

曾志雄, 罗毅智, 余乔东, 等. 基于时间序列和多元模型的集约化猪舍温度预测 [J]. 华南农业大学学报, 2021, 42(3): 111-118.
ZENG Zhixiong, LUO Yizhi, YU Qiaodong, et al. Temperature prediction of intensive pig house based on time series and multivariate models[J]. Journal of South China Agricultural University, 2021, 42(3): 111-118.

基于时间序列和多元模型的集约化猪舍温度预测

曾志雄¹, 罗毅智¹, 余乔东¹, 蔡 任¹, 吕恩利¹, 夏晶晶^{1,2}

(1 华南农业大学 工程学院, 广东 广州 510642; 2 广东机电职业技术学院, 广东 广州 510515)

摘要:【目的】从挖掘猪舍历史环境参数数据时序信息角度出发, 提出基于时间序列模型和多元模型序列的猪舍温度预测模型。【方法】采取缺失部分环境因子统计预测, 评估猪舍环境中相对湿度、二氧化碳浓度、氧气浓度等环境因子对温度预测的影响程度。针对猪舍温度时间序列进行数据预处理, 滤除错误值和缺失值, 采用时间序列模型构建基于门控循环单元网络 (Gated recurrent unit, GRU) 的猪舍温度预测模型, 采用多元模型建立基于梯度提升决策树 (Extreme gradient boosting, XGBoost) 缺失值重要程度的猪舍温度预测模型。将该预测模型用于预测广东省某集约化猪场母猪分娩舍温度, 并与循环神经网络 (Recurrent neural network, RNN) 模型、反向神经网络 (Back propagation neural network, BPNN) 模型进行对比试验。【结果】对比温度预测值与实测值发现, 基于 GRU 模型对应的猪舍温度均方根误差和平均绝对误差分别为 0.25 和 0.19 ℃, 平均绝对百分比误差为 0.65%; 基于 XGBoost 多元模型的猪舍温度均方根误差和平均绝对误差分别为 1.21 和 0.71 ℃, 平均绝对百分比误差为 2.50%。在时间序列的温度预测模型中, GRU 模型表现出更优的预测效果; 在多元模型的温度预测中, XGBoost 模型的预测效果更优。【结论】本研究使用的 GRU 模型在时间维度上对母猪分娩舍温度的变化起到了预警作用, 确定了各种环境参数对温度的影响程度, 为养殖环境的精细调控提供了参考。

关键词: 分娩舍; 温度预测; 时间序列模型; 多元特征模型; 特征重要性
中图分类号: S23 **文献标志码:** A **文章编号:** 1001-411X(2021)03-0111-08

Temperature prediction of intensive pig house based on time series and multivariate models

ZENG Zhixiong¹, LUO Yizhi¹, YU Qiaodong¹, CAI Ren¹, LÜ Enli¹, XIA Jingjing^{1,2}
(1 College of Engineering, South China Agricultural University, Guangzhou 510642, China;
2 Guangdong Mechanical & Electrical Polytechnic, Guangzhou 510515, China)

Abstract: 【Objective】 On account of history environmental factors such as temperature and humidity in farrowing houses, based on the data of a farrowing house of a modern pig farm in Guangdong Province, the temperature prediction model of pig house was studied based on time series and multivariate models. 【Method】 The effects of relative humidity, carbon dioxide concentration, oxygen concentration and other environmental factors on temperature prediction in pigsty was evaluated by statistical prediction of missing part environmental factors. Data preprocessing was carried out for the time series of piggy temperature to filter out

收稿日期: 2020-10-13 网络首发时间: 2021-03-03 13:44:13
网络首发地址: <https://kns.cnki.net/kcms/detail/44.1110.s.20210303.1134.004.html>
作者简介: 曾志雄 (1989—), 男, 实验师, 博士, E-mail: zhixzeng@scau.edu.cn; 通信作者: 吕恩利 (1979—), 男, 副教授, 博士, E-mail: enlilv@scau.edu.cn
基金项目: 国家重点研发计划 (2018YFD0701002); 国家自然科学基金 (31901736, 31971806); 广东省普通高校“服务乡村振兴计划”重点领域专项 (2019KZDZX2040); 广东省重点领域研发计划 (2019B020225001); 广东省畜禽疫病防治研究重点实验室开放课题 (YDWS1904)

the error value and missing value. The time series model was used to construct the temperature prediction model of piggery based on gated recurrent unit (GRU). Based on extreme gradient boosting (XGBoost), the temperature prediction model of missing value importance degree in pigsty was established using multivariate model. The prediction model was applied to predict the temperature of a farrowing house of an intensive pig farm in Guangdong Province, and compared with the recurrent neural network (RNN) model and back propagation neural network (BPNN) model. 【Result】 Comparing the predicted value with the measured value, root mean square error (RMSE), mean absolute error (MAE) and mean absolute percent error (MAPE) of the GRU model were 0.25 °C, 0.19 °C and 0.65% respectively, while RMSE, MAE and MAPE of the XGBoost multicomponent model were 1.21 °C, 0.71 °C and 2.50% respectively. For temperature prediction based on time series models, GRU model showed better prediction effect. The XGBoost model was better in temperature prediction for multivariate model. 【Conclusion】 The CRU model used in this study plays an early warning role on the temperature change of the farrowing house in time dimension, and the influence degree of various environmental parameters on temperature is also determined, which provides a reference for the fine regulation of the breeding environment.

Key words: farrowing house; temperature prediction; time series model; multivariate model; feature importance

集约化猪舍是一个复杂的小气候环境,受内、外部环境的影响,猪舍内部温度具有波动、滞后、强干扰、多耦合等特征。而温度是养猪生产过程中的重要影响因素^[1]。猪舍内部环境过热、过冷、潮湿等不良环境条件,可导致猪体不适、生产能力下降、成活率降低,甚至诱发感冒、腹泻或呼吸道疾病等^[2]。据美国国家研究委员会 (National research council, NRC) 报道,在一定范围内,气温每升高 1 °C,猪采食量减少 40 g;在持续高温 (28~35 °C) 环境下,生猪日采食量较常温下降 14%,日增质量下降 21%,料重比增加 0.23。为降低能耗、减少饲料消耗、优化猪舍环境,前人研究了生猪养殖环境下相对湿度、二氧化碳浓度、氧气浓度、氨气浓度、硫化氢浓度等环境因子对温度预测的影响^[3-4],进而精准预测猪舍温度,并提出环境调控与预警方案^[5-6],提高了生猪养殖存活率,对我国养猪产业具有重要的实际意义。

目前,国内外学者主要从理化统计模型和算法分析 2 个方面对养殖场温度预测进行相关研究^[7-14]。Norton 等^[15] 和 Andonov 等^[16] 探索了自然通风下畜舍温度的变化规律,但受限於应用场景,模型稳定性尚不明确,预测精度还不能满足应用。随着计算机性能的发展,算法分析展现其稳定性和预测精度高的特点。杨柳^[17] 提出基于历史温湿度序列进行多尺度分析,得到温湿度的变化规律曲线,并通过支持向量机回归预测出下一时刻猪舍内温湿度值,该算法发挥了支持向量机对线性数据的拟合特性,但是对温度的预测精度及准确性较低。前人基于模糊方法技术,利用组态软件研究了猪舍内环境因子的相互影响规律^[18-21],但模型并未对环境因素进行

针对性调控^[22]。陈昕等^[23] 基于神经网络和二次聚类,进一步对温室温度进行二次预测,相比于未聚类的 BP 神经算法,提高了温度预测的准确度。杨亮等^[24] 和 Rodriguez 等^[25] 提出基于 EMD-LSTM 的猪舍氨气浓度预测模型,将氨气进行尺度分解,能够较好地拟合猪舍内环境参数与氨气浓度之间的关系。

本文采用时间序列模型^[20],预测未来 10 min 的温度数据;基于多元模型,研究环境因素中的各个特征对温度的影响程度,同时添加 BP 神经网络模型^[26]、循环神经网络^[27] 等,开展对比研究。结合猪舍多点多源环境监测系统,建立基于时间序列和多元模型的猪舍温度预测模型,为猪舍环境的精细调控提供参考。

1 研究方法

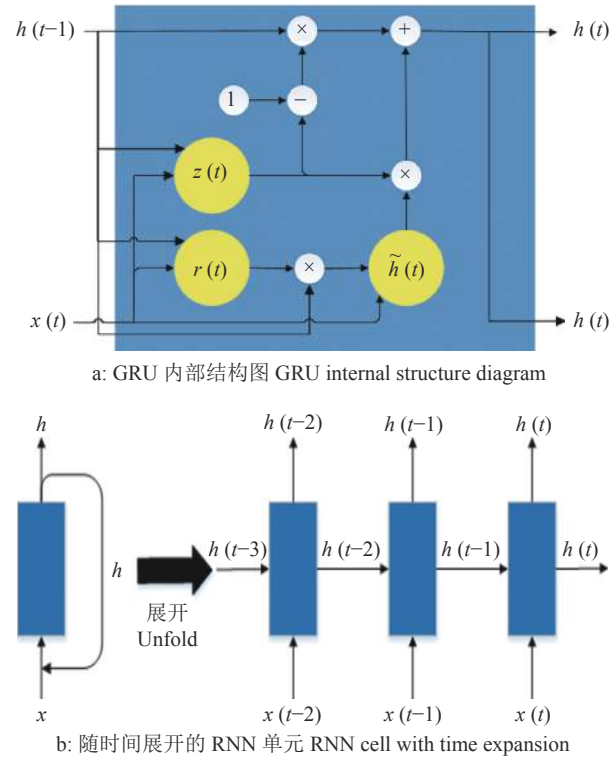
1.1 门控循环单元网络算法

门控循环单元网络 (Gated recurrent unit, GRU) 内部结构如图 1a 所示。GRU 是一种在长短期神经网络 (Long short-term memory, LSTM) 基础上进行改进的循环神经网络 (Recurrent neural network, RNN) 算法^[28-29],图 1b 为随时间展开的 RNN 单元,该算法能够克服传统 RNN 算法梯度消失的问题。与 LSTM 相比,减少了训练所需的参数、降低模型复杂度、提升训练效率。

GRU 算法具体公式如下:

$$r(t) = \text{sigmoid}[U_r h(t-1) + W_r x(t)], \tag{1}$$

$$z(t) = \text{sigmoid}[U_z h(t-1) + W_z x(t)], \tag{2}$$



$h(t-1)$ 为上一个传输下来的状态; $x(t)$ 为当前节点的输入; $z(t)$ 为控制更新的门控; $r(t)$ 为控制重置的门控; $\tilde{h}(t)$ 为重置后的数据; $h(t)$ 为当前的隐藏状态

$h(t-1)$ is the last state transferred down; $x(t)$ is the input for the current node; $z(t)$ is a gating control for controlling updates; $\tilde{h}(t)$ is the reset data; $h(t)$ is the current hidden state

图 1 循环神经网络结构图

Fig. 1 Cyclic neural network structure diagram

$$\tilde{h}(t) = \text{ReLU}\{U[r(t) \odot h(t-1)] + Wx(t)\}, \tag{3}$$

$$h(t) = [1 - z(t)]h(t-1) + z(t)\tilde{h}(t), \tag{4}$$

式中, U_r 和 W_r 分别是各个门的隐藏状态权重和输入变量权重, $\odot$ 是 Hadamard 积, $h(t)$ 表示 t 时刻的隐藏状态, $\tilde{h}(t)$ 为重置后的数据, $x(t)$ 为 t 时刻的输入, $y(t)$ 是 t 时刻的输出, 同时也作为 $t+1$ 时刻的输入。

1.2 极端梯度提升

极端梯度决策树 (Extreme gradient boosting, XGBoost) 是梯度提升 (Gradient boosting) 的一种高效系统实现, 其对梯度提升决策树 (Gradient boosting decision tree, GBDT) 进行了优化改进, 能够并行计算、近似建树和对稀疏数据进行有效处理, 对 CPU 和内存的使用进行了优化, 使其在机器学习领域展现了很好的效果^[30]。其目标函数为

$$\text{Obj}(\varphi) = \sum_{i=1}^n l(\hat{y}_i, y_i) + \sum_{k=1}^g \Omega(f_k), \tag{5}$$

$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \omega^2, \tag{6}$$

式中, $\hat{y}_i$ 为预测值, y_i 为真实值, l 为损失函数, $\Omega(f_k)$ 为正则化项, f_k 为一棵决策树, γ 和 λ 为超参数, T 为决策树中的叶子节点数, ω 为决策树中叶子节点得分。XGBoost 还使用了二阶泰勒展开, 设第 t 次的损失函数为

$$L^{(t)} = \sum_{i=1}^n [l(y_i, \hat{y}_i^{(t-1)}) + f_t(x_i)] + \Omega(f_t), \tag{7}$$

对 $L^{(t)}$ 做二阶泰勒展开

$$L^{(t)} \cong \sum_{i=1}^n \left[l(y_i, \hat{y}_i^{(t-1)}) + g_i f_t(x_i) + \frac{1}{2} g_i f_t^2(x_i) \right] + \Omega(f_t), \tag{8}$$

式中, $g_i = \partial_{\hat{y}_i^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)})$ 。

1.3 基于时间序列和多元模型的集约化猪舍温度预测模型

猪舍温度数据序列受猪舍内其他环境因子影响, 例如相对湿度、空间位置、氧气浓度、二氧化碳浓度等。本研究在同一时刻, 探讨环境因子缺失对猪舍温度数据预测的影响, 构建基于时间序列与多元模型的猪舍温度预测模型 (图 2)。

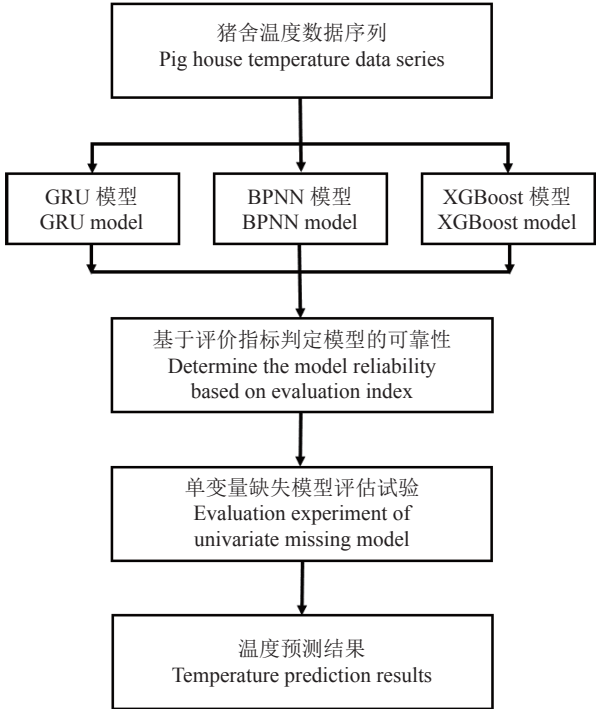


图 2 基于时间序列和温度序列的猪舍温度预测流程图

Fig. 2 A flow chart of temperature prediction of piggery based on time series and temperature series

具体步骤如下:

1) 数据采集与处理。利用无线多点多源远程监测系统^[31], 采集母猪分娩舍内温度、氧气浓度、二氧化碳浓度、相对湿度等参数, 然后对系统采集的数

据进行预处理, 去除错误值, 最后归一化, 并对数据进行预处理分析。

2) 基于 GRU 时间序列猪舍温度预测模型的可靠性检测。采用 GRU 算法, 构建基于温度的时间序列预测模型, 模型的输入参数是分娩母猪舍二氧化碳浓度、相对湿度和氧气浓度, 模型的输出参数是分娩母猪舍未来 10 min 的舍内温度值。首先, 初始化 GRU 预测模型参数; 随后, 不断优化调整预测模型参数, 提高分娩舍内温度预测的准确度。

3) 基于单变量缺失的模型评估检测。删减单个环境因子, 构建缺失单个变量的模型, 采用步骤 2) 的方法预测分娩舍内温度, 定义评价单变量缺失模型的评价指标, 根据评价指标, 得到各个环境因子的重要程度。首先, 基于 GRU 温度预测模型, 将模型的输入参数删减某一个环境因子, 模型的输出参数是分娩舍未来 10 min 的舍内温度预测值; 其次, 根据评价指标, 计算该环境因子对猪舍温度预测模型的影响程度; 最后, 根据统计结果, 分析各个环境因子的重要程度。

1.4 集约化猪舍温度预测试验

1.4.1 数据采集 以广东省云浮市某集约化种猪场的母猪分娩舍作为研究对象, 该猪舍长度 1 990 cm、宽度 1 100 cm, 屋顶采用平顶式设计, 地板类型为半漏缝式地板, 通风方式为机械通风。利用无线多点多源远程监测系统进行数据采集^[32], 采集频率为每 10 min 采集 1 次。该场区猪舍配备降温湿帘、屋顶小窗及侧墙负压风机等环控设施。试验开始前, 提前进入猪舍内进行勘察踩点, 详细记录猪舍的猪只类型和数量、建筑尺寸、通风方式、环控原理等。绘制猪舍平面布局图并进行监测区域划分, 为节点设备的合理布置提供参考。试验时间为 2018 年 7 月 15—17 日, 除去设备布置、清理节点、电池充电等耗时, 总有效测试时长约 72 h。选取舍内 41 台无线监测节点(下文统称“节点”)采集到的 16 981 条有效环境信息数据作为原始试验数据, 包括温度、相对湿度、二氧化碳浓度、氧气浓度、数据包上传时间、节点号等各类试验环境信息。试验完成后, 下载试验数据并分类编排整理, 对原始数据进行预处理, 剔除异常数据包, 进行统计学分析并绘图。

1.4.2 节点工作原理 图 3 为无线多点多源远程监测系统工作原理的拓扑结构, 舍内各个节点测得的环境数据经 DTU 上传至服务器, 工作人员可登录网站或者移动端查看。

1.4.3 节点布置方法 为避免产后母猪的应激反应对哺乳仔猪造成影响, 在猪场管理人员的许可下

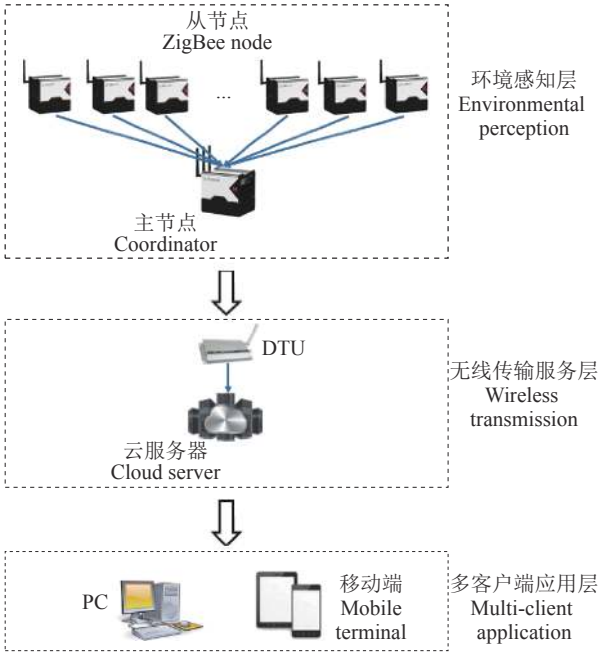


图 3 监测系统工作原理

Fig. 3 The operating principle of monitoring system

进行了产前分娩舍的试验测试。试验分娩舍内存有妊娠期 107 d(产前 7 d) 的待产母猪共 30 头。建筑外墙一侧设置有降温湿帘, 舍内通风采用屋顶小窗进风, 侧墙负压风机排风方式。屋顶小窗为横向排列, 每排 5 个小窗共 2 排。侧墙负压风机为纵向排列, 风机包括 1 台变频风机 (80 cm) 和 3 台恒速风机 (100 cm)。试验猪舍长度 1 990 cm、宽度 1 100 cm、高度 260 cm, 过道宽度 104 cm; 每个哺乳母猪栏单格长度 240 cm、宽度 182 cm, 总计 30 格。单格内限位栏宽度 70 cm, 仅容纳 1 头哺乳母猪。猪舍地面下方为水泡粪地沟, 沟内设有地沟风机进行排风。

试验分娩舍内部空间适中, 待产母猪的限位栏间隔较宽, 3 排栏体单格均为纵向排列。母猪头对头、尾对尾分布, 待产母猪腹部一侧留有仔猪活动空间。根据实地勘察结果, 结合猪只分布及舍内通风特点, 选取部分限位栏单格内的母猪及仔猪猪头区域作为节点设备的主要监测点, 并选取舍内出风口区域及舍外区域作为试验对照组。

将节点设备进行合理布置划分, 以研究试验分娩舍内各环境因素的分布特性。为此, 将舍内监测区域详细划分为母猪呼吸区域高度平面 (0.70 m)、仔猪呼吸区域高度平面 (0.35 m)、出风口区域高度平面 (0.70 m) 及舍外区域, 分别挂设 15、9、3 和 2 台从节点设备, 设置每 10 min 采集 1 次数据。主节点置于入口信号较好区域, 确认主、从节点位置编号并开启电源进行 72 h 不间断监测。试验过程中通过后台定时查看各节点设备数据上传记录。试

验分娩舍平面结构图及监测区域主、从节点分布如图4所示。

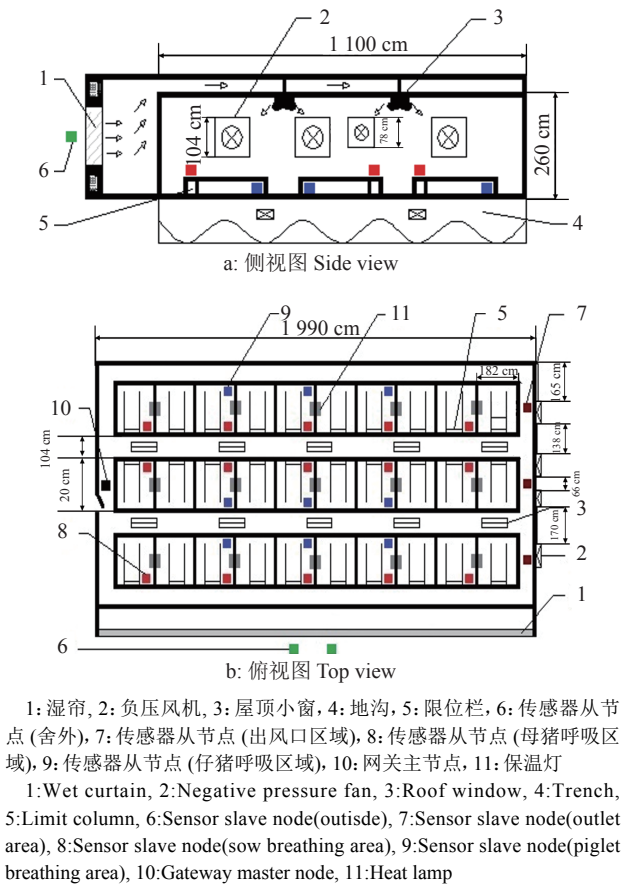


图4 试验分娩舍侧、俯视图及各区域节点分布

Fig. 4 Side and top views of test farrowing pig house and node distribution in each region

1.4.4 数据预处理 参照文献[33]对系统采集到的数据进行预处理, 分为3个步骤:

1) 异常数据处理。本次采集到的数据中, 数据异常情况主要为缺失值和错误值2种。因间歇断电造成的数据异常、少数数据丢失, 即为缺失值; 本试验中数据的错误值主要是指温度或者相对湿度值, 小于或等于0的数据, 由试验条件可知显然这是不合理的数据。对这2类异常数据使用拉格朗日插值法, 以异常点两侧的正常数据取平均值填充, 计算公式如下:

$$\begin{cases} X_n = \text{Error} - \text{Value} \\ X'_n = \frac{X_{n-1} + X_{n+1}}{2} \end{cases}, \quad (9)$$

式中, Error 为异常值, Value 为正常值, X_n 和 X'_n 分别为异常数据值及替代值, n 表示样本所在的时刻序列号。

2) 按节点划分数据集进行时间序列化处理。由于时间序列的试验需要在同一观测节点, 因此按节点将数据分成了41组, 并根据时间形成序列。根据程序设定, 按照时间序以2:1的比例划分为训练集

和预测集。

3) 归一化处理。由于数据中各种变量的量纲不同, 为提高预测的准确性, 本文对数据进行归一化处理, 计算公式如下:

$$X^* = \frac{X_n - X_{\min}}{X_{\max} - X_{\min}}, \quad (10)$$

式中, X^* 为归一化后的数值, X_n 为观测值, $X_{\max}$ 和 $X_{\min}$ 分别表示最大值和最小值。

1.4.5 结果评价指标 为了判断时间序列是否具有预测价值, 基于统计学方法对时间序列进行平稳性检验和白噪声检验, 指标为单位根检验 P 值, 当平稳性检验的 P 小于0.05时, 该序列是平稳序列, 反之为非平稳序列; 当白噪声检验的 P 小于0.05时, 该序列是非白噪声序列, 反之为白噪声序列。从而判断该组时间序列是否具有预测价值。

为了比较不同预测方法的预测结果, 采用3种误差指标: 平均绝对误差 (Mean absolute error, MAE)、平均绝对百分比误差 (Mean absolute percent error, MAPE) 和均方根误差 (Root mean square error, RMSE) 来衡量各模型的预测结果, 指标数值越小, 预测方法的精度越高, 说明预测效果越好^[34]。三者的表达式如下:

$$\text{MAE} = \frac{1}{n} \sum_{t=1}^n |\text{Observed}_t - \text{Predicted}_t|, \quad (11)$$

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{\text{Observed}_t - \text{Predicted}_t}{\text{Observed}_t} \right| \times 100\%, \quad (12)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (\text{Observed}_t - \text{Predicted}_t)^2}, \quad (13)$$

式中, Observed_t 和 Predicted_t 分别表示第 t 个观测值和预测值, n 表示样本规模。

选取最优多元模型, 采用单变量缺失预测温度。分别使用XGBoost、反向传播神经网络 (Back propagation neural network, BPNN)^[35]和线性回归 (Linear regression, LR)^[36]建立多元模型, 构建缺失单个变量的模型 (以下称为“缺失变量预测模型”), 该模型误差越大, 说明对应的缺失变量对预测结果的误差越大, 定义某一自变量对预测结果的贡献度得分为 Score_i , Score_i 越高, 对应的缺失变量对预测结果越重要。计算公式如下:

$$\text{Score}_i = \lg \text{MAPE}_i, \quad (14)$$

式中, MAPE_i 指的是第 i 个缺失变量的MAPE。

1.4.6 对比模型 为了评估GRU模型和XGBoost模型的效果, 基于时间序列模型, 添加RNN模型^[37]

进行对比;基于多元模型,添加 BPNN 和 LR 模型进行对比。

2 结果与分析

2.1 基于 GRU 时间序列的猪舍温度预测模型可靠性测试结果

针对猪舍温度时间序列数据特点,采取基于 GRU 检验温度时间序列,验证其是否具有预测价值。首先,基于统计学方法对时间序列进行平稳性检验和白噪声检验^[38];然后,采用基于 GRU 神经网络算法对模型进行训练和测试,预测未来 10 min 分娩舍温度的数值;最后,添加 RNN 模型,选取最优模型。

对分娩舍温度时间序列数据集进行统计,得出其原始序列平稳性检验、一阶差分序列平稳性值和一阶差分序列白噪声检验的 P 值分别为 0.13、 2.07×10^{-15} 和 2.77×10^{-49} 。可见,温度的一阶差分序列为平稳非白噪声序列,该时间序列的一阶差分序列有较强的短期相关性,可对其进行时间序列建模。使用 GRU 和 RNN 预测模型进行对比分析,得到对应模型的 MAE、MAPE 和 RMSE 3 个指标,结果见图 5。图 5a、5b 和 5c 的结果表明,相对于 RNN 模型,GRU 模型的 RMSE、MAE 和 MAPE 均较小;图 5d 可以看出,相比于 RNN 模型,GRU 模型的数据预测值曲线与观测值曲线更加接近。

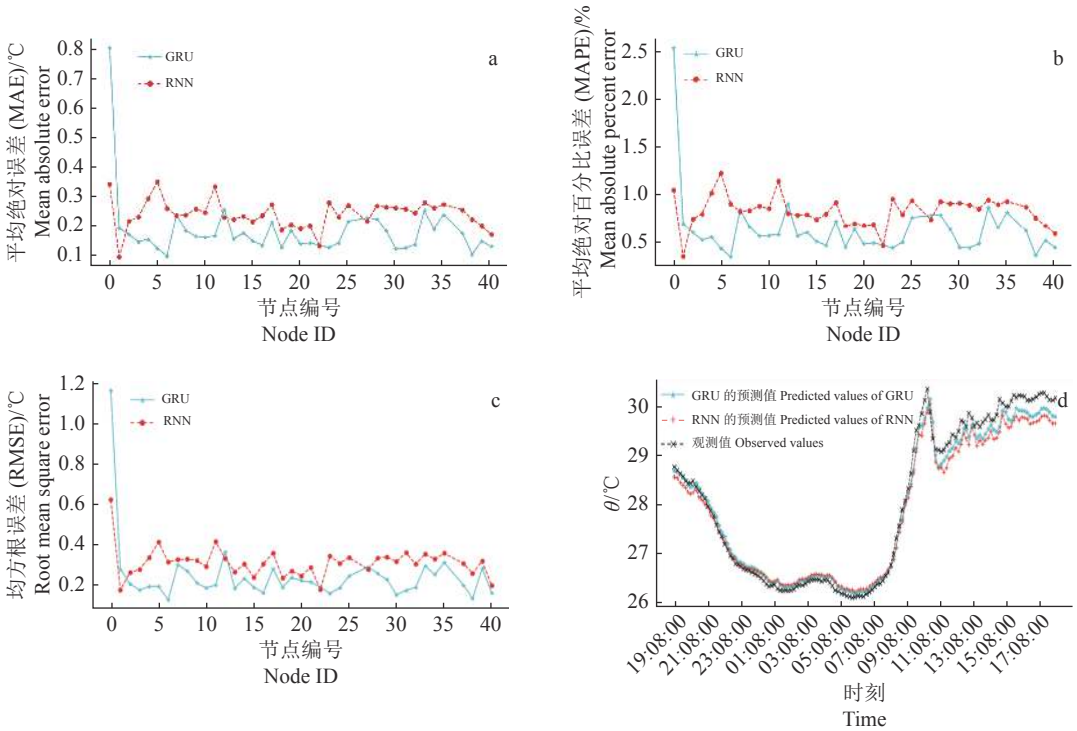


图 5 GRU 和 RNN 时间序列模型温度预测结果对比

Fig. 5 Comparison of temperature prediction results of GRU and RNN time series models

基于上述试验结果,对每个模型的 3 项指标进行计算,得到全部节点的指标平均值,结果见表 1。

表 1 温度预测试验中 GRU 和 RNN 时间序列模型的指标¹⁾
Table 1 Indexes of GRU and RNN time series models in temperature prediction test

模型 Model	RMSE/°C	MAE/°C	MAPE/%
GRU	0.25	0.19	0.65
RNN	0.31	0.24	0.84

1)RMSE: 均方根误差, MAE: 平均绝对误差, MAPE: 平均绝对百分比误差

1)RMSE: Root mean square error, MAE: Mean absolute error, MAPE: Mean absolute percent error

由表 1 可见,对于不同的误差指标 (RMSE、MAE 和 MAPE), GRU 模型取得的值都比 RNN 模型的要小。表明在时间序列预测试验中, GRU 模型的性能优于 RNN 模型。所以,选取 GRU 模型,进行基于单变量缺失模型的评估检测。

2.2 基于单变量缺失模型评估检测结果

为研究环境因子缺失对猪舍温度预测的影响,采用单变量缺失的多元模型和 GRU 预测模型,得到温度预测的结果,进而得到该环境因子对猪舍温度预测的贡献度。首先,选取经过预处理的数据集,在不区分节点的情况下,按照 2:1 的比例随机划分为训练集和测试集;然后,选取 BPNN 和 LR 等回

归模型, 进行对比分析, 模型的输出是温度序列, 模型的输入是剩余的环境因子。测试集验证结果 (表 2) 表明, XGBoost 和 BPNN 模型的 3 项指标均优于 LR 模型。XGBoost 模型的 MAPE 和 MAE 最低, XGBoost 模型的 RMSE(1.21 °C) 略高于 BPNN 模型 (0.89 °C), 这可能是因为 RMSE 对于个别离群点的敏感性以及较小的数据量对预测结果造成了影响。

表 2 温度预测试验中多元模型的指标¹⁾
Table 2 Indexes of multivariate model in temperature prediction test

模型 Model	RMSE/°C	MAE/°C	MAPE/%
XGBoost	1.21	0.71	2.50
BPNN	0.89	0.72	2.54
LR	1.25	0.89	3.14

1)RMSE: 均方根误差, MAE: 平均绝对误差, MAPE: 平均绝对百分比误差

1)RMSE: Root mean square error, MAE: Mean absolute error, MAPE: Mean absolute percent error

由表 2 可知, XGBoost 和 BPNN 模型相比 LR 模型较为可靠, 为了避免单个模型的预测结果的偶然性, 选择这两者同时用于构建缺失变量预测模型, 结果如表 3 所示。由表 3 可知, 根据 XGBoost 和 BPNN 模型的预测结果, 变量重要程度 (Score) 排名得出了相同的结果, 2 个模型的 Score 排序由高到低均为节点编号 (即空间位置)>相对湿度>二氧化碳浓度>氧气浓度。XGBoost 和 BPNN 的模型得到的结果基本一致, 保证了该方法和排名结果的可靠性。

表 3 温度预测试验中变量缺失的 XGBoost 和 BPNN 多元模型的 Score 值
Table 3 Score values of XGBoost and BPNN multivariate models with missing variables in temperature prediction test

缺失变量 Missing variable	Score	
	XGBoost	BPNN
CO ₂ 浓度 CO ₂ content	0.415	0.435
O ₂ 浓度 O ₂ content	0.407	0.426
相对湿度 Relative humidity	0.428	0.475
节点编号 Node ID	0.433	0.481

3 结论

针对母猪分娩舍的温度预测问题, 提出了基于时间序列模型的纵向预测融合多元模型的横向预

测方法。1) 基于时间序列模型, 对猪舍温度序列进行纵向预测, 试验结果表明 GRU 模型优于 RNN 模型, MAPE 分别为 2.50 和 0.031, 在分娩舍的实际环境参数调控管理中属于可接受的误差范围, 在时间维度上对分娩舍温度的变化起到了预警作用。2) 基于多元模型, 对猪舍温度序列进行横向预测, 本文从环境因素之间相互影响这一事实出发研究其他因素对温度的影响, 提出一种基于 XGBoost 模型的猪舍温度预测方法, 并与 BPNN 和 LR 模型的效果进行对比。结果表明, XGBoost 和 BPNN 模型优于 LR 模型, 但在不同的评价指标上, XGBoost 和 BPNN 模型性能各有优劣。从 RMSE 来看, BPNN 模型的误差小于 XGBoost 模型; 但从 MAE 和 MAPE 来看, XGBoost 模型的误差小于 BPNN 模型。从 RMSE、MAE 和 MAPE 的公式来看, RMSE 受离群点的影响较大, 本文的数据集也存在着部分的噪声且采集的数据不够多等问题, 这些因素对预测结果造成了影响。总体来看, XGBoost 模型有较大的应用潜力。3) 在探究某一自变量对预测结果贡献度的试验中, 确定了各种环境参数对温度的影响程度。其中, 节点对试验结果影响最大, 即空间位置对温度影响最大, 建议在猪舍温度的调控中, 可以在不同的空间位置采用不同措施, 精细化调控养殖环境。

本研究试验数据在时间和空间的数量不足, 今后应采用有线数据采集系统补充全年度全过程的数据, 同时研究季节性变化对于分娩母猪猪舍温度变化的影响, 并根据机器学习预测模型, 补充温度调控的试验。

参考文献:

[1] 汪开英, 苗香雯, 崔绍荣, 等. 猪舍环境温湿度对育成猪的生理及生产指标的影响[J]. 农业工程学报, 2002, 18(1): 99-102.

[2] 冯霞, 王思珍, 曹颖霞, 等. 饮水温度对断奶仔猪生产性能和养分表观消化率的影响[J]. 家畜生态学报, 2011, 32(3): 40-44.

[3] 秦琳琳, 马国旗, 储著东, 等. 基于灰色预测模型的温室温湿度系统建模与控制[J]. 农业工程学报, 2016, 32(S1): 233-241.

[4] KÖNIG M, HEMPEL S, JANKE D, et al. Variabilities in determining air exchange rates in naturally ventilated dairy buildings using the CO₂ production model[J]. Biosystems Engineering, 2018, 174: 249-259.

[5] 王德福, 黄会男, 张洪建, 等. 生猪养殖设施工程技术研究现状与发展分析[J]. 农业机械学报, 2018, 49(11): 1-14.

[6] 杨飞云, 曾雅琼, 冯泽猛, 等. 畜禽养殖环境调控与智能养殖装备技术研究进展[J]. 中国科学院院刊, 2019, 34(2): 163-173.

[7] 黄凯, 唐倩, 沈丹, 等. 冬季猪舍内温湿度与有害气体分

- 布规律研究[J]. 畜牧与兽医, 2019, 51(7): 35-41.
- [8] XIE Q, NI J, BAO J, et al. A thermal environmental model for indoor air temperature prediction and energy consumption in pig building[J]. *Building and Environment*, 2019, 161: 106238.
- [9] 韩建军, 南少伟, 李建平, 等. 基于随机森林算法的粮堆机械通风温度预测及控制研究[J]. 河南工业大学学报(自然科学版), 2019, 40(5): 107-113.
- [10] RONG L, AARNINK A J A. Development of ammonia mass transfer coefficient models for the atmosphere above two types of the slatted floors in a pig house using computational fluid dynamics[J]. *Biosystems Engineering*, 2019, 183: 13-25.
- [11] BEDI J, TOSHNIWAL D. Empirical mode decomposition based deep learning for electricity demand forecasting[J]. *IEEE Access*, 2018, 6: 49144-49156.
- [12] XIE Q, NI J, SU Z. A prediction model of ammonia emission from a fattening pig room based on the indoor concentration using adaptive neuro fuzzy inference system[J]. *Journal of Hazardous Materials*, 2017, 325: 301-309.
- [13] DASKALOV P, ARVANITIS K, SIGRIMIS N, et al. Development of an advanced microclimate controller for naturally ventilated pig building[J]. *Computers and Electronics in Agriculture*, 2005, 49(3): 377-391.
- [14] LITAGO J, BAPTISTA F J, MENESES J F, et al. Statistical modelling of the microclimate in a naturally ventilated greenhouse[J]. *Biosystems Engineering*, 2005, 92(3): 365-381.
- [15] NORTON T, GRANT J, FALLON R, et al. Optimising the ventilation configuration of naturally ventilated livestock buildings for improved indoor environmental homogeneity[J]. *Building and Environment*, 2010, 45(4): 983-995.
- [16] ANDONOV K, DASKALOV P, MARTEV K. A new approach to controlled natural ventilation of livestock buildings[J]. *Biosystems Engineering*, 2003, 84(1): 91-100.
- [17] 杨柳. 基于多尺度信息融合的猪舍环境控制系统设计[D]. 西安: 陕西科技大学, 2019.
- [18] 李立峰, 武佩, 麻硕士, 等. 基于组态软件和模糊控制的分娩母猪舍环境监控系统[J]. *农业工程学报*, 2011, 27(6): 231-236.
- [19] 谢秋菊, 苏中滨, NI J Q, 等. 猪舍环境适宜性模糊综合评价[J]. *农业工程学报*, 2016, 32(16): 198-205.
- [20] HEMPEL S, KÖNIG M, MENZ C, et al. Uncertainty in the measurement of indoor temperature and humidity in naturally ventilated dairy buildings as influenced by measurement technique and data variability[J]. *Biosystems Engineering*, 2018, 166: 58-75.
- [21] WANG Y, DONG H, ZHU Z, et al. CH₄, NH₃, N₂O and NO emissions from stored biogas digester effluent of pig manure at different temperatures[J]. *Agriculture, Ecosystems & Environment*, 2016, 217: 1-12.
- [22] DE SOUZA G B J, ROSSI L A, MENEZES D S Z. PID temperature controller in pig nursery: Spatial characterization of thermal environment[J]. *International Journal of Biometeorology*, 2018, 62(5): 773-781.
- [23] 陈昕, 唐湘璐, 李想, 等. 二次聚类与神经网络结合的日光温室温度二步预测方法[J]. *农业机械学报*, 2017, 48(S1): 353-358.
- [24] 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(S1): 353-360.
- [25] RODRIGUEZ M R, LOSADA E, BESTEIRO R, et al. Evolution of NH₃ concentrations in weaner pig buildings based on setpoint temperature[J]. *Agronomy*, 2020, 10(1): 107.
- [26] FERREIRA P M, FARIA E A, RUANO A E. Neural network models in greenhouse air temperature prediction[J]. *Neurocomputing*, 2002, 43(1): 51-75.
- [27] 陈英义, 程倩倩, 方晓敏, 等. 主成分分析和长短时记忆神经网络预测水产养殖水体溶解氧[J]. *农业工程学报*, 2018, 34(17): 183-191.
- [28] 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.
- [29] CHEN C, ZHU W, STEIBEL J, et al. Recognition of aggressive episodes of pigs based on convolutional neural network and long short-term memory[J]. *Computers and Electronics in Agriculture*, 2020, 169: 105166.
- [30] FERREIRA L B, DA CUNHA F F. New approach to estimate daily reference evapotranspiration based on hourly temperature and relative humidity using machine learning and deep learning[J]. *Agricultural Water Management*, 2020, 234: 106113.
- [31] 曾志雄, 董冰, 吕恩利, 等. 猪舍环境无线多点多源远程监测系统设计与试验[J]. *农业机械学报*, 2019, 51(2): 1-10.
- [32] 俞守华, 区晶莹, 张洁芳. 猪舍有害气体测定与温度智能控制算法[J]. *农业工程学报*, 2010, 26(7): 290-294.
- [33] 郑旭曼. 基于集成学习的 O₃ 浓度逐小时预测模型研究[D]. 上海: 华东师范大学, 2018.
- [34] 邹伟东, 张百海, 姚分喜, 等. 基于改进型极限学习机的日光温室温湿度预测与验证[J]. *农业工程学报*, 2015, 31(24): 194-200.
- [35] 李婷, 季宇寒, 张漫, 等. 基于 PLSR 和 BPNN 方法的番茄光合速率预测比较(英文)[J]. *农业工程学报*, 2015, 31(S2): 222-229.
- [36] LIANG J, LI W, BRADFORD S, et al. Physics-informed data-driven models to predict surface runoff water quantity and quality in agricultural fields[J]. *Water*, 2019, 11(2): 200.
- [37] 朱虹, 李爽, 郑丽敏, 等. 基于粒子群算法的生猪养殖物联网节点部署优化研究[J]. *农业机械学报*, 2016, 47(5): 254-262.
- [38] 左志宇, 毛罕平, 张晓东, 等. 基于时序分析法的温室温度预测模型[J]. *农业机械学报*, 2010, 41(11): 173-177.