环境科学
(HUANJING KEXUE)
ENVIRONMENTAL SCIENCE
第44卷 第7期
2023年7月15日

目 次

Eco-Environmental Knowledge Web

环 境 科 学
Environmental Science

第44卷第7期 2023年7月
Vol.44, No.7 Jul., 2023

基于机器学习的长江流域农田氮径流流失负荷估算

张育福[1,2], 潘哲祺[1], 陈丁江[1,2,3]*

(1. 浙江大学环境与资源学院, 杭州 310058; 2. 浙江大学环境修复与生态健康教育部重点实验室, 杭州 310058; 3. 浙江省农业资源与环境重点实验室, 杭州 310058)

摘要: 定量解析长江流域农田氮径流流失特征是实现长江及其河口氮污染有效控制的关键科学基础. 基于收集的长江流域570个旱地和434个水田田间氮径流流失数据组,采用相关性分析、结构方程模型、方差分解和机器学习方法,探究了影响旱地和水田总氮径流流失强度的主要因素,建立了基于机器学习的长江流域旱地和水田总氮径流流失强度预测模型,量化了长江流域农田总氮径流流失负荷. 结果表明,径流深、施氮量和土壤氮含量是影响旱地总氮径流流失强度的主要因素; 径流深和施氮量是水田总氮径流流失强度的主要影响因素. 与分类与回归树、多元线性回归和增强回归树方法相比,采用随机森林算法构建的长江流域旱地和水田总氮径流流失强度预测模型具有更高的精度(R^2 为 0.65 ~ 0.94). 基于随机森林算法的预测模型估算的2013年长江流域农田总氮径流流失负荷(以N计)为 0.47 $Tg \cdot a^{-1}$(旱地: 0.25 $Tg \cdot a^{-1}$; 水田: 0.22 $Tg \cdot a^{-1}$),中下游地区贡献了58%的流失负荷. 模型预测5种防治情景下的长江流域农田氮流失负荷可削减2.4% ~ 9.3%,其中减少径流量的削减效果最为显著. 长江流域农田氮面源污染防治必须协同加强氮肥精准管理、减少农田径流量和提高土壤氮利用,且应将重点放在中下游地区. 所发展的基于机器学习建模方法克服了氮径流流失强度与影响因素之间函数关系难以确定的问题,为估算区域或流域农田氮流失负荷提供了简便且可靠的方法.

关键词: 长江流域; 农田; 径流; 氮流失; 机器学习; 面源污染

中图分类号: X522 **文献标识码**: A **文章编号**: 0250-3301(2023)07-3913-10 **DOI**: 10.13227/j.hjkx.202208129

Estimation of Cropland Nitrogen Runoff Loss Loads in the Yangtze River Basin Based on the Machine Learning Approaches

ZHANG Yu-fu[1,2], PAN Zhe-qi[1], CHEN Ding-jiang[1,2,3]*

(1. College of Environmental & Resource Sciences, Zhejiang University, Hangzhou 310058, China; 2. Ministry of Education Key Laboratory of Environment Remediation and Ecological Health, Zhejiang University, Hangzhou 310058, China; 3. Zhejiang Provincial Key Laboratory of Agricultural Resource and Environment, Zhejiang University, Hangzhou 310058, China)

Abstract: A quantitative understanding of cropland nitrogen (N) runoff loss is critical for developing efficient N pollution control strategies. Using correlation analysis, a structural equation model, variance decomposition, and machine learning methods, this study identified the primary influencing factors of total N (TN) runoff loss from uplands (n =570) and paddy (n =434) fields in the Yangtze River Basin (YRB) and then developed a machine learning-based prediction model to quantify cropland N runoff loss load. The results indicated that runoff depth, soil N content, and fertilizer addition rate were the major influencing factors of TN runoff loss from uplands, whereas TN runoff loss rate from paddy fields was mainly regulated by runoff depth and fertilizer addition rate. Among the four used machine learning methods, the prediction models based on the random forest algorithm presented the highest accuracy (R^2 =0.65-0.94) for predicting upland and paddy field TN runoff loss rates. The random forest algorithm based model estimated a total cropland TN loss load in the YRB of 0.47 $Tg \cdot a^{-1}$ (upland: 0.25 $Tg \cdot a^{-1}$; paddy field: 0.22 $Tg \cdot a^{-1}$) in 2013, with 58% of TN runoff loss load derived from the midstream and downstream regions. The models predicted that TN runoff loss loads from croplands in YRB would decrease by 2.4%-9.3% for five scenarios, with higher TN load reductions occurring from scenarios with decreased runoff amounts. To mitigate cropland N nonpoint source pollution in YRB, it is essential to integrate efficient water, fertilizer, and soil nutrient managements as well as to consider the midstream and downstream regions as the high priority area. The machine learning-based modeling method developed in this study overcame the difficulty of identifying the functional relationships between cropland TN loss rate and multiple influencing factors in developing relevant prediction models, providing a reliable method for estimating regional and watershed cropland TN loss load.

Key words: Yangtze River Basin; cropland; runoff; nitrogen loss; machine learning; nonpoint source pollution

河流、湖库和河口等水体过量活性氮引发的富营养化、藻类暴发和死氧区等问题已成为当前国内外水环境治理领域普遍关注的焦点之一[1]. 有研究表明,农田氮径流流失是造成许多地区地表水体氮过量的重要原因[2]. 定量解析流域尺度农田氮径流流失特征是实现水体氮污染高效控制的关键科学基础.

流域/区域农田氮径流流失负荷的估算主要依赖于田间监测和模型模拟方法. 显然,由于农田氮径流流失存在明显的时空异质性[3,4],因此,基于少数典型田块监测结果难以刻画流域/区域尺度农田氮径流流失特征. 尽管输出系数法、SWAT 和 SPARROW 等数学模型已被广泛应用于流域/区域尺度农田氮径流流失强度/负荷估算、关键源区识别等,但在实际的应用中仍面临较大挑战[5~7].

收稿日期: 2022-08-15; **修订日期**: 2022-09-24

基金项目: 浙江省重点研发计划项目(2019C02047); 国家自然科学基金项目(41877465, 42177352); 国家重点研发计划项目(2021YFD1700802)

作者简介: 张育福(1995 ~),男,博士,主要研究方向为农业非点源氮磷污染,E-mail: zhangyufu@zju.edu.cn

* 通信作者,E-mail: chendj@zju.edu.cn

SWAT 等分布式模型往往需要大量水文水质、土地利用方式和污染源等数据进行充分校正和验证,导致很多流域因数据不足而难以准确应用[8]. 输出系数法等集总式模型由于往往缺乏实际应用流域/区域的关键参数值(常采用已有研究报道值),导致估算结果存在较大的不确定性[9]. 近年来,国内外学者采用荟萃分析方法估算流域/区域农田总氮径流流失负荷,即基于收集文献报道的研究流域/区域农田氮流失监测数据,通过相关统计分析方法,建立基于多影响因素的农田氮径流流失强度多元统计模型,从而实现流域/区域的农田氮流失负荷估算[3,10,11]. 然而,由于农田氮径流流失强度的影响因素众多且各影响因素之间存在自相关性,导致难以科学确定各影响因素与农田氮径流流失强度之间的多元函数关系,使得估算结果可能存在较大的不确定性. 机器学习作为一种融合了统计学、概率和代数等多种数学方法的科学算法[12],通过反复训练学习从大量候选结果中筛选出自变量与因变量之间的最佳函数关系,克服了常规多元统计分析中自变量与众多因变量之间函数关系难以确定的难题[13]. 因此,机器学习方法已被广泛应用于生物信息学、生物化学、水产养殖和气候学等领域的模拟预测[14]. 然而,目前尚鲜见应用机器学习方法估算流域/区域农田氮径流流失负荷的研究报道.

近十几年以来,长江流域水体及河口富营养化、缺氧和藻华等问题频发[15,16]. 据报道,由于氮含量过高,2000 ~ 2010 年,长江口及其邻近沿海水域发生了 500 多起的藻华事件[17]. 农业面源污染已成为长江流域河流、湖库和河口水体氮过量的主要原因[18]. 尽管不少学者应用输出系数模型[7]、数据驱动的升级模型[10]和荟萃分析方法[3]估算了长江流域农业面源氮污染负荷,但是由于缺少足够数据对模型进行有效校验,导致以往研究估算的长江流域面源氮污染负荷(以 N 计,下同)存在较大的不确定性($0.54 \sim 3.40\ Tg \cdot a^{-1}$)[3,7,19,20]. 因此,需要进一步发展更加精准的长江流域农田氮径流流失强度/负荷方法.

本研究利用 1998 ~ 2017 年报道的长江流域农田径流氮流失数据,识别影响旱地和水田氮径流流失的主要因素,构建基于机器学习的长江流域旱地和水田径流氮流失强度预测模型,估算旱地和水田总氮径流流失负荷及空间分布特征,预测不同情景下农田总氮流失负荷的定量变化,从而提出相应的管控对策建议. 研究结果提升了对长江流域农田氮径流流失的定量认识,以期为推进农业面源污染防治提供科学支撑.

1 材料与方法

1.1 研究区概况

长江流域(90° ~ 122°E, 24° ~ 35°N)是我国面积最大的流域(面积为 180 万 km^2,图 1),横跨我国东部、中部和西部地区的三大经济区. 流域气候以亚热带季风为主,多年平均降水量为1 045 mm,大部分降水集中在夏季(5 ~ 9 月). 由于空间跨度广、地形复杂,流域年降雨量存在较强的时空异质性,流域多年平均降雨量从上游到下游介于 859 ~ 1 528 mm 之间[21]. 长江流域耕地面积占全国耕地面积的 27%,施用的肥料占全国的 32%,流域人口约为全国人口的 35%[3]. 长江流域旱地以种植蔬菜、玉米、小麦、油菜和棉花等作物为主,水田则主要种植水稻(表 1).

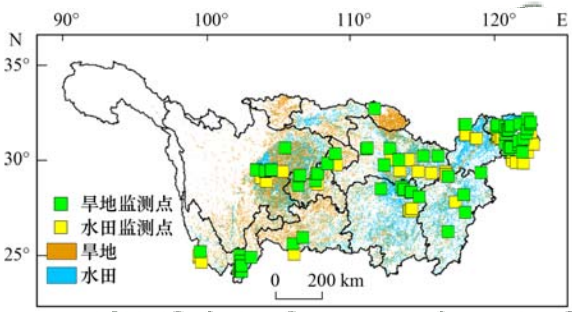


图 1 长江流域旱地和水田径流氮流失监测点分布

Fig. 1 Monitoring sites for nitrogen runoff loss from upland and paddy fields in the Yangtze River Basin

表 1 长江流域水田和旱地种植作物类型及其收集的氮流失强度数据组数

Table 1 Planted crop types of upland and paddy fields as well as data numbers of collected nitrogen loss rates in the Yangtze River Basin

空间分布	土地利用类型	作物种植类型	数据组数
上游	旱地	蔬菜	21
		玉米	24
		小麦	5
		油菜	5
		小麦和油菜轮作	27
		小麦和玉米轮作	72
		其它	83
	水田	水稻	48
中游	旱地	蔬菜	12
		玉米	45
		小麦和油菜轮作	5
		小麦和玉米轮作	4
		油菜和棉花轮作	10
		其它	51
	水田	水稻	82
下游	旱地	蔬菜	90
		玉米	11
		棉花	5
		小麦和玉米轮作	1
		小米和油菜轮作	3
		其它	96
	水田	水稻	304

1.2　数据来源

本研究以“氮流失”“径流”“旱地”“水田”“农田”为关键词，运用 Web of Science、Google Scholar 和中国知网等数据库检索获取长江流域农田总氮径流流失强度数据资料. 收集的文献数据不包含综述、实验室模拟和流域模拟等文献数据，缺失的数据通过查阅文献补充. 最终收集了来自 160 篇文献的1 004个长江流域农田总氮径流流失强度数据组（旱地和水田监测点数分别为 93 个和 73 个，图 1 和表 1）. 数据指标包括总氮径流流失强度、施氮强度、土壤理化性质、耕作管理措施和试验现场状况等信息.

农田径流氮流失预测所需数据包括土壤氮含量、施氮强度、径流深、土地利用和耕地面积（图 1 和图 2）. 土壤氮含量数据来自联合国粮农组织[22] 和 He 等[23] 的研究；施氮量数据来源于中华人民共和国国家统计局[24]；径流数据利用 SCS-CN 模型估算得到[25]，估算径流所需的数据来源于中国气象数据网[26]、中国水资源公报[27]、中国水利年鉴[28]、Zhou 等[29] 和 Wang 等[30] 的研究；土地利用数据来自中国科学院资源环境科学与数据中心[31]；耕地面积数据来自中国统计年鉴[32]. 由于收集的径流深数据最近的年份为 2013 年，因此，本研究将 2013 年作为长江流域农田氮径流流失负荷的估算年.

1.3　基于机器学习算法的长江流域农田总氮径流流失强度预测模型构建方法

首先，基于收集的长江流域农田总氮径流流失强度数据，协同相关性分析、结构方程模型、方差分解方法分别识别影响旱地和水田总氮径流流失强度的主要因素；然后，运用随机抽样的方法，按照 75% 和 25% 的比例将收集的长江流域旱地和水田总氮径流流失强度数据组分别划分为训练集和测试集，分别构建基于各机器学习算法的长江流域旱地和水田总氮径流流失强度预测模型，并采用决定系数（R^2）、均方根误差（RMSE）和平均绝对误差（MAE）等指标评价各模型预测精度；最后，应用精度最佳的预测模型估算长江流域旱地和水田总氮径流流失负荷及空间分布特征.

本研究选取的建模方法为 4 种常用机器学习算法，即多元线性回归、随机森林、增强回归树和分类与回归树[33]. 多元线性回归是研究一个变量与多个解释变量的回归方法[33]. 随机森林是由多个决策树组成的森林的集成算法，其将每棵决策树都看作

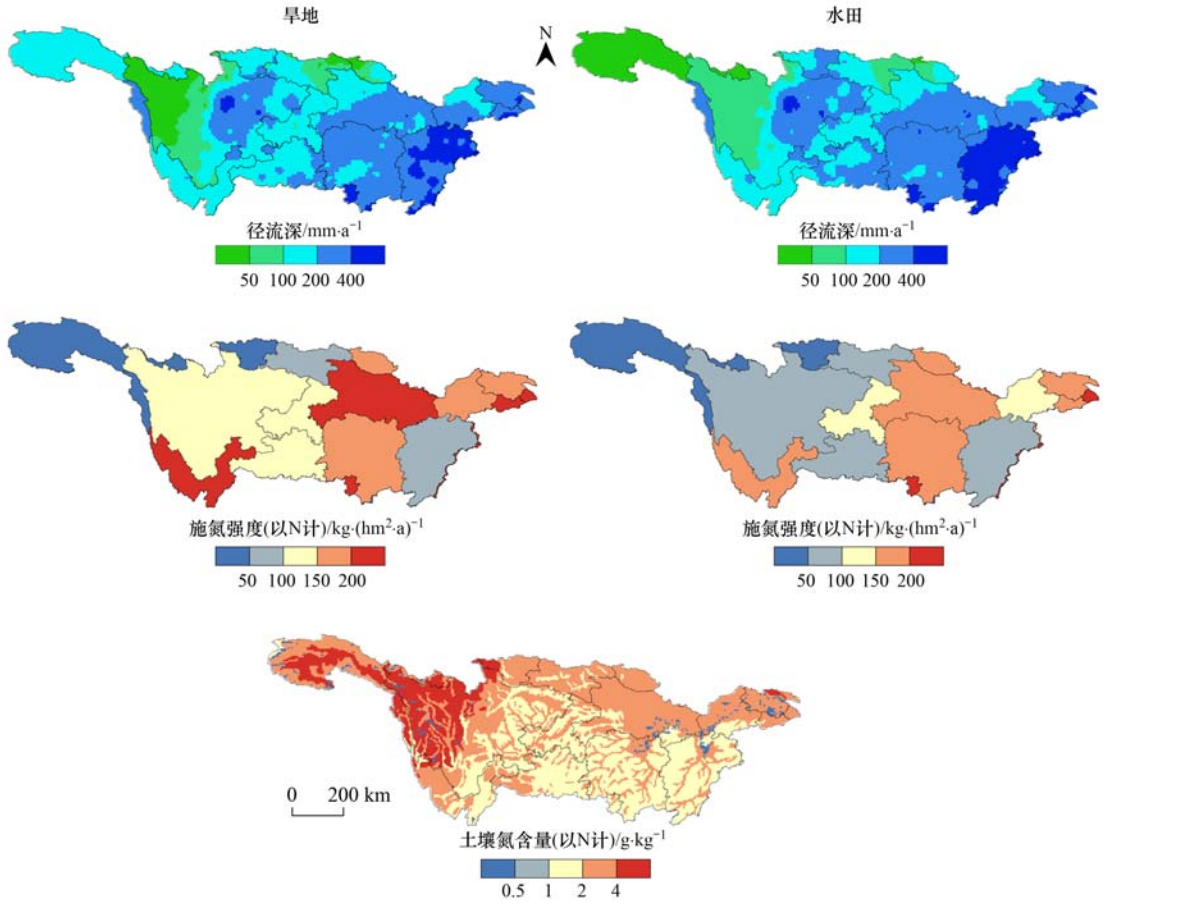


图 2　长江流域农田总氮流失预测所需数据

Fig. 2　Required data for prediction of total nitrogen loss from cropland in the Yangtze River Basin

是一个回归器,在所有的决策树输出相应的结果之后,投票选择出最合适的结果作为随机森林模型的最终预测,使用多个决策树共同完成学习任务,可以有效解决单一学习器训练结果不准确、容易过拟合等问题[34]. 增强回归树是包含分类回归树模型构建及其预测能力提升的集成算法,分类回归树是一种二元递归分解方法,预测能力提升是通过对多个回归树进行加权平均来进行预测以提高单个模型的预测精度[35]. 分类与回归树是一种典型的监督学习算法,其基本工作原理是对由自变量和因变量构成的训练集进行循环二分形成二叉树结构,主要采用 CINI 系数作为节点的分裂依据[36].

1.4 情景预测方法

本研究利用构建的长江流域农田总氮径流流失强度预测模型,预测了 5 种防治策略情景下的长江流域农田氮径流流失负荷. 已有的研究表明,施氮量、径流深和土壤氮含量是影响农田氮径流流失强度的主要因素[3,10,11]. 采用土壤-作物系统综合管理策略,可降低施氮量而不影响作物产量[37]; 采用节水灌溉措施可削减 6% ~ 36% 的径流量[38]; 通过实施套作/间作、适当增加作物种植密度等方式可增加作物对土壤氮的吸收利用[39],从而有效降低土壤氮含量. 因此,本研究设置了“减少 15% 化肥”“减少 15% 土壤氮”“减少 15% 径流深”“减少 15% 化肥和减少 15% 径流深”“减少 15% 土壤氮和减少 15% 径流深”等 5 种情景.

1.5 统计分析方法

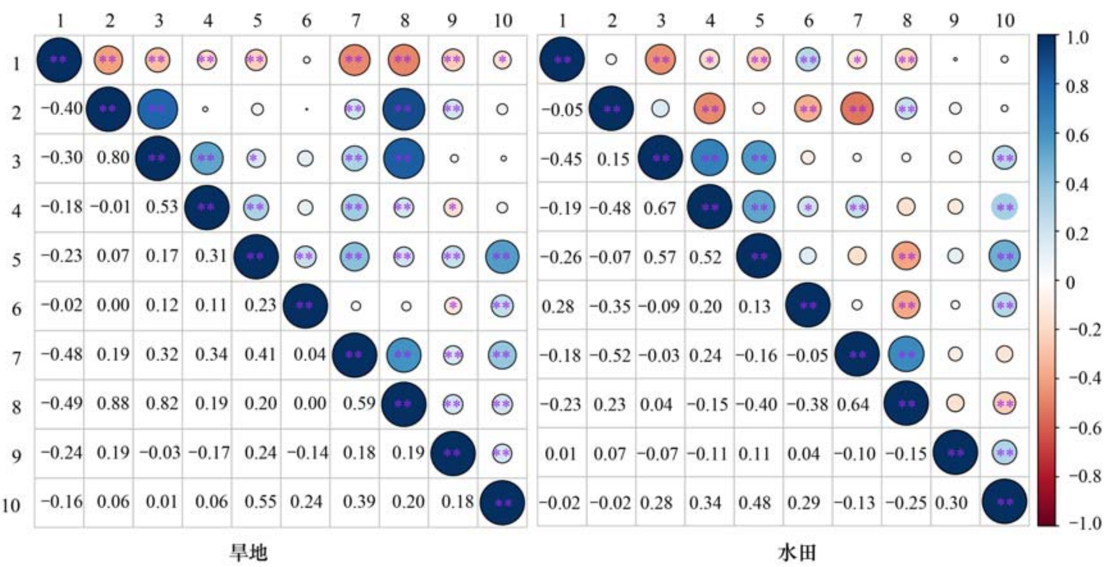
旱地和水田总氮径流流失强度与其影响因素之间的相关性分析采用 Pearson 相关系数法. 结构方程模型分析、方差分解分析和预测模型的构建均在 R 语言软件中进行,其中结构方程模型分析采用 lavaan 包,而方差分解分析和预测模型的构建均采用 caret 包. 所有统计检验采用 SPSS 18.0 软件(SPSS, Inc., USA)分析. 所有栅格计算均在 ArcGIS 10.2 软件(ESRI, Inc., USA)中进行. 各类图像采用 Excel 2017 (Microsoft, Inc., USA) 和 Prism 8.0 (GraphPad Software, Inc., USA)软件绘制.

2 结果与讨论

2.1 长江流域农田总氮径流流失强度的影响因素

收集的 1998 ~ 2017 年长江流域旱地和水田总氮径流流失强度数据(以 N 计,下同)[旱地:0.01 ~ 85.9 $kg\cdot(hm^2\cdot a)^{-1}$, $n = 570$; 水田:0.04 ~ 59.5 $kg\cdot(hm^2\cdot a)^{-1}$, $n = 434$]均呈现 3 个数量级变异性,意味着农田氮流失受到众多因素的影响. 相关性分析结果表明(图 3),旱地($r = 0.55$)和水田($r = 0.48$)总氮径流流失强度均与径流深呈显著($P < 0.01$)正相关关系,这是由于降雨和灌溉产生的径流是农田氮流失的主要驱动力[40]. 总氮径流流失强度与旱地($r = 0.24$)和水田($r = 0.29$)施氮量呈现显著正相关关系,表明肥料氮是农田径流氮流失的重要来源[3]. 旱地($r = 0.18$)和水田($r = 0.30$)氮径流流失强度与土壤黏粒含量($P < 0.05$)呈现显著正相关性,这可能是由于较高的黏粒含量限制了氮垂直淋溶且增加了地表径流量[41],从而促进了氮径流流失.

旱地总氮径流流失强度还与土壤氮含量($r =$



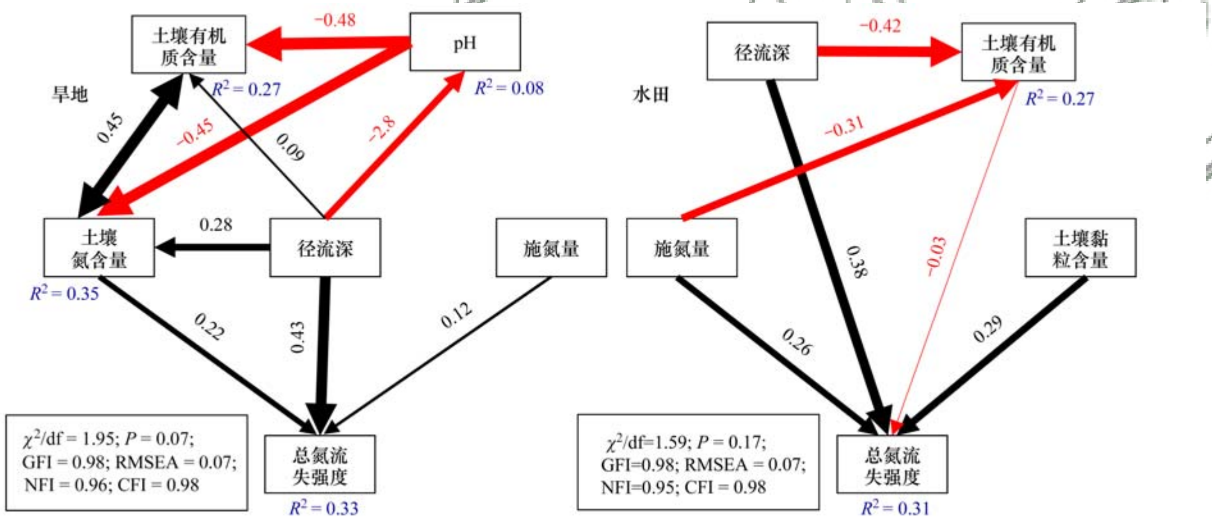
1. pH, 2. 碳氮比, 3. 碳磷比, 4. 氮磷比, 5. 径流深, 6. 施氮量, 7. 土壤氮含量, 8. 土壤有机质含量, 9. 土壤黏粒含量, 10. 总氮径流流失强度; 色柱表示相关系数 r; * 表示 $P < 0.05$, ** 表示 $P < 0.01$

图 3 长江流域旱地和水田总氮径流流失强度与影响因素的相关性

Fig. 3 Correlations between upland and paddy field total nitrogen runoff loss rates and relevant influencing factors in the Yangtze River Basin

0.39)、有机质含量(r=0.20)和 pH(r= −0.16)呈显著(P<0.05)相关关系,表明土壤遗留氮是旱地总氮径流流失的重要来源[42]. 旱地总氮径流流失强度与土壤 pH 的负相关关系(r= −0.16)可能是因为较低的 pH 会增强氮矿化作用酶的活性[43],从而增加氮的流失. 不同于旱地,水田总氮径流流失强度与土壤氮含量和 pH 无关(P>0.05). 但是,水田总氮流失强度与土壤有机质含量(r= −0.25)呈显著(P<0.05)负相关关系,这可能是由于厌氧条件会促进有机质在土壤中的累积,进而增强了反硝化作用[10,42]. 水田总氮径流流失强度与碳磷比(r=0.28)、氮磷比(r=0.34)呈显著(P<0.05)正相关关系,这表明土壤磷含量的增加会限制农田氮流失,这可能是由于土壤磷含量的增加会提高藻类多样性,从而增强土壤-水界面上微生物聚集体对氮的固定作用[44].

结构方程模型分析表明,旱地氮径流流失强度不仅受径流深、土壤氮含量和施氮量的直接影响,而且受径流深、土壤有机质含量和 pH 值的间接影响(图 4). 径流深在一定程度上可以反映降雨量大小[45],提高了土壤含水量,从而有利于土壤有机氮累积[46]. 另一方面,降雨还会引发水土流失使黏粒大量淋失,降低土壤对酸的缓冲性及土壤 pH,从而促进旱地氮径流流失[47]. 与旱地相似,水田总氮径流流失强度会受径流深、施氮量、土壤有机质含量和土壤黏粒含量的直接影响,受径流深和施氮量的间接作用(图 4). 径流深和施氮量对土壤有机质含量存在重要影响,从而间接影响水田氮径流流失强度(图 4). 由于化肥施用有利于有机质的矿化,削弱了土壤对氮的固持能力[48,49],从而促进氮的径流流失. 方差分解分析结果进一步表明(图 5),土壤有机质含量对旱地总氮径流流失变异性的贡献可被径流深和土壤氮含量解释,而氮磷比和碳磷比对水田总氮径流流失变异性的贡献可被径流深和施氮量解释. 以上相关性、结构方程模型和方差分解分析结



箭头上的数字表示标准化路径系数,黑色和红色箭头分别表示正、负效应,箭头粗细表示影响效应的大小;GFI 表示拟合优度指数,RMSEA 表示网络近似均方根误差,NFI 表示标准拟合指数,CFI 表示比较拟合指数

图 4　长江流域旱地和水田氮径流流失强度的影响因素作用路径

Fig. 4　Interaction paths of factors affecting total nitrogen runoff loss rates of upland and paddy fields in the Yangtze River Basin

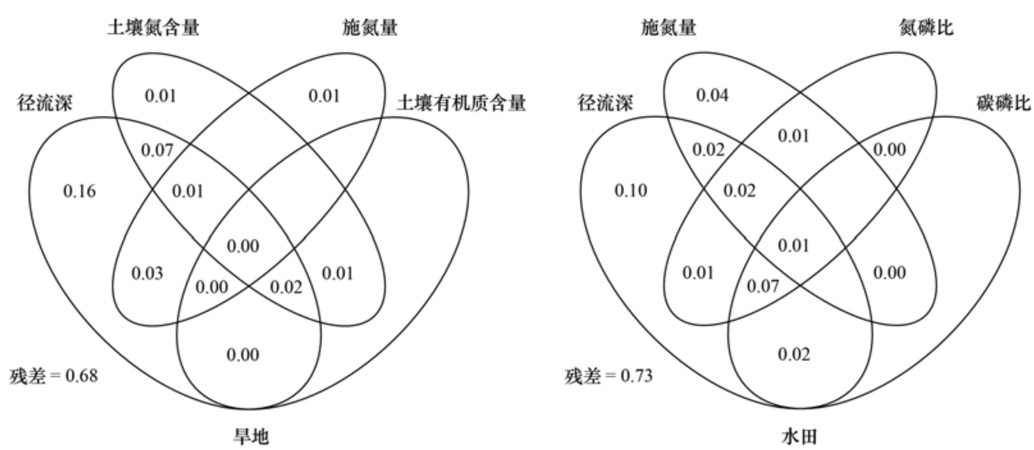


图 5　长江流域旱地和水田总氮径流流失强度影响因素的方差分解

Fig. 5　Variance decompositions of influencing factors on total nitrogen runoff loss rates from upland and paddy fields in the Yangtze River Basin

果表明,径流深、土壤氮和施氮量是旱地总氮径流流失强度的主要影响因素,而径流深和施氮量是水田总氮径流流失强度的主要影响因素.

2.2 基于机器学习的长江流域农田总氮径流流失强度预测模型构建

基于以上长江流域旱地和水田总氮径流流失强度的主要影响因素分析结果,构建并比较了基于随机森林、分类与回归树、增强回归树和多元线性回归的农田总氮径流流失强度预测模型.对比4种模型的预测精度,无论是在训练集还是测试集中随机森林算法都具有最高的预测精度(R^2 为0.65~0.94),增强回归树算法次之(R^2 为0.50~0.91),而分类与回归树(R^2 为0.05~0.55)和多元线性回归算法(R^2 为0.39~0.52)的预测精度较低(表2).已有的研究表明,农田总氮径流流失强度与影响因素之间存在非线性的响应关系[9].基于随机森林和增强回归树算法的模型预测精度较高,意味着随机森林和增强回归树算法对处理农田总氮径流流失强度与影响因素之间非线性关系的良好适用性[3].分类与回归树算法也呈现较低的精度可能是由于该方法对数据存在过度拟合情况[50].

表2 基于4种机器学习方法的长江流域农田总氮径流流失强度预测模型精度

Table 2 Accuracies of the four machine learning-based methods models for predicting total nitrogen runoff loss rates from upland and paddy field in the Yangtze River Basin

土地利用	数据集	机器学习方法	R^2	RMSE	MAE
旱地	训练集	随机森林	0.94	2.2	1.4
		分类与回归树	0.55	5.4	4.0
		多元线性回归	0.52	5.6	3.8
		增强回归树	0.91	2.4	1.7
	测试集	随机森林	0.65	5.2	3.2
		分类与回归树	0.46	6.4	4.7
		多元线性回归	0.39	6.9	4.4
		增强回归树	0.67	5.0	3.3
水田	训练集	随机森林	0.91	2.5	1.7
		分类与回归树	0.47	5.9	4.8
		多元线性回归	0.48	5.9	4.5
		增强回归树	0.74	4.2	3.0
	测试集	随机森林	0.68	5.0	3.1
		分类与回归树	0.05	8.7	6.3
		多元线性回归	0.39	6.6	4.8
		增强回归树	0.50	6.1	3.9

根据流域尺度面源污染的模型精度评价标准(R^2 >0.50表示模型精度可接受[51]),基于随机森林算法构建的模型(R^2 为0.65~0.94,RMSE为2.2~5.2)精度总体满意,能较为准确地预测长江农田总氮径流流失强度(图6).相比于以往采用多元非线性回归等方法构建的长江流域农田氮径流流失强度预测模型(R^2 为0.57~0.85,RMSE为4.9~9.1)[9~11],本研究构建的模型精度总体更高.基于机器学习的建模方法克服了农田氮径流流失强度与影响因素之间函数关系难以确定的问题,为区域或流域尺度的农田氮流失强度及负荷空间分布特征估算提供了较为可靠的方法.

2.3 长江流域农田氮径流流失负荷估算及情景预测

基于以上构建的基于随机森林算法的旱地和水田氮径流流失强度预测模型和预测所需数据(图2),估算的2013年长江流域农田总氮径流流失负荷为0.47 Tg·a^{-1}(旱地:0.25 Tg·a^{-1};水田:0.22 Tg·a^{-1}).以往研究估算的1980~2010年长江流域农田总氮径流流失总负荷介于0.54~3.4 Tg·a^{-1}之间[3,7,19,20],总体高于本研究估算的2013年结果,这可能与"测土配方施肥"政策的大面积推广有关,导致长江流域施氮量在2011~2017年期间减少了10%[3].此外,长江流域耕地面积减少也是导致2013年总氮径流流失负荷偏低的原因.1995~2017年期间,长江流域耕地面积减少了4.6%[3].

2013年长江流域农田平均总氮径流流失强度为12.4 kg·(hm^2·a)$^{-1}$[旱地:13.5 kg·(hm^2·a)$^{-1}$;水田:11.4 kg·(hm^2·a)$^{-1}$].旱地[2.3~28.1 kg·(hm^2·a)$^{-1}$]和水田[1.5~23.6 kg·(hm^2·a)$^{-1}$]总氮径流流失强度均存在较大的空间变异性(图7).由于长江中下游较高的年均降雨量(上游:1 045 mm;中下游:1 429 mm)[3],农田总氮径流流失热点区域主要分布在中游和下游地区(中游和下游地区旱地和水田平均总氮径流流失强度分别比上游的高40%和7%,图7).具体而言,广西、浙江、湖南、广东的旱地和广西、湖南、福建、上海的水田的总氮

流失强度显著高于其他省份(图 8). 长江上游地区的旱地面积大约占全流域旱地面积的62%,贡献了54%的旱地总氮流失总负荷,而中游和下游地区分别贡献了36%和10%的总氮流失负荷. 上游、中游和下游地区的水田分别贡献了24%、41%和35%的总氮流失负荷. 因此,应将中下游地区(如广西、浙江、湖南、广东的旱地和广西、湖南、福建、上海的水田)作为长江流域农田总氮流失防治的重点区域.

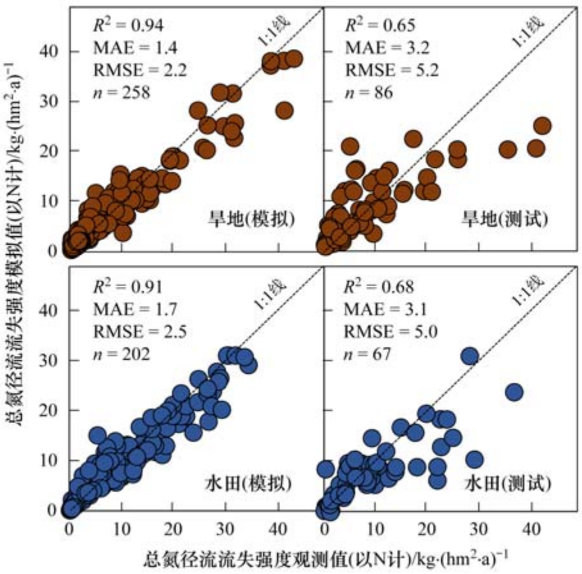


图 6　基于随机森林算法的长江流域旱地和水田总氮径流流失强度预测模型模拟值与观测值比较

Fig. 6　Comparisons of simulated and measured total nitrogen runoff loss rates from upland and paddy field in the Yangtze River Basin using the random forest algorithm based model

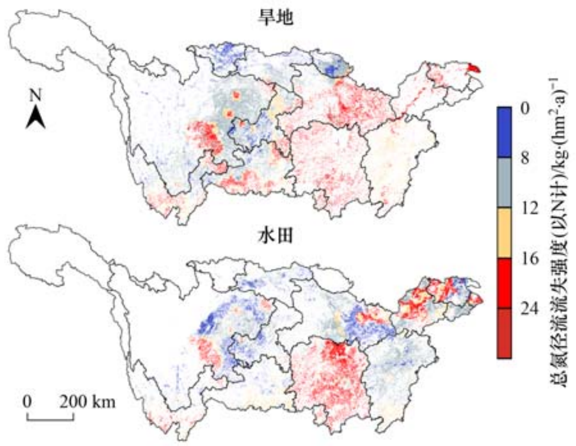


图 7　2013 年长江流域旱地和水田总氮径流流失强度空间分布特征

Fig. 7　Spatial distribution of total nitrogen loss rate via surface runoff in the Yangtze River Basin in 2013

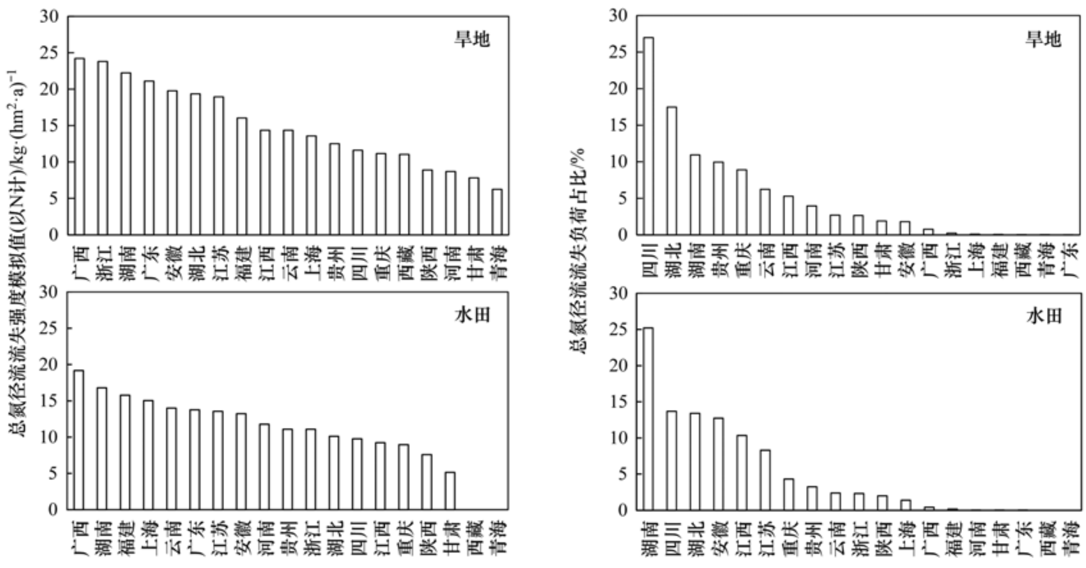
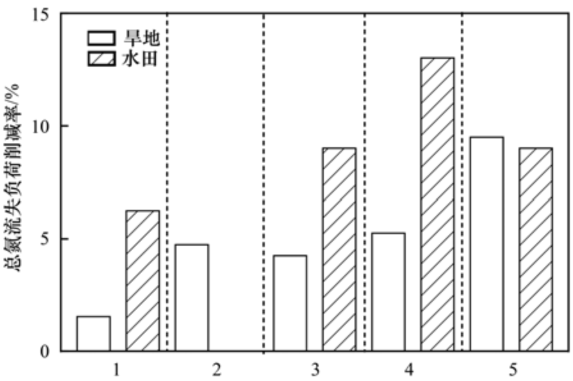


图 8　2013 年长江流域各省农田总氮径流流失强度及负荷占比

Fig. 8　Provincial-level distribution of total nitrogen loss rate and percentage of loss load in the Yangtze River Basin in 2013

模型预测结果表明,5 种情景下可以削减 2.4% ~9.3%的长江流域农田总氮流失负荷(图 9). 在单一防治措施中,“减少 15%径流深”情景对农田总氮流失控制效果最佳,可削减 4.3% (1.1 万 t) 旱地总氮流失负荷和 9.0% (2.0 万 t) 水田总氮流失负荷; 而“减少 15%化肥”情景和“减少 15%土壤氮”情景可使旱地和水田总氮流失负荷分别降低 1.5% ~4.7% (0.4 ~1.2 万 t) 和 0% ~6.3% (0 ~1.4 万 t). 在组合防治措施中,“减少 15%化肥和减少 15%径流深”情景可使旱地和水田总氮流失负荷分别削减 5.3% (1.3 万 t) 和 13.0% (2.9 万 t); 在“减少 15%土壤氮和减少 15%径流深”情景下,旱地和水田总氮流失负荷分别削减了 9.5% (2.4 万 t) 和 9.0% (2.0 万 t). 以上结果表明,农田总氮径流流失负荷对径流深的变化最敏感,这可能是由于较高的径流深会产生更强的驱动力,以及使更多的氮通过灌溉和降水(大气沉降)的方式输入到农田中[中国通过灌溉输入到农田的氮占总氮输入的 1.8% ~

7.9%[52],长江流域年均大气氮沉降输入强度为 20.7 kg·(hm^2·a)$^{-1}$[53]].以上预测结果表明,控制旱地总氮径流流失应侧重于降低径流量和土壤氮含量,而水田则应同时减低径流量和施氮量.为了减少农田径流量,应加大推广节水灌溉技术(例如,旱地采用滴灌和微喷灌溉,水田采用浅湿灌溉等[38])和循环灌溉方法[54].由于土壤氮可能是历年过量施用化肥和有机肥造成的,因此应考虑将土壤氮作为作物生产的潜在氮源予以再利用[55],并进一步推广应用测土配方施肥技术,减少土壤氮进一步过量积累及流失[56].综上所述,长江流域农田氮面源污染防治应加强农田土壤养分、水分和施肥的综合管理.



1. 减少 15% 化肥,2. 减少 15% 土壤氮,3. 减少 15% 径流深,4. 减少 15% 化肥和减少 15% 径流深,5. 减少 15% 土壤氮和减少 15% 径流深

图 9 5 种情境下基于随机森林算法模型预测长江流域旱地和水田总氮径流流失负荷削减率

Fig. 9 Predicted reduction percentages in upland and paddy field nitrogen runoff loss loads in the Yangtze River Basin using the random forest algorithm based model under five scenarios

3 结论

(1)长江流域旱地总氮径流流失强度的主要影响因素是径流深、施氮量和土壤氮含量,而径流深和施氮量是调控水田总氮径流流失强度的主要因素.

(2)所建立的基于随机森林算法的长江流域旱地和水田氮径流流失强度预测模型具有较高的精度,克服了氮径流流失强度与影响因素之间函数关系难以确定的问题,为定量识别流域/区域农田总氮径流流失负荷及空间分布特征提供了可靠方法.

(3)2013 年长江流域农田总氮径流流失负荷为 0.47 Tg·a^{-1},其中旱地、水田总氮径流流失负荷分别为 0.25 Tg·a^{-1}和 0.22 Tg·a^{-1}.长江流域旱地(广西、浙江、湖南和广东)、水田(广西、湖南、福建和上海)总氮径流流失热点区主要分布在中下游地区.

(4)为有效控制长江流域农田氮面源污染,应协同加强氮肥精准管理、减少农田径流量和提高土壤氮利用,且应将重点放在中下游地区.

参考文献:

[1] Richter A, Burrows J P, Nüß H, *et al*. Increase in tropospheric nitrogen dioxide over China observed from space[J]. Nature, 2005, **437**(7055): 129-132.

[2] Cui Z L, Wang G L, Yue S C, *et al*. Closing the N-use efficiency gap to achieve food and environmental security[J]. Environmental Science & Technology, 2014, **48**(10): 5780-5787.

[3] Zhang Y F, Wu H, Yao M Y, *et al*. Estimation of nitrogen runoff loss from croplands in the Yangtze River Basin: a meta-analysis[J]. Environmental Pollution, 2021, **272**, doi: 10.1016/j.envpol.2020.116001.

[4] 房志达, 苏静君, 赵洪涛, 等. 红壤丘陵区小流域典型土地利用的面源氮磷输出特征[J]. 环境科学, 2021, **42**(11): 5394-5404.
Fang Z D, Su J J, Zhao H T, *et al*. Output characteristics of nitrogen and phosphorus from non-point source pollution of typical land use in a micro-watershed in hilly red soil region[J]. Environmental Science, 2021, **42**(11): 5394-5404.

[5] Shen J, Zhao Y. Combined Bayesian statistics and load duration curve method for bacteria nonpoint source loading estimation[J]. Water Research, 2010, **44**(1): 77-84.

[6] Shrestha S, Kazama F, Newham L T H. A framework for estimating pollutant export coefficients from long-term in-stream water quality monitoring data[J]. Environmental Modelling & Software, 2008, **23**(2): 182-194.

[7] 陈亚荣, 阮秋明, 韩凤翔, 等. 基于改进输出系数法的长江流域面源污染负荷估算[J]. 测绘地理信息, 2017, **42**(1): 96-99, 104.
Chen Y R, Ruan Q M, Han F X, *et al*. Estimation of non-point source pollution load of Yangtze watershed based on improved export coefficient model[J]. Journal of Geomatics, 2017, **42**(1): 96-99, 104.

[8] Chen D J, Dahlgren R A, Shen Y N, *et al*. A Bayesian approach for calculating variable total maximum daily loads and uncertainty assessment[J]. Science of the Total Environment, 2012, **430**: 59-67.

[9] Hou X K, Zhou F, Leip A, *et al*. Spatial patterns of nitrogen runoff from Chinese paddy fields[J]. Agriculture, Ecosystems & Environment, 2016, **231**: 246-254.

[10] Hou X K, Zhan X Y, Zhou F, *et al*. Detection and attribution of nitrogen runoff trend in China's croplands[J]. Environmental Pollution, 2018, **234**: 270-278.

[11] 夏永秋, 杨旺鑫, 施卫明, 等. 我国集约化种植业面源氮发生量估算[J]. 生态与农村环境学报, 2018, **34**(9): 782-787.
Xia Y Q, Yang W X, Shi W M, *et al*. Estimation of non-point source N emission in intensive cropland of China[J]. Journal of Ecology and Rural Environment, 2018, **34**(9): 782-787.

[12] 张玉超, 钱新, 钱瑜, 等. 基于机器学习方法的太湖叶绿素 a 定量遥感研究[J]. 环境科学, 2009, **30**(5): 1321-1328.
Zhang Y C, Qian X, Qian Y, *et al*. Quantitative retrieval of chlorophyll a concentration in Taihu Lake using machine learning methods[J]. Environmental Science, 2009, **30**(5): 1321-1328.

[13] Jordan M I, Mitchell T M. Machine learning: trends, perspectives, and prospects[J]. Science, 2015, **349**(6245): 255-260.

[14] Fang K, Shen C P, Kifer D, *et al*. Prolongation of SMAP to spatiotemporally seamless coverage of continental U. S. using a deep learning neural network[J]. Geophysical Research Letters, 2017, **44**(21): 11030-11039.
[15] Liu W F, Yang H, Liu J G, *et al*. Global assessment of nitrogen losses and trade-offs with yields from major crop cultivations[J]. Science of the Total Environment, 2016, **572**: 526-537.
[16] 柴超, 俞志明, 宋秀贤, 等. 长江口水域富营养化特性的探索性数据分析[J]. 环境科学, 2007, **28**(1): 53-58.
Chai C, Yu Z M, Song X X, *et al*. Exploratory data analysis to the study of eutrophication in the Yangtze River Estuary, China [J]. Environmental Science, 2007, **28**(1): 53-58.
[17] Shen L, Xu H P, Guo X L, *et al*. Characteristics of large-scale harmful algal blooms (HABs) in the Yangtze River Estuary and the adjacent East China Sea (ECS) from 2000 to 2010[J]. Journal of Environmental Protection, 2011, **2**(10): 1285-1294.
[18] Tong Y D, Bu X G, Chen J Y, *et al*. Estimation of nutrient discharge from the Yangtze River to the East China Sea and the identification of nutrient sources [J]. Journal of Hazardous Materials, 2017, **321**: 728-736.
[19] Bao X, Watanabe M, Wang Q X, *et al*. Nitrogen budgets of agricultural fields of the Changjiang River basin from 1980 to 1990[J]. Science of the Total Environment, 2006, **363**(1-3): 136-148.
[20] Wang L J, Zheng H, Zhao H, *et al*. Nitrogen balance dynamics during 2000-2010 in the Yangtze River Basin croplands, with special reference to the relative contributions of cropland area and synthetic fertilizer N application rate changes[J]. PLoS One, 2017, **12**(7), doi: 10.1371/journal.pone.0180613.
[21] Chen J, Wu X D, Finlayson B L, *et al*. Variability and trend in the hydrology of the Yangtze River, China: annual precipitation and runoff[J]. Journal of Hydrology, 2014, **513**: 403-412.
[22] Food and Agriculture Organization of the United Nations. Harmonized world soil database v 1.2[EB/OL]. https://www.fao. org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/, 2022-09-22.
[23] He W T, Jiang R, He P, *et al*. Estimating soil nitrogen balance at regional scale in China's croplands from 1984 to 2014[J]. Agricultural Systems, 2018, **167**: 125-135.
[24] 中华人民共和国国家统计局. 年度数据[EB/OL]. https://data.stats.gov.cn/easyquery.htm? cn = C01, 2022-09-22.
[25] Zhang Y F, Jiao X Y, Wei Y H, *et al*. Long-term (1990-2013) changes and spatial variations of cropland runoff across China [J]. Water, 2022, **14**(18), doi: 10.3390/w14182918.
[26] 国家气象科学数据中心. 中国地面气候资料日值数据集(V3.0)[EB/OL]. http://data.cma.cn/, 2022-09-22.
[27] 中华人民共和国水利部. 中国水资源公报[EB/OL]. http://www.mwr.gov.cn/sj/tjgb/szygb/, 2022-09-22.
[28] 《中国水利年鉴》编纂委员会. 中国水利年鉴[EB/OL]. https://navi.cnki.net/knavi/yearbooks/YAGUJ/detail, 2022-09-22.
[29] Zhou F, Bo Y, Ciais P, *et al*. Deceleration of China's human water use and its key drivers[J]. Proceedings of the National Academy of Sciences of the United States of America, 2020, **117**(14): 7702-7711.
[30] Wang J X, Zhu Y Y, Sun T H, *et al*. Forty years of irrigation development and reform in China [J]. Australian Journal of Agricultural and Resource Economics, 2020, **64**(1): 126-149.
[31] 徐新良, 刘纪远, 张树文, 等. 中国多时期土地覆被遥感监测数据集(CNLUCC)[EB/OL]. https://www.resdc.cn/data.aspx? DATAID = 347, 2022-09-22.
[32] 中华人民共和国国家统计局. 中国统计年鉴[EB/OL]. http://www.stats.gov.cn/tjsj/ndsj/, 2022-09-22.
[33] Knoll L, Breuer L, Bach M. Large scale prediction of groundwater nitrate concentrations from spatial data using machine learning[J]. Science of the Total Environment, 2019, **668**: 1317-1327.
[34] 陈鹤, 陶晔, 毛振镀, 等. 环境微生物研究中机器学习算法及应用[J]. 微生物学报, 2022, **62**(12): 4646-4662.
Chen H, Tao Y, Mao Z D, *et al*. A review of machine learning algorithms for environmental microbiology [J]. Acta Microbiologica Sinica, 2022, **62**(12): 4646-4662.
[35] Elith J, Leathwick J R, Hastie T. A working guide to boosted regression trees[J]. Journal of Animal Ecology, 2008, **77**(4): 802-813.
[36] Choubin B, Moradi E, Golshan M, *et al*. An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines[J]. Science of the Total Environment, 2019, **651**: 2087-2096.
[37] Cui Z L, Zhang H Y, Chen X P, *et al*. Pursuing sustainable productivity with millions of smallholder farmers[J]. Nature, 2018, **555**(7696): 363-366.
[38] Zhuang Y H, Zhang L, Li S S, *et al*. Effects and potential of water-saving irrigation for rice production in China [J]. Agricultural Water Management, 2019, **217**: 374-382.
[39] Du X B, Wang Z, Lei W X, *et al*. Increased planting density combined with reduced nitrogen rate to achieve high yield in maize[J]. Scientific Reports, 2021, **11**(1), doi: 10.1038/S41598-020-79633-Z.
[40] Leuzinger S, Körner C. Rainfall distribution is the main driver of runoff under future CO_2 - concentration in a temperate deciduous forest[J]. Global Change Biology, 2010, **16**(1): 246-254.
[41] Udeigwe T K, Wang J J, Zhang H L. Predicting runoff of suspended solids and particulate phosphorus for selected Louisiana soils using simple soil tests [J]. Journal of Environmental Quality, 2007, **36**(5): 1310-1317.
[42] Wu L, Tang S R, He D D, *et al*. Conversion from rice to vegetable production increases N_2O emission via increased soil organic matter mineralization [J]. Science of the Total Environment, 2017, **583**: 190-201.
[43] Elrys A S, Ali A, Zhang H M, *et al*. Patterns and drivers of global gross nitrogen mineralization in soils[J]. Global Change Biology, 2021, **27**(22): 5950-5962.
[44] Liu J Z, Sun P F, Sun R, *et al*. Carbon-nutrient stoichiometry drives phosphorus immobilization in phototrophic biofilms at the soil-water interface in paddy fields[J]. Water Research, 2019, **167**, doi: 10.1016/j.watres.2019.115129.
[45] Linsley R K. The relation between rainfall and runoff: review paper[J]. Journal of Hydrology, 1967, **5**: 297-311.
[46] Ou Y, Rousseau A N, Wang L X, *et al*. Spatio-temporal patterns of soil organic carbon and pH in relation to environmental factors—A case study of the Black Soil Region of Northeastern China[J]. Agriculture, Ecosystems & Environment, 2017, **245**: 22-31.
[47] 周碧青, 邱龙霞, 张黎明, 等. 基于灰色关联-结构方程模型的土壤酸化驱动因子研究[J]. 土壤学报, 2018, **55**(5): 1233-1242.
Zhou B Q, Qiu L X, Zhang L M, *et al*. Study on driving factors of soil acidification based on grey correlation-structure equation

model[J]. Acta Pedologica Sinica, 2018, **55**(5): 1233-1242.

[48] Jin Q, Wang C, Sardans J, *et al*. Effect of soil degradation on the carbon concentration and retention of nitrogen and phosphorus across Chinese rice paddy fields[J]. CATENA, 2022, **209**, doi: 10.1016/j.catena.2021.105810.

[49] 李霞, 田光明, 朱军, 等. 不同磷肥用量对水稻土有机碳矿化和细菌群落多样性的影响[J]. 土壤学报, 2014, **51**(2): 360-372.
Li X, Tian G M, Zhu J, *et al*. Effects of rate of phosphorus fertilizer on organic carbon mineralization and bacterial community diversity in paddy soil[J]. Acta Pedologica Sinica, 2014, **51**(2): 360-372.

[50] 胡贵贵, 杨粉莉, 杨联安, 等. 基于主成分和机器学习的土壤有机质含量空间预测建模[J]. 干旱区地理, 2021, **44**(4): 1114-1124.
Hu G G, Yang F L, Yang L A, *et al*. Spatial prediction modeling of soil organic matter content based on principal components and machine learning[J]. Arid Land Geography, 2021, **44**(4): 1114-1124.

[51] van Liew M W, Arnold J G, Garbrecht J D. Hydrologic simulation on agricultural watersheds: choosing between two models[J]. Transactions of the ASABE, 2003, **46**(6): 1539-1551.

[52] Yuan S, Peng S B. Exploring the trends in nitrogen input and nitrogen use efficiency for agricultural sustainability[J]. Sustainability, 2017, **9**(10), doi: 10.3390/su9101905.

[53] Xing G X, Zhu Z L. Regional nitrogen budgets for China and its major watersheds[J]. Biogeochemistry, 2002, **57**(1): 405-427.

[54] Hama T, Nakamura K, Kawashima S, *et al*. Effects of cyclic irrigation on water and nitrogen mass balances in a paddy field[J]. Ecological Engineering, 2011, **37**(10): 1563-1566.

[55] Chen D J, Shen H, Hu M P, *et al*. Legacy nutrient dynamics at the watershed scale: principles, modeling, and implications[J]. Advances in Agronomy, 2018, **149**: 237-313.

[56] Li H, Huang G, Meng Q, *et al*. Integrated soil and plant phosphorus management for crop and environment in China. A review[J]. Plant and Soil, 2011, **349**(1-2): 157-167.

HUANJING KEXUE

Environmental Science (monthly)

Vol. 44 No. 7

Jul. 15, 2023

CONTENTS