

DOI: 10.13718/j.cnki.xdzk.2025.10.020

王耀东, 燕振刚, 宋伟涛, 等. 牛舍甲烷浓度时序预测模型研究 [J]. 西南大学学报(自然科学版), 2025, 47(10): 234-244.

牛舍甲烷浓度时序预测模型研究

王耀东, 燕振刚, 宋伟涛, 杨发发

甘肃农业大学 信息科学技术学院, 兰州 730070

摘要: 甲烷浓度对牛生长发育影响较大, 对牛舍甲烷浓度进行预测可以为牛舍环境精准控制提供科学参考。构建了基于门控循环神经网络、改进麻雀搜索算法(ISSA)和反向传播算法(BP)相结合的牛舍甲烷预测模型。首先利用门控循环单元(GRU)模型提取牛舍环境数据中的非线性特征, 随后通过 ISSA 对 GRU 模型的超参数进行优化, 得到最优 GRU 模型, 从而提升模型对非线性特征的拟合能力, 最后采用 BP 模型对经过 ISSA-GRU 模型优化后的预测残差特征进行进一步拟合, 提升模型预测精度。利用 2024 年 4 月 29 日至 6 月 30 日在试验牛舍采集的数据进行训练和测试, 结果表明: 该模型能够有效拟合牛舍多环境参数。与 BP、GRU、GRU-BP、SSA-GRU、ISSA-GRU、ISSA-GRU-ARIMA 等模型相比, 本文提出的 ISSA-GRU-BP 模型具有较高的预测精度, 其 R^2 、RMSE 和 MAPE 分别为 0.934、 0.899×10^{-6} 和 9.638%。

关键词: 牛舍; 甲烷浓度预测; 门控循环神经网络; 改进麻雀搜索

算法; 反向传播算法; 残差

中图分类号: TP391

文献标识码: A

开放科学(资源服务)标识码(OSID):



文章编号: 1673-9868(2025)10-0234-11

A Study of Time-Series Prediction Model of Methane Concentration in Cattle Barns

WANG Yaodong, YAN Zhengang, SONG Weitao, YANG Fafa

College of Information Science and Technology, Gansu Agricultural University, Lanzhou 730070, China

Abstract: Methane concentration has a large impact on cattle growth and development, and the prediction of methane concentration in cattle barns can provide a scientific reference for the precise control of cattle barn environment. A methane prediction model based on the combination of gated recurrent neural network, improved sparrow search algorithm (ISSA) and back propagation algorithm (BP) was constructed

收稿日期: 2025-02-17

基金项目: 甘肃省重点研发计划项目(21YF5FA095); 甘肃省高等学校创新基金项目(2021A-057); 甘肃省财政厅项目(GSCZZ-20160909-03); 国家自然科学基金项目(31660347)。

作者简介: 王耀东, 硕士研究生, 研究方向为智慧农业、农业温室气体。

通信作者: 燕振刚, 博士, 教授。

for cattle barn. Firstly, the Gated Recurrent Unit (GRU) model was used to extract the nonlinear features from environmental data of the barn, and then the hyperparameters of the GRU model were optimized by ISSA to obtain the optimal GRU model, so as to improve the model's ability to fit the nonlinear features. Finally, the BP model was used to further fit the predicted residual features after optimization of the ISSA-GRU model, so as to improve the model prediction accuracy. Using the data collected in the experimental barn from April 29 to June 30, 2024 for training and testing, the results showed that the model was able to effectively fit the multi-environmental parameters of the barn. Compared with BP, GRU, GRU-BP, SSA-GRU, ISSA-GRU and ISSA-GRU-ARIMA, the ISSA-GRU-BP model proposed in this paper has a high prediction accuracy, with R^2 , $RMSE$, and $MAPE$ of 0.934, 0.899×10^{-6} , and 9.638%, respectively.

Key words: cattle barn; methane concentration prediction; gated recurrent neural network; improved sparrow search algorithm; back propagation algorithm; residuals

在全球范围内,甲烷(CH_4)是影响全球温室效应的一种重要温室气体,其排放带来的增温潜力是二氧化碳的 28 倍,而反刍动物是甲烷排放的重要贡献者^[1]。在我国,牛养殖是畜牧业中甲烷的最大来源之一^[2-3]。牛舍内空气质量较差会使动物的呼吸系统产生疾病^[4]。饲料在牛消化道中被微生物发酵,会产生甲烷并排出体外。甲烷本身对牛没有任何毒害作用,但牛舍内甲烷量的增大会使氧气相对减少,造成慢性缺氧,同时也会存在引起火灾甚至爆炸的危险^[5]。研究发现,通风率、舍内外温度、相对湿度对温室气体浓度均有显著影响^[6-8]。因此,监测并控制牛舍甲烷浓度的变化非常重要。

目前,国内外针对畜舍内环境的预测模型主要集中在畜舍二氧化碳和氨气浓度的预测上,对牛舍甲烷浓度预测研究较少。这些模型主要可以分为两大类。一类是基于数理化学统计模型,这类模型通常通过生成原理或影响因素来构建,往往依赖于大量的试验数据支持,其预测性能存在一定的局限性,例如:文献[9]利用肉牛和奶牛的干物质摄入量 and 营养成分建立了线性和非线性模型来预测甲烷产量;文献[10]和文献[11]通过对不同结构的牛舍建筑进行研究,得到牛舍的温室气体和氨气排放量。这类模型往往比较复杂,而且受畜舍环境、建筑结构等因素的影响。另一类基于机器学习智能算法的预测模型可以解决机理复杂且未知的非线性建模问题^[12],相对简单,而且对非线性函数具有良好的拟合能力,例如:文献[13]利用 LightGBM-SSA-ELM 组合模型结合新疆羊舍的环境数据,对羊舍的二氧化碳进行了准确预测;文献[14]和文献[15]分别构建了 SSA-LSTM 和 LSTM-AT 模型来预测温室环境。这类模型能够更好地提取气体浓度特征,同时预测模型精度也高。

通过采集系统采集的牛舍环境数据具有时序性、非线性等特点。门控循环单元(GRU)模型广泛应用于时间序列数据的预测^[16-17],解决了传统 RNN 梯度消失和梯度爆炸的问题,能够提取牛舍采集数据中的非线性关系,关注数据长期依赖的问题,而且训练速度较快。传统的 GRU 模型在训练过程中往往需要依靠人工经验手动调节参数,其性能受到很大影响。如何选择最佳参数组合以获得更好的预测效果,是 GRU 预测模型必须解决的问题之一。麻雀搜索算法(SSA)是由文献[18]提出的群智能优化算法,模拟了麻雀群觅食并逃避捕食者的行为。加入混沌序列、混沌扰动和高斯变异的改进麻雀搜索算法(ISSA)解决了 SSA 算法容易陷入局部最优的问题^[19]。BP 模型是一种多层前馈神经网络,学习过程包括正向传播和反向传播两部分。网络的第一层为输入层,最后一层为输出层,中间为隐藏层,通过不断调整网络的权重和阈值,使误差达到最小,具有强大的非线性映射能力^[20]。

本文使用牛舍环境监测数据作为输入量,通过 ISSA 算法对 GRU 模型的超参数进行优化调整,以提升 GRU 模型的非线性拟合能力,然后融合 BP 反向传播算法进一步提取 ISSA-GRU 模型预测残差序列的有用信息,构建 ISSA-GRU-BP 预测模型,提升模型预测精度,为牛舍甲烷控制提供可行参考。

1 材料与方法

1.1 试验区域

试验数据区域为甘肃省张掖市临泽县的一个试验牛舍(长约 36 m, 宽约 13 m)。舍内饲养了 50 头 6 月龄的西门塔尔犊牛(199 ± 1.07 kg, 雄性), 每天于 06:30 和 14:30 饲喂干草苜蓿和补充精料, 全天自由饮水。牛舍东南西北四面墙正中各带有一扇 3×4 m 的大门, 东西两面墙带有四扇 1×2 m 的窗户, 对称屋顶有中央脊通风口, 全天打开。由连接南北门的一条水泥过道将牛舍分为东西两侧, 临近过道为饲喂区。

1.2 试验数据采集

牛舍内安装了基于物联网的牛舍环境实时监测系统, 实时获取牛舍主体区域的环境数据并传输至监测平台。系统拓扑架构如图 1 所示, 系统实物图如图 2 所示。该系统以 STM32F103 控制芯片和 Zigbee-CC2530 通信芯片为核心, 获取牛舍实时的甲烷浓度、舍内温湿度、舍外温湿度和流动风速等数据, 数据以 1 min 为采样间隔进行记录, 并通过 Modbus 协议传输, 利用 4G 模块将采集的信息上传至物联网云平台。

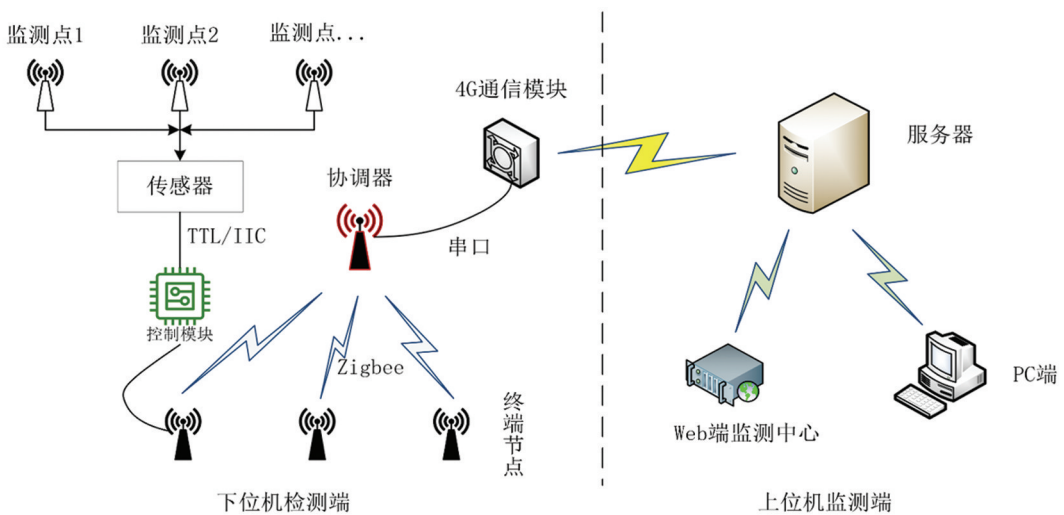


图 1 监测系统拓扑架构

1.3 试验数据预处理

1.3.1 数据处理

将采集的 2024 年 4 月 29 日至 2024 年 6 月 30 日的牛舍环境数据作为试验数据。由于 1 h 内牛舍环境参数变化不大, 将牛舍环境数据以 1 min 的时间间隔采集后, 进行平均化处理, 转化为 60 min 的时间间隔数据, 总计生成了 1 492 组环境数据。

1.3.2 异常数据处理

考虑到训练数据和测试数据中可能存在缺失部分, 为了避免预处理中出现数据分布不一致的问题, 本文使用了向前填充、向后填充与均值填充相结合的方式保证填充数据的覆盖率。同时采取拉伊达准则将试验数据中的异常值替换为其相邻数据的平均值。

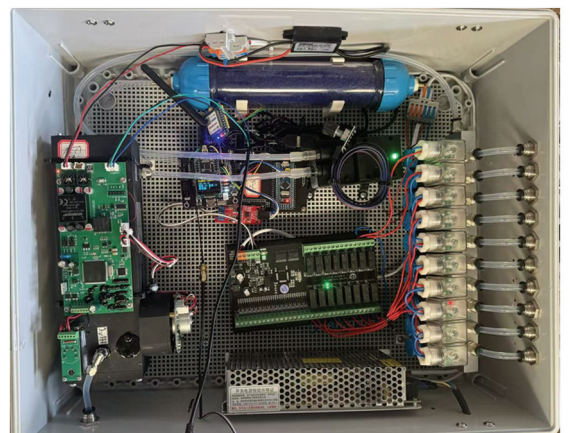


图 2 检测模块部分实物图

1) 更新门 z_t

更新门 z_t 决定了当前的隐藏状态应该由上一时刻的隐藏状态保留多少信息, 以及由当前输入生成多少新信息, 从而捕捉序列数据的短期依赖关系, 公式如下:

$$z_t = \sigma(\mathbf{W}_z \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_z) \quad (2)$$

2) 重置门 r_t

重置门 r_t 决定了前一时间步的隐藏状态在多大程度上被抛弃, 以便更好提取序列数据的长期依赖关系, 公式如下:

$$r_t = \sigma(\mathbf{W}_r \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_r) \quad (3)$$

3) 候选隐藏状态 $\hat{\mathbf{h}}_t$

候选隐藏状态 $\hat{\mathbf{h}}_t$ 是通过将当前输入和前一隐藏状态结合, 经过激活函数得到的潜在隐藏状态, 公式如下:

$$\hat{\mathbf{h}}_t = \tanh(\mathbf{W}_h \cdot [\mathbf{r}_t \cdot \mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_h) \quad (4)$$

4) 最终隐藏状态 \mathbf{h}_t

最终隐藏状态 \mathbf{h}_t 是通过更新门对上一时刻的隐藏状态和候选隐藏状态进行加权得到的, 公式如下:

$$\mathbf{h}_t = (1 - z_t) \cdot \mathbf{h}_{t-1} + z_t \cdot \hat{\mathbf{h}}_t \quad (5)$$

式(2) - (5) 中, σ 是 sigmoid 激活函数; \mathbf{x}_t 是当前神经元的输入; \mathbf{h}_{t-1} 是上一个神经元的隐藏状态; \mathbf{W}_z 、 \mathbf{W}_r 和 \mathbf{W}_h 分别是更新门、重置门和候选隐藏状态的权重矩阵; \mathbf{b}_z 、 \mathbf{b}_r 和 \mathbf{b}_h 分别是更新门、重置门和候选隐藏状态的偏置值。

本文选择均方误差(MSE)作为 GRU 模型的损失函数, 并利用 Adam 优化器优化权重, 以逐步减小误差, 获得高精度的预测模型。

2.2 ISSA 改进麻雀搜索算法

由于 GRU 预测模型所需超参数众多, 靠人工经验进行调参难度大且费时。通过优化算法对 GRU 预测模型超参数进行寻优, 计算出最优超参数组合, 以获得最优的 GRU 模型。SSA 麻雀搜索算法是一种智能群体优化算法, 能够快速进行全局搜索并及时收敛^[24], 但容易陷入局部最优解。而 Tent 混沌序列具有随机性、遍历性等特点^[19], 能够显著提升优化算法的寻优性能, 同时增加 SSA 算法种群的多样性, 使种群分布更加均匀从而加快收敛速度^[25], 还可以对局部极值通过混沌扰动扩大搜索空间, 并引入高斯变异对原始解周围进行重点搜索, 从而使搜索算法跳出局部最优^[16]。因此, 通过引入 Tent 混沌序列、混沌扰动和高斯变异得到改进后的麻雀搜索算法。在 ISSA 算法中, 每只麻雀对应一个解, 麻雀的数量则表示解的总数。一只麻雀包含所需要优化的超参数, 分别为时间窗口、学习率、批量大小、GRU 子层每层神经元数量。ISSA 算法步骤如下:

1) 使用公式(6)生成 Tent 混沌序列 S 并初始化种群, 然后将其映射到搜索空间产生解, 以增强种群的多样性和分布均匀性

$$S_{i,u} = (2S_i) \bmod 1 + \text{rand}(0, 1) \quad (6)$$

式中: $S_{i,u}$ 为种群中第 i 个个体 S_i 的第 u 维变量; $\text{rand}(0, 1)$ 为区间 $[0, 1]$ 内的随机数。

2) 本试验中通过均方误差来计算种群中所有个体的适应度 f_i , 根据适应度来评价个体的优劣, 并找到种群当前的全局最优位置 X_{best} 和最差位置 X_{worst} 。改进麻雀搜索算法会根据适应度调整种群个体的位置, 以逐步逼近全局最优解。

3) 随机选取种群中一定比例的个体作为发现者、捕食者和警戒者。根据适应度值和算法规则分别使用公式(7)、(8)对发现者位置和捕食者位置进行更新, 并使用公式(9)对警戒者位置进行更新。

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \exp\left(\frac{-i}{\alpha i_{\max}}\right) & R_2 < S_T \\ X_{i,j}^t + QL & R_2 \geq S_T \end{cases} \quad (7)$$

式中: $X_{i,j}^{t+1}$ 表示第 i 个个体在第 t 次迭代中的第 j 维的当前位置; i_{\max} 为设定的最大迭代次数; $\alpha \in (0, 1]$ 为一个随机数; $R_2 \in [0, 1]$ 为安全值; $S_T \in [0.5, 1]$ 为警戒值, 当 $R_2 < S_T$ 时, 代表当前位置是安全的, 发现者可以继续搜索, 当 $R_2 \geq S_T$ 时, 当前位置被认为存在风险, 警戒者发出警报, 所有麻雀需要立即更换位置; Q 表示服从正态分布的随机数; L 用于计算种群的分布情况(一个所有元素均为 1 的 $1 \times u$ 的矩阵)。

$$X_{i,j}^{t+1} = \begin{cases} X_P^{t+1} + |X_{i,j}^t - X_P^{t+1}| C^+ L & 0 \leq i \leq \frac{n}{2} \\ Q \exp\left(\frac{X_{\text{worst}}^t - X_{i,j}^t}{i^2}\right) & i > \frac{n}{2} \end{cases} \quad (8)$$

式中: X_P 为当前发现者的最优位置, X_{worst}^t 表示第 t 次迭代时的全局最差位置, C 为一个所有元素均为 1 或 -1 的 $1 \times u$ 矩阵, $C^+ = C^T(CC^T)^{-1}$ 。

$$X_{i,j}^{t+1} = \begin{cases} X_{\text{best}}^t + \beta(X_{i,j}^t - X_g^t) & f_i > f_{\text{best}} \\ X_{i,j}^t + K \frac{X_{i,j}^t - X_{\text{worst}}^t}{f_i - f_{\text{worst}} + \epsilon} & f_i = f_{\text{best}} \end{cases} \quad (9)$$

式中: X_{best}^t 表示第 t 次迭代时的全局最优位置; f_{best} 为全局最佳适应度, $f_i > f_{\text{best}}$ 表明麻雀处于危险位置, $f_i = f_{\text{best}}$ 表示警戒者发现天敌, 所有麻雀需要转移到安全区域; β 为步长控制参数(服从均值为 0、方差为 1 的正态分布的随机数); K 为 $(-1, 1)$ 上的随机数, 用来确定麻雀移动的方向; f_{worst} 表示全局最差适应度; ϵ 表示一个无限接近 0 的常数, 用于避免分母为 0, 并在每次迭代中更新。

4) 每完成一轮迭代后, 需要计算当前迭代的最小适应度, 以及 GRU 模型超参数寻优后的结果, 更新全局最优位置 X_{best} 和最差位置 X_{worst} , 并重新计算调整后种群适应度值。

5) 当个体的适应度值 f_i 小于种群的平均适应度 f_{avg} 时, 使用公式(10) 通过高斯扰动对较优的解进行局部优化, 探索其周围的解空间并进一步寻找当前最优解, 然后将变异后的解和原始解进行比较, 保留适应度最优的解。

$$X_i^{t+1} = X_i^t + \alpha \cdot (X_{\text{best}}^t - X_i^t) + \beta \cdot N(0, \sigma^2) \quad (10)$$

式中: X_i^t 为当前个体位置; X_{best}^t 为当前全局最优位置; α 为学习因子, 控制个体向最优解靠近的程度; β 为随机因子, 控制高斯扰动的幅度; $N(0, \sigma^2)$ 表示均值为 0、方差为 σ^2 的高斯分布随机数。

6) 在种群更新后, 可以确定当前的全局最优适应度及其对应的最优位置。一旦达到预设的最大迭代次数则停止运行; 否则继续重复执行步骤 3)。

2.3 BP 反向传播算法

尽管 ISSA-GRU 模型显著提高了 $f_i \geq f_{\text{avg}}$ 甲烷浓度预测的精度, 但在对训练集 2 进行预测时生成的残差序列 D_{CH_4} 仍然具有一定的特征, 可以进一步提取。而 BP 神经网络具有较强的拟合能力, 通过残差序列构建的数据集建立 BP 模型并进行迭代, 从而进一步提高预测模型的精度。

BP 神经网络(Back Propagation Neural Network)是一种多层前馈神经网络, 由输入层、隐含层和输出层 3 部分构成^[26](图 3)。BP 的核心机制包含两个步骤: 输入信号的正向传递与误差的反向传递, 用于持续优化网络中的权重和偏置。在正向传递过程中, 输入的样本数据依次通过隐含层并传递至输出层, 生成对应的输出结果。当预测结果与真实结果存在偏差时, 模型进入误差的反向传递阶段^[27]。反向传递将输出误差逐层反传至输入层, 并依据每一层接收到的误差调整网络的权重和阈值, 直至误差最终收敛^[28-29]。计算过程如下^[30]:

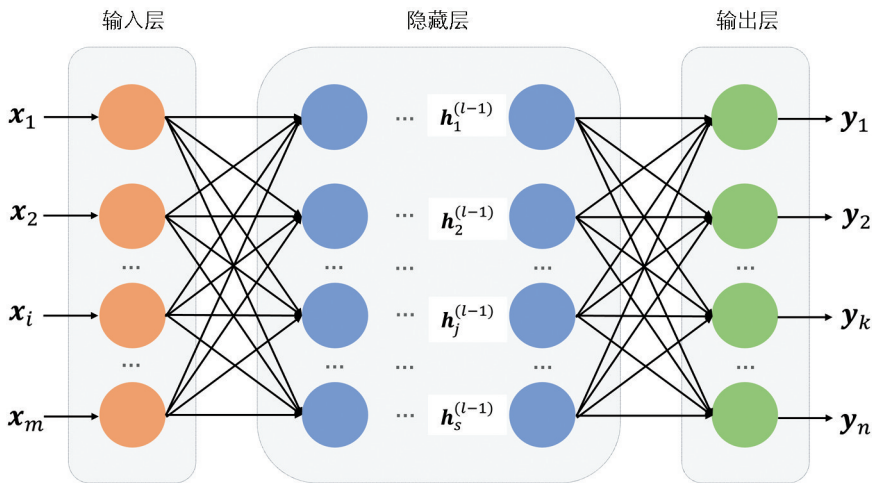


图 4 BP 神经网络结构图

1) 前向传播

加权求和与激活函数的计算公式如下所示:

$$z_i^{(l)} = \sum_j w_{ij}^{(l)} a_j^{(l-1)} + b_i^{(l)} \quad (11)$$

$$a_i^{(l)} = f(z_i^{(l)}) \quad (12)$$

式中: $z_i^{(l)}$ 表示第 l 层中第 i 个神经元的加权输入; $w_{ij}^{(l)}$ 为连接第 $(l-1)$ 层第 j 个神经元到第 l 层第 i 个神经元的权重; $a_j^{(l-1)}$ 表示第 $(l-1)$ 层第 j 个神经元的输出值(即激活值); $b_i^{(l)}$ 为第 l 层第 i 个神经元的偏置; $a_i^{(l)}$ 为第 l 层第 i 个神经元的输出值; f 代表激活函数。

2) 反向传播

输出层误差计算公式如下所示:

$$\delta_k^{(L)} = a_k^{(L)} - y_k \quad (13)$$

式中: $\delta_k^{(L)}$ 为输出层第 k 个神经元的误差; $a_k^{(L)}$ 为输出层第 k 个神经元的预测输出值; y_k 为目标值; L 表示输出层。

隐藏层的误差计算公式如下所示:

$$\delta_i^{(l)} = \sum_k \delta_k^{(l+1)} w_{ki}^{(l+1)} f'(z_i^{(l)}) \quad (14)$$

式中: $\delta_i^{(l)}$ 表示第 l 层第 i 个神经元的误差; $\delta_k^{(l+1)}$ 为第 $(l+1)$ 层第 k 个神经元的误差; $w_{ki}^{(l+1)}$ 为连接第 l 层第 i 个神经元到第 $(l+1)$ 层第 k 个神经元的权重; $f'(z_i^{(l)})$ 为激活函数 f 对 $z_i^{(l)}$ 的导数; $z_i^{(l)}$ 为第 l 层第 i 个神经元的加权输入。

权重更新和偏置更新计算公式如下所示:

$$w_{ij}^{(l)} = w_{ij}^{(l)} - \eta \delta_i^{(l)} a_j^{(l-1)} \quad (15)$$

$$b_i^{(l)} = b_i^{(l)} - \eta \delta_i^{(l)} \quad (16)$$

式中: $w_{ij}^{(l)}$ 表示连接第 $(l-1)$ 层第 j 个神经元到第 l 层第 i 个神经元的权重; η 代表学习率, 用于控制梯度下降的步长; $\delta_i^{(l)}$ 为第 l 层第 i 个神经元的误差; $a_j^{(l-1)}$ 为第 $(l-1)$ 层第 j 个神经元的激活值; $b_i^{(l)}$ 为第 l 层第 i 个神经元的偏置。

2.4 模型评价指标

为对本文建立的 ISSA-GRU-BP 牛舍甲烷浓度预测模型及对比模型进行评价, 本文使用均方根误差 (RMSE)、平均绝对百分比误差 (MAPE) 和决定系数 (R^2) 作为评价指标。

2.5 试验环境

本文模型开发环境为 AMD EPYC 9754 处理器, Nvidia RTX4090D 显卡, Windows 11 操作系统, py-

thon 3.7 编程语言, Anaconda 3 集成开发环境, Keras 框架和 SKlearn 框架组合。

3 结果与分析

3.1 牛舍环境数据及相关性分析

选取 2024 年 4 月 29 日至 2024 年 6 月 30 日的牛舍环境数据作为试验数据, 统计结果如表 1 所示。牛舍内甲烷浓度受到多种环境因素的影响, 将这些环境因素作为外部因素融入甲烷浓度预测模型, 有助于模型更全面地学习预测目标与环境变量之间的关系, 提升模型预测精度。对环境因素做 Pearson's 系数相关性分析, 结果如表 2 所示。

表 1 牛舍环境数据统计

环境因子	最大值	最小值	平均值	标准差
舍外温度	36.440	1.230	21.731	6.288
舍外相对湿度	90.447	5.256	30.752	19.877
甲烷	17.549	1.750	6.262	2.796
舍内温度	38.610	1.186	21.451	6.275
舍内相对湿度	82.066	9.169	40.918	20.261
流动风速	7.751	0.102	2.569	1.348

表 2 皮尔逊系数相关性分析

环境因子	甲烷	环境因子	甲烷
舍外温度	-0.62	舍内温度	-0.78
舍外相对湿度	0.36	舍内相对湿度	0.53
甲烷	1.00	流动风速	-0.47

牛舍内甲烷浓度与舍内外湿度具有正相关性, 舍内湿度相关系数为 0.53, 舍外湿度相关系数为 0.36; 甲烷浓度与舍内温度、舍外温度、流动风速呈负相关, 相关系数分别为 -0.78、-0.62 和 -0.47。相关性结果表明牛舍甲烷浓度与牛舍环境因素具有相关性, 构建模型时加入这些因素可以使模型更好地学习预测目标与环境因素的关系, 提升模型预测精度。因此, 本文选取舍外温度、舍外相对湿度、舍内温度、舍内相对湿度和流动风速作为牛舍甲烷浓度的外部影响因素。

3.2 试验过程

为了提高甲烷浓度预测模型的精度, 将 ISSA、GRU 和 BP 模型有机结合, 构建基于 ISSA-GRU-BP 的牛舍甲烷浓度预测模型, 首先应通过预试验确定 GRU 子层的层数, 以便于使用 ISSA 算法对建立的 GRU 模型超参数进行寻优。从 1 层到 10 层依次测试, 当 GRU 子层层数为 3 时, 甲烷浓度预测的 RMSE 最小, 为 10.880×10^{-6} , 所以在本试验中将 GRU 子层数目设置为 3。试验结果如图 5 所示。

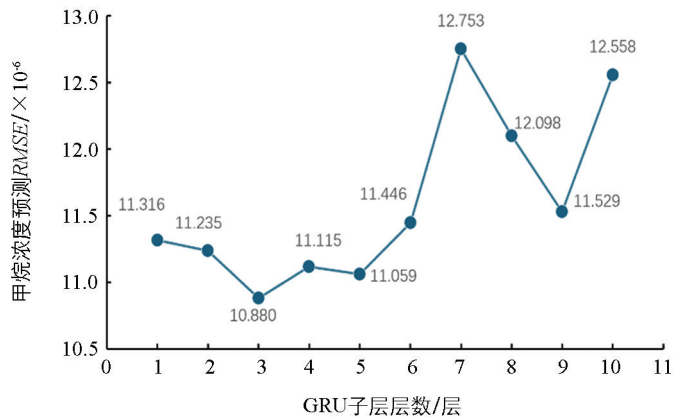


图 5 不同 GRU 子层数时模型预测的 RMSE

在确定 GRU 子层层数后, 通过 ISSA 算法对 GRU 神经网络的 6 个超参数进行寻优, 得到最优的 GRU 模型超参数组合, 寻优范围为: 时间窗口 1~10、学习率 0.0001~0.001、批数目 32~300、GRU 每层神经元个数 1~600。在 GRU 神经网络模型优化过程中, 要对 ISSA 算法的参数进行设置, 以达到算法的理想状态^[16,19,25], 参数设置为: 种群数 15, 最大迭代次数 50, 发现者比例 80%, 警戒者比例 20%, 预警值 0.8。GRU 模型的超参数优化范围和优化结果如表 3 所示。

表 3 GRU 模型超参数寻优范围及结果

寻优参数	寻优范围	寻优结果
时间窗口	1~10	2
学习率	0.0001~0.001	0.0001
批数目	32~300	89
GRU 子层 1 神经元个数/个	1~600	138
GRU 子层 2 神经元个数/个	1~600	284
GRU 子层 3 神经元个数/个	1~600	565

通过 ISSA 算法对 GRU 模型的超参数进行寻优后, 用 ISSA-GRU 预测模型对训练集 2 进行预测, 得到甲烷浓度的实际值与预测值之间的残差序列 D_{CH_4} , 残差序列如图 6 所示。

将残差序列 D_{CH_4} 作为数据集, 用于构建 BP 模型, 以进一步提取 ISSA-GRU 模型生成的残差序列特征。随后, 在测试集上对构建的 ISSA-GRU-BP 模型进行精度评估。通过图 7 可以看出, ISSA-GRU-BP 模型预测的牛舍甲烷浓度值变化趋势与实际值的变化趋势非常接近, 预测值与实际值的波动范围和变化模式基本一致, 误差较小, 表明模型能够很好地捕捉甲烷浓度的变化特征。在少数时间点(例如峰值或快速下降的区域), 预测值与实际值之间存在轻微的偏差, 但总体来看误差较小, 对模型的整体预测精度影响不大。甲烷浓度随着时间变化呈现一定的波动性。ISSA-GRU-BP 模型能够较好地跟随这种动态变化, 适合处理非线性且时序性明显的的数据。这也表明, 模型能优化 GRU 非线性拟合能力, 且在利用 BP 提取残差特征方面也有较好效果。

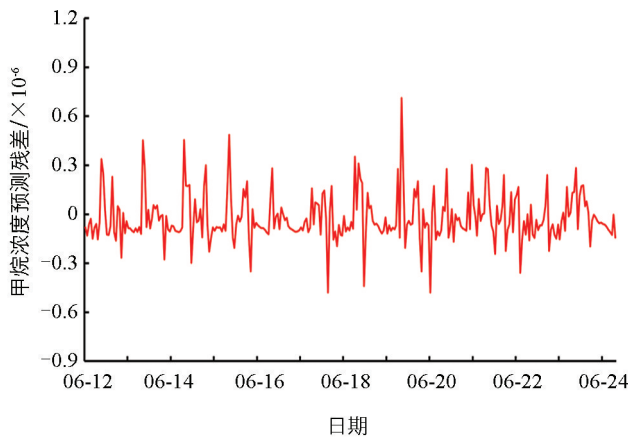


图 6 残差序列曲线

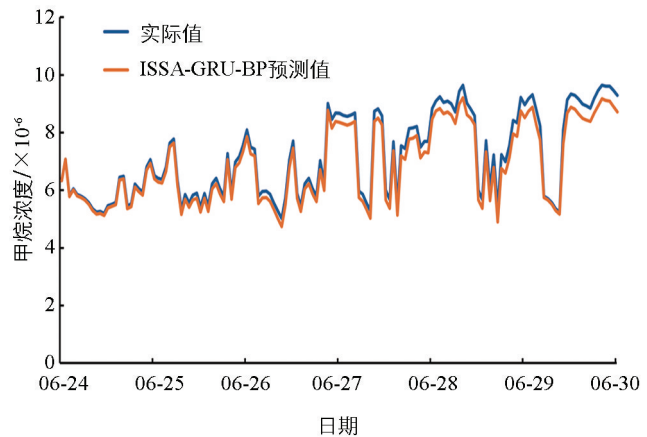


图 7 模型预测值对比

3.3 不同模型预测结果对比分析

为验证本文模型的预测效果, 将 BP、GRU、GRU-BP、SSA-GRU、ISSA-GRU、ISSA-GRU-ARIMA 与本文提出的 ISSA-GRU-BP 预测模型进行对比试验。不同模型相同模块参数设置相同。试验结果如表 4 所示。

表 4 不同模型预测精度对比

模型	R^2	$RMSE/10^{-6}$	$MAPE/\%$
BP	0.831	0.979	10.605
GRU	0.851	0.963	10.417
GRU-BP	0.895	0.929	10.012
SSA-GRU	0.885	0.937	10.098
ISSA-GRU	0.919	0.911	9.774
ISSA-GRU-ARIMA	0.913	0.915	9.833
ISSA-GRU-BP	0.934	0.899	9.638

由表 4 可见, 本文所构建的 ISSA-GRU-BP 预测模型精度最优, 甲烷浓度预测的 R^2 、 $RMSE$ 和 $MAPE$ 分别为 0.934、 0.899×10^{-6} 和 9.638%, 说明通过改进麻雀搜索算法 ISSA 优化 GRU 模型来提取牛舍环境因子的非线性特征并通过 BP 神经网络模型进行拟合可以提高模型的预测精度。

4 结论

采用 ISSA 算法对 GRU 模型的超参数进行优化, 增强了 GRU 模型的非线性拟合能力, 从而有效提升了甲烷浓度预测模型的精度。试验证明, 本文提出的 ISSA-GRU-BP 模型能够有效拟合牛舍多环境参数, 提高甲烷浓度预测, 预测精度高于 GRU、BP 等模型, 为牛群生长环境改善提供了新的思路和方法。

未来可以增加数据样本量和时间跨度, 涵盖更多的季节变化和极端气候条件, 以提高模型的泛化能力, 并可以进一步加入其他影响因素, 例如饲料成分、牛舍通风方式、气压等, 以提升模型对复杂关系的捕捉能力。

参考文献:

- [1] KUMARI S, FAGODIYA R, HILOIDHARI M, et al. Methane Production and Estimation from Livestock Husbandry: A Mechanistic Understanding and Emerging Mitigation Options [J]. *Science of the Total Environment*, 2020, 709: 136135.
- [2] 管林森. 加强碳中和目标下畜牧业绿色低碳发展 [EB/OL]. (2022-03-04) [2024-02-11]. <https://www.cnfood.cn/article?id=1499630621445816322>.
- [3] 中华人民共和国生态环境部. 中华人民共和国气候变化第四次国家信息通报 [EB/OL]. (2023-12-29) [2024-07-26]. <https://www.mee.gov.cn/ywdt/hjywnews/202312/W020231229717234502302.pdf>.
- [4] PANEL ON ANIMAL HEALTH AND WELFARE (AHAW) E F S A. Scientific Opinion on the Welfare of Cattle Kept for Beef Production and the Welfare in Intensive Calf Farming Systems [J]. *EFSA Journal European Food Safety Authority*, 2012, 10(5): 2669.
- [5] 刘明, 张恩平, 宋宇轩. 牛舍有害气体排放规律及减除措施研究进展 [J]. *家畜生态学报*, 2019, 40(5): 76-81.
- [6] HUANG DD, GUO H Q. Diurnal and Seasonal Variations of Greenhouse Gas Emissions from a Naturally Ventilated Dairy Barn in a Cold Region [J]. *Atmospheric Environment*, 2018, 172: 74-82.
- [7] QU Q B, GROOT J C J, ZHANG K Q, et al. Effects of Housing System, Measurement Methods and Environmental Factors on Estimating Ammonia and Methane Emission Rates in Dairy Barns: A Meta-Analysis [J]. *Biosystems Engineering*, 2021, 205: 64-75.
- [8] HEMPEL S, SAHA C K, FIEDLER M, et al. Non-Linear Temperature Dependency of Ammonia and Methane Emissions from a Naturally Ventilated Dairy Barn [J]. *Biosystems Engineering*, 2016, 145: 10-21.
- [9] WANG Y D, SONG W T, WANG Q, et al. Predicting Enteric Methane Emissions from Dairy and Beef Cattle Using

- Nutrient Composition and Intake Variables [J]. *Animals*, 2024, 14(23): 3452.
- [10] 赵婉莹, 许立新, 王朝元, 等. 不同地面形式自然通风奶牛舍冬季温室气体和氨气排放量 [J]. *中国农业大学学报*, 2020, 25(1): 142-151.
- [11] TABASE R K, NÆSS G, LARRING Y. Ammonia and Methane Emissions from Small Herd Cattle Buildings in a Cold Climate [J]. *Science of the Total Environment*, 2023, 903: 166046.
- [12] 王文广, 赵文杰. 基于 GRU 神经网络的燃煤电站 NO_x 排放预测模型 [J]. *华北电力大学学报(自然科学版)*, 2020, 47(1): 96-103.
- [13] 尹航, 吕佳威, 陈耀聪, 等. 基于 LightGBM-SSA-ELM 的新疆羊舍 CO₂ 浓度预测 [J]. *农业机械学报*, 2022, 53(1): 261-270.
- [14] 祖林禄, 柳平增, 赵妍平, 等. 基于 SSA-LSTM 的日光温室环境预测模型研究 [J]. *农业机械学报*, 2023, 54(2): 351-358.
- [15] 张观山, 丁小明, 何芬, 等. 基于 LSTM-AT 的温室空气温度预测模型构建 [J]. *农业工程学报*, 2024, 40(18): 194-201.
- [16] 谢秋菊, 马超凡, 王圣超, 等. 猪舍氨气与二氧化碳浓度变化时序预测模型优化 [J]. *农业机械学报*, 2023, 54(7): 381-391.
- [17] 马超凡, 谢秋菊, 王圣超, 等. 融合统计学习与深度学习的猪舍环境预测模型 [J]. *山西农业大学学报(自然科学版)*, 2022, 42(6): 24-32.
- [18] XUE JK, SHEN B. A Novel Swarm Intelligence Optimization Approach: Sparrow Search Algorithm [J]. *Systems Science & Control Engineering*, 2020, 8(1): 22-34.
- [19] 吕鑫, 慕晓冬, 张钧, 等. 混沌麻雀搜索优化算法 [J]. *北京航空航天大学学报*, 2021, 47(8): 1712-1720.
- [20] 刘春红, 杨亮, 邓河, 等. 基于 ARIMA 和 BP 神经网络的猪舍氨气浓度预测 [J]. *中国环境科学*, 2019, 39(6): 2320-2327.
- [21] 李建丽, 程雪, 王娟. 基于 CFD 的冬季房舍式牛舍湿热环境模拟及优化 [J]. *中国农机化学报*, 2024, 45(10): 69-75, 81.
- [22] CHO K, VAN MERRIENBOER B, GULCEHRE C, et al. Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation [C] // *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha: Association for Computational Linguistics, 2014.
- [23] 李英豪, 施爱平, 范欣竹. 基于 Attention-GRU 模型的酿酒葡萄蒸腾量预测 [J]. *排灌机械工程学报*, 2024, 42(11): 1181-1188.
- [24] 张永芳, 王芳. 基于 SSA-RBF 网络的日光温室温湿度预测模型研究 [J]. *河北农业大学学报*, 2021, 44(3): 115-121.
- [25] 张娜, 赵泽丹, 包晓安, 等. 基于改进的 Tent 混沌万有引力搜索算法 [J]. *控制与决策*, 2020, 35(4): 893-900.
- [26] 孙益壮. 基于激光扫描的综放工作面放煤量自适应监测技术 [D]. 徐州: 中国矿业大学, 2022.
- [27] 王秀康, 邢英英, 张富仓. 膜下滴灌施肥番茄水肥供应量的优化研究 [J]. *农业机械学报*, 2016, 47(1): 141-150.
- [28] 王小川, 史峰, 郁磊, 等. MATLAB 神经网络 43 个案例分析 [M]. 北京: 北京航空航天大学出版社, 2013.
- [29] 朱凤磊, 张立新, 胡雪, 等. 基于蝙蝠优化 BP-PID 算法的精准施肥控制系统研究 [J]. *农业机械学报*, 2023, 54(S1): 135-143, 171.
- [30] 李沐阳, 高建虎, 雍学善, 等. 卷积神经网络的半监督层位追踪方法 [J]. *石油地球物理勘探*, 2024, 59(5): 938-947.
- [31] 汤岩, 王福林, 焦扬. 黑龙江省农业机械总动力预测分析——基于 ARIMA 模型 [J]. *农机化研究*, 2014, 36(2): 44-47.