doi:10.6041/j.issn.1000-1298.2022.05.039

融合多环境参数的鸡粪氨气排放预测模型研究

丁露雨[1]　吕　阳[2]　李奇峰[1,2]　王朝元[3]　余礼根[1]　宗伟勋[4]

(1. 清远市智慧农业农村研究院, 清远 511500; 2. 国家农业信息化工程技术研究中心, 北京 100097;
3. 中国农业大学水利与土木工程学院, 北京 100083; 4. 清远市农业科学研究所, 清远 511500)

摘要: NH_3是影响舍内肉鸡生长发育的主要有害气体,对其排放量的准确测量与预测有助于建立鸡舍环境调控模型,提升畜禽福利化养殖的水平。生产中,NH_3 监测多采用电化学传感器,精度差且寿命短,较难直接获取 NH_3 排放量。结合 NH_3 产生和释放的机理过程,选择相对较易获取的 CO_2排放量(E_{CO_2})和 H_2O 排放量(E_{H_2O})等环境参数建立 NH_3 排放量的预测模型。建立了肉鸡厚垫料养殖模式下,舍内鸡粪气体排放的模拟试验装置,连续多日向试验装置内投入等量鸡粪以模拟鸡舍每日粪便生成,监测温度、相对湿度以及 CO_2、H_2O、NH_3 排放量数据。基于多种机器学习方法和环境参数,构建了 NH_3 排放量预测模型,并运用特征和排列重要性探究参数重要程度,运用部分依赖图和个体条件期望图探究模型对参数的依赖关系。依据氨气排放预测相关知识,将温度和相对湿度计算为水汽压差(VPD),对比引入 VPD 后,不同参数组合方式对最优模型的影响。结果表明极限随机树模型预测 NH_3 排放量的效果最好,其 R^2为 0.916 7、均方根误差为 0.289 7 mg/(kg·h)、平均绝对百分比误差为 10.82%。分析各模型参数,该模型对 E_{H_2O}的依赖性最大,引入 VPD 对极限随机树的预测能力没有提升。基于温度、相对湿度、E_{H_2O}、E_{CO_2}建立的极限随机树模型可较好地预测肉鸡垫料饲养工艺下粪便的 NH_3 排放量。

关键词: 鸡粪; 氨气; 厚垫料饲养; 排放速率; 机器学习; 预测模型

中图分类号: S831.4　　**文献标识码**: A　　**文章编号**: 1000-1298(2022)05-0366-10　　**OSID**:

Prediction Model of Ammonia Emission from Chicken Manure Based on Fusion of Multiple Environmental Parameters

DING Luyu[1]　LÜ Yang[2]　LI Qifeng[1,2]　WANG Chaoyuan[3]　YU Ligen[1]　ZONG Weixun[4]

(1. *Qingyuan Intelligent Agricultural Research Institute*, *Qingyuan* 511500, *China*
2. *National Engineering Research Center for Information Technology in Agriculture*, *Beijing* 100097, *China*
3. *College of Water Resources and Civil Engineering*, *China Agricultural University*, *Beijing* 100083, *China*
4. *Qingyuan Agricultural Science Research Institute*, *Qingyuan* 511500, *China*)

Abstract: NH_3 is a major harmful gas that affects the growth of broilers in a chicken house. The accurate measurement and prediction of its emissions will help establish environmental regulation model and improve the welfare in the chicken house. The electrochemical sensors are commonly used in the real practice for measuring NH_3 concentration, which shows a low accuracy and short life time and makes it difficult to measure NH_3 emissions directly. Combined with the mechanism process of NH_3 released from manure, CO_2 and H_2O emissions that are relatively easier and cheaper obtained are selected to predict NH_3 emissions. The gaseous emissions from chicken manure housed in a deep litter system were experimentally simulated. The same amount of chicken manure was injected into the experimental setup for multiple days to simulate the daily manure generated in a chicken house, and to monitor the temperature, relative humidity and the emission of CO_2, H_2O and NH_3 from manure. The prediction model for the NH_3 emission was developed based on a variety of machine learning methods and environmental parameters. The importance of features and permutation were analyzed to explore the

收稿日期: 2021-06-17　修回日期: 2021-09-02

基金项目: 广东省科技计划项目(2019B090905006)、北京市农林科学院科技创新能力建设专项(KJCX20200421、KJCX20211007)和北京市农林科学院青年科研基金项目(QNJJ201913)

作者简介: 丁露雨(1988—),女,高级工程师,博士,主要从事畜舍环境智能监测与调控、畜禽智慧养殖研究,E-mail: dingly@nercita.org.cn

通信作者: 余礼根(1985—),男,副研究员,博士,主要从事畜牧业信息化关键技术与智慧养殖工程研究,E-mail: yulg@nercita.org.cn

importance of parameters, and the partial dependence graph as well as the individual condition expectation graph were analyzed to explore the dependence of the model on the parameters. Water pressure difference (VPD) was calculated using the temperature and relative humidity and introduced in modeling according to the knowledge of mechanism process of ammonia emissions. Comparisons were made to investigate the influence of different parameters on the optimal model after the introduction of VPD. The model based on extreme random tree showed the best performance in predicting NH_3 emissions, with R^2 of 0.916 7, RMSE of 0.289 7 mg/(kg·h), and MAPE of 10.82%. The most important parameter in the model was the H_2O emission, and the extreme random tree model had the greatest dependence on H_2O emission. The introduction of VPD did not improve the prediction ability of extreme random trees. Therefore, the optimal model was the extreme random tree model established based on T, H, E_{H_2O}, E_{CO_2} to predict the NH_3 emission from broiler manure in a deep litter system.

Key words: chicken manure; ammonia; deep litter system; emission rate; machine learning; prediction model

0　引言

氨气(NH_3)是密闭鸡舍主要的有害气体,生产中由于通风不足或通风不及时,常出现有害气体浓度累积、空气质量差的问题,严重制约了畜禽健康生产[1-4]。随着我国畜牧业的不断发展,数字化、福利化养殖成为了畜牧业升级转型的发展方向[5],有效预测舍内 NH_3 浓度可辅助通风量需求计算、评估舍内空气质量,及时、合理地调控舍内 NH_3 水平,对畜禽健康生产具有重要意义。

目前,畜禽舍 NH_3 预测模型的研究大致包括两类,第一类是直接预测舍内 NH_3 浓度的模型。例如,文献[6]结合长短时记忆神经网络(LSTM),使用随机森林(RF)对影响 NH_3 浓度的环境变量进行重要性排序,然后构建鸡舍 NH_3 浓度预测模型。文献[7]提出基于粒子群算法-优化深度神经网络算法(PSO-DNN)预测冬季平养鸡舍 NH_3 浓度,并证明该模型的预测精度高于现有的随机森林模型。舍内 NH_3 浓度受其排放量和舍内通风量的影响,现有研究中建立的 NH_3 浓度预测模型大都没有考虑环境调控设备运行状态和气体排放量水平,且数据量有限,所建立的模型泛化能力和鲁棒性较差,无法直接对设备运行状态予以调控指导和决策指令。第二类是预测 NH_3 产生量或排放量的模型[8-11]。这类模型通过排放机理或统计分析预测 NH_3 排放水平,结合动物环境需求可以获得环境调控参数(如通风量),指导设备进行环境调控。然而,现有模型存在参数数量多或没有考虑动物个体及环境因素不确定性对 NH_3 排放影响的问题,不同生产条件下的预测误差大。

NH_3 排放过程中常伴随着其他气体的释放,相同条件下,畜禽生产中各类气体的排放具有相关性,可以考虑结合气体产生与释放的机理过程,引入其他易测气体的排放量,用于预测 NH_3 排放[12-14],这有利于代替部分不确定因素的影响,使得预测结果更具有可靠性。本文选取相对易于测量且能代表动物及环境变化不确定性影响的变量参数(温度、相对湿度、H_2O 排放量、CO_2 排放量)作为输入特征,利用机器学习方法建立肉鸡粪 NH_3 排放预测模型,对比不同模型类别对 NH_3 排放的预测效果,以期为日常畜禽舍环境管理提供模型参考。

1　材料与方法

地面垫料平养是国内常见的一种肉鸡饲养方式,该模式投资少、成本低、易于管理,但粪便在舍内堆积时间长,舍内 NH_3 浓度较高,尤其是冬季通风量较小时,易诱发呼吸道疾病。本文在实验室条件下模拟平养肉鸡舍冬季 NH_3 排放过程,测试粪便发酵过程中 NH_3、CO_2、H_2O 浓度及排放量,用于建立 NH_3 排放估计模型的数据集,利用不同的输入参数和模型开展对比研究,构建 NH_3 排放量估计模型,试验流程如图 1 所示。

1.1　气体排放测试与数据采集

1.1.1　试验装置

肉鸡舍气体排放试验装置如图 2 所示。该试验装置尺寸(长×宽×高)为 0.86 m×0.45 m×0.65 m,参照动态箱原理设计,由亚克力材料制成,内部设置有直径 0.34 m、高 0.2 m 圆柱型鸡粪容器以及鸡粪搅动耙、手工操作窗口等设施,便于模拟肉鸡对鸡粪的踩踏及鸡粪与垫料的混合。箱体两侧各设 9 个进出气孔,一侧为进气面,对侧作为出气面连接气泵模拟机械通风,通气量为 0.618 m^3/h,出气口管路中接流量计监测箱内通气量。为防止出气孔气体影响模拟环境各项参数数值,气泵出气口设置延长管将废气排至室外。

1.1.2　数据采集与计算

肉鸡舍气体排放模拟试验地点位于中国农业大

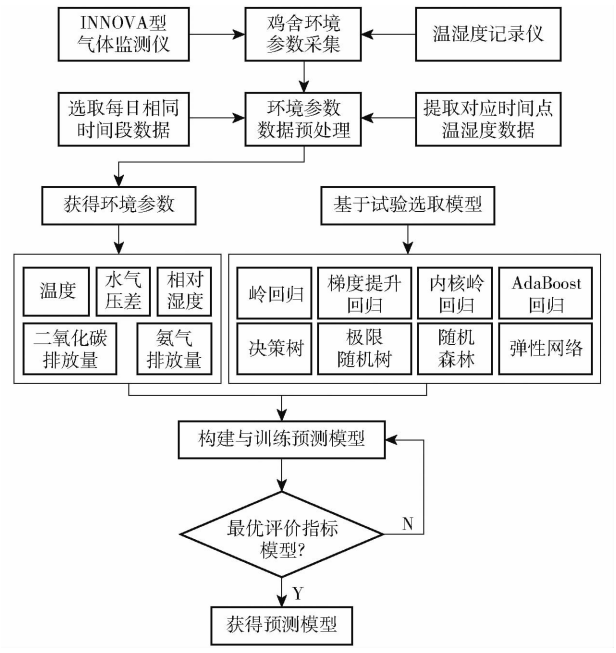

图 1 试验流程图

Fig. 1 Flow chart of experiment

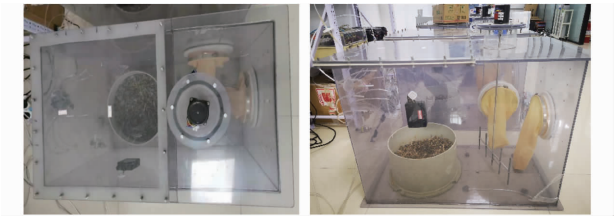

图 2 试验装置实物图

Fig. 2 Demonstration of experimental setup

学上庄试验站内,试验共计 12 d。试验初始时,装置内铺加垫料并放置 3 000 g 鸡粪,之后每日向试验箱内添加鸡粪 150 g,每日采集进、出气口 NH_3、CO_2、H_2O 等气体浓度(质量比)。箱内设置温湿度自动记录仪(艾普瑞(上海)精密光电有限公司)记录箱内温度 T 及相对湿度,采集频率 2 次/min。试验装置进、出气口设有气体采样点,通过 INNOVA 1403 型多路器(LumaSense Technologies 公司,美国)及 INNOVA 1512i 型光声谱多气检测仪(LumaSense Technologies 公司,美国)进行气体连续、循环采样及浓度检测。每次测量约 1 min,前 20 s 为管道冲洗,后 40 s 为气体浓度分析,每个点测 3 次,进、出口采样点交替采集,去掉第 1 次测量结果,取后 2 次平均值为该采样点当前气体浓度。

每日开盖添加鸡粪对排放系统有一定影响,取到达稳态时的相同时间阶段(每日 15:30 至次日 09:25)进行数据分析,共 444 组数据计算气体排放量。排放量计算公式[15]为

$$E = \frac{(C_o - C_i)V}{M} \tag{1}$$

式中 E——目标气体排放量,mg/(kg·h)

C_o——试验装置出气口采样点气体浓度,mg/m^3

C_i——试验装置进气口采样点气体浓度,mg/m^3

V——通气量,m^3/h

M——每日鸡粪量,kg

1.2 NH_3 排放预测模型

1.2.1 参数选择

(1)CO_2 和 H_2O

从 NH_3 产生机理考虑,鸡粪的尿酸分解是鸡舍内 NH_3 产生的主要途径,因为家禽的肝脏中没有精氨酸酶和氨甲酰磷酸合成酶,不能通过尿素循环把体内代谢产生的氨合成尿素,只能在肝脏和肾脏中合成嘌呤,并在黄嘌呤氧化酶的作用下生成尿酸[16]。尿酸在多种微生物酶的作用下水解成尿素和乙醛酸,最后尿素在脲酶的作用下产生 NH_3 和 CO_2,尿酸生成 NH_3 的反应式如图 3 所示[17-18]。

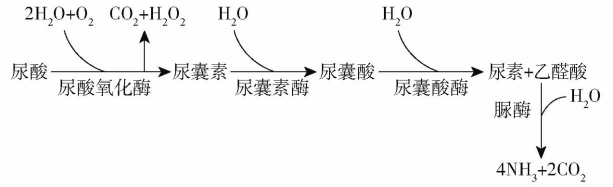

图 3 尿酸分解变成氨气的过程

Fig. 3 Decomposition of uric acid into ammonia

鸡粪中 40% ~90% 的有机氮会转化为氨,并以铵根离子(NH_4^+)和游离氨(NH_3)的形式存在[19]。从传质过程考虑,NH_3 的排放实质上是将游离氨气化后从粪污表面转移到大气的过程,受粪便与排放表面空气层之间的浓度差或分压力差影响。根据文献[20],NH_3 排放需要经液膜层传输到气膜层,最终进入外界空气,该过程中必然会伴随着 H_2O 挥发。

分析鸡粪产生 NH_3 的机理及其产生后排放的机理可以看出,NH_3 的产生伴随着 CO_2 的生成,而 NH_3 挥发与 H_2O 挥发密不可分。相比于 NH_3 浓度,CO_2 浓度和相对湿度更容易实现低成本和高精度的测量,且传感器寿命相对于传统的 NH_3 浓度电化学传感器长,维护成本更低。因此,选取 H_2O 排放量(E_{H_2O})及 CO_2 排放量(E_{CO_2})作为环境参数参与模型构建。

(2)温度 T 和相对湿度 H

温度、相对湿度是影响畜禽粪便 NH_3 产生与挥发的关键环境参数。大量研究冬季畜禽舍内温湿度与有害气体分布规律的结果表明,NH_3 浓度与温、湿度具有显著相关性[20-21]。当处于冬季低温条件时,鸡舍的保温需求会与通风形成较为明显的矛盾,通风不足时舍内的 NH_3 浓度和相对湿度会产生相互

影响。文献[21]研究冬季温、湿度对平养鸡舍 NH_3 浓度影响时发现：NH_3 浓度与 T 有较强的正线性关系，当舍内环境温度上升时，大部分情况下 NH_3 浓度也会上升；当处于不通风阶段时，舍内 NH_3 浓度会与相对湿度形成负线性相关，随相对湿度的下降而升高，这可能是由于 NH_3 易溶于水[22]，湿度较大时，会有部分 NH_3 随之溶解。

1.2.2　模型选择

机器学习以数据为研究对象，监督学习作为机器学习的重要组成部分，可以从标签数据中学习模型。标签数据表示输入、输出的对应关系，预测模型对给定的输入产生相应的输出，监督学习的本质是学习输入到输出的映射规律统计[23]。

以与鸡粪 NH_3 排放相关的环境参数为特征数据，以 NH_3 作为标签数据，目的是从 T、H、E_{H_2O}、E_{CO_2} 等特征数据中学习对应关系，使其能够预测标签数据，该种模型是典型的监督学习模型。其中，输入变量（T、H、E_{H_2O}、E_{CO_2}）和输出变量（氨气排放量 E_{NH_3}）均为连续变量，这类预测问题成为回归问题，对应模型为非概率模型。因此，采用监督学习的方法建模，考虑到数据量只有 444 条，不适用于大规模的参数训练，所以使用统计学习——即统计机器学习的方法进行模型构建。

鸡粪 NH_3 排放预测模型由数据直接学习决策函数，所以属于监督学习中的判别模型。依据模型特点，选择较为常用的判别模型对鸡粪 NH_3 排放进行预测，包括广义线性回归模型、内核岭回归、决策树及集成方法（随机森林、极限随机树、AdaBoost 回归、梯度提升回归）8 个模型，对比各个模型在鸡粪 NH_3 排放预测中的决定系数 R^2。

1.2.3　模型训练

采用 Python 语言 sklearn 库进行模型训练，训练过程中的数据集划分方式、超参数选择对模型训练结果至关重要。

（1）数据集划分

模型训练采取 10 折随机排列交叉验证的方法进行数据集划分，即对数据集随机抽样选出训练集和测试集，进行 10 次的评估和测试，最终对 10 次的结果取均值，得到最终的模型评价结果。该方法可以增强模型结果的可信度、减少数据集划分对模型训练的影响[24]。

（2）超参数

训练模型中，决策树及集成模型的训练涉及超参数选择问题。超参数的选择关系到预测结果的优劣，好的模型参数会提升模型的效果，所以模型参数的选择至关重要。超参数不能在模型中自动学习，需要在训练过程中结合多种评价指标进行选择。使用网格追踪法寻找超参数的最优值，并结合所采用的学习机方法进行微调，选择一组最优超参数。本文涉及到的主要超参数如表 1 所示。

表 1　本文模型中涉及的主要超参数

Tab.1　Hyperparameters involved in different models

参数名称	参数说明	参数对应模型
max_depth	树的最大深度，如果没有指定参数，训练直到所有叶子都是同一类别，或者直到所有叶子都包含少于 min_samples_split 个样本	决策树、随机森林、极限随机树、梯度提升回归
min_samples_leaf	叶节点处需要的最小样本数，只有当任何深度的分割点在左分支和右分支中的每个分支上至少留下 min_samples_leaf 个训练样本时，才考虑该点	决策树、随机森林、极限随机树
min_samples_split	拆分内部节点所需的最少样本数	决策树、随机森林、极限随机树、梯度提升回归
splitter	用在每个节点上选择拆分的策略，包括“optimum”（选择最佳拆分）、“random”（选择最佳随机拆分）两种策略	决策树
max_samples	从训练集中抽取样本的数量，用于训练每个基估计器	随机森林、极限随机树
n_estimators	树模型中“tree（树）”的数量	随机森林、极限随机树、AdaBoost 回归、梯度提升回归
loss	每次增强迭代后更新权重时使用的损失函数，“square”表示均方误差	AdaBoost 回归
subsample	用于拟合各个基础学习者的样本比例	梯度提升回归

1.3　模型评价指标

使用预测模型的决定系数 R^2、均方根误差（RMSE）、平均绝对百分比误差（MAPE）作为评价指标[10]，测试不同模型对鸡粪 NH_3 排放量的预测效果。

设模型预测值 $\hat{y}=\{\hat{y}_1,\hat{y}_2,\cdots,\hat{y}_n\}$，排放速率实测值 $y=\{y_1,y_2,\cdots,y_n\}$，实测值均值为 $\bar{y}$，任意一组实测值和预测值分别为 y_i 和 $\hat{y}_i$。决定系数 R^2 可以表征模型预测值对样本观测值的拟合程度，模型越好 R^2 越靠近 1，相反，模型效果越差越靠近 0。

2 结果与讨论

2.1 气体排放测试数据

气体排放测试期间,动态箱内平均温度、相对湿度分别为(6.1±1.1)℃、(69.9±5.4)%,期间的平均 NH_3、CO_2 及 H_2O 排放量分别为(3.2±0.9) mg/(kg·h)、(147.5±61.6) mg/(kg·h)、(184.1±43.9) mg/(kg·h)。测试期间,T、相对湿度、E_{NH_3}、E_{CO_2}和 E_{H_2O}的变化趋势如图 4 所示。试验开始时,鸡粪进行发酵反应,前 4 d 时 NH_3 排放量随时间上升并达到最高点,然后大幅下降,之后每日随鸡粪不断添加,重复以上趋势。同时,通过与其它环境参数曲线进行对比观察发现,NH_3 排放量的数量级远小于 CO_2 和 H_2O,但 3 种气体排放量的变化趋势几乎一致,这说明 E_{H_2O}、E_{CO_2}比 E_{NH_3}相对更容易实现准确检测,研究中选择的 T、H、E_{H_2O}、E_{CO_2}这 4 个自变量对 E_{NH_3}预测有较好的可解释性和代表性。

2.2 不同预测模型对比

结合已有超参数的选择方法、各模型超参数的数量及范围,在确保准确度的前提下,使用网格追踪法寻找左右超参数并进行微调。不同模型最佳超参数选择后的结果如表 2 所示。

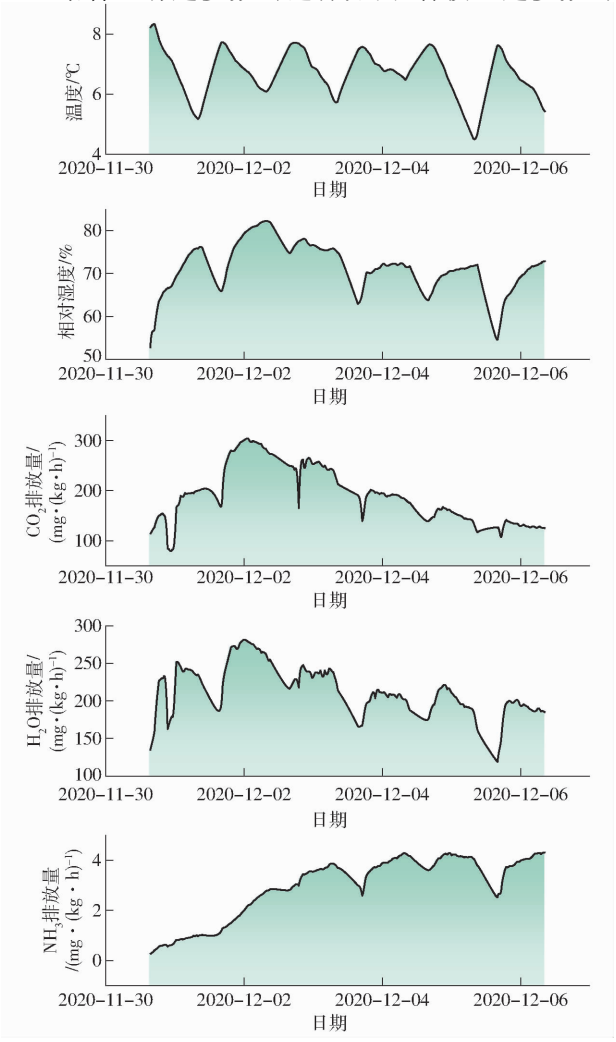

图 4 鸡粪气体排放测试数据

Fig. 4 Chicken manure gas emission test data

表 2 不同树模型选用的超参数值

Tab. 2 Hyperparameter values selected by different tree models

参数	决策树	随机森林	极限随机树	AdaBoost 回归	梯度提升回归
max_depth	15	10	20	—	10
min_samples_leaf	1	1	1	—	—
min_samples_split	3	2	2	—	10
splitter	random	—	—	—	—
max_samples	—	0.8	1	—	—
n_estimators	—	300	200	100	300
loss	square	square	square	square	square
subsample	—	—	—	—	0.6

注:"—"表示模型没有该参数的设置功能。

在最佳超参数选择的基础上,分析各模型以 T、H、E_{H_2O}、E_{CO_2}为自变量预测 E_{NH_3}的评价指标(表 3)发现,岭回归、弹性网络、内核岭回归预测氨气排放量的效果很差,决定系数均不足 0.2,预测误差大,而树模型和同质集成的树模型的结果较好,其中极限随机树模型表现最优。

表 3 不同模型评价指标

Tab. 3 Evaluation indexes of different models

模型	R^2	RMSE/(mg·(kg·h)$^{-1}$)	MAPE/%
岭回归	0.123 3	0.907 4	44.42
弹性网络	0.129 4	0.909 9	44.51
内核岭回归	0.112 4	0.908 0	44.60
决策树	0.768 5	0.538 8	12.53
随机森林	0.840 4	0.381 8	15.25
极限随机树	0.916 7	0.289 7	10.82
AdaBoost 回归	0.728 1	0.523 5	19.22
梯度提升回归	0.872 2	0.381 9	11.42

极限随机树也是一种集成方法,使用多个树来预测样本,图 5 为集成模型中的一颗子树。该子树是一颗关于所选参数的决策树,该树只显示了前 4 层,记录着分割的特征和对应的分割点。计算分割点的方法中,进一步增强了分割点的随机性,利用的特征是候选特征的随机集合。根据所有候选特征任意生成分割点,在生成的分割值中选择最佳的点进行分割。这种方法通常可以减小模型的离散度,所以极限随机树在多个模型中表现最好。

选择表 3 中性能较好的树模型作为预测模型对测试集数据进行验证分析,预测结果与实测结果对比如图 6 所示。使用 T、H、E_{H_2O}、E_{CO_2}作为影响因子

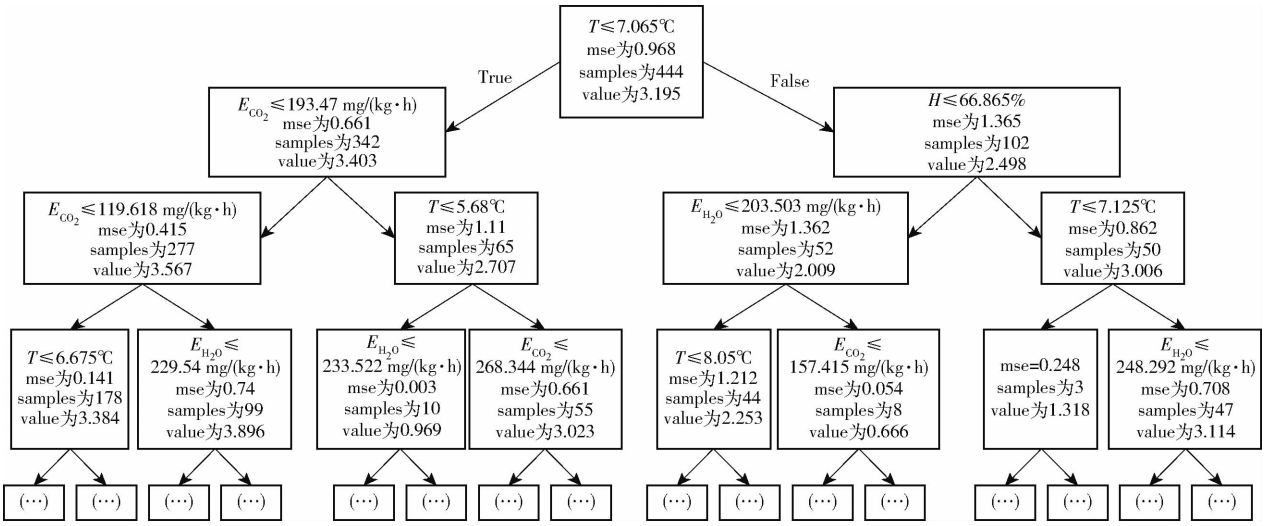

图 5　树模型结构图

Fig. 5　Structure diagram of tree model

参与建模时，决策树、随机森林、极限随机树、AdaBoost 回归、梯度提升回归的平均残差百分比分别为(3 ± 4.6)%、(4 ± 3.7)%、(3 ± 3.5)%、(6 ± 4.8)%、(5 ± 4.5)%。可见，各模型预测 NH_3 排放量时多表现为高估，其中决策树和极限随机树平均残差百分比最低，但决策树模型的残差百分比标准差更高，故而模型的稳定性不如极限随机树，极限随机树模型的预测效果最佳。

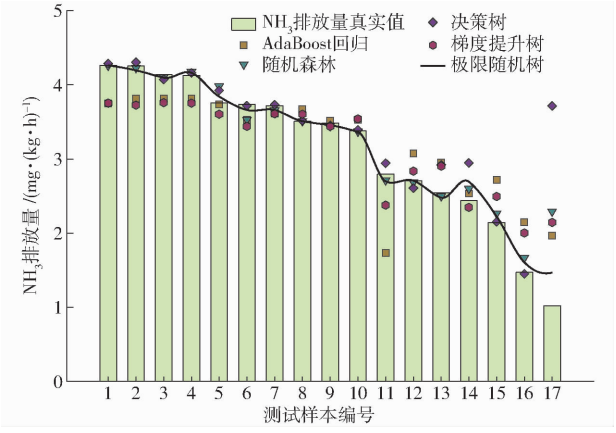

图 6　各类树模型预测结果对比

Fig. 6　Comparison of prediction results of various tree models

2.3　参数

2.3.1　参数重要性

特征重要性和排列重要性分析可以判别各参数对模型的重要程度。特征重要性是指特征(即输入参数)对预测结果的影响，是模型对每个特征重要程度的描述。排列重要性是随机打乱某个特征之后，衡量该特征对模型的影响，从而得出特征的重要性。

选取表现最好的极限随机树作为分析对象，利用排列重要性和特征重要性分析 T、H、E_{H_2O}、E_{CO_2} 在建模过程中的重要程度，探究 NH_3 排放量预测的机理知识与模型重要性的联系。

从图 7 可以看出，E_{H_2O} 在 NH_3 排放预测中起着重要作用，这可能是因为 NH_3 易溶于水，在今后的研究中可以重点关注 E_{H_2O} 与 E_{NH_3} 之间的相互作用关系。T 和 E_{CO_2} 在特征重要性和排列重要性两方面评价中，影响程度并不对应相同，虽然排名有所变化，但不能一定说明 E_{CO_2} 比 T 的影响程度大。因为测试时环境温度相对比较稳定，变化幅度较小(变异系数 18.0%)，而 E_{CO_2} 的变化相对更大(变异系数 41.8%)，因此，随机打乱 T 和 E_{CO_2} 两个特征后预测的结果变化幅度不同。此外，粪便发

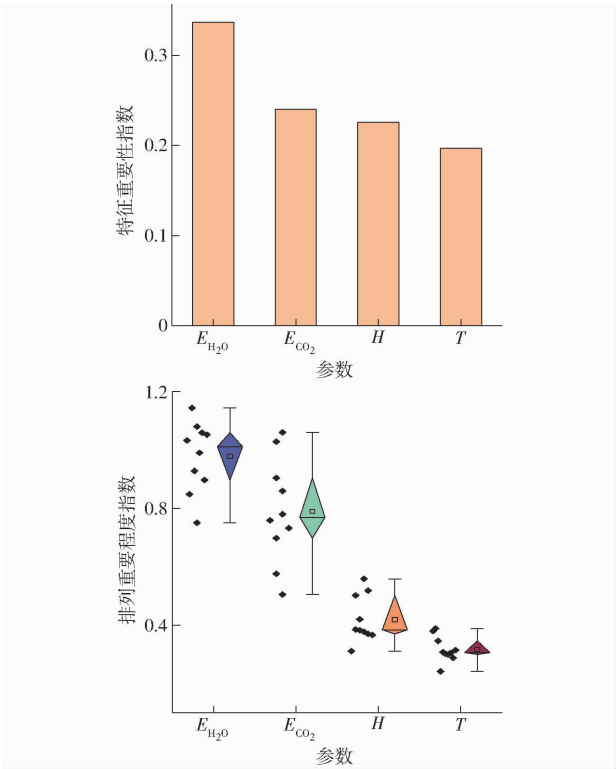

图 7　特征重要性和排列重要性图

Fig. 7　Feature importance and permutation importance charts

酵过程中 CO_2 的产生具有温度依赖性，T 变化后，E_{H_2O} 也会随之变化。

综上所述，如果同一个特征在不同的评价方法中都重要，那么该特征就应该重点关注。在本次试验中，两种评价方法中 E_{H_2O} 的得分都很高，因此，不同参数组合时，模型中都需要考虑 E_{H_2O}。

2.3.2 模型对特征变量的依赖性

通过个体条件期望图（Individual conditional expectation，ICE）和部分依赖图进行依赖性分析，即某一特征变量的值变化对 NH_3 排放量预测值的影响。

个体条件期望图是用来可视化分析目标响应和一组输入特征变量之间的交互作用，可以分析目标函数与输入的特征变量之间的依赖性。设 x_s 为输入特征集合（即特征参数），x_c 为其补集。响应 f 在 x_s 处的部分依赖性定义为

$$p_{d_{X_s}}(x_s) = \int f(x_s, x_c) p(x_c) \mathrm{d}x_c \tag{2}$$

式中 $p(x_c)$——x_c 出现的概率

其中 $f(x_s, x_c)$ 是给定样本的响应函数，本研究中，该响应函数是极限随机树模型。x_s 是在样本集中随机选择的值。ICE 线定义为在 x_s 和 x_s 的单个样本补集 $x_c^{(i)}$ 下，响应函数的计算结果。

图 8 所示分别为 T、H、E_{H_2O}、E_{CO_2} 为研究对象的条件期望图，即变化其中一个环境参数的取值，保留其他参数取值不变，运用训练完成的极限随机树模型预测 E_{NH_3} 的结果。图中，测试数据集样本为 50 个样本，细线代表每条数据，粗线代表所有细线的均值。曲线越陡峭，说明该参数对于模型越重要，反之，对模型重要性越低。

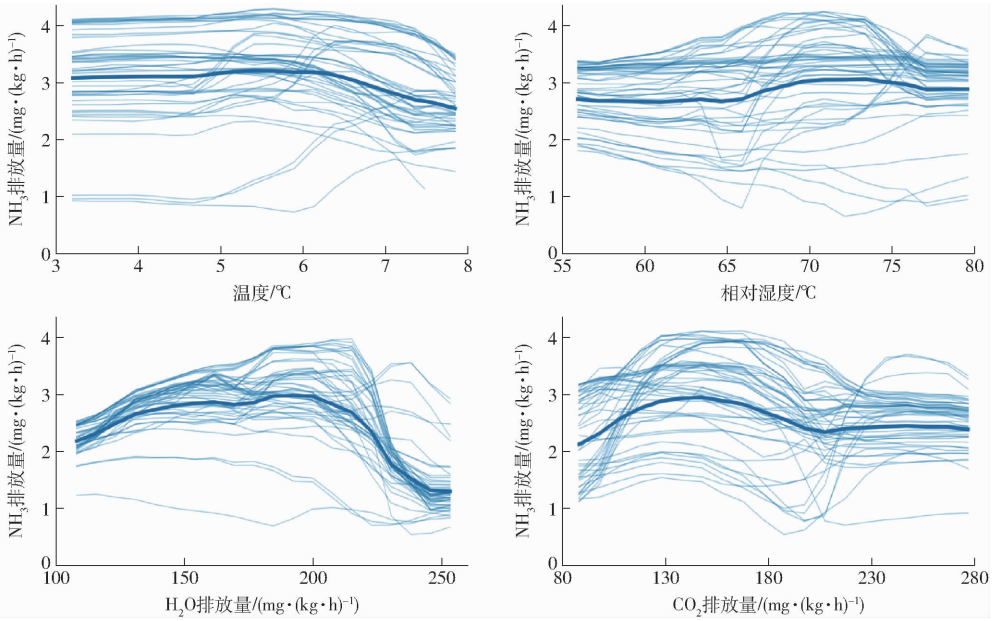

图 8 个体条件期望图（ICE）

Fig. 8 Individual conditional expectation graph

从图 8 中可以看出，当 $T<6.5$℃时，T 的 ICE 曲线变化平缓，说明 T 变化，其余值不变的情况下，E_{NH_3} 的预测值几乎不变，即 T 的变化对 E_{NH_3} 的预测结果影响不大；当 $T>6.5$℃时，T 与 E_{NH_3} 大致呈线性的负相关关系。E_{H_2O} 的变化最陡峭，说明 E_{H_2O} 在预测 E_{NH_3} 的过程中贡献度最大。

部分依赖图与个体条件期望图相似，区别在于研究对象的抽取方式不同，个体条件期望图抽取数据的方式符合数据分布特征，部分依赖图是按照等间隔方式抽取数据并计算。这能够展现出两个特征变量对模型预测影响的函数关系：近似线性关系、单调关系或者更复杂的关系。为了更加形象地观察特征组合与模型之间的交互作用，在三维空间内用两个特征的相互作用绘制部分依赖图。

由图 9 可以看出，相比于单个变量，同时考虑两个变量的影响时，变量与模型之间的交互变得更为复杂。以 E_{H_2O} 和 E_{CO_2} 为例，在 E_{H_2O} 小于 200 mg/(kg·h) 时，E_{CO_2} 对 E_{NH_3} 的作用与单变量时的趋势大致相同，但是当 E_{H_2O} 大于 200 mg/(kg·h) 时，E_{H_2O} 的影响占据主导地位，E_{CO_2} 在该段的影响几乎可以忽略。这也体现出变量影响之间的一个原则：模型中表现强的变量会掩盖掉表现弱的变量。

2.4 环境参数融合对模型的影响

2.4.1 水汽压差

T 和 H 度二者间存在互作关系，当空气中 H 不变时，T 升高会使得 H 降低。NH_3 排放研究中，常使用水汽压差（VPD）来表征 T 和 H 二者对 NH_3 排放的共同影响[25]。因此，将模型参数进一步简化，利

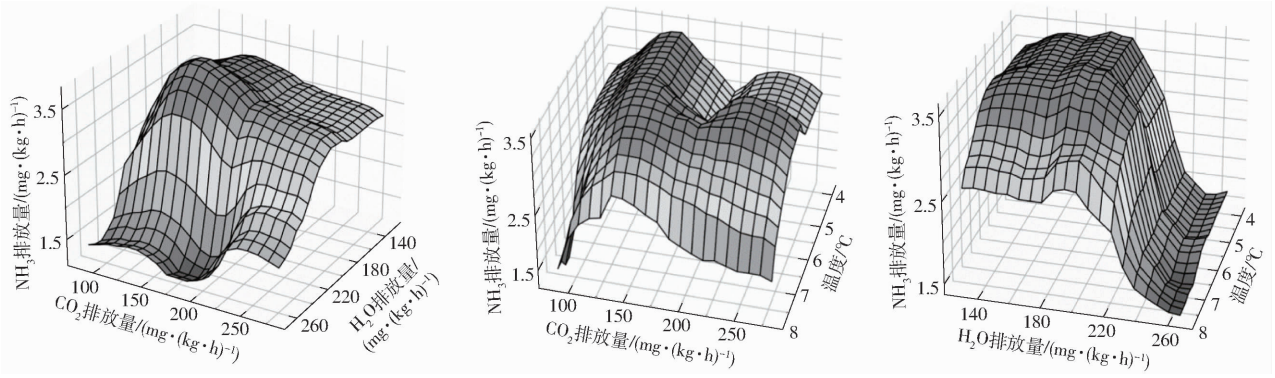

图 9　三维空间内的部分依赖图

Fig. 9　Partial dependence diagram in three dimensional space

用 VPD 代替 T 和 H，探索模型的预测效果。

VPD 是空气中水汽分压力与饱和水汽压之间的差值，可通过 T 和 H 计算获得[26]，公式为

$$V = 0.611(1 - H)\mathrm{e}^{\frac{17.502T}{T+240.97}} \tag{3}$$

式中　V——水汽压差，kPa

计算后，气体排放测试时的 VPD 变化曲线如图 10 所示。

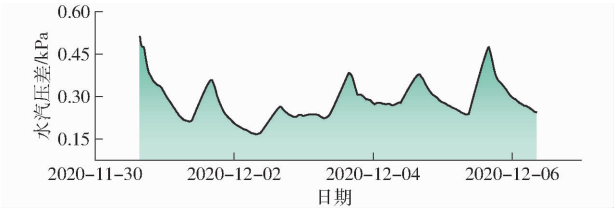

图 10　鸡粪气体排放测试期间的 VPD 变化曲线

Fig. 10　VPD during experiment of gaseous emissions from chicken manure

2.4.2　水汽压差对模型的影响

为了探索衍生参数对模型准确率的影响，选用 VPD、E_{H_2O}、E_{CO_2} 等 3 个自变量的不同组合进行建模尝试，具体组合方式为：VPD、E_{H_2O}、E_{CO_2}；VPD、E_{H_2O}；VPD、E_{CO_2}。

对比包含 VPD 的参数组合与 T、H、E_{H_2O}、E_{CO_2} 的参数组合，不同特征变量利用极限随机树模型预测 E_{NH_3} 的决定系数 R^2 如图 11 所示。

从参数数量上看，相同模型随着参数数量的变少，模型的 R^2 呈降低的趋势。各模型的预测效果并没有随着引入 VPD 而有所提升，甚至有所降低。这可能是因为本研究中所选用的模型参数较少，树模型可选的参数较为单一，最后构建出的树模型差异性不大，影响了树的多样性，进而导致了参数数量与评价指标得分之间成正比的结果。从最优模型的选择角度来看，不同参数组合时，仍然是各类树模型的拟合效果更好，最佳模型仍然是极限随机树模型，也体现了极限随机树模型的预测准确性。

综合不同模型种类和参数组合的建模效果可

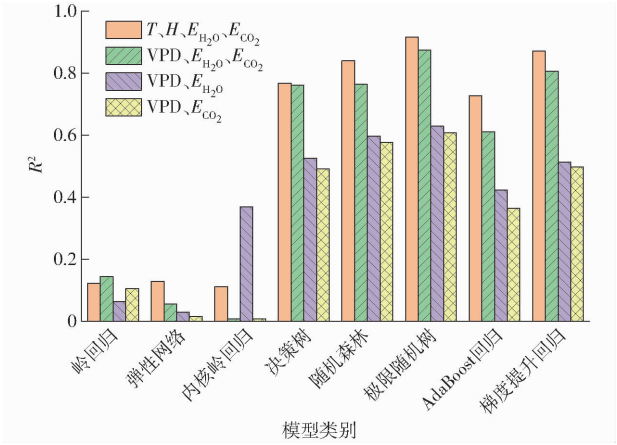

图 11　不同参数组合特征输入时模型的决定系数

Fig. 11　R^2 of model when different parameter combination features were input

知，采用极限随机树模型且以 T、H、E_{H_2O}、E_{CO_2} 为模型特征变量输入时，E_{NH_3} 的预测效果最好。

2.5　模型应用

（1）模型应用条件

应用中，上述模型的特征参数 E_{H_2O}、E_{CO_2} 为粪便挥发所产生[27-28]，因此，该模型不同条件的应用方式不同。当模型应用场景的气体排放源仅为粪便时可直接使用模型估计 NH_3 排放量，如粪便贮存、堆肥过程中 NH_3 排放量估计、缺失值填充等。当模型应用场景的气体排放源包含粪便和动物时，需对测得的 E_{H_2O}、E_{CO_2} 进行校正，扣除肉鸡自身产生的 CO_2 和水汽后再输入模型进行测算，如肉鸡舍氨气排放量预测、以改善空气质量为目标的通风量测算等。根据国际农业工程学会国际农业工程委员会（CIGR），肉鸡生产中排放的 CO_2 和水汽量可通过肉鸡产生的显热、潜热和总热量，利用已有模型进行计算[29-30]。

（2）模型应用优化

本文中参与模型构建的数据量及各参数变化范围有限，模型应用中可能会存在一些局限。由 2.3.1 节可知，当 E_{H_2O} 大于 200 mg/(kg·h) 时，E_{NH_3}

预测值与 E_{H_2O} 之间有很强的依赖性。本研究中 E_{H_2O} 的范围在 110 ~ 270 mg/(kg·h)之间,当 E_{H_2O} 大于 270 mg/(kg·h)时模型的准确度可能会有所变化。T 是影响 CO_2、NH_3、H_2O 排放的重要环境参数,本研究中 T 的变化范围相对较小(4 ~ 8℃),进而出现了模型对 T 依赖性最弱的结果。从机理上看,T 是影响气体排放最主要的环境因素之一,作用程度远大于相对湿度;当 T 从 10℃以内上升到 20℃以上后养殖场的气体排放会增加数十倍[31-32]。当 T 大于 10℃,如果直接使用前述模型可能会出现较大的误差。

本研究构建的模型具有一定的自适应能力,鉴于上述局限的存在,应用中可通过增加数据量进行优化训练,提高模型普适性。数据量的增加包括两方面:一是扩大数据分布的范围,二是增加各范围数据的数量。文献[33]在分类模型泛化能力的研究中发现,样本数量的增加和样本分布范围的扩大,可以有效提升模型普适性,而且数据分布范围的影响大于数据增量的影响[34]。因此,未来应用中优先考虑扩大数据分布范围。例如,可以增加 T 大于 10℃以上的环境数据和气体排放数据,与原始数据混合训练极限随机树模型,提高模型的普适性。

3 结论

(1)考虑到畜禽生产及粪便发酵过程中,NH_3 浓度及排放量存在数值小、传感器腐蚀性高、不易采集的情况,本文结合 NH_3 产生的机理和机器学习建模技术,利用相对较易采集的环境参数作为影响因子,对比了 8 种不同模型的预测效果,建立了垫料饲养条件下肉鸡粪 NH_3 排放量预测模型,为垫料饲养肉鸡舍通风调控奠定模型基础,也为粪便管理中肉鸡粪 NH_3 排放量预测提供模型支撑。

(2)通过预测模型结果对比发现,使用 T、H、E_{H_2O}、E_{CO_2}作为特征输入,以极限随机树为预测模型时,肉鸡粪 NH_3 排放量的预测效果最优,R^2 为 0.916 7,E_{NH_3} 预测的平均残差百分比为(3 ± 3.5)%。各特征参数中,E_{H_2O}对模型的影响最为重要。引入 VPD 代替 T 和 H 并不能得到更好的模型效果,甚至会降低模型预测的准确度。应用中,当排放源同时包括粪便和动物时,应当扣除肉鸡自身产生的 CO_2 和 H_2O,对 E_{H_2O}、E_{CO_2}校正后再输入模型中进行预测。

参 考 文 献

[1] 李俊. 规模化畜禽养殖场消毒技术[J]. 畜牧兽医科学,2020(13):155 - 156.
LI Jun. Disinfection technology of large-scale livestock and poultry farms[J]. Animal Husbandry and Veterinary Science, 2020(13):155 - 156. (in Chinese)

[2] 朱春红,章双杰,陶志云,等. 禽舍内氨气产生、危害及其有效减排措施[J]. 家畜生态学报,2019,40(6):72 - 77.
ZHU Chunhong, ZHANG Shuangjie, TAO Zhiyun, et al. Ammonia generation, harm and effective emission reduction measures in poultry houses[J]. Journal of Livestock Ecology, 2019,40(6): 72 - 77. (in Chinese)

[3] SOLIMAN E S, MOAWED S A, HASSAN R A. Influence of micro climatic ammonia levels on productive performance of different broilers' breeds estimated with univariate and multivariate approaches[J]. Vet World,2017,10(8): 880 - 887.

[4] NASEEM S, KING A J. Ammonia production in poultry houses can affect health of humans, birds, and the environment-techniques for its reduction during poultry production[J]. Environ. Sci. Pollut. R., 2018,25(16): 15269 - 15293.

[5] 李奇峰,王文婷,余礼根,等. 信息技术在畜禽养殖中的应用进展[J]. 中国农业信息,2018,30(2):15 - 23,41.
LI Qifeng, WANG Wenting, YU Ligen, et al. Application progress of information technology in livestock and poultry breeding [J]. China Agricultural Information, 2018,30(2): 15 - 23,41. (in Chinese)

[6] 杨亮, 刘春红, 郭昱辰,等. 基于 EMD - LSTM 的猪舍氨气浓度预测研究[J]. 农业机械学报, 2019, 50(增刊):360 - 367.
YANG Liang, LIU Chunhong, GUO Yuchen, et al. Prediction of ammonia concentration in fattening piggery based on EMD - LSTM[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(Supp.):360 - 367. (in Chinese)

[7] 邹修国,宋圆圆,徐泽颖,等. 基于 PSO - DNN 的平养鸡舍冬季氨气浓度预测模型研究[J]. 南京农业大学学报,2021,44(1):184 - 193.
ZOU Xiuguo, SONG Yuanyuan, XU Zeying, et al. Study on prediction model of ammonia concentration in flat chicken house in winter based on PSO - DNN[J]. Journal of Nanjing Agricultural University, 2021,44(1): 184 - 193. (in Chinese)

[8] 孙瑞锋,李同树,步长英,等. 影响肉鸡鸡粪产氨量相关因素的线性分析与氨气生成量的预测[A]. 中国畜牧兽医学会家禽学分会. 中国家禽科学研究进展——第十四次全国家禽科学学术讨论会论文集[C]// 中国畜牧兽医学会家禽学分会:中国畜牧兽医学会,2009:5.

[9] 朱海生,董红敏. 猪舍氨气排放预测模型的研究现状[J]. 农业工程学报,2006,22(增刊 2):187 - 192.
ZHU Haisheng, DONG Hongmin. Research status of ammonia emission prediction model in pig house [J]. Journal of Agricultural Engineering, 2006,22(Supp.2): 187 - 192. (in Chinese)

[10] 耿红,高芳,陈霖,等. 畜禽源氨气排放因子估算方法研究[J]. 环境科学学报,2017,37(8):3077 - 3084.
GENG Hong, GAO Fang, CHEN Lin, et al. Study on estimation method of ammonia emission factors from livestock and poultry[J]. Journal of Environmental Sciences, 2017,37(8): 3077 - 3084. (in Chinese)

[11] CHEN Yongxing, XIN Hongwei, LI Hong, et al. Comparison of ammonia emissions from poultry houses based on diurnal

integration vs. daily means of gas concentration and duilding ventilation rate[C] // The Ninth International Environment Symposium. ASABE Valencia Conference Center Valencia, 2012.
[12] DING Luyu, CAO Wei, SHI Zhengxiang, et al. Carbon dioxide and methane emissions from the scale model of open dairy lots [J]. J. Air. Waste Manag. Assoc., 2016, 66(7):715 – 725.
[13] BJERG B, ZHANG G, MADSEN J, et al. Methane emission from naturally ventilated livestock buildings can be determined from gas concentration measurements[J]. Environmental Monitoring and Assessment, 2012, 184(10):5989 – 6000.
[14] RONG Li, LIU Dezhao, PEDERSEN E F, et al. Effect of climate parameters on air exchange rate and ammonia and methane emissions from a hybrid ventilated dairy cow building[J]. Energy & Buildings, 2014, 82:632 – 643.
[15] 丁露雨. 奶牛运动场温室气体排放规律与排放过程模拟研究[D]. 北京：中国农业大学, 2016.
DING Luyu. Study on greenhouse gas emission law and emission process simulation of dairy farm[D]. Beijing: China Agricultural University, 2016. (in Chinese)
[16] 李季,王同心,姚卫磊,等. 畜禽舍氨气排放规律及对畜禽健康的危害[J]. 动物营养学报,2017,29(10):3472 – 3481.
LI Ji, WANG Tongxin, YAO Weilei, et al. Ammonia emission law in livestock and poultry houses and its harm to livestock and poultry health[J]. Journal of Animal Nutrition, 2017,29(10): 3472 – 3481. (in Chinese)
[17] 李柱,时建忠,顾宪红. 控制猪氨气排放技术研究进展[J]. 家畜生态学报,2010,31(5):4 – 7.
LI Zhu, SHI Jianzhong, GU Xianhong. Research progress on ammonia emission control technology of pigs[J]. Journal of Livestock Ecology, 2010,31(5): 4 – 7. (in Chinese)
[18] 步长英. 丝兰皂甙、芽孢杆菌对肉鸡生产性能、氮代谢和排泄物氨气散发量的影响[D]. 泰安：山东农业大学,2008.
[19] KOERKAMP P W G G. Review on emissions of ammonia from housing systems for laying hens in relation to sources, processes, building design and manure handling[J]. Journal of Agricultural Engineering Research,1994,59(2):73 – 87.
[20] 李奇峰,田丰,丁露雨,等. 家禽粪便管理过程中氨气减排技术研究进展[J]. 家畜生态学报,2020,41(9):1 – 6.
LI Qifeng, TIAN Feng, DING Luyu, et al. Research progress of ammonia emission reduction technology in poultry manure management[J]. Journal of Livestock Ecology, 2020,41(9): 1 – 6. (in Chinese)
[21] 张蕾,冯婧,韩静,等. 冬季地面平养蛋鸡舍温湿度对氨气浓度的影响[J]. 畜禽业,2019,30(11):1 – 5.
ZHANG Lei, FENG Jing, HAN Jing, et al. Effects of temperature and humidity on ammonia concentration of ground level layer houses in winter[J]. Livestock and Poultry Industry, 2019,30(11): 1 – 5. (in Chinese)
[22] 张静,刘双红,孙斌. 猪场舍内氨气对猪的危害[J]. 养殖技术顾问,2014(4):26.
ZHANG Jing, LIU Shuanghong, SUN Bin. Harm of ammonia gas in pig farms to pigs[J]. Breeding Technical Consultant, 2014(4): 26. (in Chinese)
[23] 李航. 统计学习方法[M]. 2 版. 北京:清华大学出版社,2019:10 – 34.
[24] XU Lu, FU H Y, GOODARZI M, et al. Stochastic cross validation[J]. Chemometrics & Intelligent Laboratory Systems, 2018, 175:74 – 81.
[25] DING Luyu, LI Qifeng, WANG Chaoyuan, et al. Determination of the mass transfer coefficient of ammonia emissions from dairy open lots using a scale model[J]. Biosystems Engineering, 2020, 190:145 – 156.
[26] 张静,王力. 黄土塬区苹果园蒸散与环境因素的关系[J]. 林业科学,2018,54(3):29 – 38.
ZHANG Jing, WANG Li. Relationship between evapotranspiration of apple orchard and environmental factors in Loess Plateau [J]. Forestry Science, 2018,54(3): 29 – 38. (in Chinese)
[27] 彭才望,贺喜,孙松林,等. 斗式黑水虻处理猪粪有机肥取料机设计与试验[J]. 农业机械学报,2021,52(2):145 – 156.
PENG Caiwang, HE Xi, SUN Songlin, et al. Design and experiment on shoveling device of pig manure organic fertilizer by hermetia illucen transforming based on bucket-wheel mechanism[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021,52(2): 145 – 156. (in Chinese)
[28] 袁兴茂,李霄鹤,吴海岩,等. 规模化猪场固粪好氧快速发酵工艺与设备研究[J]. 农业机械学报,2021,52(9):355 – 360.
YUAN Xingmao, Ll Xiaohe, WU Haiyan, et al. Design and experiment of rapid compost process and equipment by using solid manure on scale pig farm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021,52(9):355 – 360. (in Chinese)
[29] International Commission of Agricultural Engineering. Climatization of animal houses heat and moisture production at animal and house levels[R]. Horsens: Research Centre Bygholm, Danish Institute of Agricultural Sciences, 2002.
[30] 马承伟,苗香雯. 农业生物环境工程[M]. 北京:中国农业出版社,2010:240 – 243.
[31] DING Luyu, ZHANG Guoqiang, KRISTENSEN S, et al. CO_2 and CH_4 Emissions from the ground level of dairy open lots[J]. Journal of the Air & Waste Management Association, 2016, 66(7): 715 – 725.
[32] DING Luyu, LU Qikun, XIE Lina, et al. Greenhouse gas emissions from dairy open lots and manure stockpile in Northern China: a case study[J]. Journal of the Air & Waste Management Association, 2016, 66(3): 267 – 279.
[33] HENDRICKSON T, PERFORS A, NAVARRO D J, et al. Sample size, number of categories and sampling assumptions: exploring some differences between categorization and generalization[J]. Cognitive Psychology,2019, 111: 80 – 102.
[34] 马新宇,范意兴,郭嘉丰,等. 关于短文本匹配的泛化性和迁移性的研究分析[J]. 计算机研究与发展,2022,59(1):118 – 126.
MA Xinyu, FAN Yixing, GUO Jiafeng, et al. Research and analysis on generalization and migration of short text matching [J]. Computer Research and Development,2022,59(1):118 – 126. (in Chinese)