Article Information	Fill in Information in Each Box Below
Article Type	Research article
Article Title (within 20 words without abbreviations)	Short-term forecasting of indoor CO_2 concentration in commercial pig barns using time-series models: A walk-forward validation and out-of-sample testing framework
Running Title (within 10 words)	Short-term forecasting of indoor CO_2 concentration in commercial pig barns
Author	Mohammad Ismail[1], Dae Yeong Kang[2], Oluwasegun Moses Ogundele[1], Seong Hun Kim[2], Reginald Gbenga Azuatalam[1], D. L. Vindya Sathsarani Dasanayaka[2], Usama Hanif[1], Pirlepesov Daniyar Maratovich[2], Myeong Yong Kang[2], Jayanta Kumar Basak[3], Hyeon Tae Kim[4*]
Affiliation	*Corresponding author: Hyeon Tae Kim (Email: bioani@gnu.ac.kr; Tel: +82-55-772-1896) [1]Department of Bio-systems Engineering, Gyeongsang National University, Jinju, 52828, Republic of Korea. [2]Department of Smart Farm, Gyeongsang National University, Jinju, 52828, Republic of Korea. [3]Department of Environmental Science and Disaster Management, Noakhali Science and Technology University, Noakhali 3814, Bangladesh [4]Department of Bio-industrial Machinery Engineering, Gyeongsang National University, Institute of Smart Space Agriculture (ISSA), Jinju, 52828, Republic of Korea
ORCIDs	Mohammad Ismail: https://orcid.org/0009-0000-8031-7240 Dae Yeong Kang: https://orcid.org/0000-0002-7481-9276 Oluwasegun Moses Ogundele: https://orcid.org/0000-0003-3595-2186 Seong Hun Kim: https://orcid.org/0009-0001-2771-7170 Reginald Gbenga Azuatalam: https://orcid.org/0009-0005-2141-7678 D. L. Vindya Sathsarani Dasanayaka: https://orcid.org/0009-0006-3826-7087 Usama Hanif: https://orcid.org/0009-0009-5050-7284 Pirlepesov Daniyar Maratovich: https://orcid.org/0009-0003-0328-4934 Myeong Yong Kang: https://orcid.org/0000-0003-3564-3340 Jayanta Kumar Basak: https://orcid.org/0000-0002-0718-1444 Hyeon Tae Kim: https://orcid.org/0000-0003-0788-1536
Competing interests	No potential conflict of interest relevant to this article was reported.

ACCEPTED

Funding sources	This work was supported by the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry (IPET) and the Korea Smart Farm R&D Foundation (KosFarm) through the Smart Farm Innovation Technology Development Program and funded by the Ministry of Agriculture, Food and Rural Affairs (MAFRA), the Ministry of Science and ICT (MSIT), and the Rural Development Administration (RDA) under grant number RS-2025-02312734.
Acknowledgements	This work was supported by the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry (IPET) and the Korea Smart Farm R&D Foundation (KosFarm) through the Smart Farm Innovation Technology Development Program and funded by the Ministry of Agriculture, Food and Rural Affairs (MAFRA), the Ministry of Science and ICT (MSIT), and the Rural Development Administration (RDA) under grant number RS-2025-02312734.
Availability of data and material	Upon reasonable request, the datasets of this study can be available from the corresponding author.
Authors' contributions **Please specify the authors' role using this form.**	Conceptualization: Ismail M, Kim HT Data curation: Kang DY, Azuatalam RG, Hanif U Formal analysis: Ismail M, Azuatalam RG Methodology: Ismail M, Ogundele OM, Dasanayaka DLVS Software: Kang DY, Kim SH, Kim HT Validation: Kang DY, Ogundele OM, Kim HT Investigation: Kang MY, Basak JK, Kim HT Writing - original draft: Ismail M, Kim SH, Maratovich PD Writing - review & editing: Ismail M, Kang DY, Ogundele OM, Kim SH, Azuatalam RG, Dasanayaka DLVS, Hanif U, Maratovich PD, Kang MY, Basak JK, Kim HT.
Ethics approval and consent to participate	Experiments were approved by the Gyeongsang National University Institutional Animal Care Committee (GNU-150508-R0029) and conducted in accordance with its guidelines.
Declaration of generative AI	During the preparation of this work, ChatGPT, QuillBot, and Grammarly were used solely for language editing and improving the clarity and readability of the manuscript. After using this tool, the manuscript was reviewed and edited as needed and authors will assume full responsibility for the publication.

ACCEPTED

ABSTRACT

Short-term forecasting of CO_2 concentration in commercial pig barns (CPBs) is essential for effective indoor air quality management. However, previous studies have often used simple train–test splitting for CO_2 forecasting, which may not adequately capture real-world temporal variability in CPBs. Therefore, this study integrated walk-forward (WF) validation with continuous model updating and out-of-sample (OOS) testing to assess forecasting performance on unseen observations. Hourly-resampled data were used to generate one-step-ahead forecasts for the subsequent hour. Four time-series models, ARIMA, SARIMA, SARIMAX, and Prophet, were evaluated using indoor environmental data from Barn A (piglets) and Barn B (sows), while Barn C (control) was included for comparison of indoor environmental conditions across the three barn types. Model performance was assessed using root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), residual diagnostics, and the entropy-weighted residual quality index (EW-RQI). Significant differences in indoor air quality parameters were observed among the barns ($p < 0.001$). During WF validation, SARIMAX achieved the best overall forecasting performance in Barn A, reducing RMSE, MAE, and MAPE by 4.58–55.37%, 3.50–61.73%, and 2.54–63.83%, respectively, compared with the other models, whereas performance differences in Barn B were relatively small. During OOS testing, SARIMAX remained the best-performing model in Barn B, reducing RMSE, MAE, and MAPE by 15.67–42.04%, 7.10–41.31%, and 6.67–40.71%, respectively. In Barn A, SARIMAX achieved lower MAE and MAPE, although ARIMA produced the lowest RMSE. Moreover, SARIMA and SARIMAX exhibited the better overall residual quality, achieving the highest EW-RQI values across both WF validation and OOS testing. These findings may provide valuable insights for monitoring CO_2 concentrations and support improved indoor atmospheric management in CPBs.

Keywords: Carbon dioxide; Time-series models; Commercial pig barns; Animal welfare.

ACCEPTED

INTRODUCTION

Livestock production accounts for approximately 14.5% of global anthropogenic greenhouse gas (GHG) emissions, equivalent to 7.10 Gt CO_2-eq annually [1,2]. Among livestock sectors, pig production contributes approximately 13% of livestock-related GHG emissions, and its environmental impact is expected to increase with rising global pork demand, which accounts for nearly 45% of global meat consumption and is projected to grow by more than 8% by 2033 [3]. These emissions mainly arise from feed production, manure management, energy use, and

enteric processes. At the farm level, indoor CO_2 concentration is a key indicator of air quality, influenced by animal respiration, microbial decomposition, ventilation efficiency, and environmental conditions such as temperature and relative humidity [3,4]. Inadequate ventilation can elevate indoor CO_2 levels, adversely affecting animal health, feed efficiency, welfare, and the health of farm workers [5]. However, the temporal variability of indoor CO_2 remains poorly understood, highlighting the need for accurate monitoring and forecasting to support effective air quality management and climate change mitigation.

Several studies have applied machine learning (ML) and deep learning (DL) to monitor CO_2 concentrations and environmental conditions in pig production systems. Basak et al. [3] predicted CO_2 emissions in pig barns using seven regression models based on pig body mass and feed intake, where exponential regression performed best ($R^2 > 0.78$ for training, 0.75 for testing) and lasso performed worst ($R^2 < 0.65$, RMSE > 20 ppm). However, feed intake and body mass may not consistently represent CO_2 emissions, which may limit model reliability. Similarly, Besteiro et al. [4] showed that wavelet neural network (WNN) models can accurately predict CO_2 dynamics in weaner pig buildings, with the autoregressive model outperforming the explanatory model (RMSE: 26.33 vs. 154.36 ppm). In addition, the hybrid OTDBO–TCN–GRU model achieved highly accurate prediction (MAE: 0.047 ppm, MSE: 0.004 ppm, R^2: 0.987), outperforming DBO–TCN–GRU and conventional models (e.g., GRU, LSTM) with up to 37.20% and 51.60% reductions in MAE, 66.70% and 81.20% reductions in MSE, and R^2 improvements of up to 5.50%, while maintaining errors below 0.30 mg m^{-3} [6]. However, validation under experimental barn conditions may limit its generalizability to commercial systems. In addition to CO_2 prediction, many studies have applied ML and DL techniques to predict other environmental parameters, such as temperature, $PM_{2.5}$, H_2S, NO_3, CH_4, humidity, PM_{10}, and NH_3 in pig barn environments [7–11]. However, the high complexity of these models may limit their direct adoption in commercial large-scale pig barn environments. Additionally, most existing studies rely on feature-based forecasting approaches rather than historical time-series data [3,12]. This reliance may limit their applicability at the barn level in resource- and data-scarce settings where extensive environmental and operational measurements are often unavailable.

Alongside ML and DL techniques, classical time-series models remain important for environmental and indoor air quality prediction due to their ability to capture temporal patterns using historical data alone [13–15]. Among these, ARIMA is commonly used for univariate forecasting, SARIMA extends it by introducing seasonality, SARIMAX further integrates

ACCEPTED

exogenous variables to improve predictive performance, and Prophet provides a decomposable framework for trend and seasonal modeling [14,16–19]. These models are widely used as baseline approaches in environmental forecasting. For example, Shin et al. [14] showed that ARIMA outperforms Prophet for predicting CO_2 concentration in indoor microclimates. However, its performance declines with low-frequency data, limiting its suitability for environments where high-resolution measurements are not available. Manigandan et al. [17] evaluated SARIMA and SARIMAX models for forecasting U.S. natural gas production and consumption (million cubic feet) and found that SARIMAX significantly outperformed SARIMA, with lower errors in both production (RMSE: 184.53 versus 3106.47; MAPE: 15.93% versus 125.66%) and consumption (RMSE: 254.45 versus 1636.48; MAPE: 24.36% versus 69.45%). Elamin and Fukushige [20] proposed a SARIMAX framework incorporating both main and interaction effects of exogenous variables for hourly electricity demand forecasting in Japan, achieving improved accuracy with reductions of 22.2% in MAPE, 21.3% in MAE, and 21.8% in RMSE compared with the standard SARIMAX model. This observation aligns with Ward et al. [21], who emphasized that incorporating trend, autocorrelation, and seasonal components can significantly reduce forecasting errors.

However, existing studies have primarily explored large-scale, outdoor, or experimental datasets, with limited focus on high-resolution, localized indoor environments such as CPBs. These environments exhibit dynamic and nonlinear CO_2 patterns driven by animal activity, ventilation control, and microclimatic conditions, making accurate forecasting particularly challenging [4]. Moreover, systematic comparisons of classical time-series models under realistic validation frameworks remain limited. While a conventional train–test split evaluates models on unseen data, it provides only a fixed assessment and does not capture performance under continuously evolving time-series data. In contrast, combining walk-forward (WF) validation with out-of-sample (OOS) testing provides a more rigorous assessment of model robustness under evolving temporal conditions and performance on unseen future observations [22]. Accordingly, this study aims to (1) examine whether significant differences exist in environmental conditions (CO_2, relative humidity, and temperature) across three CPBs using statistical analysis, and (2) develop and systematically compare four classical time-series models (ARIMA, SARIMA, SARIMAX, and Prophet) for short-term CO_2 forecasting using WF validation and OOS testing, together with signal decomposition and autocorrelation diagnostics to identify temporal patterns. The findings of this study provide valuable insights for improving indoor air quality monitoring and management in CPBs.

ACCEPTED

MATERIALS AND METHODS

Description of the pig barns

This study was conducted in three identical CPBs (Barns A, B, and C) at Daerim Farm, Goseong-gun, South Korea. Each barn had internal dimensions (45 m × 12 m × 3.20 m), galvanized steel roofs insulated with 5–7 cm layers and 6 cm thick polystyrene-insulated side walls. The barns had partially slatted concrete floors (60–65 m^2) elevated 0.80 m above ground, allowing manure drainage into an under-floor slurry system. Ventilation was provided by a negative-pressure system with four axial exhaust fans installed 144 cm above ground, adjustable rear-wall inlets, and an air damper (Auto-Damper 250, Sanison Co., Ltd., Daegu, Korea) installed at 177 cm to maintain airflow (0.16 m^3 s^{-1}). Barn A housed 84 crossbred Yorkshire piglets (24 ± 5 kg, 49 days old), Barn B contained 24 breeding sows (180–220 kg), and Barn C remained unoccupied as a control. All barns followed standard commercial practices, including automated dry feeding and drinking systems with *ad libitum* access, and feeding was provided three times daily in Barns A and B.

Data collection

This study followed a longitudinal observational design under commercial pig farming conditions to capture short-term CO_2 dynamics. Data were collected over 66 days, including a 5-day adaptation period from 27 to 31 May 2025, followed by 61 days of continuous monitoring from 1 June to 31 July 2025. Data collected during the adaptation period were excluded from all analyses. Each barn was equipped with an identical Livestock Environmental Monitoring System (LEMS; AgriRobo Tech Co. Ltd., Incheon, Korea), installed at comparable locations to ensure measurement consistency (Supplementary Fig. 1). No modifications were made to routine farm operations, allowing the data to reflect natural commercial conditions. CO_2 concentration (ppm), temperature (°C), and relative humidity (%) were recorded at 10-min intervals. The CO_2 sensor had a measurement range of 0–5000 ppm with ±5% accuracy, while the temperature and relative humidity sensors had accuracies of ±0.40°C and ±3%, respectively. The system was calibrated before installation and periodically according to the manufacturer's recommendations.

Data preprocessing and exploration

The datasets comprised three environmental parameters recorded at 10-min intervals in Barns A, B, and C. No missing or invalid values were detected across the datasets, and therefore no

ACCEPTED

imputation was required. Timestamps were converted to datetime format and used as the index. Data were resampled to 60-min intervals using mean aggregation to reduce fluctuations, improve temporal stability, reduce dimensionality, and enhance computational efficiency [23]. Temporal variations are presented in Fig. 1. Exploratory data analysis, including summary statistics, boxplots, and one-way ANOVA, was conducted to evaluate parameter distributions and inter-barn differences. All processing and analyses were performed using Python version 3.11.4 in a Jupyter Notebook environment.

Model selection

This study employed four classical time-series models, namely ARIMA, SARIMA, SARIMAX, and Prophet, to forecast CO_2 concentrations in CPBs. These models were selected based on their capability to characterize temporal dependencies, trends, seasonality, and external influences within environmental time-series data [18,21]. ARIMA was used to capture autoregressive and moving-average relationships in non-seasonal patterns, whereas SARIMA was applied to account for periodic seasonal variations. SARIMAX extended the SARIMA framework by integrating exogenous environmental variables that influence CO_2 fluctuations. Prophet was selected for its ability to model nonlinear trends, multiple seasonal patterns, and irregular changes in time-series observations [14]. These models were selected to provide a computationally efficient and interpretable forecasting framework suitable for practical CO_2 monitoring in pig barns.

Model execution

All models were developed and evaluated using a consistent framework to compare their capabilities and identify the most suitable approach for accurate forecasting of CO_2 concentrations. Model-specific hyperparameters are summarized in Supplementary Table 1. Outliers were detected using the interquartile range (IQR) method to reduce the influence of extreme observations. Datasets from Barn A and Barn B were subsequently divided into an initial 70% training period (1 June–15 July 2025; 1,080 hourly observations), followed by a 20% WF validation period (16–25 July 2025; 240 hourly observations) and a 10% OOS testing period (26–30 July 2025; 120 hourly observations) to maintain temporal integrity and prevent data leakage. During WF validation, the forecasting origin was advanced sequentially and the models were updated using newly available observations, whereas the OOS period was retained as a completely unseen future period for final evaluation. For ARIMA, SARIMA, and SARIMAX, model orders were selected using stationarity analysis together with

ACCEPTED

autocorrelation function (ACF) and partial autocorrelation function (PACF) diagnostics (Supplementary Fig. 2). The ADF test indicated that both datasets from Barn A and Barn B were non-stationary at the original level ($p > 0.05$). The corresponding ADF statistics were −0.311 for Barn A and −0.903 for Barn B. After first differencing, both datasets achieved stationarity, with highly significant results ($p < 0.001$). The ADF statistics after first differencing were −10.891 for Barn A and −12.199 for Barn B. Therefore, first-order differencing (d = 1) was selected. Based on the identified 24-h CO_2 periodicity, seasonal differencing at a 24-h lag (d = 1) was applied to account for the daily seasonal structure. A single seasonal difference was selected to capture this seasonal component without introducing higher-order seasonal differencing. The final ARIMA, SARIMA, and SARIMAX configurations were (2,1,2) for Barn A and (1,1,1) for Barn B, with seasonal orders of (2,1,2,24) and (1,1,1,24), respectively, to capture the dominant 24-h CO_2 periodicity (Supplementary Table 1, Supplementary Fig. 2). Additionally, SARIMAX incorporated temperature, humidity, lagged CO_2 variables (lags 1 and 24), rolling statistics (window = 3), interaction terms, and time-based features as external predictors, all of which were standardized using Z-score normalization. Prophet was configured with barn-specific changepoint prior scales (0.04 for Barn A and 0.05 for Barn B), Fourier orders (4 and 5, respectively), seasonal components, and external regressors (Supplementary Table 1). All hyperparameters reported in this table were selected using the training data. During walk-forward validation, model parameters were estimated using only the data available at each forecasting origin. For SARIMAX, StandardScaler was fitted using only the corresponding training data at each forecasting origin. All models were evaluated using a WF expanding window validation strategy described in Eq. (1) [24].

$$\hat{y}_{t+1} = f(y_{1:t}) \tag{1}$$

where $y_{1:t}$ denotes all observed values from time step 1 to t, $f(\cdot)$ represents the forecasting model trained on the expanding historical window, and $\hat{y}_{t+1}$ is the one-step-ahead prediction generated at each iteration. This process is repeated sequentially by updating the training set as Eq. (2).

$$D_{train}^{(t+1)} = D_{train}^{(t)} \cup \{y_{t+1}\} \tag{2}$$

where $D_{train}^{(t)}$ represents the training dataset at time step t, y_{t+1} denotes the newly observed value at the subsequent time step, and $\cup$ is the union operator indicating the sequential addition of new observations to the training set. Unlike traditional ratio-based train-test splitting, which

ACCEPTED

partitions data into fixed subsets such as D_{train} and D_{test}, the WF approach updates the training data iteratively by incorporating each new observation while preserving temporal order and enabling repeated model retraining [22]. At each forecasting origin, preprocessing, scaling, and feature generation were performed using only observations available up to that period. No future information was incorporated into the model development process. This strategy is particularly important for time-series forecasting as it prevents data leakage, ensures a realistic evaluation of predictive performance, and allows the training window to expand progressively with each new observation [22,24,25].

Model evaluation metrics

The performance of the forecasting models was evaluated using mean absolute error (MAE), root mean squared error (RMSE), and mean absolute percentage error (MAPE), computed under both WF validation and OOS testing. These MAE, RMSE, and MAPE were calculated using Eqs. (3), (4), and (5), respectively.

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t| \tag{3}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2} \tag{4}$$

$$MAPE = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \tag{5}$$

where y_t represents the observed CO_2 concentration at time t, $\hat{y}_t$ denotes the corresponding predicted value, and n is the total number of observations.

Model residual diagnostic analysis

Residual diagnostic analysis was performed to evaluate the adequacy, robustness, and stability of the applied models under both WF validation and OOS testing. Residuals were first calculated by subtracting predicted values from observed values, with positive and negative residuals indicating underprediction and overprediction, respectively. The residuals were then examined to assess their distribution and error characteristics. The ACF was used to detect any remaining temporal dependence, where significant lag correlations indicated that the models had not fully captured the underlying serial structure [24]. Seasonal-trend decomposition using

ACCEPTED

LOESS (STL) was then applied to identify residual trend, seasonal, and irregular components, while fast Fourier transform (FFT) analysis was performed to detect hidden periodicities in the frequency domain [26]. The diagnostic metrics obtained from ACF, STL, and FFT were integrated into a residual quality index (RQI), which was further extended to an entropy-weighted residual quality index (EW-RQI) by assigning data-driven weights based on information entropy [27]. Finally, model stability and generalization capability were evaluated by comparing residual diagnostic metrics between the WF validation and OOS testing frameworks.

RESULTS

Variability of environmental parameters across pig barns

Environmental conditions varied markedly among the three pig barns (Table 1; Fig. 2), indicating distinct microclimatic environments. The highest mean temperature was observed in Barn A (26.14 °C), followed by Barn B (23.93 °C) and Barn C (20.65 °C). Relative humidity remained consistently high across all barns, with slightly higher levels in Barn B (87.89%) than in Barn A (85.41%) and Barn C (85.08%). Mean CO_2 concentration was highest in Barn A (754.60 ± 267.24 ppm), intermediate in Barn B (610.52 ± 87.71 ppm), and lowest in Barn C (398.87 ± 27.70 ppm), indicating that Barn C maintained the most stable CO_2 conditions. These patterns are consistent with the distributions shown in Fig. 2, where ANOVA indicated significant differences among the three barn conditions observed in this study for all environmental parameters ($p < 0.001$). Moreover, the quartile distributions (Q1, Q2, and Q3) showed generally right-skewed distributions, especially for CO_2, reflecting occasional high-concentration events. Therefore, forecasting models capable of accurately capturing episodic concentration peaks are essential for effective indoor air quality management.

ACCEPTED

Performance assessment of time-series forecasting models

Table 2 summarizes the predictive performance of ARIMA, SARIMA, SARIMAX, and Prophet models under WF validation and OOS testing across Barn A and Barn B. In Barn A, the WF validation stage RMSE, MAE, and MAPE values ranged from 68.07 to 152.45 ppm, 48.52 to 126.85 ppm, and 4.98% to 13.77%, respectively. However, OOS errors were notably higher, ranging from 166.82 to 211.05 ppm, 112.56 to 173.84 ppm, and 10.22% to 13.98%, respectively. Across Barn B, WF validation RMSE, MAE, and MAPE values ranged from 37.12 to 38.57 ppm, 25.49 to 26.60 ppm, and 4.03% to 4.19%, respectively, while OOS values

ranged from 40.13 to 69.25 ppm, 33.75 to 57.49 ppm, and 4.34% to 7.32%, respectively, across the evaluated models.

Overall, SARIMAX achieved the better performance in Barn A during WF validation and OOS testing phase. In Barn A WF validation, SARIMAX achieved the lowest RMSE, reducing errors by 11.30%, 4.58%, and 55.37% compared to ARIMA, SARIMA, and Prophet models, respectively. Similarly, it reduced MAE by 11.91%, 3.50%, and 61.73%, and MAPE by 11.23%, 2.54%, and 63.83% compared to ARIMA, SARIMA, and Prophet, respectively. For OOS testing phase in Barn A, ARIMA achieved the lowest RMSE, whereas SARIMAX achieved the lowest MAE and MAPE. Compared with SARIMAX, ARIMA reduced RMSE by 6.28% but increased MAE by 7.41% and MAPE by 8.81%. However, SARIMAX performed substantially better than SARIMA and Prophet, reducing MAE by 26.65% and 35.25%, RMSE by 11.59% and 15.66%, and MAPE by 15.19% and 26.89%, respectively. Similarly, SARIMAX also demonstrated better overall performance during OOS testing phase in Barn B, while differences during WF validation remained relatively small. In WF validation, SARIMAX reduced RMSE by 1.45% and 1.30% compared to ARIMA and SARIMA, respectively, although Prophet achieved the lowest RMSE overall. For MAE, SARIMAX performed slightly worse than the other models, while MAPE values were nearly similar between SARIMAX and Prophet (4.03% and 4.05%, respectively). During OOS testing, SARIMAX achieved the best overall performance. Compared with ARIMA, SARIMA, and Prophet, SARIMAX reduced RMSE by 42.04%, 15.67%, and 25.29%, MAE by 41.30%, 7.10%, and 25.82%, and MAPE by 40.71%, 6.67%, and 26.06%, respectively. Overall, SARIMAX performed better across both barns during OOS evaluation, while performance variations during WF validation were comparatively small across all models.

Error structure and residual distribution of the time-series models

Residual distributions revealed clear differences in model error structures during WF validation and OOS testing (Fig. 3). Mean residuals and their corresponding 5th–95th percentile intervals are reported to characterize prediction bias and the central 90% of residual errors. In Barn A, SARIMAX produced an almost unbiased residual distribution during WF validation (−0.01 ppm; −102.08 to 114.16 ppm) and maintained the smallest residual bias during OOS testing (−26.79 ppm; −178.90 to 86.91 ppm). In contrast, Prophet exhibited substantially larger residual bias and wider error distributions during both WF validation (32.49 ppm; −182.23 to 285.87 ppm) and OOS testing (45.73 ppm; −134.16 to 210.64 ppm). A similar pattern was

ACCEPTED

observed in Barn B, where SARIMAX showed near-zero residuals during WF validation (2.04 ppm; −35.59 to 39.22 ppm), while SARIMA produced the smallest residual bias during OOS testing (19.43 ppm; −21.24 to 57.77 ppm), closely followed by SARIMAX (25.80 ppm; −18.40 to 62.60 ppm). In contrast, Prophet again exhibited the largest residual bias and widest error distribution during WF validation (−4.25 ppm; −70.71 to 76.32 ppm) and OOS testing (91.64 ppm; −13.25 to 190.13 ppm). Collectively, SARIMAX and SARIMA consistently exhibited the most stable and least biased residual distributions across both barns, whereas Prophet showed the greatest residual bias and dispersion, particularly under OOS conditions.

Residual diagnostics and signal decomposition analysis (ACF, STL, and FFT)

Residual diagnostics presented in Table 3 and Supplementary Figs. 3–6 provide a comprehensive assessment of model adequacy through temporal dependence (Max ACF), structural decomposition (trend, seasonality, and noise ratios), frequency-domain characteristics (FFT dominance), and integrated residual quality indices (RQI and EW-RQI). In Barn A, SARIMA and SARIMAX exhibited the most favorable residual characteristics during WF validation, with the lowest residual autocorrelation (Max ACF = 0.19), negligible trend components (0.01–0.02), relatively low seasonality (0.34–0.35), high noise contributions (0.56–0.57), and minimal FFT dominance (0.02), indicating residuals that closely approximated white noise. In contrast, Prophet retained strong residual autocorrelation (0.88), greater trend and seasonal components, lower noise contribution (0.13), and higher FFT dominance (0.09), resulting in the lowest RQI and EW-RQI. During OOS testing, residual quality deteriorated across all models. However, SARIMAX maintained the lowest residual autocorrelation (0.55), highest noise contribution (0.41), lowest FFT dominance (0.06), and highest EW-RQI (0.99), indicating better generalization. In contrast, Prophet and ARIMA exhibited stronger temporal dependence, greater residual seasonality, lower noise dominance, and substantially lower RQI and EW-RQI.

A similar trend was observed in Barn B, although the optimal model differed under OOS testing. During WF validation, SARIMAX exhibited the most random residual structure, characterized by low residual autocorrelation (Max ACF = 0.16), minimal trend (0.02), relatively low seasonality (0.27), highest noise contribution (0.60), low FFT dominance (0.02), and highest EW-RQI (0.87), closely followed by SARIMA (0.86). However, under OOS testing, SARIMA demonstrated superior residual characteristics, with the lowest residual autocorrelation (0.53), lower FFT dominance (0.15), highest RQI (0.92), and highest EW-RQI (0.71), whereas

ACCEPTED

SARIMAX exhibited the strongest residual autocorrelation (0.96), highest FFT dominance (0.32), and a substantially lower EW-RQI (0.37), indicating reduced generalization despite its superior WF validation performance. Although SARIMAX achieved favorable OOS prediction errors, its high Max ACF and low EW-RQI indicate that substantial temporal structure remained in the residuals. Prophet and ARIMA also retained pronounced residual autocorrelation and stronger seasonal structure, which contributed to lowest RQI and EW-RQI values. These findings demonstrate that SARIMAX performed better across most validation scenarios, whereas SARIMA was superior in Barn B under OOS testing, indicating the influence of barn-specific data characteristics and the evaluation strategy on model performance.

Model stability and robustness assessment

The model stability assessment in Fig. 4 quantified changes in residual diagnostic characteristics between WF validation and OOS testing across Barn A and B. In Barn A, SARIMAX exhibited the highest stability, with the smallest increases in residual autocorrelation (+0.36), trend (+0.11), seasonality (+0.01), and FFT dominance (+0.04), along with a moderate reduction in noise ratio (−0.16), indicating consistent residual behavior during OOS testing. Conversely, ARIMA and SARIMA showed larger increases in residual dependence and greater reductions in noise ratio, whereas Prophet maintained high autocorrelation despite relatively smaller diagnostic changes. In Barn B, SARIMA demonstrated the most stable performance, with minimal increases in ACF (+0.38), trend (+0.04), and FFT dominance (+0.13), although the increased seasonality (+0.41) indicated sensitivity to structural changes. SARIMAX preserved seasonal characteristics but exhibited a substantial increase in ACF (+0.80) and a reduction in noise ratio (−0.58), indicating stronger residual dependence. Overall, SARIMAX showed better stability in Barn A, whereas SARIMA performed best in Barn B, highlighting the influence of barn-specific datasets on model stability.

DISCUSSION

Environmental variability and air quality dynamics in pig barns

Indoor livestock environments are continuously influenced by animal respiration, manure decomposition, ventilation, stocking density, animal activity, environmental control strategies, and interactions with external climatic conditions [3,28]. Consistent with these interacting influences, environmental parameters differed significantly among the investigated pig barns

ACCEPTED

($p < 0.001$; Fig. 2), demonstrating substantial variability associated with housing conditions. The piglet barn recorded the highest average temperature (26.14 °C), followed by the sow barn (23.93 °C) and the empty barn (20.65 °C) (Table 4). These values align with the ranges reported by other studies in pig production systems (17.23–29.4 °C; Table 4). For example, Besteiro et al. [4] reported 27.79 °C in a commercial swine building, Yeo et al. [29] observed approximately 29.40 °C in experimental pig housing, whereas Basak et al. [3] measured lower temperatures of 17.23–18.72 °C. The elevated temperature in the piglet barn of this study was likely attributed to greater stocking density and associated metabolic heat production. Relative humidity also differed significantly among barns ($p < 0.001$; Fig. 2), with mean values of 85.41%, 87.89%, and 85.08% in Barns A, B, and C, respectively. These levels exceeded the values reported in previous studies (approximately 59–68%), including 59.92% [4], 59.58–67.85% [3] and approximately 65% [30,31], suggesting reduced moisture removal under the prevailing seasonal and operational conditions. Persistently high humidity can promote microbial survival, condensation, and deterioration of indoor air quality while increasing respiratory health risks for both livestock and farm workers [32–34]. The greater humidity variability observed in the empty barn further suggests stronger indoor–outdoor atmospheric exchange in the absence of livestock-derived heat buffering.

In addition to thermal and moisture conditions, CO_2 concentrations also differed significantly among the barns ($p < 0.001$; Fig. 2), with the highest mean concentration observed in the piglet barn (754.60 ppm), followed by the sow barn (610.52 ppm), while the empty barn maintained near-ambient levels (398.87 ppm). These concentrations were lower than those reported by Besteiro et al. [4], Yeo et al. [29], and Deb et al. [30], but comparable to those reported by Wu et al. [31] and Basak et al. [3], indicating moderate CO_2 accumulation in the investigated commercial barns. The higher CO_2 concentration in the piglet barn may have been associated with animal occupancy and respiration, although ventilation and other contributing factors were not directly measured. Elevated CO_2 concentrations may impair air quality and adversely affect animal health and productivity [3]. As measurements were conducted during the summer (June–July), seasonal conditions likely contributed to the elevated temperature and humidity, while the higher CO_2 concentration in the piglet barn primarily reflected greater animal occupancy and respiratory activity. Overall, these findings suggest that barn function, stocking density, and ambient climate collectively influence the indoor environment, underscoring the importance of barn-specific environmental forecasting for precision livestock management.

Comparative evaluation of time-series models for indoor CO_2 forecasting in pig barns

ACCEPTED

The significant microclimatic variability observed among the investigated pig barns highlights the need for reliable forecasting models capable of capturing continuously changing indoor CO_2 dynamics. To evaluate forecasting robustness under these heterogeneous conditions, four classical time-series models were assessed using both WF validation and OOS testing. During WF validation, model performance was relatively consistent in Barn B, where MAE and RMSE differed by only 4.36% and 3.91%, respectively, whereas Barn A showed much greater variability (161.48% for MAE and 124.00% for RMSE), largely due to the poorer performance of Prophet (Table 2). Across both barns, forecasting accuracy declined substantially during OOS testing, with prediction errors increasing by 26.90–205.10%. This deterioration may reflect both the difference between the two evaluation strategies and differences in the difficulty of the evaluated time periods. WF validation continuously updates model parameters using an expanding training window, allowing adaptation to newly observed patterns, whereas OOS testing evaluates completely unseen data without updating model parameters [22,24,35]. Therefore, the higher OOS errors cannot be attributed solely to the absence of model updating, as the final OOS period may also contain temporal patterns or microclimatic conditions that were more difficult to forecast. However, WF adaptive learning cannot be achieved using the conventional train–test split, where models are trained once on a fixed dataset and evaluated on unseen observations without updating their parameters.

Among the evaluated models, SARIMAX demonstrated relatively consistent forecasting performance across both WF validation and OOS testing (Table 2). Although it did not achieve the lowest error for every metric during WF validation, it showed good performance on unseen OOS data. In Barn A, SARIMAX reduced OOS MAE by 6.90% and 26.65% relative to ARIMA and SARIMA, respectively, although its RMSE remained 6.70% higher than that of ARIMA. In Barn B, SARIMAX achieved the lowest MAE, RMSE, and MAPE during OOS testing, outperforming both ARIMA and SARIMA across all metrics. In contrast, Prophet showed comparable performance during WF validation in Barn B but exhibited larger forecasting errors during OOS testing, indicating that model performance varied with barn-specific CO_2 dynamics. The consistent OOS performance of SARIMAX may be associated with its use of additional predictors and temporal features, including environmental variables, lagged and rolling features, interaction terms, and time-related patterns. These features may have helped the model capture temporal and environmental variations associated with indoor CO_2 dynamics [15]. In contrast, ARIMA and SARIMA rely mainly on historical observations and assume relatively stable temporal relationships, making them less adaptable when

ACCEPTED

environmental conditions change rapidly [18]. CPBs rarely satisfy these assumptions because ventilation rates fluctuate with thermal demand, animal metabolic heat production varies with age and activity, and manure decomposition continuously alters emission intensity [4,7]. Consequently, incorporating external predictors provides a more realistic representation of the processes governing indoor CO_2 dynamics [36]. Similar improvements have been reported in other forecasting applications. Elamin and Fukushige [20] observed approximately 22% reductions in MAE, RMSE, and MAPE after incorporating exogenous variables into electricity demand forecasting. Likewise, Manigandan et al. [17] reported substantial reductions in RMSE and MAPE for natural gas production forecasting using SARIMAX, while Alharbi and Csala [15] demonstrated improved renewable energy forecasting after including meteorological variables. Similar benefits have also been reported for livestock environments, where environmental variables enhanced predictions of indoor air quality and gas emissions compared with purely autoregressive models [10,37,38]. Overall, these findings indicate that multivariable statistical frameworks are better suited for practical pig barn CO_2 forecasting under dynamic environmental conditions

Residual diagnostic analysis showed that differences among the forecasting models extended beyond conventional accuracy metrics (Fig. 3; Table 3). During WF validation, SARIMAX and SARIMA achieved the highest EW-RQI values in both barns, whereas Prophet consistently ranked lowest (Table 3). Under OOS testing, SARIMAX remained the highest-ranked model in Barn A (EW-RQI = 0.99), while SARIMA performed best in Barn B (EW-RQI = 0.71), indicating superior preservation of residual structure under temporal distributional shifts. These differences further suggest that model selection should be considered barn-specific, as forecasting performance may depend on the autoregressive structure and environmental variability of indoor CO_2 dynamics, which may partly explain the relatively lower performance of Prophet's additive trend and seasonality framework [21,39]. This interpretation is consistent with the residual ACF, STL, and FFT analyses (Supplementary Figs. 3–6), which showed that SARIMAX and SARIMA generally produced residuals with weaker autocorrelation, reduced remaining seasonal structure, and less pronounced spectral peaks than Prophet. Similar findings were also reported by Shin et al. [14], who showed that ARIMA outperformed Prophet for indoor CO_2 forecasting, whereas Besteiro et al. [4] demonstrated that rapid changes in ventilation operation and animal activity can generate highly dynamic indoor gas concentrations that may be difficult to represent using additive trend and seasonality models. However, the present model comparison does not establish these processes as the primary

ACCEPTED

447 determinants of indoor CO_2 dynamics, but rather suggests them as potential explanations for
448 the observed differences in forecasting performance.

449 The higher forecasting error during OOS evaluation exceeded that reported for the hybrid mass
450 balance–BNN model of Ahn et al. [40], in which CVRMSE increased by approximately 77–
451 122% as the prediction horizon extended from 5 to 15 min, likely reflecting the greater
452 environmental heterogeneity and operational variability of commercial pig production systems
453 [41]. Although SARIMAX exhibited the better robustness among the evaluated classical
454 models, substantially lower prediction errors have been reported for advanced deep learning
455 approaches, including RF–PSO–LSTM (RMSE = 75.422 μg m^{-3}, MAE = 51.839 μg m^{-3}, R^2 =
456 0.992), OTDBO–TCN–GRU (MAE = 0.047 ppm, MSE = 0.004 ppm, R^2 = 0.987), and LSTM–
457 EMD–Attention (RMSE = 0.467, MAE = 0.295) [6,41,42]. However, these approaches
458 typically require extensive feature engineering, hyperparameter optimization, large training
459 datasets, and greater computational resources, limiting their practical deployment. Moreover,
460 conventional train–test splitting used in many previous studies may overlook temporal
461 distributional shifts and overestimate predictive performance. By contrast, the integrated WF
462 validation and OOS testing framework adopted in this study provides a more rigorous
463 assessment of forecasting robustness under continuously evolving barn environments. Overall,
464 the results indicate that interpretable SARIMAX and SARIMA models can provide a practical
465 balance between forecasting accuracy, computational efficiency, and operational applicability
466 for continuous CO_2 forecasting in CPBs. However, the most suitable model may vary
467 depending on barn characteristics and forecasting requirements.

468 **CONCLUSIONS**

469 This study compared ARIMA, SARIMA, SARIMAX, and Prophet models for short-term CO_2
470 forecasting in CPBs using a framework integrating WF validation and OOS testing.
471 Performance was evaluated using RMSE, MAE, and MAPE, together with residual diagnostic
472 analyses to assess error structure, temporal dependence, and spectral characteristics. The results
473 revealed substantial inter-barn variability, highlighting the importance of considering barn-
474 specific characteristics when developing CO_2 forecasting models. SARIMAX demonstrated
475 relatively consistent forecasting performance across the evaluation stages, particularly during
476 WF validation in Barn A and OOS testing in both barns, while ARIMA achieved the lowest
477 OOS RMSE in Barn A, Prophet showed the lowest WF MAE and RMSE in Barn B, and
478 SARIMA provided the highest OOS EW-RQI in Barn B. These mixed results indicate that no

ACCEPTED

single model consistently outperformed the others across all barns, evaluation stages, and performance criteria. Residual diagnostic analyses further showed differences in the remaining temporal dependence and spectral structure among models, with SARIMAX and SARIMA generally exhibiting weaker residual structure than Prophet. The forecasting framework may provide useful information for monitoring indoor CO_2 dynamics. However, the study is limited by the relatively short observation period, limited number of barns, lack of barn-level replication, single-sensor spatial representation, one summer period, and limited exogenous environmental and operational variables. Future studies should incorporate longer multi-season datasets, multiple barns for each animal or housing condition, spatially distributed sensing, and additional environmental and operational variables, including ventilation conditions and animal activity, to further evaluate the robustness and practical applicability of CO_2 forecasting for precision livestock farming.

REFERENCES

1. FAO. Livestock solutions for climate change. Food Agric Organiszations United Nations [Internet]. 2010; https://doi.org/http://ieeexplore.ieee.org/document/1574227/

2. Eska N, Lee JS, Park KH. Greenhouse gas emissions from livestock: sources, estimation, and mitigation. J Anim Sci Technol. 2024;66:1083–98. https://doi.org/10.5187/jast.2024.e86

3. Basak JK, Kim NE, Shahriar SA, Paudel B, Moon BE, Kim HT. Applicability of statistical and machine learning–based regression algorithms in modeling of carbon dioxide emission in experimental pig barns. Air Qual Atmos Heal [Internet]. Springer; 2022 [cited 2026 Apr 3];15:1899–912. https://doi.org/10.1007/s11869-022-01225-9

4. Besteiro R, Arango T, Ortega JA, Rodríguez MR, Fernández MD, Velo R. Prediction of carbon dioxide concentration in weaned piglet buildings by wavelet neural network models. Comput Electron Agric [Internet]. Elsevier; 2017 [cited 2026 Apr 3];143:201–7. https://doi.org/10.1016/J.COMPAG.2017.10.025

5. Zhang L, Mao Y, Chen Z, Hu X, Wang C, Lu C, et al. A systematic review of life-cycle GHG emissions from intensive pig farming: Accounting and mitigation. Sci Total Environ [Internet]. Elsevier B.V.; 2024;907:168112. https://doi.org/10.1016/j.scitotenv.2023.168112

6. Guo Z, Yin Z, Lyu Y, Wang Y, Chen S, Li Y, et al. Research on Indoor Environment Prediction of Pig House Based on OTDBO–TCN–GRU Algorithm. Animals. 2024;14:1–20. https://doi.org/10.3390/ani14060863

7. Shen X, Ming S, Wu W, Xie D, Yao H, Vishnupriya V, et al. Short-term time series predictions of greenhouse gas emission from dairy houses based on diverse machine learning

ACCEPTED

models. Comput Electron Agric [Internet]. Elsevier; 2026 [cited 2026 Apr 3];247:111699. https://doi.org/10.1016/J.COMPAG.2026.111699
8. Basak JK, Arulmozhi E, Moon BE, Bhujel A, Kim HT. Modelling methane emissions from pig manure using statistical and machine learning methods. Air Qual Atmos Heal [Internet]. Springer; 2022 [cited 2026 Apr 3];15:575–89. https://doi.org/10.1007/s11869-022-01169-0
9. Basak JK, Arulmozhi E, Khan F, Okyere FG, Park J, Kim HT. Modeling of ambient environment and thermal status relationship of pig's body in a pig barn. Indian J Anim Res [Internet]. Agricultural Research Communication Centre; 2020 [cited 2026 Apr 3];54:1049–54. https://doi.org/10.18805/ijar.B-1186
10. Dai X, Ni JQ, Pan Q, Wang K, Heber AJ. Monitoring of tempreature, humidity and air quality inside pig weaner house in eastern China. Nongye Jixie Xuebao/Transactions Chinese Soc Agric Mach. Chinese Society of Agricultural Machinery; 2016;47:315–22. https://doi.org/10.6041/J.ISSN.1000-1298.2016.07.043
11. Bus JD, Franchi GA, Boumans IJMM, te Beest DE, Webb LE, Jensen MB, et al. Short-term associations between ambient ammonia concentrations and growing-finishing pig performance and health. Prev Vet Med. Elsevier B.V.; 2025;242. https://doi.org/10.1016/J.PREVETMED.2025.106555
12. Deb NC, Basak JK, Arulmozhi E, Kang DY, Tamrakar N, Seo EW, et al. Estimation of ammonia and methane concentrations in pig barns based on daily activities, physiological parameters and various growth phases of pigs. Air Qual Atmos Heal 2025 188 [Internet]. Springer; 2025 [cited 2026 May 17];18:2199–211. https://doi.org/10.1007/S11869-025-01761-0
13. Krovi TS, Shanmugasundaram J. Forecasting Future CO2 Levels (Ppm) using the SARIMAX Model. 2024;13.
14. Shin SH, Deb NC, Arulmozhi E, Tamrakar N, Ogundele OM, Kook J, et al. Prediction of Carbon Dioxide Concentrations in Strawberry Greenhouse by Using Time Series Models. Agric [Internet]. Multidisciplinary Digital Publishing Institute; 2024 [cited 2026 Apr 8];14. https://doi.org/10.3390/agriculture14111895
15. Alharbi FR, Csala D. A Seasonal Autoregressive Integrated Moving Average with Exogenous Factors (SARIMAX) Forecasting Model-Based Time Series Approach. Inventions. 2022;7. https://doi.org/10.3390/inventions7040094
16. Alsharif MH, Younes MK, Kim J. Time Series ARIMA Model for Prediction of Daily and Monthly Average Global Solar Radiation: The Case Study of Seoul, South Korea. Symmetry 2019, Vol 11, Page 240 [Internet]. Multidisciplinary Digital Publishing Institute; 2019 [cited

ACCEPTED

2026 Apr 27];11:240. https://doi.org/10.3390/SYM11020240
17. Manigandan P, Alam MS, Alharthi M, Khan U, Alagirisamy K, Pachiyappan D, et al. Forecasting natural gas production and consumption in united states-evidence from sarima and sarimax models. Energies. 2021;14:1–17. https://doi.org/10.3390/en14196021
18. Kwarteng S, Andreevich P. Comparative Analysis of ARIMA, SARIMA and Prophet Model in Forecasting. Res Dev [Internet]. Science Publishing Group; 2024 [cited 2026 Apr 3];5:110–20. https://doi.org/10.11648/j.rd.20240504.13
19. Szostek K, Mazur D, Drałus G, Kusznier J. Analysis of the Effectiveness of ARIMA, SARIMA, and SVR Models in Time Series Forecasting: A Case Study of Wind Farm Energy Production. Energies [Internet]. Multidisciplinary Digital Publishing Institute; 2024 [cited 2026 Apr 27];17:4803. https://doi.org/10.3390/en17194803
20. Elamin N, Fukushige M. Modeling and forecasting hourly electricity demand by SARIMAX with interactions. Energy [Internet]. Elsevier Ltd; 2018;165:257–68. https://doi.org/10.1016/j.energy.2018.09.157
21. Ward MP, Iglesias RM, Brookes VJ. Autoregressive models applied to Time-Series data in veterinary science. Front Vet Sci. Frontiers Media S.A.; 2020;7:604. https://doi.org/10.3389/fvets.2020.00604
22. Bergmeir C, Hyndman RJ, Koo B. A note on the validity of cross-validation for evaluating autoregressive time series prediction. Comput Stat Data Anal [Internet]. North-Holland; 2018 [cited 2026 Apr 27];120:70–83. https://doi.org/10.1016/J.CSDA.2017.11.003
23. Bayat AR, Mäntysaari P, Stefański T, Negussie E, Lidauer M, Huhtanen P. Comparing carbon dioxide, oxygen, and methane exchanges, and heat production measured using GreenFeed and respiration chambers. J Dairy Sci [Internet]. Elsevier; 2025 [cited 2026 Apr 27];108:12328–39. https://doi.org/10.3168/JDS.2025-26682
24. Hyndman RJ, Athanasopoulos G. Forecasting: Principles and Practice (3rd ed.) [Internet]. South. Econ. J. 2021 [cited 2026 Apr 27]. p. 442. https://otexts.com/fpp3/%0Ahttps://otexts.com/fpp2/. Accessed 27 Apr 2026
25. Tashman LJ. Out-of-sample tests of forecasting accuracy: an analysis and review. Int J Forecast [Internet]. Elsevier; 2000 [cited 2026 Apr 27];16:437–50. https://doi.org/10.1016/S0169-2070(00)00065-0
26. Percival DB, Walden AT. Wavelet Methods for Time Series Analysis. Wavelet Methods Time Ser [Internet]. Cambridge University Press; 2000 [cited 2026 Apr 27]; https://doi.org/10.1017/CBO9780511841040
27. Liu DJ, Li L. Application study of comprehensive forecasting model based on entropy

ACCEPTED

weighting method on trend of PM2.5 concentration in Guangzhou, China. Int J Environ Res Public Health. 2015;12:7085–99. https://doi.org/10.3390/ijerph120607085
28. Kang DY, Moon BE, Kang MY, Kook JH, Deb NC, Tamrakar N, et al. Analysis of Pig Tendencies to Stay Specific Sections Within the Pig Barn According to Environmental Parameters and Facilities Features. Agric. 2025;15:1–19. https://doi.org/10.3390/agriculture15121282
29. Yeo UH, Jo SK, Kim SH, Park DH, Jeong DY, Park SJ, et al. Applicability of Machine-Learned Regression Models to Estimate Internal Air Temperature and CO2 Concentration of a Pig House. Agronomy. 2023;13. https://doi.org/10.3390/agronomy13020328
30. Deb NC, Basak JK, Karki S, Arulmozhi E, Kang DY, Tamrakar N, et al. Prediction of digestible energy requirement in growing finishing stage of pigs using machine learning models. J Agric Food Res. 2025;19. https://doi.org/10.1016/j.jafr.2025.101700
31. Wu Z, Xu K, Chen Y, Liu Y, Song W. A new prediction model based on deep learning for pig house environment. Sci Rep. 2024;14:1–16. https://doi.org/10.1038/s41598-024-82492-7
32. Buoio E, Cialini C, Costa A. Air Quality Assessment in Pig Farming: The Italian Classyfarm. Animals. 2023;13. https://doi.org/10.3390/ani13142297
33. Kim KY, Ko HJ. Indoor distribution characteristics of airborne bacteria in pig buildings as influenced by season and housing type. Asian-Australasian J Anim Sci. 2019;32:742–7. https://doi.org/10.5713/ajas.18.0415
34. XIONG Y, MENG Q shi, GAO J, TANG X fang, ZHANG H fu. Effects of relative humidity on animal health and welfare. J Integr Agric. 2017;16:1653–8. https://doi.org/10.1016/S2095-3119(16)61532-0
35. Feng Y, Zhang Y, Wang Y. Out-of-sample volatility prediction: Rolling window, expanding window, or both? J Forecast. 2024;43:567–82. https://doi.org/10.1002/for.3046
36. Uzal SS, Ilhan F. Effects of hourly, daily and seasonal variation of hazardous gases and climatic factors on the welfare of sheep housed in solid-floor confinement barns. S Afr J Anim Sci [Internet]. South African Bureau for Scientific Publications; 2019 [cited 2026 Apr 27];49:441–54. https://doi.org/10.4314/SAJAS.V49I3.5
37. Blanes V, Pedersen S. Ventilation Flow in Pig Houses measured and calculated by Carbon Dioxide, Moisture and Heat Balance Equations. Biosyst Eng [Internet]. Academic Press; 2005 [cited 2026 Apr 27];92:483–93. https://doi.org/10.1016/J.BIOSYSTEMSENG.2005.09.002
38. Ortega JA, Losada E, Besteiro R, Arango T, Ginzo-Villamayor MJ, Velo R, et al. Validation of an AutoRegressive Integrated Moving Average model for the prediction of animal zone temperature in a weaned piglet building. Biosyst Eng [Internet]. Academic Press;

ACCEPTED

2018 [cited 2026 Apr 27];174:231–8. https://doi.org/10.1016/J.BIOSYSTEMSENG.2018.07.012

39. Seyfi SU, Ilhan F. Effects of hourly, daily and seasonal variation of hazardous gases and climatic factors on the welfare of sheep housed in solid-floor confinement barns. S Afr J Anim Sci [Internet]. South African Society for Animal Science (SASAS); 2019 [cited 2026 Apr 27];49:441–54. https://doi.org/10.4314/SAJAS.V49I3.5

40. Ahn KU, Kim DW, Cho K, Cho D, Cho HM, Chae CU. Hybrid Model for Forecasting Indoor CO2 Concentration. Buildings [Internet]. MDPI; 2022 [cited 2026 Apr 3];12:1540. https://doi.org/10.3390/buildings12101540

41. Cen H, Yu L, Pu Y, Li J, Liu Z, Cai Q, et al. A Method to Predict CO2 Mass Concentration in Sheep Barns Based on the RF-PSO-LSTM Model. Animals. 2023;13:1–20. https://doi.org/10.3390/ani13081322

42. Guo J, Zhang B, Lin L, Xu Y, Zhou P, Luo S, et al. Multi-model fusion method for predicting CO2 concentration in greenhouse tomatoes. Comput Electron Agric [Internet]. Elsevier; 2024 [cited 2026 Apr 3];227:109623. https://doi.org/10.1016/J.COMPAG.2024.109623

ACCEPTED

List of Tables

Table 1. Descriptive statistics of environmental parameters across pig Barns A, B and C.

Parameters	Mean	Min	Q1	Q2	Q3	Max	STD
Temperature (Barn A)	26.14	23.71	25.43	26.00	26.78	30.65	1.05
Temperature (Barn B)	23.93	17.82	22.33	23.88	24.97	33.13	2.49
Temperature (Barn C)	20.65	14.72	19.27	20.83	21.94	27.48	2.14
Humidity (Barn A)	85.41	64.96	78.67	86.97	92.67	97.00	7.70
Humidity (Barn B)	87.89	61.25	81.52	89.79	94.98	99.48	8.07
Humidity (Barn C)	85.08	50.97	79.39	88.12	92.76	96.27	9.00
CO2 (Barn A)	754.60	451.86	548.31	619.96	923.01	1536.06	267.24
CO2 (Barn B)	610.52	402.17	547.92	586.00	659.88	903.67	87.71
CO2 (Barn C)	398.87	358.09	379.89	392.94	410.10	619.48	27.70

Note: Q1, Q2, and Q3 represent the 25th, 50th (median), and 75th percentiles, respectively; STD = standard deviation.

Table 2. Performance evaluation metrics of the time-series models under WF validation and OOS testing in pig Barn A.

Dataset	WF validation			OOS testing		
Metrics	MAE	RMSE	MAPE (%)	MAE	RMSE	MAPE (%)
Barn A						
ARIMA	55.08	76.74	5.61	120.90	166.82	11.12
SARIMA	50.28	71.34	5.11	153.46	201.34	12.05
SARIMAX	48.52	68.07	4.98	112.56	178.00	10.22
Prophet	126.85	152.45	13.77	173.84	211.05	13.98
Barn B						
ARIMA	25.80	38.57	4.19	57.49	69.25	7.32
SARIMA	26.04	38.51	4.10	36.33	47.59	4.65
SARIMAX	26.60	38.01	4.03	33.75	40.13	4.34
Prophet	25.49	37.12	4.05	45.49	53.72	5.87

ACCEPTED

640 **Table 3.** Residual diagnostic metrics and entropy-weighted model ranking.

Barn	Approach	Model	Max ACF	Trend Ratio	Seasonality Ratio	Noise Ratio	FFT Dominance	RQI	EW-RQI	Rank
A	WF	SARIMAX	0.19	0.02	0.34	0.57	0.02	0.99	0.99	1
A	WF	SARIMA	0.19	0.01	0.35	0.56	0.02	0.98	0.98	2
A	WF	ARIMA	0.24	0.01	0.51	0.43	0.02	0.84	0.79	3
A	WF	Prophet	0.88	0.17	0.66	0.13	0.09	0.00	0.00	4
A	OOS	SARIMAX	0.55	0.13	0.35	0.41	0.06	0.99	0.99	1
A	OOS	SARIMA	0.80	0.34	0.32	0.19	0.10	0.28	0.32	2
A	OOS	Prophet	0.89	0.18	0.56	0.23	0.11	0.23	0.21	3
A	OOS	ARIMA	0.87	0.13	0.62	0.22	0.12	0.19	0.16	4
B	WF	SARIMAX	0.16	0.02	0.27	0.60	0.02	0.93	0.87	1
B	WF	SARIMA	0.15	0.03	0.27	0.58	0.02	0.93	0.86	2
B	WF	ARIMA	0.19	0.03	0.30	0.56	0.02	0.82	0.70	3
B	WF	Prophet	0.58	0.14	0.24	0.45	0.03	0.10	0.23	4
B	OOS	SARIMA	0.53	0.07	0.68	0.16	0.15	0.92	0.71	1
B	OOS	SARIMAX	0.96	0.31	0.12	0.02	0.32	0.10	0.37	2
B	OOS	Prophet	0.84	0.09	0.83	0.09	0.23	0.48	0.34	3
B	OOS	ARIMA	0.88	0.14	0.83	0.07	0.22	0.41	0.29	4

641

642 **Table 4.** Comparison of indoor environmental conditions (mean CO_2 concentration, relative
643 humidity, and temperature) reported in previous pig barn studies.

Housing type	CO_2 (ppm)	Relative Humidity (%)	Temperature (°C)	Reference
Commercial swine building	1498.18	59.92	27.79	[4]
Experimental pig house	568.08	64.68	23.71	[31]
Experimental pig house	986.30	—	29.40	[29]
Experimental pig house	121.92	59.58	17.23	[3]
Experimental pig house	797.55	67.85	18.72	[3]
Experimental pig house	896.46	64.74	20.12	[30]
Commercial pig barn (A)	754.60	85.41	26.14	Present study
Commercial pig barn (B)	610.52	87.89	23.93	Present study
Commercial pig barn (C)	398.87	85.08	20.65	Present study

644

645

ACCEPTED

List of Figures

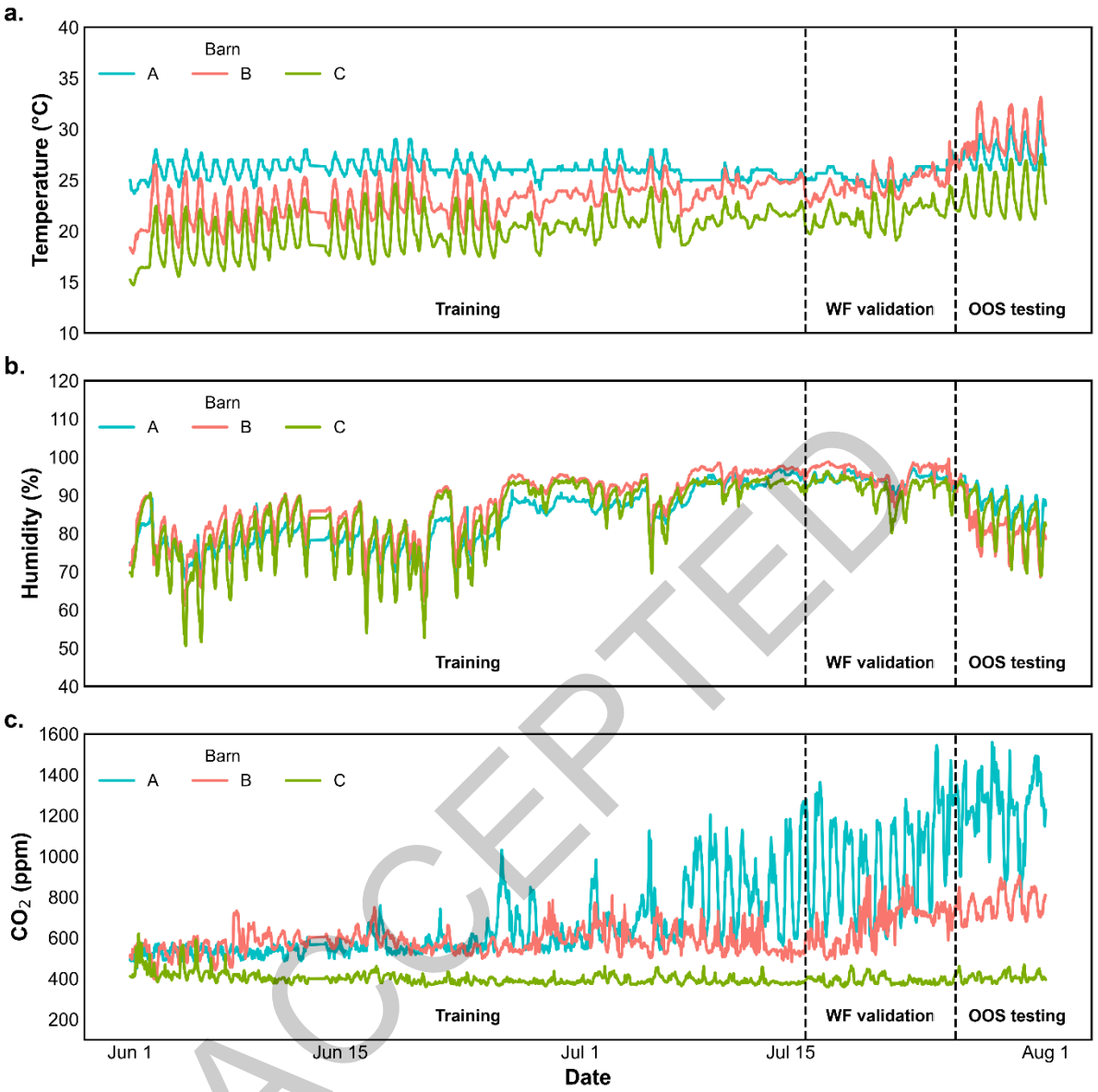
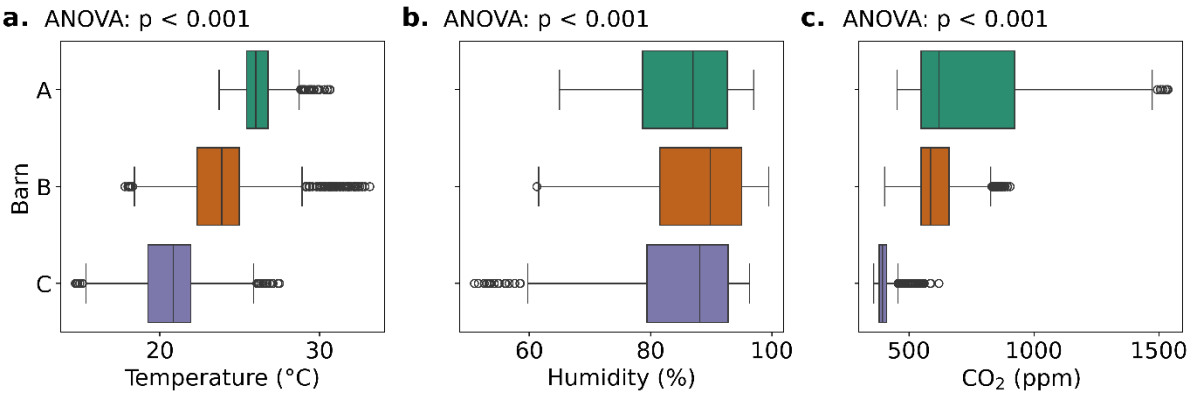


Fig. 1. Temporal variation of environmental parameters across three pig barns. (a) Temperature (°C), (b) Relative humidity (%), and (c) CO_2 concentration (ppm). The vertical dashed lines indicate the boundaries between the training, WF validation, and OOS testing periods.

ACCEPTED



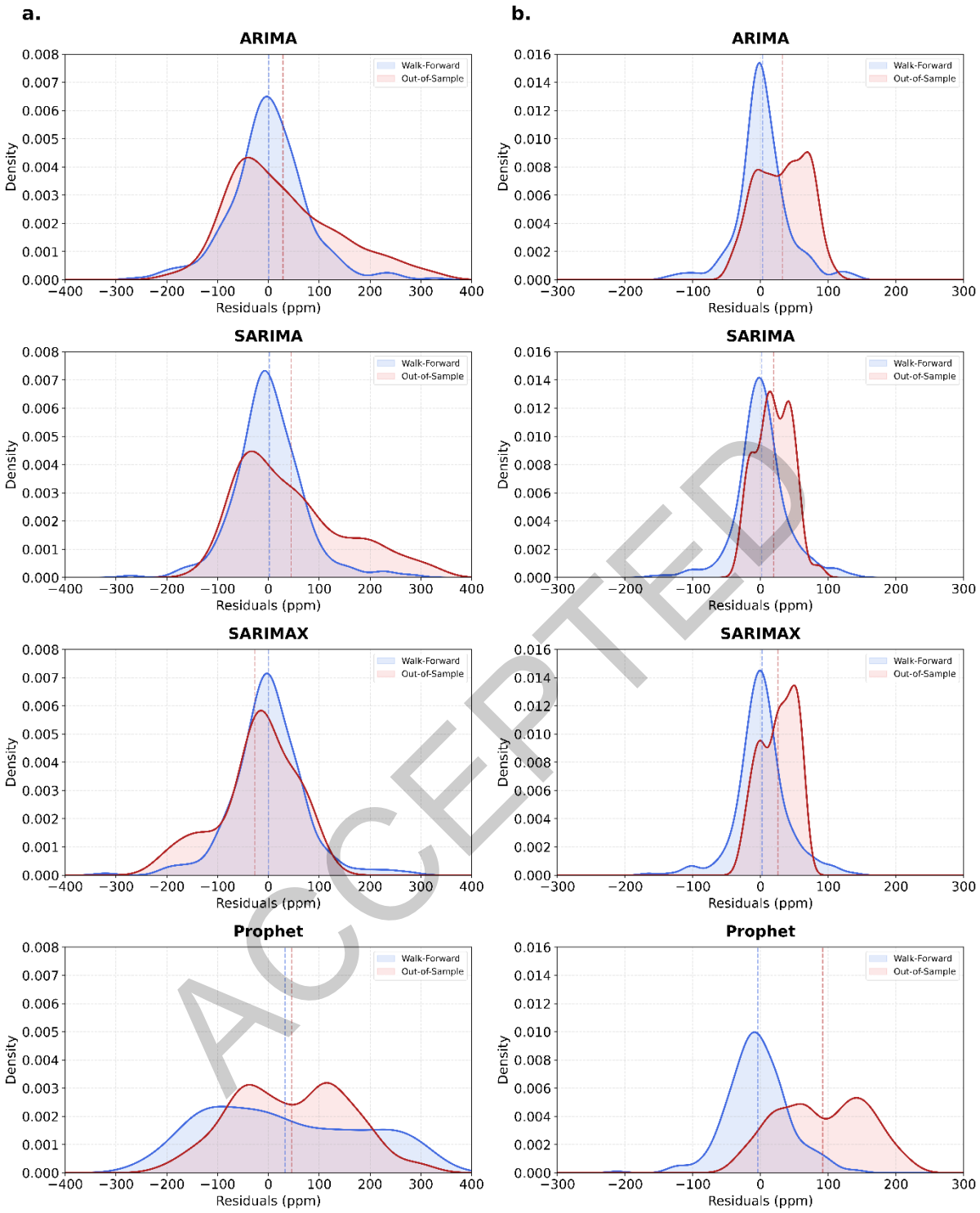
652

653 **Fig. 2.** Comparison of environmental conditions across three pig Barns. (a) Temperature (°C),

654 (b) Relative humidity (%), and (c) CO_2 concentration (ppm).

655

ACCEPTED



ACCEPTED

Fig. 3. Residual distributions of the ARIMA, SARIMA, SARIMAX, and Prophet models WF validation and OOS testing across (a) Barn A and (b) Barn B.

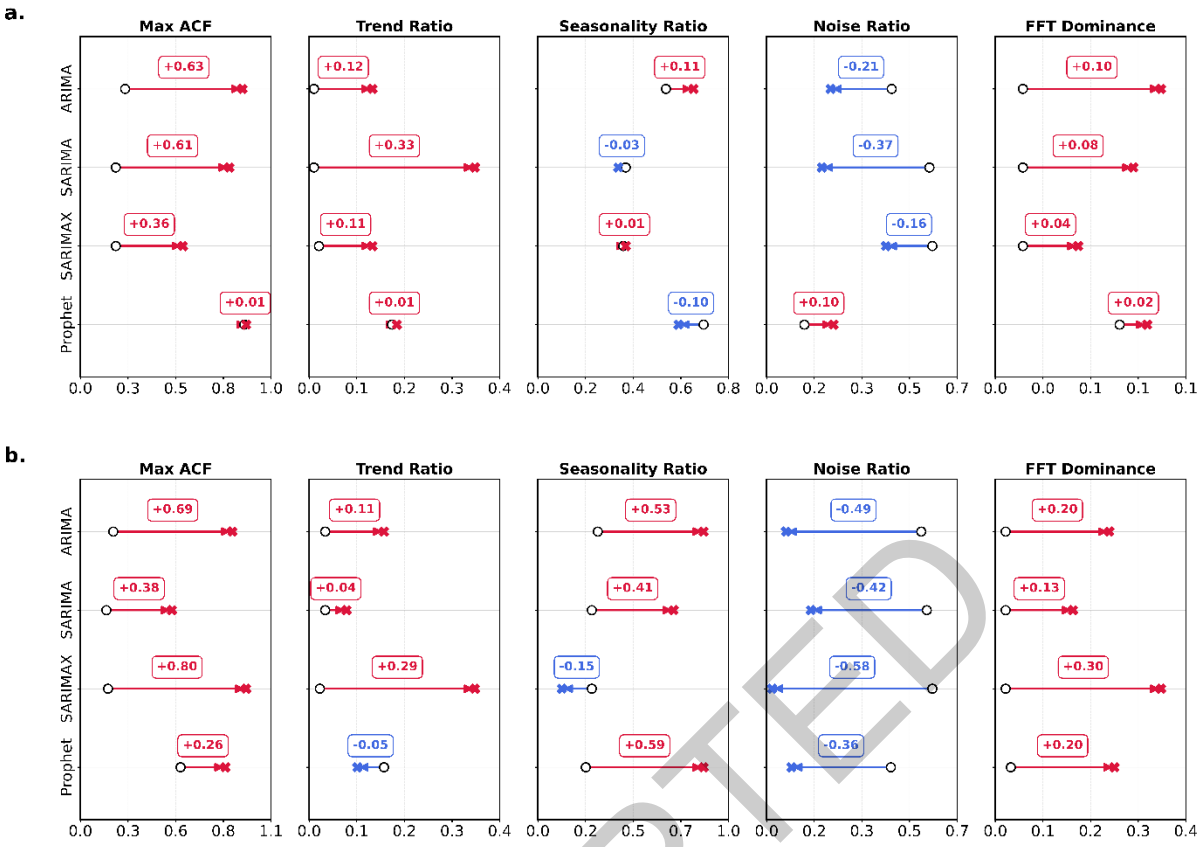


Fig. 4. Model stability assessment based on residual diagnostic shifts from WF validation to OOS testing across (a) Barn A and (b) Barn B. Open circles (○) indicate WF values, crosses (×) indicate OOS values, and horizontal lines show the shift (Δ = OOS − WF). Red (+) and blue (−) indicate increases and decreases, respectively; shorter lines indicate greater stability.

ACCEPTED