安徽农业大学学报, 2026, 53(4): 663-672
Journal of Anhui Agricultural University [DOI] 10.13610/j.cnki.1672-352x.20260915.006

陈冠玉，崔玉环. 基于可解释集成学习的农业流域地表水总氮浓度预测及关键驱动因子识别：以杭埠河流域为例[J]. 安徽农业大学学报, 2026, 53(4): 663-672.

Chen Guanyu, Cui Yuhuan. Interpretable ensemble learning-based prediction of surface-water total nitrogen concentration and identification of key driving factors in an agricultural watershed: a case study of the Hangbu River Watershed[J]. Journal of Anhui Agricultural University, 2026, 53(4): 663-672.

基于可解释集成学习的农业流域地表水总氮浓度预测及关键驱动因子识别：以杭埠河流域为例

陈冠玉，崔玉环*

(安徽农业大学资源与环境学院，安徽 合肥 230036)

摘 要：【目的】为提高农业流域地表水总氮（total nitrogen，TN）浓度预测精度并识别关键驱动因子，以巢湖重要支流杭埠河流域为研究对象开展分析。【方法】基于 2024 年 5 月—2025 年 9 月研究区 62 个采样点的监测数据，共采集 558 份水样；经缺失值筛查后，保留 535 个有效样本，并构建包含气象、水分、植被、土壤和景观格局信息的解释变量数据集。比较岭回归（ridge）、支持向量回归（support vector regression，SVR）、随机森林（random forest，RF）、极端随机树（extremely randomized trees，ET）和极端梯度提升（extreme gradient boosting，XGBoost）等模型，并进一步构建基于折外预测（out-of-fold，OOF）的 Stacking 集成模型；采用 SHapley Additive exPlanations（SHAP）方法识别关键驱动因子。【结果】不同模型对杭埠河流域地表水 TN 浓度的预测效果存在明显差异，其中 Stacking 模型的测试集平均决定系数（R^2）、均方根误差（RMSE）和平均绝对误差（MAE）分别为 0.810 6 ± 0.047 6、0.897 4 ± 0.096 4 和 0.625 6 ± 0.033 3，该模型在 R^2、RMSE 及结果稳定性方面表现较优。SHAP 分析表明，归一化植被指数（NDVI）、温度、降水、蒸散量和土壤湿度是影响 TN 浓度预测的主要因子。【结论】杭埠河流域地表水 TN 浓度可通过集成学习方法实现较高精度预测，其预测结果与植被状态、气象条件及流域水分环境密切相关。研究结果可为流域水质预测、关键变量识别及农业面源污染防控提供参考。

关键词：农业流域；总氮浓度；关键驱动因子；可解释集成学习；SHAP

中图分类号：TP18；X52 **文献标志码：**A **文章编号：**1672−352X (2026)04−0663−10

Interpretable ensemble learning-based prediction of surface-water total nitrogen concentration and identification of key driving factors in an agricultural watershed: a case study of the Hangbu River Watershed

Chen Guanyu, Cui Yuhuan

(School of Resources and Environment, Anhui Agricultural University, Hefei 230036, China)

Abstract: [**Objective**] To improve the prediction accuracy of surface-water total nitrogen (TN) concentration in agricultural watersheds and identify key driving factors, the Hangbu River Watershed, an important tributary of Chaohu Lake, was selected as the study area. [**Method**] Based on monitoring data collected from 62 sampling sites in the study area from May 2024 to September 2025, a total of 558 water samples were obtained. After screening for missing values, 535 valid samples were retained, and a dataset containing explanatory variables related to meteorology, water conditions, vegetation, soil, and landscape pattern was constructed. Ridge regression (Ridge), support vector regression (SVR), random forest (RF), extremely randomized trees (ET), and extreme gradient boosting (XGBoost) were compared, and an out-of-fold (OOF)-based Stacking ensemble model was further developed. SHapley Additive exPlanations (SHAP) was used to identify key driving factors. [**Result**] The prediction performance of different models for surface-water TN concentration in the Hangbu River Watershed

收稿日期：2026-03-10；**录用日期：**2026-05-08；**网络出版日期：**2026-09-16

基金项目：安徽省自然科学基金项目（2308085MD114）；国家自然科学基金项目（32171573）；安徽省高校自然科学研究项目（2023AH051002）

作者简介：陈冠玉，硕士研究生，E-mail：guanyuchen@stu.ahau.edu.cn

* **通信作者：**崔玉环，博士，副教授，E-mail：cuiyh@ahau.edu.cn

varied considerably. Among them, the Stacking model achieved an average test coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE) of 0.810 6 ± 0.047 6, 0.897 4 ± 0.096 4, and 0.625 6 ± 0.033 3, respectively, showing relatively better performance in terms of R^2, RMSE, and result stability. SHAP analysis indicated that normalized difference vegetation index (NDVI), temperature, precipitation, evapotranspiration, and soil moisture were the major factors affecting TN concentration prediction. [**Conclusion**] Surface-water TN concentration in the Hangbu River Watershed can be predicted with relatively high accuracy using ensemble learning methods, and its variation is mainly driven by vegetation status and meteorological-hydrological conditions. The results can provide a reference for watershed water-quality prediction, identification of key driving factors, and control of agricultural non-point source pollution.

Keywords: agricultural watershed; total nitrogen concentration; key driving factors; interpretable ensemble learning; SHAP

农业流域地表水总氮（total nitrogen，TN）是表征水体营养盐负荷水平、评估富营养化风险及水环境质量变化的重要指标之一[1-2]。在农业流域中，TN 的形成、迁移与输出受气象水文条件、植被覆盖变化、土壤理化性质以及土地利用与景观格局等多因素共同驱动，通常表现出显著的季节波动和空间异质性[2-3]。尤其是在农业活动强度较高的区域，耕地分布格局、景观破碎化程度及降水径流过程的变化，会影响氮素的来源输入、地表迁移和河道汇输效率，从而改变河流 TN 的时空分布特征[3-4]。因此，开展农业流域地表水 TN 浓度预测及关键驱动因素识别，对于揭示流域氮素迁移规律、优化监测布设、识别重点源区及支撑分区控污具有重要的科学意义和管理价值。

近年来，随着多源环境数据的积累与计算方法的发展，机器学习因其较强的非线性拟合能力和多变量处理能力，在水质预测领域得到广泛应用[5]。已有研究表明，支持向量机、随机森林、梯度提升树等模型在河流、湖泊及河口等水体水质参数预测中均表现出较好的适用性[5-6]。与此同时，可解释人工智能方法也逐步被引入环境模型研究，其中 SHapley Additive exPlanations（SHAP）可用于量化输入变量对模型预测结果的贡献程度，从而增强模型解释性并提升研究结果的可用性[7]。总体来看，现有研究已为复杂环境条件下的水质模拟与驱动因子解析提供了较为丰富的方法基础。

然而，农业流域地表水 TN 变化过程具有多源驱动、强耦合和高异质性的特点，现有研究仍存在一些不足：一是单一机器学习模型虽能较好拟合非线性关系，但其预测结果易受样本划分、噪声扰动及变量相关性的影响，模型稳定性和泛化能力仍有待提高；二是部分研究更侧重于预测精度的提升，对模型结果的环境解释与关键驱动因子识别关注不足；三是针对农业流域这类受自然过程与人类活动共同影响且样本规模相对有限的研究对象，如何在保证预测性能的同时兼顾模型稳健性与可解释性，仍是有待进一步解决的问题[8-9]。

针对上述问题，本文以巢湖重要支流杭埠河流域为研究对象，整合气象水文、植被、土壤及景观格局等多源解释变量，构建农业流域地表水 TN 浓度预测与关键驱动因子识别框架。在借鉴农业流域地表水营养盐机器学习预测研究的基础上[9]，比较岭回归（ridge）、支持向量回归（support vector regression，SVR）、随机森林（random forest，RF）、极端随机树（extremely randomized trees，ET）和极端梯度提升（extreme gradient boosting，XGBoost）等模型，并进一步构建基于折外预测（out-of-fold，OOF）的 Stacking 集成模型，结合 SHAP 方法识别关键驱动因子及其相对贡献。本文重点关注以下问题：杭埠河流域地表水 TN 浓度是否能够实现较高精度预测；哪些环境因子对 TN 变化起主导作用；集成学习方法是否能够提升预测结果的综合性能与稳健性；研究结果可为农业流域水质预测、关键驱动因子识别及面源污染管理提供参考。

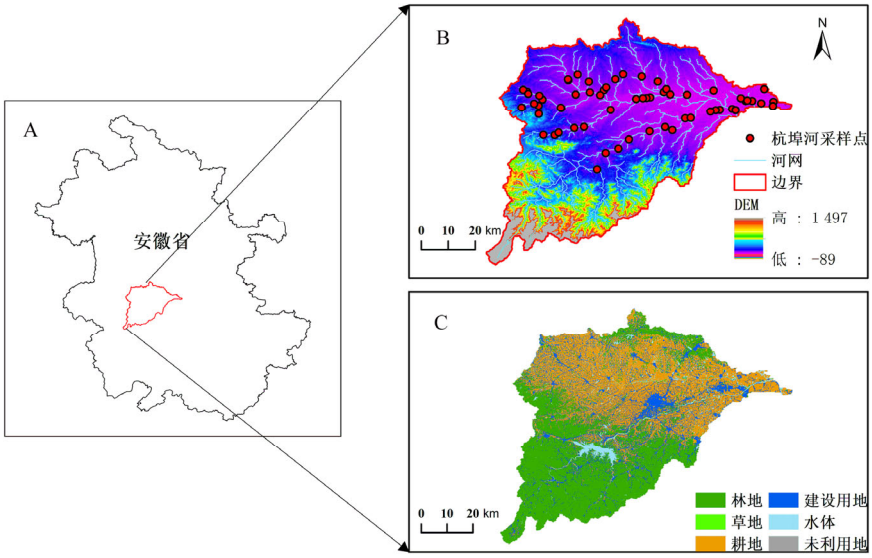
1 材料与方法

1.1 研究区域与样本点设计

杭埠河流域位于安徽省中西部（116°26′～117°15′E，31°01′～31°34′N），流域面积约 4 246 km^2，隶属长江水系巢湖流域，行政范围涉及岳西、舒城、霍山及六安市金安区，以及合肥市肥西、庐江等地（图 1）。杭埠河源自大别山岳西县主簿塬，流经舒城县晓天镇、山七镇，最终汇入巢湖，干流长度约 70.12 km。

流域地势总体呈“西南高、东北低”格局：山区海拔可达 1 539 m，中部丘陵区约 400 m，东北部平原区多低于 100 m；成土母质以残积、坡积物为主，主要土壤类型为黄棕壤与水稻土。流域属北亚

热带湿润季风气候，年均降水量约 1 350.5 mm、年均气温约 15.6 ℃；水资源较丰富，流域内分布有龙河口水库等多座水库，年径流波动明显。作为巢湖重要入湖补给区，流域来水量约占巢湖总水量的 30%，主要支流包括丰乐河、龙潭河、南港河与清水河等。为兼顾水质的空间差异与季节变化，本研究在杭埠河流域布设 62 个地表水采样点，并开展 9 期现场监测（2024-05-18、2024-07-07、2024-09-25、2024-11-12、2025-01-09、2025-03-24、2025-05-20、2025-07-09、2025-09-16），共采集 558 份水样；采样点位分布见图 1B。采样后水样立即密封冷藏并送实验室分析；TN 采用碱性过硫酸钾消解—紫外分光光度法测定（HJ 636-2012）[10]。使用 Trimble GPS（GeoXH 2008）记录各采样点经纬度，并为每个点位分配唯一编号（ID）以关联不同监测期的重复观测。



底图来源于安徽省行政区划矢量数据，未作修改，审图号为 GS（2024）0650 号。

A.研究区在安徽省内的位置；B.研究区地形高程、河网及采样点分布；C.研究区土地利用类型分布。

图 1　研究区位置及环境背景特征

Figure 1　Location and environmental background characteristics of the study area

表 1　预测变量、含义、单位与数据来源

Table 1　Predictive variables, definitions, units, and data sources

类别	符号	含义	单位	数据源	原始分辨率/时间尺度	本研究时间尺度/处理
水文气象	T	气温（空气温度）	℃	Google Earth Engine（GEE）气象数据集	逐日/逐时	月尺度汇总；提取前重采样/配准
水文气象	P_{mm}	降水量	mm	Google Earth Engine（GEE）气象数据集	逐日/逐时	月尺度汇总；提取前重采样/配准
水文气象	ET	蒸散量	mm	GEE 平台蒸散产品（与 T、P 一致处理）	逐日/逐时	月尺度汇总；提取前重采样/配准
水文气象	SM	土壤湿度	$m^3 \cdot m^{-3}$	Google Earth Engine（GEE）气象数据集	逐日/逐时	月尺度汇总；提取前重采样/配准
植被	NDVI	归一化植被指数	—	Sentinel-2 Level-2A（L2A）地表反射率（GEE）	多光谱影像	云/阴影剔除后按月中值合成
土壤	CEC	阳离子交换量	$cmol(+) \cdot kg^{-1}$	中国土壤数据集（第二版），国家青藏高原科学数据中心	约 90 m	静态；重采样至分析网格
景观	Cropland	耕地面积占比	%	Dynamic World 土地利用（GEE）	土地利用栅格	按月或与监测期匹配；计算后重采样/配准
景观	MESH	有效网格面积（破碎化）	km^2	Dynamic World 派生景观指标（GEE 计算）	土地利用栅格	按月或与监测期匹配；计算后重采样/配准
景观	FRAC_MN	平均分形维数	—	Dynamic World 派生景观指标（GEE 计算）	土地利用栅格	按月或与监测期匹配；计算后重采样/配准
景观	IJI	交错与并置指数	%	Dynamic World 派生景观指标（GEE 计算）	土地利用栅格	按月或与监测期匹配；计算后重采样/配准

1.2 数据来源与预测因子

为揭示杭埠河流域地表水 TN 的时空异质性及其主要环境驱动特征，本文整合多源数据构建预测因子体系，涵盖水文气象过程、植被生长状态、土壤理化性质以及景观格局特征 4 个方面（表 1）。其中，水文气象与植被指标用于表征流域水热条件及其季节波动，土壤理化性质反映下垫面对氮素吸附、缓冲与迁移转化的基础背景，景观格局指标则用于刻画子流域土地利用组成与空间结构对面源氮输出的潜在调控作用[11-12]。

为确保多源信息具有可比性并与水质监测期严格对齐，本研究对随时间变化的变量统一在月尺度上进行聚合与匹配；对栅格类数据在属性提取前进行空间一致化处理（投影统一、像元对齐与重采样）。随后，以子流域为基本空间单元汇总景观与环境因子，并与采样点（或其所属子流域）的 TN 观测值进行关联，形成用于模型训练与解释分析的预测因子数据集[13]。

1.3 技术路线

本研究的技术路线如图 2 所示。首先获取杭埠河流域 TN 监测数据并构建多源环境预测因子体系，随后进行数据预处理与模型构建，通过多模型比较确定最优模型，并结合 SHAP 方法识别 TN 变化的关键变量及其相对贡献。

1.4 数据预处理

为保证模型构建的可靠性，首先对建模涉及的目标变量 TN、月份变量及解释变量进行数值化处理，并对数据完整性进行检查。研究共采集 558 份水样，对目标变量 TN 及 10 个解释变量（T、ET、NDVI、CEC、Pmm、SM、Cropland、MESH、FRAC_MN 和 IJI）进行缺失值筛查，剔除上述字段中任一项存在缺失值的样本后，最终保留 535 个有效样本用于模型训练与分析。鉴于各解释变量量纲和取值范围存在差异，采用 Min–Max 归一化方法对输入特征进行处理，以降低量纲差异对模型训练的影响。为避免测试集信息泄露，全部预处理过程均仅基于训练集拟合，并将所得变换规则应用于测试集。

1.5 模型构建

为比较不同模型在流域地表水 TN 预测中的适用性，本研究构建了由单一基线模型与集成模型组成的比较框架，包括 Ridge、SVR、RF、ET、XGBoost 及 Stacking 集成模型，其中 Stacking 模型采用 OOF 策略构建。

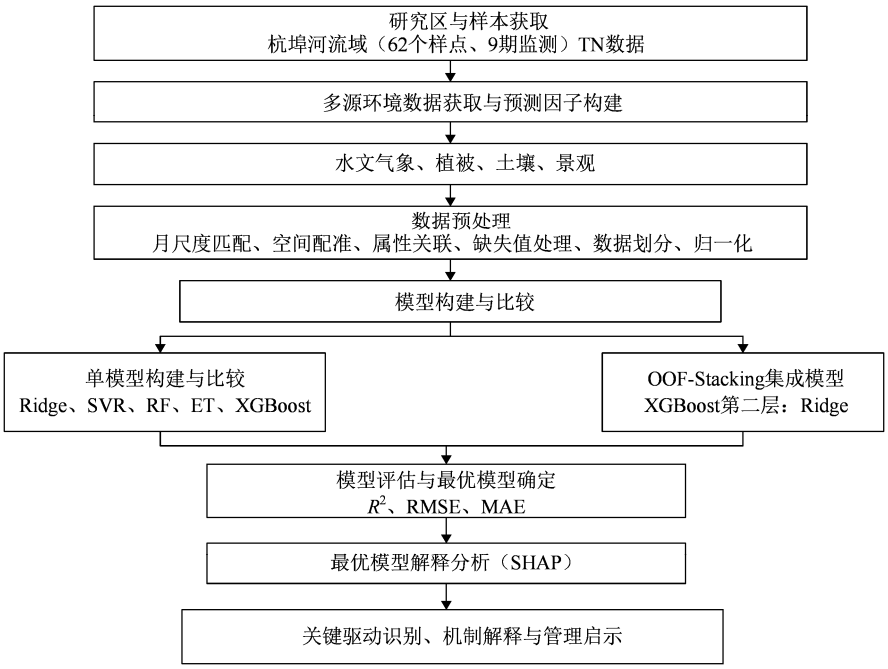


图 2 技术路线图

Figure 2 Technical framework of the study

1.5.1 基线模型 本文选取 Ridge、SVR、RF、ET 和 XGBoost 作为基线模型，以比较不同类型算法在农业流域地表水 TN 预测中的适用性。上述模型分别代表线性回归、核方法、Bagging 树模型和 Boosting 树模型，能够从不同角度表征 TN 与多源环境因子之间的复杂关系。其中，Ridge 用于提供线性基线，SVR 适用于高维、小样本和非线性问题，RF 通过多树集成提高模型稳健性，ET 在随机树框

架下进一步增强随机性并兼顾训练效率与泛化能力，XGBoost 则基于梯度提升策略，能够较好刻画复杂非线性关系及局部特征[14-15]。

1.5.2　Stacking 集成模型　在基线模型比较的基础上，本研究进一步构建基于折外预测（out-of-fold，OOF）的 Stacking 集成模型，以发挥不同学习器之间的互补优势（图 3）。基学习器的选择主要遵循 3 项原则：预测性能较优、建模机制互补以及在重复随机划分条件下表现相对稳定。比较结果表明，RF、ET 和 XGBoost 均能够较好刻画 TN 与多源环境因子之间的非线性关系，且分别代表 Bagging 树模型、强随机化树模型和 Boosting 树模型三类不同的集成学习机制，具有较好的结构互补性。相比之下，SVR 虽具有一定预测能力，但与树模型的互补优势不够突出，且在多次随机划分条件下未表现出显著优于 RF 的稳定增益，因此未纳入第一层集成。第二层采用 Ridge 回归作为元学习器，主要考虑其形式简单、抗共线性能力较强，能够在样本量相对有限的条件下对多个基学习器输出进行稳定整合，从而提高模型整体稳健性。为避免信息泄露，本研究采用 OOF-Stacking 策略：在每个随机种子下，首先按 8∶2 划分训练集和测试集，并基于月份变量进行分层抽样。随后在训练集内部进行 5 折分层交叉验证，生成训练集 OOF 元特征，并对测试集预测结果取平均以构建测试集元特征；最终由 Ridge 元学习器输出测试集预测结果。进一步地，采用 5 个随机种子进行重复评估，并以测试集性能指标的均值和标准差作为最终结果，以提高评价的稳健性。

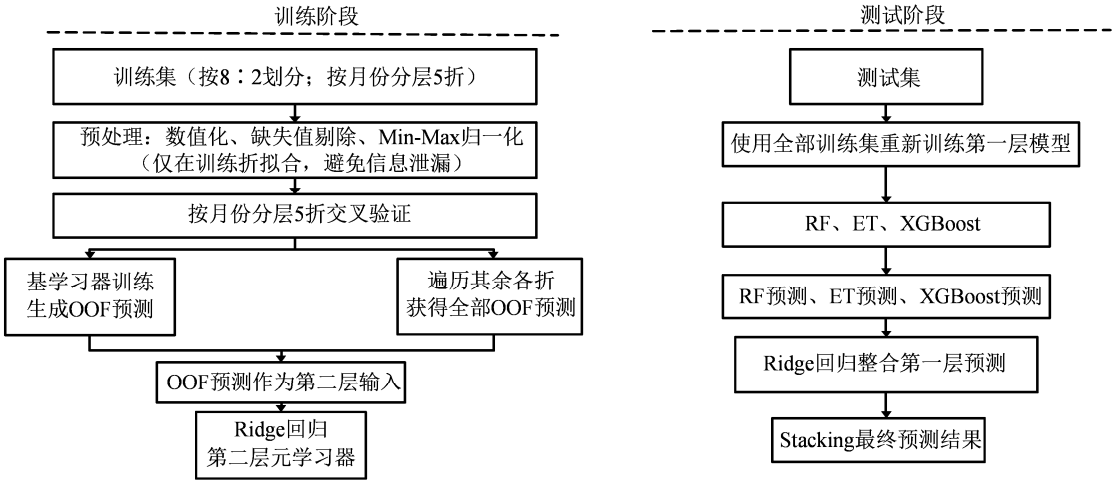


图 3　OOF-Stacking 集成预测框架示意图

Figure 3　Schematic diagram of the OOF-Stacking ensemble prediction framework

1.6　模型评估与解释

为综合评价模型性能，本文采用决定系数（R^2）、均方根误差（RMSE）和平均绝对误差（MAE）作为评价指标。其中，R^2 用于衡量模型对 TN 变化的解释能力，RMSE 用于表征预测值与观测值之间的整体偏差，对较大误差更为敏感，MAE 则反映预测误差的平均绝对水平。3 项指标从拟合优度和误差大小两个方面对模型预测能力进行综合评价。设第 i 个样本的观测值为 Y_i，预测值为 $\hat{Y}_i$，样本量为 n，观测均值为 $\bar{Y}$，则 3 项指标定义如下：

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2} \quad (1)$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|Y_i - \hat{Y}_i| \quad (2)$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2} \quad (3)$$

为提高模型结果的可解释性，本文采用 SHAP 方法对最优模型进行解释分析[7]。具体利用 SHAP 全局重要性条形图识别对模型预测贡献较大的关键变量，并结合蜜蜂群图（beeswarm plot）展示各特征在不同取值条件下对模型输出的影响方向及影响强度，从而揭示主要环境因子对 TN 变化的相对贡献及其潜在非线性作用关系。

2　结果与分析

2.1　TN 的时空变异特征

图 4 展示了研究期内杭埠河流域地表水 TN 浓度的时序箱线分布特征。总体来看，不同采样期 TN 浓度差异明显，表现出较强的时间波动性。2024 年 5 月、7 月、9 月和 11 月各期 TN 浓度整体较低，

箱线图中位数均处于较低水平，且箱体相对较窄，表明该阶段 TN 浓度总体偏低、波动幅度有限。进入 2025 年后，TN 浓度整体升高，其中 2025 年 1 月中位数最高，离散程度也相对较大，说明该时期流域 TN 水平显著上升且样点间差异增强。随后，2025 年 3 月、5 月、7 月和 9 月 TN 浓度较 2025 年 1 月有所回落，但总体仍高于 2024 年同期水平。不同采样期箱体高度、上下须长度及离群点数量存在差异，进一步表明 TN 浓度在年际和季节尺度上均具有明显的动态变化特征。

从空间分布来看，杭埠河流域 TN 浓度具有显著的空间异质性，且不同采样期的高值区与低值区分布格局存在明显变化。2024 年各期整体以低值为主，仅在局部区域出现零散高值斑块，空间差异相对较弱；而 2025 年 1 月和 3 月流域大部分区域 TN 浓度整体升高，高值区范围明显扩大，显示出较强的面状分布特征。2025 年 5 月以后，TN 高值范围有所收缩，空间格局由大范围偏高逐步转为局部差异主导，但部分区域仍持续表现出相对较高或较低浓度水平，说明 TN 空间分布在动态变化的同时也具有一定稳定性。总体而言，杭埠河流域地表水 TN 浓度在时间上表现为先低后高、随后波动调整的变化趋势，在空间上则表现出由局部斑块差异向较大范围高值分布、再向局部化分异演变的特征[16]。

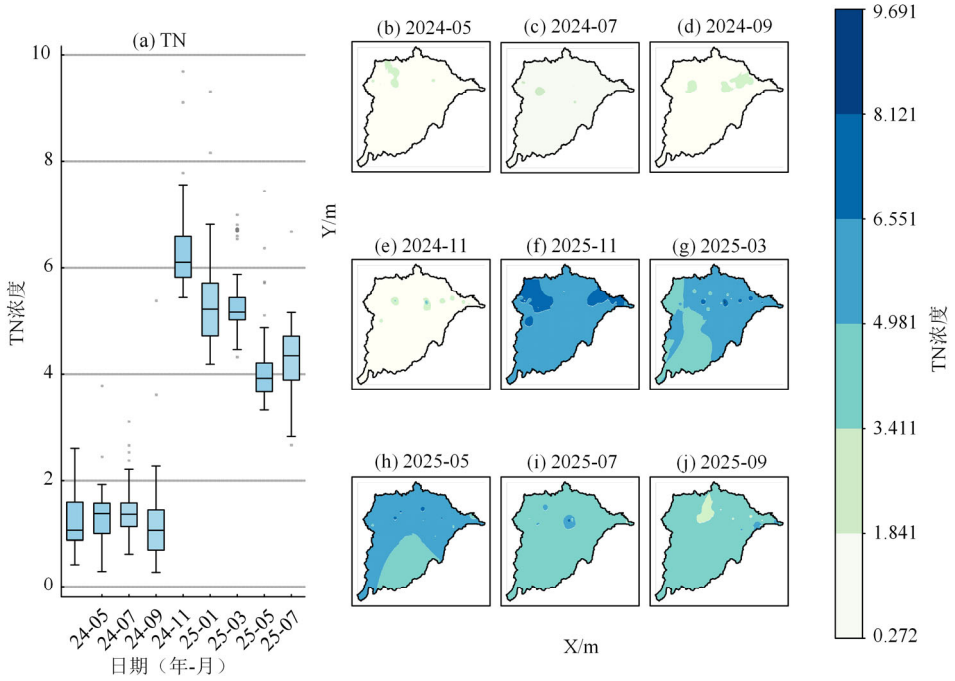


图 4 杭埠河流域地表水 TN 浓度的时间变化箱线图及空间分布特征

Figure 4 Temporal boxplots and spatial distribution characteristics of TN concentration in surface water of the Hangbu River Basin

2.2 模型性能比较

表 2 展示了不同模型在测试集上的平均预测性能（均值±标准差）。为降低单次随机划分带来的偶然性，本研究采用 5 个随机种子对不同模型进行重复评估，并以测试集性能指标的均值及标准差表征模型表现。总体来看，不同模型在 TN 预测中的表现存在明显差异，其中树模型及 Stacking 集成模型整体优于线性模型，说明非线性建模框架更有利于刻画多源环境因子与 TN 之间的复杂关系。

表 2 不同模型在测试集上的平均预测性能比较（均值±标准差）

Table 2 Comparison of average predictive performance of different models on the test set (mean ± standard deviation)

模型	平均测试集 R^2	平均测试集 RMSE	平均测试集 MAE
OOF-Stacking（RF+ET+XGBoost，Meta=Ridge）	0.810 6 ± 0.047 6	0.897 4 ± 0.096 4	0.625 6 ± 0.033 3
RF	0.796 6 ± 0.071 3	0.922 9 ± 0.138 3	0.598 8 ± 0.061 8
SVR	0.791 9 ± 0.019 0	0.947 7 ± 0.048 5	0.700 6 ± 0.038 7
ET	0.717 7 ± 0.015 8	1.105 5 ± 0.066 1	0.882 3 ± 0.039 3
XGBoost	0.653 2 ± 0.033 4	1.223 2 ± 0.062 7	1.006 3 ± 0.039 2
Ridge	0.276 8 ± 0.008 3	1.769 4 ± 0.081 8	1.550 5 ± 0.059 2

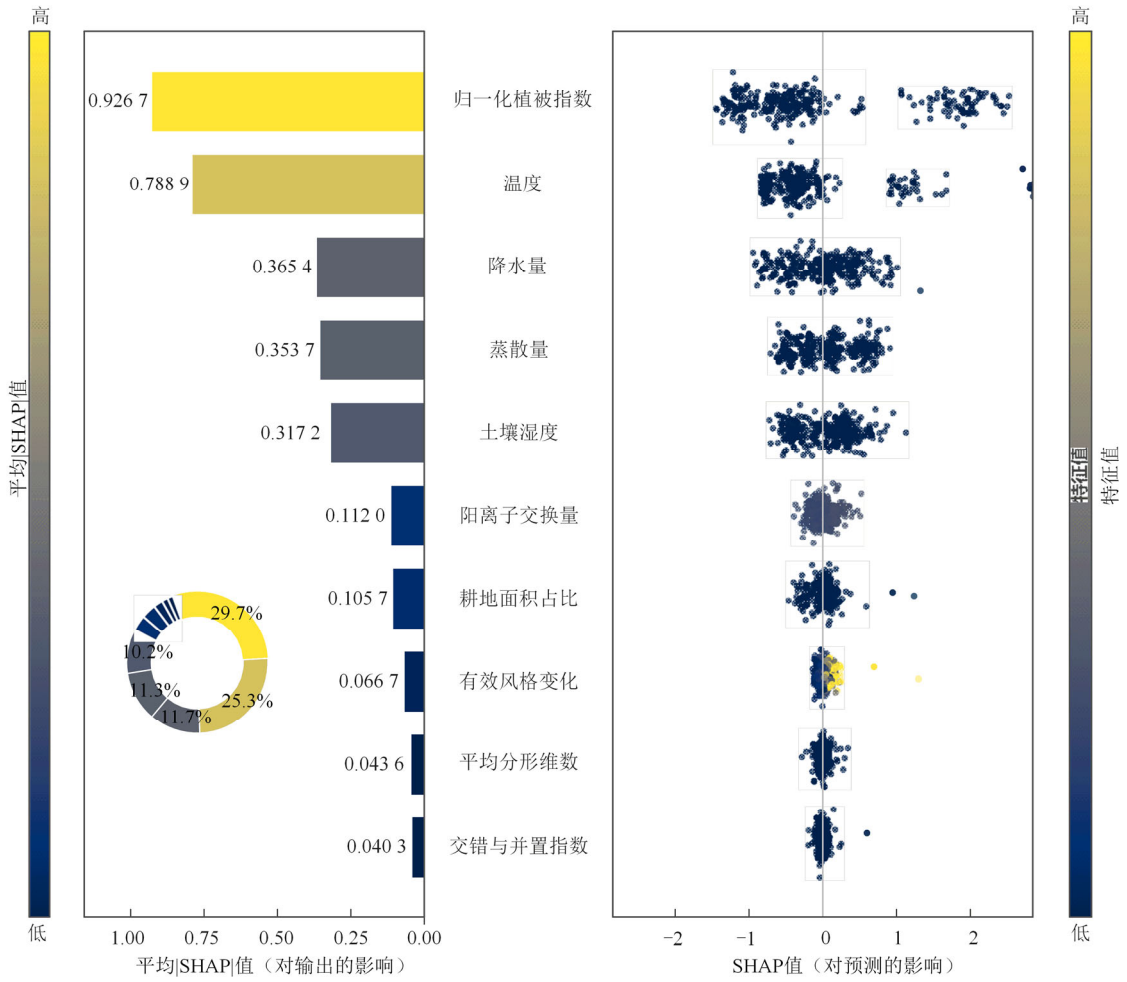


图 5　最优 Stacking 集成模型的 SHAP 特征重要性及局部解释结果

Figure 5　SHAP feature importance and local explanation results of the optimal Stacking ensemble model

结果表明，OOF-Stacking 集成模型在多随机种子重复评估下表现出较好的综合性能。以 RF、ET 和 XGBoost 作为基学习器、以 Ridge 回归作为元学习器构建的 Stacking 模型，其平均测试集 R^2、RMSE 和 MAE 分别为 0.810 6 ± 0.047 6、0.897 4 ± 0.096 4 和 0.625 6 ± 0.033 3。相比之下，单模型中 RF 表现相对最优，平均测试集 R^2 为 0.796 6 ± 0.071 3，RMSE 和 MAE 分别为 0.922 9 ± 0.138 3 和 0.598 8 ± 0.061 8；SVR 次之，平均测试集 R^2 为 0.791 9 ± 0.019 0。ET 和 XGBoost 的预测精度相对较低，而 Ridge 模型表现最弱，表明线性模型难以充分表征 TN 与多源环境因子之间的复杂非线性关系。

进一步以综合表现最优的单模型 RF 为参照进行比较可见，Stacking 模型的平均测试集 R^2 由 0.796 6 提高至 0.810 6，增加 0.014 0（1.76%）；RMSE 由 0.922 9 降低至 0.897 4，减少 0.025 5（2.76%）；而 MAE 则由 0.598 8 上升至 0.625 6，增加 0.026 8（4.48%）。这表明，Stacking 模型在拟合优度和整体误差控制方面优于最优单模型 RF，但在 MAE 指标上并未进一步改善。另一方面，Stacking 模型在 5 次随机种子重复评估中的 R^2、RMSE 和 MAE 标准差分别较 RF 下降 33.2%、30.3%和 46.1%，说明其结果波动更小、对数据划分的敏感性更低，模型稳定性更强。综合来看，Stacking 模型在 R^2、RMSE 及结果稳定性方面表现更优，但在平均绝对误差控制方面并未全面占优，因此本文将其界定为综合性能最优模型，而非所有单项指标均最优的模型。考虑到 RMSE 对较大误差更为敏感，而 MAE 反映整体平均误差水平，Stacking 模型在 RMSE 上的优势说明其对较大偏差具有更强的抑制能力；而 MAE 略高于 RF 则表明，其在平均绝对偏差控制方面仍有进一步优化空间。

2.3　基于 SHAP 的驱动因子归因分析

为进一步揭示最优模型预测结果背后的关键驱动因素，本文基于前文测试集表现最优的 Stacking 集成模型开展 SHAP 解释分析。图 5 给出了主要环境变量的全局重要性排序及其局部影响分布。结果表明，NDVI、温度和降水量是影响 TN 预测结果的

前 3 位驱动因子，其平均绝对 SHAP 值分别为 0.926 7、0.788 9 和 0.365 4，显著高于其他变量，说明这三类因子对模型输出贡献最大。蒸散量和土壤湿度的重要性紧随其后，平均绝对 SHAP 值分别为 0.353 7 和 0.317 2，也表现出较强的解释作用。相比之下，阳离子交换量、耕地面积占比、有效网格面积、平均分形维数和交错与并置指数等变量的平均绝对 SHAP 值相对较低，对模型预测的边际贡献较小。总体来看，在模型解释层面，杭埠河流域地表水 TN 预测结果主要与植被状态和气象因子相关，而土壤属性及景观格局因子的影响相对有限[17]。

从 beeswarm 图展示的局部解释结果来看，不同驱动因子对 TN 的影响方向和作用强度存在明显差异。NDVI 对应的 SHAP 值分布范围最宽，且高取值样本更多集中于正向贡献区域，表明较高 NDVI 往往对应更高的 TN 预测值，是影响模型输出的首要驱动因子。温度也表现出较强影响，其高取值样本整体偏向正向 SHAP 区域，说明升温条件下 TN 浓度更可能升高。降水量的 SHAP 值分布于零值两侧，反映出其作用方向并不单一，即在不同环境背景下既可能促进 TN 增加，也可能表现为抑制效应，说明降水过程对氮素迁移、汇流输送与稀释具有复杂影响。蒸散量和土壤湿度的 SHAP 值同样呈现一定的双向波动，但总体分布较前 3 位因子更集中，说明其对 TN 变化具有较稳定但相对次要的影响。阳离子交换量、耕地面积占比及各景观格局指标的 SHAP 值大多分布在零值附近，表明这些变量对模型预测的贡献相对较弱。总体而言，SHAP 结果表明，杭埠河流域 TN 变化受到多因子共同控制，其中 NDVI、温度和降水量是最主要驱动因子，且不同因子的作用过程具有明显非线性特征[17]。

3 讨论

3.1 Stacking 集成模型平均性能提升的原因

本研究表明，在多随机种子重复评估条件下，Stacking 集成模型的平均测试性能优于 RF、SVR 等单模型，说明多模型融合能够在一定程度上提升流域地表水 TN 浓度预测的综合表现。这一优势主要来源于不同学习器在建模机理上的互补性。RF 属于典型的 Bagging 树模型，能够较好捕捉复杂非线性关系并保持较强稳健性；ET 在树分裂过程中引入更强随机性，有助于降低方差；XGBoost 则基于 Boosting 框架，对残差学习和局部复杂模式具有较强表达能力。不同学习器对样本特征的响应方式存在差异，经过两层集成后可在一定程度上实现误差互补，从而提升模型平均性能[8]。

同时，本文采用 OOF-Stacking 训练策略，使元学习器基于训练集内部相对独立的基学习器预测结果进行构建，在避免信息泄露的同时提高模型训练的规范性。需要指出的是，Stacking 相较于最优单模型 RF 的提升幅度有限，说明在当前样本规模和变量结构下，树集成模型本身已具备较强的预测能力，而更复杂的二层集成所带来的改进主要体现在平均意义上的性能优化，而非数量级上的显著跃升。

3.2 关键变量的环境意义解释

SHAP 结果表明，杭埠河流域地表水 TN 预测结果与植被状态、气象条件及流域水分环境密切相关。需要说明的是，SHAP 结果主要反映各变量对模型预测输出的相对贡献，并不等同于严格意义上的因果关系识别，但仍可为理解 TN 变化与环境条件之间的关系提供参考[18-19]。

植被状态对 TN 预测结果具有重要指示意义。NDVI 不仅反映流域植被覆盖状况，还可在一定程度上表征农田生长季变化及地表拦截能力。植被变化会影响坡面径流形成、泥沙输移及氮素迁移过程，从而间接影响河流水体 TN 浓度。考虑到本研究区属于典型农业流域，NDVI 的变化还可能与作物生长节律及农业管理活动同步，因此其在本研究中更多体现为农业流域地表覆盖变化及季节性水文响应的综合指示变量[18-19]。

气象条件同样是影响 TN 变化的重要因素。温度升高可能加快有机氮矿化、硝化和反硝化等过程，并通过改变微生物活性和植被生长状况影响氮素转化。降水则具有明显的双重作用：既可能通过地表径流和坡面冲刷促进氮素向河道输移，也可能因来水增加产生一定的稀释作用，因此其影响方向具有较强的情景依赖性。特别是在农业流域或高强度耕作区域，极端降水和降雨增强往往会加剧硝态氮淋失和河流氮输出[20-23]。

流域水分条件对TN迁移过程也具有一定调节作用。较高的土壤湿度可能增强产流和氮素输出，并改变土壤氧化还原条件；蒸散量则在一定程度上综合表征区域水热状况及植被耗水过程。相较之下，部分土壤和景观格局变量在本研究中更多体现为背景性影响，这可能与月尺度数据匹配方式及样本结构有关。已有研究表明，景观格局和地形条件对流域氮素输移作用通常具有多尺度性和阈值特征[24-25]。

3.3 对监测与流域管理的启示

本研究结果对流域水质监测优化和污染防控具有一定启示。首先，变量重要性排序表明，NDVI、

温度、降水量、蒸散量和土壤湿度等因子应作为 TN 监测与预警的优先关注对象。与均匀增加监测指标相比，围绕关键驱动因子开展针对性监测更有助于提高监测效率和风险识别能力。

其次，TN 在时间和空间上均表现出明显异质性，因此流域管理策略应突出时段差异与区域差异。对于高风险时期，应加强降水过程前后及水文条件快速变化阶段的监测与巡查；对于长期表现为相对高值的区域，应优先布设监测断面并强化污染源排查；对于农业活动较强的区域，可结合耕地分布和降水过程，实施更具针对性的农田氮肥管理和面源污染拦截措施[26]。

3.4　不确定性分析

本文结果仍存在一定不确定性。研究样本虽覆盖 9 期监测并保留 535 个有效样本，但对短时突发过程的表征仍可能不足。月尺度环境变量匹配有利于多源数据统一处理，但也可能平滑事件尺度的瞬时波动。尽管采用 5 个随机种子进行重复评估，数据划分方式和模型参数设置仍可能引起结果波动。此外，SHAP 解释反映的是统计归因关系，并不等同于严格的因果识别[27]。因此，本文对关键驱动因子的认识应理解为基于现有样本和变量体系条件下的阶段性判断。

3.5　局限性与未来展望

尽管本文构建的 OOF-Stacking 模型在多随机种子重复评估下取得了较好的平均预测性能，但仍存在一定局限。当前结果主要基于研究区内部随机划分条件下的模型比较，尚未开展独立外部数据集或外部流域验证，因此其跨区域泛化能力仍有待进一步检验。此外，受数据可获得性及时空配准条件限制，本文现有变量体系尚未完整纳入地形条件、人类活动强度及流量过程等直接过程变量，对复杂驱动过程及其环境响应的表征仍有进一步完善空间。

未来研究可从 3 个方面进一步拓展：一是采用留一法（leave-one-month-out）、跨年验证等更严格的时序泛化检验方式；二是通过增加监测频次和空间点位，提高模型对复杂时空异质性的表征能力；三是结合地形、水动力和人类活动代理变量，并考虑气候变化背景下极端事件增强情景，进一步提升模型对 TN 迁移转化过程的表达能力，并评估模型框架在不同区域背景下的可推广性[28-29]。

4　结论

1）本研究围绕杭埠河流域地表水 TN 浓度预测及关键驱动因子识别，对多种机器学习模型进行了比较，并在此基础上构建了基于 OOF 策略的 Stacking 集成模型。结果表明，该模型在综合预测性能与稳定性方面表现较好，可为农业流域水质预测与关键驱动因子识别提供参考。

2）杭埠河流域地表水 TN 浓度具有明显的时空变异特征。不同采样期 TN 水平及离散程度存在显著差异，空间上高值区和低值区分布范围随时间发生变化，表明 TN 变化同时具有较强的时间波动性和空间异质性。

3）SHAP 归因结果表明，NDVI、温度和降水量是影响杭埠河流域地表水 TN 预测的重要变量，蒸散量和土壤湿度次之，而土壤属性与景观格局因子的整体贡献相对较弱。相关结果表明，TN 预测结果与植被状态、气象条件及流域水分环境密切相关。

4）研究结果可为流域 TN 监测优化与污染防控提供参考，即应优先关注关键驱动因子的动态变化，并结合不同时段和不同区域的差异实施分区分类管理。未来可进一步开展 leave-one-month-out 验证、跨年验证及外部流域迁移验证，以提高模型的时序泛化能力和区域适用性。

参考文献：

[1] Conley D J, Paerl H W, Howarth R W, et al. Controlling eutrophication: nitrogen and phosphorus[J]. Science, 2009, 323(5917): 1014-1015.

[2] 嵇晓燕, 王姗姗, 杨凯, 等. 2016—2020 年中国地表水中总氮浓度时空变化特征分析[J]. 生态环境学报, 2022, 31(6): 1184-1192.

[3] 秦先燕, 吴剑雄, 郑涛, 等. 店埠河流域地表水氮素时空分布和来源解析[J]. 环境科学, 2024, 45(12): 7100-7110.

[4] 陈传保. 杭埠河舒城段水质现状评价[J]. 安徽农业科学, 2011, 39(16): 9774-9775.

[5] 徐洁, 耿伯彰, 娄英斌, 等. 基于 LSTM 模型的地表水总氮预测模型[J]. 中国环境监测, 2025, 41(2): 63-68.

[6] Zhang K, Wang X Y, Liu T L, et al. Enhancing water quality prediction with advanced machine learning techniques: an extreme gradient boosting model based on long short-term memory and autoencoder[J]. J Hydrol, 2024, 644: 132115.

[7] Lundberg S M, Lee S I. A unified approach to interpreting model predictions[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. ACM, 2017: 4768-4777.

[8] Ruan J L, Cui Y, Song Y C, et al. A novel RF-CEEMD-LSTM model for predicting water pollution[J]. Sci Rep, 2023, 13: 20901.

[9] Elsayed A, Rixon S, Levison J, et al. Machine learning models for prediction of nutrient concentrations in surface

water in an agricultural watershed[J]. J Environ Manag, 2024, 372: 123305.

[10] 水质总氮的测定碱性过硫酸钾消解紫外分光光度法: HJ 636—2012[S].

[11] Pei W, Xu Q Y, Lei Q L, et al. Interactive impact of landscape composition and configuration on river water quality under different spatial and seasonal scales[J]. Sci Total Environ, 2024, 950: 175027.

[12] Liu H Y, Meng C, Wang Y, et al. Establishing the relationship between the integrated multidimensional landscape pattern and stream water quality in subtropical agricultural catchments[J]. Ecol Indic, 2021, 127: 107781.

[13] Xu S, Li S L, Zhong J, et al. Spatial scale effects of the variable relationships between landscape pattern and water quality: Example from an agricultural karst river basin, Southwestern China[J]. Agric Ecosyst Environ, 2020, 300: 106999.

[14] Chen T Q, Guestrin C. XGBoost: a scalable tree boosting system[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2016: 785-794.

[15] Li Z Q. Extracting spatial effects from machine learning model using local interpretation method: an example of SHAP and XGBoost[J]. Comput Environ Urban Syst, 2022, 96: 101845.

[16] Rixon S, Levison J, Binns A, et al. Spatiotemporal variations of nitrogen and phosphorus in a clay plain hydrological system in the Great Lakes Basin[J]. Sci Total Environ, 2020, 714: 136328.

[17] Yuan Y, Zhou C J, Wu J W, et al. An interpretable deep learning framework for river water quality prediction: a case study of the Poyang Lake Basin[J]. Water, 2025, 17(16): 2496.

[18] Newcomer M E, Bouskill N J, Wainwright H, et al. Hysteresis patterns of watershed nitrogen retention and loss over the past 50 years in United States hydrological basins[J]. Glob Biogeochem Cycles, 2021, 35(4): e2020GB006777.

[19] Xu Y T, Li P, Ma F M, et al. Watershed landscape characteristics and connectivity drive river water quality under seasonal dynamics[J]. J Clean Prod, 2024, 473: 143533.

[20] Lu C Q, Zhang J E, Tian H Q, et al. Increased extreme precipitation challenges nitrogen load management to the Gulf of Mexico[J]. Commun Earth Environ, 2020, 1: 21.

[21] Ballard T C, Sinha E, Michalak A M. Long-term changes in precipitation and temperature have already impacted nitrogen loading[J]. Environ Sci Technol, 2019, 53(9): 5080-5090.

[22] Hess L J T, Hinckley E S, Robertson G P, et al. Rainfall intensification increases nitrate leaching from tilled but not no-till cropping systems in the U.S. Midwest[J]. Agric Ecosyst Environ, 2020, 290: 106747.

[23] Zhu X Q, Miao P, Zhu H, et al. Extreme precipitation accelerates nitrate leaching in the intensive agricultural region with thick unsaturated zones[J]. Sci Total Environ, 2024, 918: 170789.

[24] Wu J H, He S J, Lu J. Multi-scale effects of topography and landscape pattern on riverine nitrogen and phosphorus nutrients in an agricultural watershed[J]. Landsc Ecol, 2025, 40(6): 112.

[25] Meng C, Liu H Y, Wang Y, et al. Landscape pattern exhibits threshold-driven effect on nitrogen export of typical land use in subtropical hilly watershed under specific hydrological regimes[J]. J Clean Prod, 2023, 419: 138322.

[26] Lyu C J, Li X J, Yuan P, et al. Nitrogen retention effect of riparian zones in agricultural areas: a meta-analysis[J]. J Clean Prod, 2021, 315: 128143.

[27] Takefuji Y. Limitations of SHAP-based interpretations in environmental and membrane filtration applications[J]. Water Res, 2026, 288: 124766.

[28] IPCC Intergovernmental Panel on Climate Change. Climate Change 2022 - Mitigation of Climate Change: Working Group III Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change[M]. Cambridge: Cambridge University Press, 2023.

[29] Zhang J E, Lu C Q, Feng H L, et al. Extreme climate increased crop nitrogen surplus in the United States[J]. Agric For Meteor, 2021, 310: 108632.