分析, 梯度提升算法, 超参数优化

Research on the Flood Prediction Model Based on the PCA-XGBoost Algorithm

Qiayu Cai

College of Engineering, Shantou University, Shantou Guangdong

Received: Aug. 12th, 2024; accepted: Sep. 5th, 2024; published: Sep. 13th, 2024

Abstract

Floods, as a globally frequent natural disaster, are influenced by various factors in terms of frequency and severity. This study, based on a large dataset containing over one million flood events and involving 20 key indicators, accurately identifies significant factors affecting flood probability through Spearman correlation analysis and random forest classification models. To overcome issues of multicollinearity and overfitting, this paper innovatively proposes a composite prediction model

combining Principal Component Analysis (PCA) and the Gradient Boosting algorithm (XGBoost)—the PCA-XGBoost Flood Prediction Model. This model significantly reduces interdependencies among input variables through dimensionality reduction and optimizes the hyperparameters of XGBoost to enhance its predictive performance and generalization ability. With rigorous cross-validation and parameter tuning, the PCA-XGBoost model demonstrates superior performance in flood prediction accuracy and operational efficiency, providing a scientific basis and effective tools for flood warning systems in various global regions. This research not only advances the innovation of flood risk assessment models but also offers theoretical support for policy-making and disaster management.

Keywords

Flood Prediction, PCA, XGBoost, Hyperparameter Optimization

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

洪水作为全球范围内频发且影响深远的自然灾害，其发生频率和严重程度受多种因素影响，包括但不限于气候变化、土地利用模式变化以及人口增长等。随着人类活动的不断扩展和深化，特别是森林砍伐、城市化进程等人为因素的加剧，洪水灾害的风险进一步上升。科学有效地预测洪水的发生概率，成为实现灾害管理和减灾目标的重要前提。

本文研究涉及的样本数据包括超过 100 万条洪水事件，涵盖了从季风强度、地形排水到城市化和气候变化等 20 个指标的详细评分，以及每个事件的洪水发生概率。该数据集具备丰富的数据量和广泛的年份跨度，确保了数据的多样性，具有较强的代表性。通过对这些数据进行数学建模和数据分析，可以深入探讨各种因素对洪水发生概率的具体影响，为未来的预测模型构建提供理论和实证支持。此研究不仅有助于提高洪水灾害预警的准确性和时效性，还能为制定有效的防灾措施和城市规划政策提供科学依据，以应对日益严峻的洪水风险挑战。

2. 洪水关键特征识别

2.1. Spearman 相关性分析

使用斯皮尔曼相关性分析分析各指标与洪水发生概率之间的关联性，以识别出对洪水发生有显著的影响指标。斯皮尔曼相关性分析通过计算两个变量排名(秩次)之间的相关性来评估它们之间的关系[1]。这种方法不依赖于数据的正态分布，对异常值也不敏感，因此特别适用于非正态分布或存在异常值的数据集。其步骤如下：

首先整理数据，将每个变量的数据转换为排名(秩次)。如果有并列的数据，则取其平均秩次。并对每一对变量排名计算其差值。最后使用斯皮尔曼相关系数的计算公式来计算相关系数。公式为：

$$\rho = 1 - \frac{6 * \sum d^2}{n * (n^2 - 1)} \quad (1)$$

其中， ρ 是斯皮尔曼相关系数， d 是每对变量排名差值， n 是样本数量。

得出计算结果如表 1 所示；并根据结果绘制出相关系数热力图，如图 1 所示。

Table 1. Correlation coefficient between features and probability of flood occurrence
表 1. 特征与洪水发生概率的相关系数

特征名	相关系数	特征名	相关系数	特征名	相关系数
季风强度	0.1893	淤积	0.1869	流域	0.1818
地形排水	0.1876	农业实践	0.1834	基础设施恶化	0.1902
河流管理	0.1872	侵蚀	0.1786	人口得分	0.1863
森林砍伐	0.1844	无效防灾	0.1828	湿地损失	0.1835
城市化	0.1813	排水系统	0.1790	规划不足	0.1808
气候变化	0.1853	海岸脆弱性	0.1778	政策因素	0.1822
大坝质量	0.1877	滑坡	0.1849	--	--

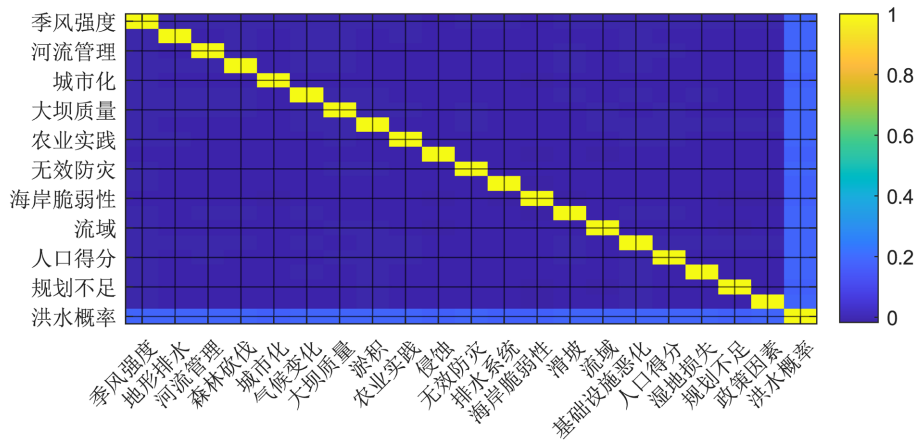


Figure 1. Correlation coefficient heatmap
图 1. 相关系数热力图

如图 1 所示，相关系数热力图中每个单元格表示两个类别之间的相关系数，颜色的深浅代表相关性的强度，颜色越深表示相关性越强，颜色越浅表示相关性越弱。通过分析，可以得出季风强度、地形排水、河流管理、气候变化、大坝质量、淤积、基础设施恶化和人口得分这些指标与洪水发生有较密切的联系。侵蚀、排水系统和海岸脆弱性与洪水发生的相关性相对较小。

2.2. 基于随机森林分类模型求解特征重要性

在深入剖析了先前提及的 20 项指标后，目标聚焦于提炼出与洪水发生概率紧密相连的关键因素。为了实现这一目标，我们将采用严谨的相关性分析与特征重要性评估策略。我们可以评估每个特征在模型预测洪水发生概率时的贡献度。通过建立的随机森林来计算特征重要性，再将建立的随机森林分类模型应用到训练、测试数据，得到模型的分类评估结果(如表 2 所示)，并将其进行可视化，绘制出混淆矩阵热力图(如图 2 所示)。

Table 2. Proportion of feature importance
表 2. 特征重要性占比

特征名称	特征重要性	特征名称	特征重要性
基础设施恶化	8.70%	森林砍伐	4.30%
地形排水	8.20%	农业实践	4.20%

续表

大坝质量	8.20%	流域	4.00%
季风强度	8.10%	规划不足	3.70%
河流管理	6.50%	政策因素	3.70%
淤积	6.00%	湿地损失	3.60%
无效防灾	5.80%	排水系统	3.20%
滑坡	5.50%	侵蚀	2.80%
气候变化	4.90%	海岸脆弱性	2.10%
人口得分	4.40%	城市化	1.90%

结合严谨的相关性分析与特征重要性评估, 我们综合得出基础设施恶化、地形排水、大坝质量、季风强度和河流管理这 5 个特征与洪水发生概率相关系数较大, 同时在随机森林分类器所计算特征的重要性较高, 说明这些指标与洪水风险的相关性较高。

3. 模型的建立与求解

3.1. 主成分分析(PCA)

为了提高洪水风险评估模型的解释能力和预测准确性, 计算不同指标的权重, 我们采用了主成分分析(PCA)方法来识别和筛选出与洪水发生概率最相关的关键指标。主成分分析是一种统计技术[2], 用于数据降维, 通过转换可能相关的多个变量为少数几个线性不相关的变量, 称为主成分。这些主成分能够捕捉原始数据中的大部分变异性, 从而简化模型而不显著损失信息。

因子分析前, 首先进行 KMO 和 Bartlett 的检验, 判断是否可以主成分分析。在因子分析中, 若拒绝原假设, 则说明可以做因子分析, 若不拒绝原假设, 则说明这些变量可能独立提供一些信息, 不适合做因子分析。

KMO 检验是 Kaiser, Meyer 和 Olkin 提出的抽样适合性检验(Measure of Sampling Adequacy)。该检验是对原始变量之间的简相关系数和偏相关系数的相对大小进行检验。对于 KMO 值, 0.8 上非常合适做主成分分析, 0.7~0.8 之间一般适合, 0.6~0.7 之间不太适合, 0.5~0.6 之间表示差, 0.5 下表示极不适合。其计算结果为:

$$KMO = \frac{\sum \sum_{i \neq j} r_{ij}^2}{\sum \sum_{i \neq j} r_{ij}^2 + \sum \sum_{i \neq j} p_{ij}^2} \quad (2)$$

Bartlett 球性检验是一种数学术语。用于检验相关阵中各变量间的相关性, 是否为单位阵, 即检验各个变量是否各自独立。对于 Bartlett 的检验, 若 P 小于 0.05, 拒绝原假设, 则说明可以做主成分分析, 若不拒绝原假设, 则说明这些变量可能独立提供一些信息, 不适合做主成分分析。

通过输出结果可得出 KMO 的值为 0.398, 同时, Bartlett 球形检验的结果显示, 显著性 P 值为 0, 水平上呈现显著性, 拒绝原假设, 各变量间具有相关性, 主成分分析有效, 程度为极不适合。

通过计算协方差矩阵来分析各特征间的相关性。并将计算结果进行可视化(如图 2 所示)。

分析结果揭示, 尽管每个主成分单独对总方差的贡献相对较小, 但它们合在一起提供了对数据结构深刻的见解。前三个主成分虽然只分别解释了约 5% 的方差, 但累积方差解释率显示前十个主成分已解释超过 50% 的总方差, 而全部 20 个主成分合计解释了近 100% 的总方差, 这表明几乎所有原始数据的信息。

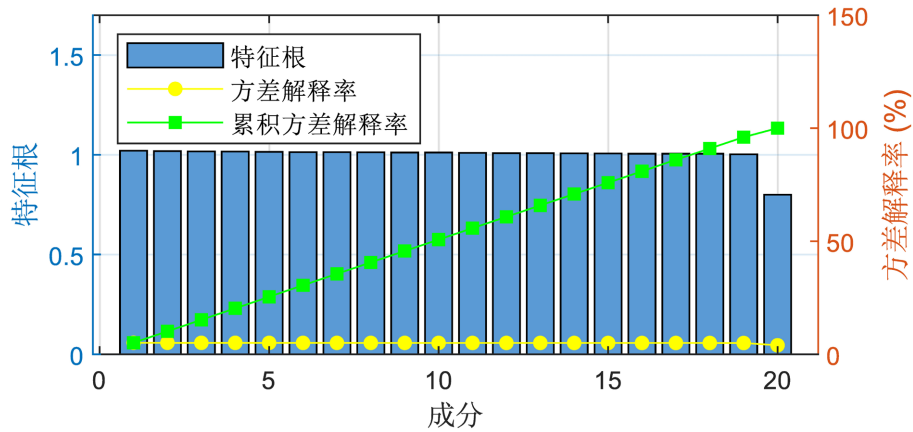


Figure 2. Visualization of total variance explanation
图 2. 总方差解释的可视化图

都被有效捕获。所以。我们可以通过分析每个主成分的贡献指标，可以更准确地理解不同因素对洪水风险的贡献。决定哪些主成分将用于进一步的洪水风险的预警评价模型的建模。可以基于累积方差解释率来选择，也可以根据主成分对特定风险因素的贡献度来选择。例如，可能会选择解释总方差超过 85% 的最少主成分数目，或选择那些具有特定水文或气象指标高负荷的主成分。使用选定的特征向量将原始数据投影到新的特征空间，形成新的数据集，即完成了降维。

在环境科学和灾害管理的研究中，评估多个影响因素的相对重要性是至关重要的。主成分分析(PCA)提供了一种有效的方法来识别哪些变量对数据集的变异贡献最大。通过分析各主成分(PCs)的方差解释率和载荷系数，我们可以量化每个变量的影响力。

3.2. PCA-XGBoost 洪水预测模型

XGBoost 主要基于梯度提升树构建的集成学习框架[3]。该方法的目的是通过组合多个简单的弱分类器来创建更高效的强分类器。与传统方法不同，XGBoost 对目标函数进行了改进，并引入树结构的复杂性作为调整项，以减少过拟合的风险。我们选择了 PCA-XGBoost 作为主要的洪水预测模型，因其在处理非线性问题和大数据集方面的高效性。模型配置如下：

- 基学习器 f 使用的是基于树的模型 `gbtree`。
- 学习率 η 设置为 0.1，控制每一步模型更新的步长。
- 树的最大深度 `max_depth` 设定为 10，旨在捕捉数据的复杂性并避免过拟合。
- 正则化项：L1 正则化 $\alpha = 0$ ，L2 正则化 $\lambda = 1$ 。
- 使用 3 折交叉验证来确保模型的稳定性和泛化能力，表示为 $k = 3$ 。

其中，训练数据集占比 70%，测试数据集占比 30%，即训练集 $n_{\text{train}} = 0.7n$ 和测试集 $n_{\text{test}} = 0.3n$ 。

特征重要性 $\text{Imp}(x_i)$ 通常基于特征 x_i 在构建决策树时的不纯度减少(例如信息增益)度量：

$$\text{Imp}(x_i) = \sum_{t \in T} \Delta i(s_t, x_i) \tag{3}$$

其中 T 是所有树的集合， s_t 是树 t 中分裂点， x_i 是考虑的特征，而 $\Delta i(s_t, x_i)$ 表示该特征在分裂点 s_t 上减少的不纯度。

使用 XGBoost 机器学习模型分析特征重要性结果，特征重要性通常用来评估在构建模型时各个特征对模型预测准确性的贡献程度。特征重要性通常基于每个特征在构建决策树时，对模型改进的贡献，比如减少不纯度的程度。模型的各个特征重要性如图 3 所示。

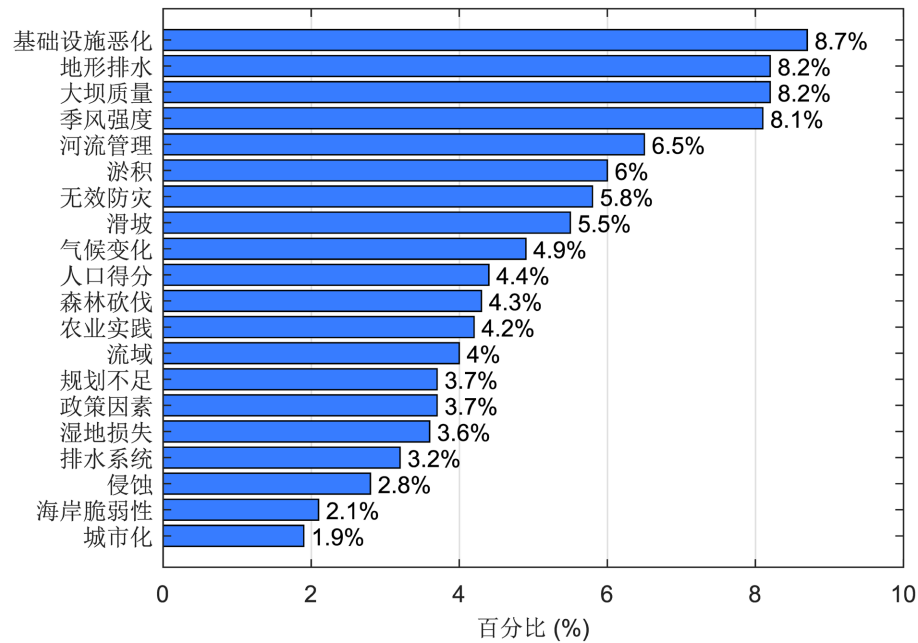


Figure 3. The importance ratio of each feature (independent variable)

图 3. 各特征(自变量)的重要性比例

模型分析显示，基础设施恶化、大坝质量和淤积是影响洪水发生的三个最重要的特征。这些变量的重要性得分远高于其他因素，表明它们在洪水预测中起着决定性作用。我们可以通过 XGBoost 内置的特征重要性评估，确定最能预测洪水风险的关键特征。

对于我们建立的 PCA-XGBoost 洪水预测模型，使用选定的特征和配置训练模型，应用交叉验证来评估模型的稳定性和准确性。我们通过量化交叉验证集、训练集和测试集的预测评价指标来衡量 XGBoost 的预测效果。

MSE (mean squared error, 均方误差)，其公式为：

$$\text{mse} = \frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2 \quad (4)$$

其中， y_i 是真实值， \hat{y}_i 表示预测值。

作为模型评估时，预测值与实际值之差平方的期望值。取值越小，模型准确度越高。

RMSE (均方根误差)：为 MSE 的平方根，取值越小，模型准确度越高，其公式为：

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (X_{\text{obs},i} - X_{\text{model},i})^2}{n}} \quad (5)$$

数学上，MAE (平均绝对误差)的计算公式可以表示为：

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (6)$$

其中：

- n 是数据点的数量。
- y_i 是第 i 个数据点的实际目标值(真实值)。
- \hat{y}_i 是模型对第 i 个数据点的预测值。

绝对误差的平均值，能反映预测值误差的实际情况。取值越小，模型准确度越高。数学上，MAPE (平均绝对百分比误差)的计算公式可以表示为：

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \tag{7}$$

是 MAE 的变形，它是一个百分比值。取值越小，模型准确度越高。

决定系数：R² (R-Square)：

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \tag{8}$$

其中，分子部分表示真实值与预测值的平方差之和，类似于均方差 MSE；分母部分表示真实值与均值的平方差之和，类似于方差 Var。

根据 R-Squared 的取值，来判断模型的好坏，其取值范围为[0, 1]：

如果结果是 0，说明模型拟合效果很差；

如果结果是 1，说明模型无错误。

一般来说，R-Squared 越大，表示模型拟合效果越好。将预测值跟只使用均值的情况下相比，结果越靠近 1 模型准确度越高。

模型评估结果如表 3 所示，在训练集上，PCA-XGBoost 洪水预测模型实现了完美拟合，MSE 为 0，而 RMSE 为 0.02 和 MAE 为 0.016，这两个低值指示模型误差极小。MAPE 为 3.19%，说明模型预测误差平均占实际值的 3.19%，R² 为 0.851，表明模型能够很好地解释训练集数据的变异。在交叉验证集中，MSE 和 RMSE 略有增加，显示模型在未见过的数据上存在泛化误差；MAE 和 MAPE 虽有所增加，但仍维持在较低水平，R²值为 0.799，较训练集有所下降，反映了模型解释能力的轻微减弱。测试集的性能指标与交叉验证集相近，进一步确认了模型的泛化能力。RMSE 和 MAE 保持稳定，表明模型在处理未知数据上的预测误差是可控的，而 R²值为 0.805，比交叉验证集略高，表明模型对测试数据具有较强的解释能力。

Table 3. Model calculation results

表 3. 模型计算结果

	MSE	RMSE	MAE	MARE	R ²
训练集	0	0.020	0.016	3.190	0.851
交叉验证集	0.001	0.023	0.018	3.726	0.799
测试集	0.001	0.023	0.018	3.690	0.805

PCA-XGBoost 洪水预测模型对洪水预测具有高准确性和可靠性，特别是在训练集上的卓越表现以及在交叉验证和测试集上良好的泛化能力，使得该模型非常适合用于建立洪水预警系统，为防灾减灾决策提供科学依据，而较高的 R²值和较低的误差指标在实际应用中意味着该模型能有效预测洪水风险，为相关工作提供准确的数据支持。

图 4 展示了 PCA-XGBoost 洪水预测模型对测试数据的预测情况。其中绿色线表示模型预测值，蓝色线表示实际观测值。通过对比两者，可以观察到模型在多个时间序列点上展现出良好的预测准确性，表明 PCA-XGBoost 模型是一个可靠的工具，适合用于洪水风险的评估和预测。

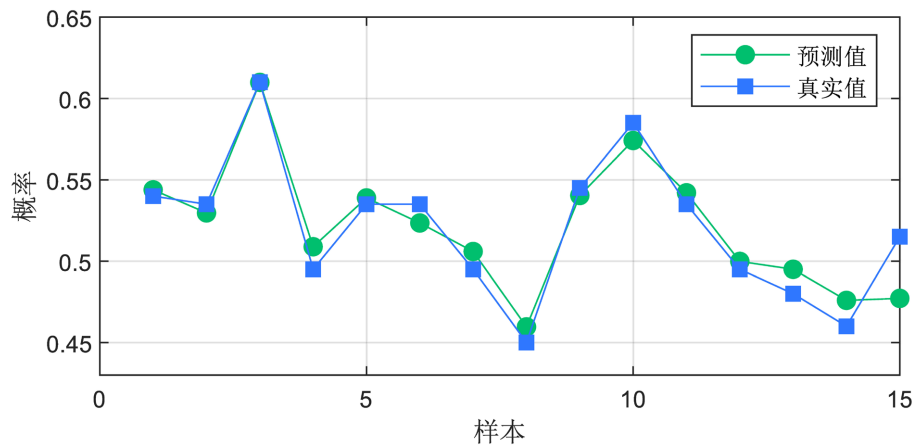


Figure 4. PCA-XGBoost flood prediction model test data prediction chart

图 4. PCA-XGBoost 洪水预测模型测试数据预测图

3.3. 模型优化

鉴于 PCA-XGBoost 在处理复杂数据和提升模型预测精度方面的优势,选择其作为本次任务的预测模型。在机器学习项目中,特别是在使用像 PCA-XGBoost 这样的强大但参数敏感的算法时,通过交叉验证(如 K 折交叉验证)和网格搜索(Grid Search)等策略来精细调优模型参数是至关重要的。这些步骤有助于找到能够在未见过的数据上表现最佳的参数组合,从而提高模型的泛化能力。

交叉验证是一种评估模型性能的技术[4],它将数据集分割成较小的子集,并使用一个子集作为验证集来评估模型在剩余子集上训练的性能。K 折交叉验证是最常见的一种,它将数据集分成 K 个大小相等的互斥子集,然后每次使用 K-1 个子集的数据来训练模型,并在剩下的一个子集上测试模型。这个过程重复 K 次,每次选择不同的子集作为测试集,最终评估指标(如 MSE 或 R^2)是所有测试结果的平均值。以下是不同 K 值的计算结果:

在三折交叉验证中, MSE 为 0.01, MAPE 为 3.726, R^2 为 0.799;

在五折交叉验证中, MSE 为 0.01, MAPE 为 3.71, R^2 为 0.801。

同时,我们可以通过遍历指定的参数网格来优化模型。对于 PCA-XGBoost 模型,可以设定学习率、树的最大深度、列采样比例等参数的多个候选值,然后网格搜索会尝试所有可能的参数组合,并使用交叉验证来评估每种组合的性能。最终,网格搜索会选择在交叉验证中表现最好的参数组合作为最优解。评估每种参数组合在交叉验证中的性能,并选择表现最好的组合作为最优参数组合。使用最优参数组合在整个训练集上重新训练 PCA-XGBoost 洪水预测模型,以准备进行最终的预测或部署。在这里,我们使用选定的参数网格和交叉验证策略来执行网格搜索,通过 scikit-learn 提供的现成的网格搜索工具,自动地遍历所有参数组合,并使用交叉验证来评估每种组合,得出了以下最优的组合(如表 4 所示)。

Table 4. Optimal combination parameters

表 4. 最优组合参数

学习率	交叉验证	树的最大深度	样本征采样率	树特征采样率	节点特征采样率
0.1	5	7	1	1	1

使用最优参数组合在整个训练集上重新训练 PCA-XGBoost 洪水预测模型,以准备进行最终的预测或部署。

4. 模型的比较与推广

4.1. 模型比较

同时通过建立随机森林回归模型来进行比较和对模型的优化。通过计算得出，交叉验证集的 MSE 为 0.02，MAPE 为 7.191， R^2 为 0.223。图 5 为测试数据预测对比图(左图是 PCA-XGBoost 洪水预测模型，右图是随机森林回归模型)。

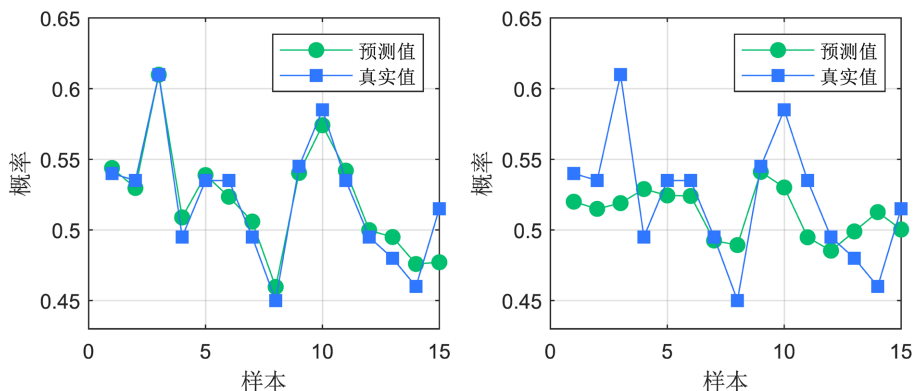


Figure 5. Prediction comparison chart for test data
图 5. 测试数据预测对比图

通过比较可以得出，PCA-XGBoost 模型展现了更高的预测准确性，特别是在对观测数据的吻合程度上，其表现优于随机森林模型。PCA-XGBoost 模型在本次洪水预测的应用中显示了其在数据处理和模型建立方面的优势，尤其是在准确追踪数据波动和预测关键转折点方面。这种高效的数据简化与先进的算法结合，使得 PCA-XGBoost 在处理大规模和复杂数据集时，具有明显的优势。

4.2. 模型的推广

1、模型的应用

跨地区洪水预警系统：可以将当前基于特定地区的洪水预测模型推广至更广泛的地理区域，通过引入更多地区的洪水数据，构建跨地区的洪水预警系统。该系统可以整合不同地区的气候、地形、水文等特征，提高洪水预测的全面性和准确性，从而更有效地支持多地区的灾害管理和应急响应。

实时洪水监测系统：将洪水预测模型与实时监测系统相结合，通过物联网技术实时采集降雨量、河流水位、土壤湿度等关键数据，实现洪水风险的实时监控和快速预警。

2、改进方向

结合气候模型：参考最新的气候研究成果，如全球气候模型(GCMs)或区域气候模型(RCMs)，将气候变化因素纳入洪水预测模型中。这可以帮助更准确地预测未来气候变化对洪水发生频率和严重程度的影响，为长期防洪规划和适应气候变化提供科学依据。

引入机器学习算法：结合深度学习、集成学习等先进的机器学习算法，进一步优化洪水预测模型的性能。这些算法能够自动学习数据中的复杂模式和特征，提高模型的预测精度和泛化能力。同时，通过对比不同算法的表现，可以选择最适合洪水预测任务的模型，实现更精准的洪水预测。

5. 结论

本研究通过整合斯皮尔曼相关性分析，在广泛数据的支持下深入探讨了洪水风险评估的多维度影响因素，并建立了 PCA-XGBoost 洪水预测模型，以提高洪水预测的准确性和操作效率。该模型首先应用

PCA 进行数据降维和特征提取，筛选出能够代表原始数据集主要变异的主成分。随后，这些主成分被用作输入，通过 XGBoost 算法进行深入的学习和精确的预测。

PCA 不仅有效地减少了数据的维度和噪声，还帮助避免了 XGBoost 模型的过拟合，从而增强了预测的泛化能力。此外，降维后的数据显著减少了模型训练和预测所需的计算资源，加快了模型的运行速度。通过分析 PCA 提取的主成分，我们能更直观地理解那些关键因素如何影响洪水的发生。

模型的验证和测试通过交叉验证和独立测试数据集完成，结果表明 PCA-XGBoost 模型在训练集上展现了优异的性能，并在新的数据集上也保持了良好的预测准确性。这一点不仅展示了模型的灵活性和扩展性，而且通过与随机森林回归模型的比较分析，进一步证实了其在洪水风险预测中的高度准确性和潜在应用价值。

PCA-XGBoost 洪水预测模型在本研究中证明了其科学重要性和应用价值。模型的建立和验证遵循了严格的数据科学原则，包括特征选择、模型训练、参数优化和交叉验证，确保了预测结果的可靠性和模型的泛化能力。该模型不仅提升了洪水预警的准确性和实用性，也为未来洪水风险管理提供了可靠的科学依据和技术路径。未来的工作将进一步优化模型的计算效率和预测范围，以便更广泛地应用于全球不同地区的洪水预测和管理中。

参考文献

- [1] Schober, P., Boer, C. and Schwarte, L.A. (2018) Correlation Coefficients: Appropriate Use and Interpretation. *Anesthesia & Analgesia*, **126**, 1763-1768. <https://doi.org/10.1213/ane.0000000000002864>
- [2] Shrestha, N. (2021) Factor Analysis as a Tool for Survey Analysis. *American Journal of Applied Mathematics and Statistics*, **9**, 4-11. <https://doi.org/10.12691/ajams-9-1-2>
- [3] Chen, T. and Guestrin, C. (2016) XGBoost. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794. <https://doi.org/10.1145/2939672.2939785>
- [4] Browne, M.W. (2000) Cross-Validation Methods. *Journal of Mathematical Psychology*, **44**, 108-132. <https://doi.org/10.1006/jmps.1999.1279>