pagepress

Journal of Agricultural Engineering

Journal of Agricultural Engineering 2026; LVII:2096
doi:10.4081/jae.2026.2096

Estimating indoor heat stress in broiler and swine facilities using high-resolution Geo-KOMPSAT-2A satellite data

Rial Arifin Rajagukguk,[1] Jinseon Park,[2] Heejou Kim,[1] Chae-rin Lee,[1] Ji-yeon Park,[1] Se-yeon Lee,[1] Taehwan Ha,[3] Se-woon Hong[1,2]

[1]Department of Rural and Bio-systems Engineering, Chonnam National University, Gwangju; [2]AgriBio Institute of Climate Change Management, Chonnam National University, Gwangju; [3]Department of Agricultural & Rural Engineering, Chungnam National University, Daejeon, Korea

Abstract

Heat stress poses a critical challenge to animal welfare and productivity in intensive livestock production systems, particularly for broilers and swine with limited thermoregulatory capacity. The temperature-humidity index (THI) is widely applied to assess thermal stress, yet conventional monitoring methods remain constrained by limited spatial coverage and scalability. This study introduces a novel satellite-based framework for estimating indoor heat stress in livestock facilities using machine learning and GEO-KOMPSAT-2A (GK2A) satellite data. The proposed outside-in approach integrates satellite-derived temperature, humidity, and solar irradiance to infer indoor environmental conditions without reliance on on-site sensors or detailed building specifications. Unlike computational fluid dynamics models, which are resource-intensive and difficult to scale, this data-driven method captures nonlinear relationships between outdoor meteorological variables and indoor microclimates. The XGBoost model demonstrated superior accuracy in estimating indoor temperature and humidity across multiple farms. When converted to THI, relative root mean square errors (rRMSE) ranged from 0.623% to 0.693% in swine farms and 0.827% to 1.332% in broiler farms, demonstrating robust performance in heat stress assessment. By leveraging geostationary satellite observations with high temporal and spatial resolution, this framework enables continuous, large-scale monitoring of thermal conditions in livestock facilities. The approach provides a practical and scalable tool to support ventilation management, cooling strategies, and animal welfare decisions under dynamic weather conditions.

Key words: heat stress; temperature–humidity index; machine learning; satellite data; livestock microclimate.

Correspondence: Se-woon Hong, Department of Rural and Bio-systems Engineering, Chonnam National University, Gwangju 61186, Republic of Korea.
E-mail: hsewoon@jnu.ac.kr

Introduction

Heat stress in intensive livestock production is a critical challenge for both animal welfare and farm productivity (Kim *et al.*, 2025; Liu *et al.*, 2025). Broilers and growing pigs are particularly vulnerable due to their limited thermoregulatory capacity and high metabolic heat production. In broiler houses, even moderate increases in ambient temperature can reduce feed intake, elevate mortality rates, and compromise meat quality (Liu *et al.*, 2020). In swine facilities, heat stress impairs fertility, slows growth, and increases susceptibility to disease (Liu *et al.*, 2022). Effective monitoring and management of heat stress are therefore essential not only for safeguarding animal health and welfare but also for maintaining production efficiency and profitability, particularly during the hot and humid summer seasons prevalent in South Korea. Thermal stress in livestock is commonly evaluated using composite indices that capture the combined effects of multiple environmental factors. Among these, the temperature–humidity index (THI) is the most widely adopted, as it integrates air temperature and relative humidity to indicate the level of heat stress experienced by animals (Niu *et al.*, 2024; Uea-Anuwong *et al.*, 2025). Higher THI values correspond to greater thermal discomfort and physiological strain. Although other factors, such as air velocity, radiant heat, and ambient pressure, also influence the microclimate inside livestock houses, temperature and humidity remain the most critical parameters for estimating animal heat stress (Herbut *et al.*, 2018). Traditional THI assessment in livestock buildings has relied mainly on in situ measurements of indoor temperature and relative humidity obtained from fixed sensors. These systems provide direct and reliable observations and are therefore widely used for farm-level heat stress monitoring and ventilation control. However, because the indoor environment of broiler and swine houses is spatially heterogeneous, point measurements from a limited number of sensors may not adequately represent the full thermal conditions experienced by animals (Deng *et al.*, 2024). Extending sensor networks across multiple locations or farms also increases the cost and effort associated with installation, calibration, and maintenance (Patrick *et al.*, 2024). To complement direct measurement, numerical and imaging-based approaches have also been investigated. Among them, computational fluid dynamics (CFD) has been extensively used to simulate airflow, temperature, and humidity distributions inside livestock houses and to evaluate

ventilation and cooling performance (Blanes-Vidal *et al.*, 2008; Lee *et al.*, 2013). CFD provides detailed spatial insight into indoor microclimates, but its application usually requires detailed building geometry, boundary conditions, and high computational resources, which limits its practicality for routine monitoring across multiple commercial farms. In addition to CFD, infrared thermography has been used to detect heat load through surface temperature patterns of animals and surrounding structures (do Amaral Vercellino *et al.*, 2025). More recently, unmanned aerial vehicles (UAVs) equipped with thermal or multispectral sensors have been explored for high-resolution environmental monitoring and hotspot detection in agricultural systems (Wan *et al.*, 2024). These innovative techniques can reveal spatial thermal variation more effectively than fixed-point measurements, but they are generally episodic rather than continuous and may be constrained by operational complexity, weather conditions, and limited applicability for routine indoor monitoring (Park *et al.*, 2025).

In recent years, machine learning has emerged as a practical and scalable alternative for estimating heat stress in livestock facilities (Chapman *et al.*, 2023; Georgiades *et al.*, 2025; Yan *et al.*, 2022). Unlike physics-based models, machine learning approaches can learn complex nonlinear relationships between outdoor environmental variables and indoor temperature and humidity without requiring detailed geometric inputs or extensive computational resources. This capability enables machine learning to estimate the key heat-stress parameters, temperature and humidity, by relying only on readily available outdoor data and the number of animals. Once trained, machine learning models can provide rapid, continuous estimations across many farms, enabling large-scale monitoring that is difficult to achieve with CFD or sensor-based systems.

Nevertheless, the accuracy of machine learning models depends heavily on the availability of reliable outdoor input data, which are not consistently accessible at every farm, especially in regions with sparse weather station networks or complex terrain. The indoor microclimate of livestock houses is strongly influenced by outdoor variables such as temperature, humidity, wind speed, and solar irradiance. While weather measurements are typically collected at fixed station locations and interpolated to estimate conditions elsewhere, interpolation often introduces substantial uncertainty. This limitation is particularly critical for solar irradiance, as cloud cover and atmospheric conditions vary rapidly across space and time, making point-based interpolation insufficient for capturing local solar variability (Rajagukguk *et al.*, 2020). Satellite observations from the GEO-KOMPSAT-2A (GK2A) platform provide an effective alternative for obtaining key outdoor meteorological inputs (Jang *et al.*, 2022). GK2A employs semi-empirical algorithms to derive surface solar irradiance and supplies weather data products with continuous, spatially consistent coverage across the Korean Peninsula. Compared with reanalysis datasets such as ERA5 (Soci *et al.*, 2024), which offer coarser spatial resolution and lower temporal frequency, or polar-orbiting satellites like MODIS that provide only a few observations per day (Yu *et al.*, 2022), GK2A delivers high-temporal-resolution geostationary measurements. Its finer spatial detail and frequent updates enable more accurate representation of rapid changes in outdoor weather conditions, which directly affect indoor thermal environments and are essential for reliable heat stress estimation.

This study proposes an outside-in estimation framework that utilizes temperature, humidity, and solar irradiance derived entirely from GK2A satellite data to estimate indoor environmental conditions and heat stress levels in broiler and swine farms. The framework employs machine learning models to capture relationships between outdoor satellite-based meteorological variables and indoor temperature and humidity. These estimated indoor parameters are then used to calculate the THI, which serves as the primary indicator of heat stress. The novelty of this study lies in two key aspects: first, the integration of satellite-derived temperature, humidity, and solar irradiance to represent the full outdoor thermal load, including radiative heating effects that are often overlooked in conventional indoor climate models; and second, the use of data-driven machine learning techniques capable of learning complex nonlinear relationships between outdoor conditions and indoor microclimates without requiring detailed physical information about building structure or ventilation. This framework offers a practical and scalable approach for estimating indoor heat stress across multiple farms without reliance on on-site sensors, thereby supporting more informed decisions on ventilation management, cooling strategies, and animal welfare under dynamic weather conditions.

Materials and Methods

Data description

Swine and broiler farms

Four livestock farms in South Korea were selected as research sites, comprising two swine farms (Bethel and Handon) and two broiler farms (Koryeo and Jangjiuk). The geographical locations of these farms, including latitude and longitude coordinates, are shown in Figure 1, which also presents exterior and interior images of each facility. Data were collected from 1 July 2024 to 30 June 2025 at 10-minute intervals. To represent overall indoor conditions, average values from all sensors installed within each farm were used for analysis. The swine farms differed in maximum capacity: Bethel housed up to 1,000 pigs per building, whereas Handon housed up to 12,000 pigs per building. The broiler farms accommodated 7,000 birds per building at Koryeo and 12,000 birds per building at Jangjiuk. The facilities also differed in ventilation and cooling strategies, which may influence indoor thermal conditions and heat transfer. Bethel used a mechanical ventilation system combined with a mist spraying system for cooling. Handon also used mechanical ventilation, together with a cooling pad system. Koryeo was operated under natural ventilation without an additional cooling system. Jangjiuk used mechanical ventilation, and its thermal control strategy included a solar-reflective coating applied to the building exterior. These differences in environmental control are relevant because ventilation mode and cooling strategy can alter sensible and latent heat exchange, indoor air movement, and the thermal response of the building envelope.

Each farm was equipped with multiple sensors positioned at different locations inside the livestock houses to continuously record indoor air temperature and relative humidity. The sensors used were HOBO MX2301A devices, which are weatherproof data loggers equipped with built-in temperature and relative humidity sensors. These devices measure temperatures from –40°C to 70°C and relative humidity from 0% to 100%, with resolutions of 0.02°C and 0.01% RH, respectively.

Satellite data

Environmental parameters were obtained from the GK2A geostationary meteorological satellite to support indoor THI estimation. GK2A is equipped with an Advanced Meteorological Imager (AMI) featuring sixteen channels across four spectral ranges: visi-

ble, near-infrared, mid-wave infrared, and long-wave infrared (Jang *et al.*, 2022). The satellite-derived environmental variables used in this study were obtained through the National Meteorological Satellite Center (NMSC) of Korea Meteorological Administration (KMA) data service and accessed *via* the official application programming interface (API). The API enables automated retrieval of GK2A products at regular temporal intervals, ensuring consistent and systematic data acquisition.

Solar irradiance data were derived from the Level 1B visible channel, which has a spatial resolution of 1 km and a wavelength range of 0.431-0.479 μm. The Level 1B data provide digital number (DN) values that must be converted into global horizontal irradiance (GHI) using a semi-empirical model. Temperature and humidity were obtained from Level 2 GK2A products, which are atmospherically corrected and physically meaningful outputs. The Level 2 dataset provides vertical profiles of temperature and humidity at 54 atmospheric levels. In this study, data from level 54, which is closest to the surface, were used to represent near-surface air temperature and humidity. These products have a spatial resolution of approximately 2 km and a temporal resolution of 10 min.

Methodology

This study aimed to estimate indoor temperature and relative humidity using machine learning models. Predicting indoor thermal conditions is inherently challenging because they are influenced by heat transfer through building materials, ventilation and air exchange processes, internal heat sources, and non-uniform air mixing. These interactions form a complex thermodynamic system that is difficult to represent accurately using purely physical models, particularly when detailed building parameters are unavailable.

Machine learning offers a practical alternative by capturing nonlinear relationships between outdoor conditions and indoor microclimates without requiring full physical characterization of the building. In this study, most input variables were derived from GK2A satellite products (Levels 1B and 2). As illustrated in Figure 2, the workflow involved selecting relevant outdoor variables, applying machine learning models to estimate indoor temperature and relative humidity, and computing THI as the indicator of heat stress.

Input data preparation and selection

Accurate selection of input variables is essential for estimating indoor environmental conditions using machine learning, as these variables must capture the primary factors influencing heat stress in livestock facilities. In this study, three types of satellite-derived environmental variables were used as predictors: temperature, humidity, and solar radiation**.** Outdoor temperature and humidity were obtained directly from the Level 2B products of the GK2A satellite, which provide spatially and temporally consistent measurements suitable for model input.

Unlike temperature and humidity, solar radiation is not directly measured by satellite sensors and therefore requires a multi-step estimation procedure. Surface solar irradiance was estimated using established semi-empirical models, including those developed by Perez *et al.*, (2002), Mueller *et al.* (2012), and Hammer *et al.* (2003). Although these models differ in their specific formulations for clear-sky index estimation, they all rely primarily on the cloud index (CI) as a key input parameter, together with additional vari-

Figure 1. Locations of the swine and broiler farms in South Korea, with latitude and longitude coordinates and representative exterior and interior images.

ables such as DN, solar zenith angle (SZA), extraterrestrial irradiance (EI), and Sun-Earth distance (SE). Clear-sky irradiance was computed using the Ineichen–Perez model (Perez *et al.*, 2002), which represents atmospheric transparency through the Linke turbidity factor (TL). Because direct turbidity measurements were not available at the study sites, turbidity values were estimated using PVLib (Stein *et al.*, 2016). The overall input variables used for semi-empirical GHI estimation are summarized in Eq. (1):

$$GHI_{sat} = f(DN, SZA, TL, CI, EI, SE) \quad \text{(Eq. 1)}$$

Surface solar irradiance consists of both direct normal irradiance (DNI) and diffuse horizontal irradiance (DHI). These components are critical because they determine how solar energy is transmitted and absorbed by the building envelope, directly influencing indoor thermal conditions. DNI represents the direct beam component of sunlight, exerting a strong heating effect on roofs and wall surfaces exposed to solar radiation, particularly under clear-sky conditions. In contrast, DHI describes the scattered portion of sunlight reaching surfaces from all directions, becoming the dominant contributor to heat gain under cloudy or partially shaded conditions. The balance between DNI and DHI affects the rate and distribution of heat accumulation in livestock houses, thereby influencing both indoor temperature and humidity. Separating GHI into its DNI and DHI components thus provides a more accurate representation of solar heat load on building structures, enhancing the reliability of machine learning models for indoor heat stress estimation. To perform this separation, a solar decomposition model was required. In this study, the YANG5 model (Yang *et al.*, 2024) was adopted, as it represents one of the most recent and accurate methods for solar separation. The model uses GHI, apparent solar time (AST), SZA, TL, and EI as input variables. The decomposition inputs are summarized in Eq. (2).

$$DNI, DHI = f(GHI_{sat}, AST, SZA, TL, EI) \quad \text{(Eq. 2)}$$

By integrating semi-empirical irradiance models with the YANG5 decomposition model, comprehensive solar radiation inputs -including GHI, DNI, and DHI- were obtained. These variables, together with temperature and humidity, were used as predictors in the machine learning framework. To further improve predictive capability, several supplementary variables were incorporated. The solar azimuth angle (SAA) was included to describe the relative position of the sun, which influences the direction and magnitude of incoming radiation on building surfaces. The number of animals (N) was also considered, as metabolic heat production contributes directly to internal heat load, altering indoor temperature and humidity. Finally, temporal features such as minute, hour, day of week, and day of month were added to capture periodic patterns and diurnal or seasonal variability in environmental conditions. These time-related factors enable the model to learn recurring fluctuations in solar radiation and microclimate behavior, thereby improving robustness and accuracy in predicting indoor heat stress.

Machine learning model

Several tree-based machine learning algorithms, including Random Forest (RF) (Lin *et al.*, 2026), Extreme Gradient Boosting (XGBoost) (Chen and Guestrin, 2016), Light Gradient-Boosting Machine (LightGBM) (Machado *et al.*, 2019), and Categorical Boosting (CatBoost) (Prokhorenkova *et al.*, 2018), were applied to

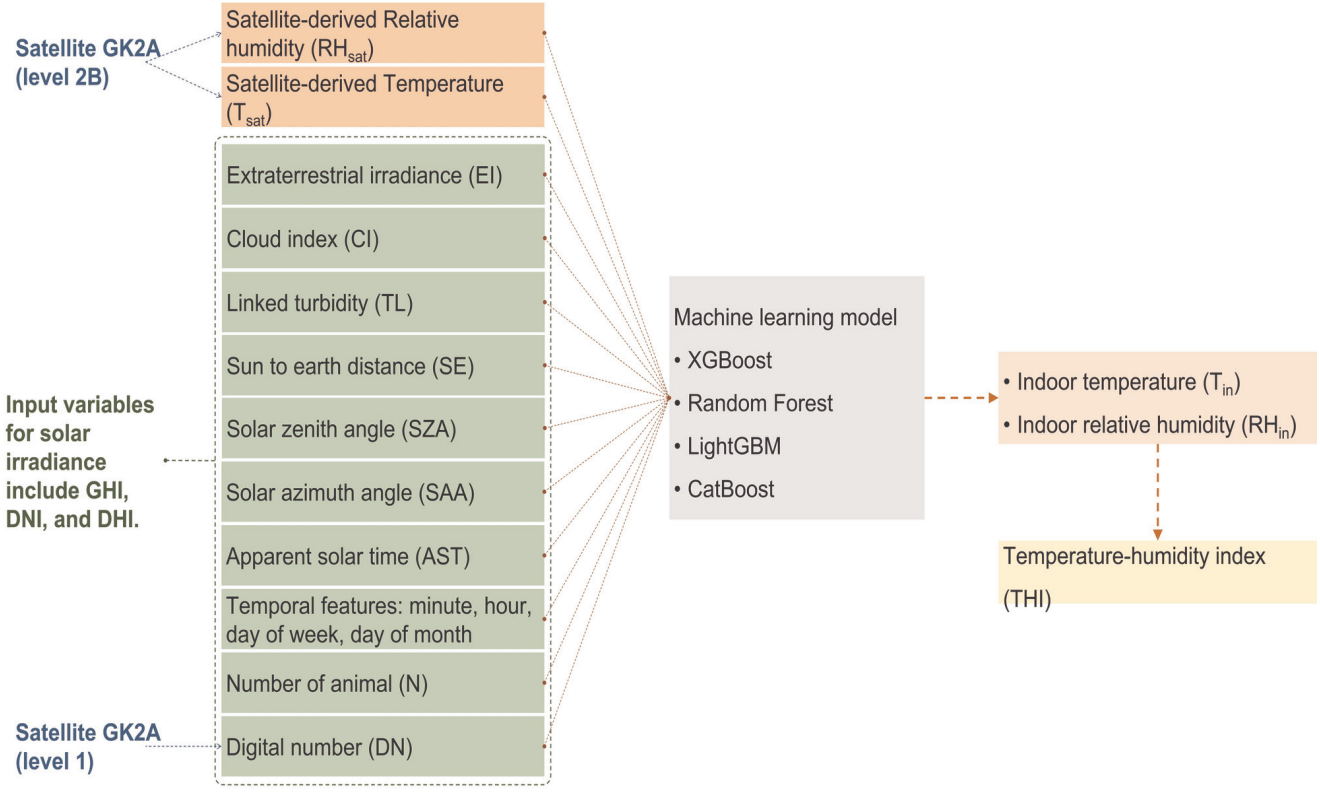

Figure 2. Workflow for input selection and application of machine learning models to estimate indoor temperature, relative humidity, and temperature-humidity index (THI).

model the relationship between outdoor environmental variables derived from satellite observations and the indoor temperature and humidity of livestock facilities. These algorithms are well-suited to heterogeneous and nonlinear datasets, particularly when the predictors include diverse physical parameters such as solar geometry, atmospheric turbidity, satellite radiance, semi-empirical irradiance estimates, and temporal features. Tree-based models effectively capture nonlinear interactions and complex dependencies among variables without imposing strict distributional assumptions or requiring extensive feature preprocessing. Their resilience to missing values and noisy satellite inputs is advantageous given the inherent variability and occasional data gaps in the GK2A dataset.

The adoption of these algorithms is further supported by their proven performance and interpretability in environmental and energy modeling research. RF serves as a robust baseline model by mitigating variance through ensemble averaging, while XGBoost and LightGBM employ gradient boosting to efficiently learn subtle feature interactions with high predictive accuracy. These models are computationally efficient, scalable to high-dimensional data, and exhibit reduced risk of overfitting when properly tuned. Such characteristics make them appropriate for real-time prediction and operational monitoring applications. To gain deeper insight into the role of each predictor, Shapley Additive Explanations (SHAP) analysis was incorporated into the modeling framework (Lundberg and Lee, 2017). SHAP provides a consistent and theoretically sound approach for evaluating the contribution of individual features to model outputs. This interpretability is vital for identifying key environmental drivers -including solar irradiance, turbidity, and satellite-derived temperature- that influence indoor conditions, thereby supporting transparent and data-informed decision-making in farm monitoring systems.

Heat stress assessment metrics

The THI is a widely used indicator for quantifying heat stress in livestock, combining the effects of air temperature and relative humidity into a single value. Originally developed to evaluate thermal comfort in humans (Al-Hadhrami, 2013), THI was later adapted for animal agriculture to represent the combined environmental load affecting livestock performance and welfare. Because humidity limits evaporative heat loss, even moderate air temperatures can cause severe stress under high humidity conditions. THI thus provides a convenient and practical measure of thermal stress intensity without requiring complex environmental measurements.

In this study, THI values were derived using Eq. (3), which combines air temperature and relative humidity to represent the level of heat stress (M'Hamdi *et al.*, 2021):

$$THI = (1.8 \times T + 32) - (0.55 - 0.0055 \times RH) \times (1.8 \times T - 26.8) \quad \text{(Eq. 3)}$$

where T is dry-bulb temperature (°C), and RH is relative humidity (%).

In broiler and swine production systems, THI is considered an essential management tool for evaluating the indoor thermal environment and guiding control measures such as ventilation, misting, or shading. Livestock houses often experience elevated temperature and humidity due to high stocking density and limited air exchange, making continuous monitoring or estimation of THI vital for maintaining animal welfare and stable production.

Table 1 presents the classification of heat-stress levels for broiler and swine farms, as defined by the National Institute of Animal Science in South Korea. These thresholds categorize thermal stress into levels ranging from "optimal" to "lethal". They were used to interpret the estimated THI values, enabling the identification of critical periods when animals were likely to experience heat stress and supporting timely management interventions.

Data splitting and hyperparameter optimization

To enhance model performance and mitigate the risk of overfitting, the machine learning algorithms in this study underwent systematic hyperparameter optimization. A combination of exhaustive search and cross-validation was employed to identify the most effective configuration of model parameters. This approach ensured that the models achieved high predictive accuracy while generalizing well to unseen data. Optimization was performed using the GridSearchCV method, which evaluates multiple hyperparameter combinations across a predefined search space. During this process, the training dataset was partitioned using five-fold cross-validation, whereby in each iteration one fold served as the validation set and the remaining folds were used for training. The algorithm systematically tested each hyperparameter configuration and selected the combination that maximized model performance metrics. This rigorous tuning procedure ensured that the final models were both robust and reliable, with hyperparameter settings tailored to the characteristics of the heterogeneous environmental dataset. The selected optimal hyperparameters for each algorithm are summarized in Table 2. To further ensure that the models captured seasonal variations in the environmental data, the dataset spanning July 2024 to June 2025 was divided into training and testing subsets based on meteorological seasons. Each season was split such that the first two-thirds of the data were allocated to

Table 1. Classification of heat-stress levels for broiler and swine farms. Source: National Institute of Animal Science, South Korea, https://chuksaro.nias.go.kr

Status	Broiler		Swine	
	THI range	Description	THI range	Description
Optimal	THI <68	Suitable environment for livestock farming.	THI <64	Suitable environment for livestock farming.
Caution	68≤ THI <72	Normal activity; beginning of temperature rise.	64≤ THI <73	Normal activity; temperature begins to increase.
Warning	68≤ THI <80	Feed intake decreases by 29%; body weight decreases by 22%.	73≤ THI <83	Mild stress; decreased feed intake, increased respiration rate and floor contact.
Danger	80≤ THI <91	Feed intake decreases by 31%; body weight decreases by 53%; respiration rate exceeds 200 breaths/min.	83≤ THI <93	Feed intake decreases by 30%; daily weight gain decreases by 35%; respiration rate and body temperature increase.
Lethal	THI ≥91	Risk of death; severe blood congestion, feather loss, difficulty standing.	THI ≥93	Risk of death; severe blood congestion, feather loss, inability to stand.

THI, temperature-humidity index.

the training set, while the remaining one-third was reserved for testing. This approach preserved the temporal structure and seasonal patterns within the data, allowing the models to learn from diverse conditions while evaluating performance on unseen periods within the same season. Specifically, the training set was composed of the early part of each season: June 2025 and July 2024 for summer, September-October 2024 for autumn, December 2024-January 2025 for winter, and March-April 2025 for spring. The test set was formed from the remaining dates within each corresponding season.

Model evaluation metrics

The accuracy of the estimation models was evaluated using several statistical metrics, including the Pearson correlation coefficient (R), root-mean-square error (RMSE), mean bias error (MBE), and mean absolute error (MAE). In addition, relative performance indicators such as relative RMSE (rRMSE), relative MBE (rMBE), and mean absolute percentage error (MAPE) were included. These relative metrics are particularly useful for assessing data that vary across different locations and time scales, as they express deviations in percentage terms, making results more interpretable and comparable across diverse conditions. The metrics were calculated using Eqs. (4)-(10):

$$R = \frac{\sum\left(y_m - \overline{y_m}\right)\left(y_e - \overline{y_e}\right)}{\sqrt{\sum\left(y_m - \overline{y_m}\right)^2\left(y_e - \overline{y_e}\right)^2}} \quad \text{(Eq. 4)}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_e - y_m)^2} \quad \text{(Eq. 5)}$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(y_e - y_m) \quad \text{(Eq. 6)}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_e - y_m\right| \quad \text{(Eq. 7)}$$

$$rRMSE = \frac{RMSE}{\overline{y_m}} \times 100\% \quad \text{(Eq. 8)}$$

$$rMBE = \frac{MBE}{\overline{y_m}} \times 100\% \quad \text{(Eq. 9)}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_e - y_m}{y_m}\right| \times 100\% \quad \text{(Eq. 10)}$$

Where y_m, y_e, $\overline{y}_m$, $\overline{y}_e$ and n represent the actual data, estimated data, mean value of actual data, mean value of estimated data, and the number of samples, respectively.

Results

Validation of satellite-derived solar irradiance, temperature, and humidity

The primary outdoor variables used for indoor heat-stress estimation were obtained from the GK2A satellite, making it necessary to evaluate the reliability of these inputs. Figure 3 presents the validation results for solar irradiance, temperature, and humidity, compared against ground-based measurements from the Korea Meteorological Administration (KMA, https://data.kma.go.kr) at the Gwangju station. The comparison indicates that GK2A products align well with in situ observations and are suitable for use as external drivers in indoor microclimate modeling. The validation yielded rRMSE values of 18.35%, 24.41%, and 26.77% for temperature, relative humidity, and GHI, respectively, confirming satisfactory performance.

Satellite-derived data provide several advantages for large-scale environmental assessment. Their wide spatial coverage and continuous temporal availability enable environmental information to be obtained for almost any location, including areas lacking ground-based sensors. Figure 4 presents example spatial maps of solar irradiance, temperature, and relative humidity at a selected time, illustrating the spatial variability that can be incorporated into the modeling framework.

Although solar irradiance can be estimated from GK2A observations, the Level 1B product provides radiometric measurements in the form of digital pixel numbers rather than direct irradiance values. Therefore, instead of using GHI directly as an input variable, the model incorporated alternative parameters derived from semi-empirical formulations, including SZA, SAA, and atmospheric turbidity. These variables are physically meaningful, consistently available from satellite products, and closely related to the radiative processes influencing indoor thermal dynamics. Using these surrogate parameters allows the model to capture solar effects without relying on additional conversion steps that may introduce uncertainty.

Temperature estimation

Table 3 presents the indoor temperature estimation results for the swine farms (Bethel and Handon) and broiler farms (Jangjiuk and Koryeo). XGBoost consistently provided the highest accuracy across all farms, with RMSE values ranging from 0.31 to 0.64°C and MAE values between 0.23 and 0.41°C. MBE values remained close to zero, indicating that the model did not systematically over- or underestimate temperature. Relative errors were small, with rRMSE values between 1.21% and 2.66 % and MAPE values mostly below 1.5%.

LightGBM showed slightly larger errors, with RMSE values between 0.37 and 0.74°C and MAE values between 0.27 and 0.48°C, though its performance remained consistent. CatBoost per-

Table 2. Optimized hyperparameters of machine learning models.

Model	Optimized hyperparameters
XGBoost	learning_rate=0.1; max_depth=10; subsample=0.9; colsample_bytree=0.7; min_child_weight=1; reg_alpha=0.1; reg_lambda=0.1;
LightGBM	learning_rate=0.1; max_depth=6; subsample=0.8; colsample_bytree=0.6; reg_alpha=0.1; reg_lambda=0.1; min_child_weight=1;
RF	max_depth = 6; min_samples_leaf = 1; min_samples_split = 2; n_estimators = 200;
CatBoost	learning_rate=0.1; depth=6; subsample=0.8; colsample_bylevel=0.8;

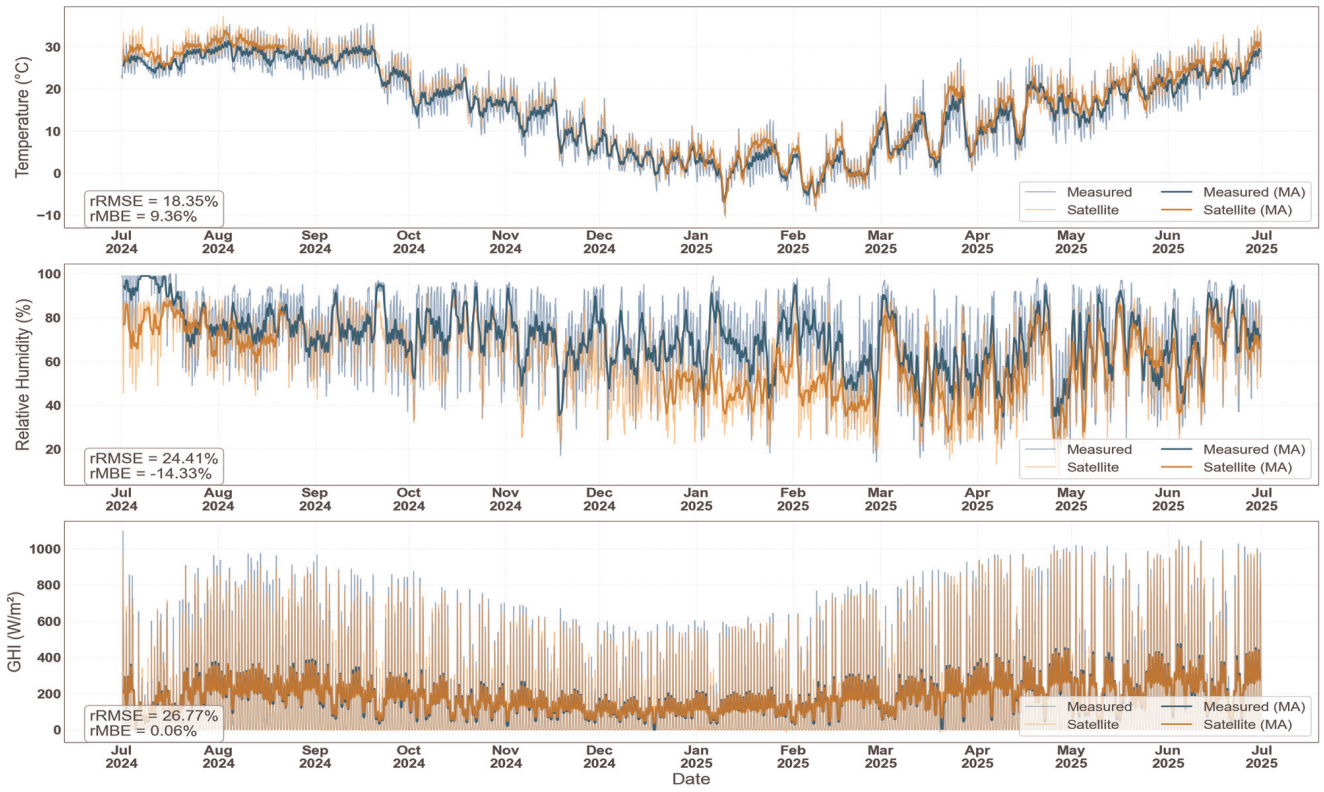

Figure 3. Validation of GK2A-derived temperature, relative humidity, and global horizontal irradiance (GHI) against ground based measurements at the Gwangju KMA station. Moving average (MA) smoothing is applied for clearer trend visualization.

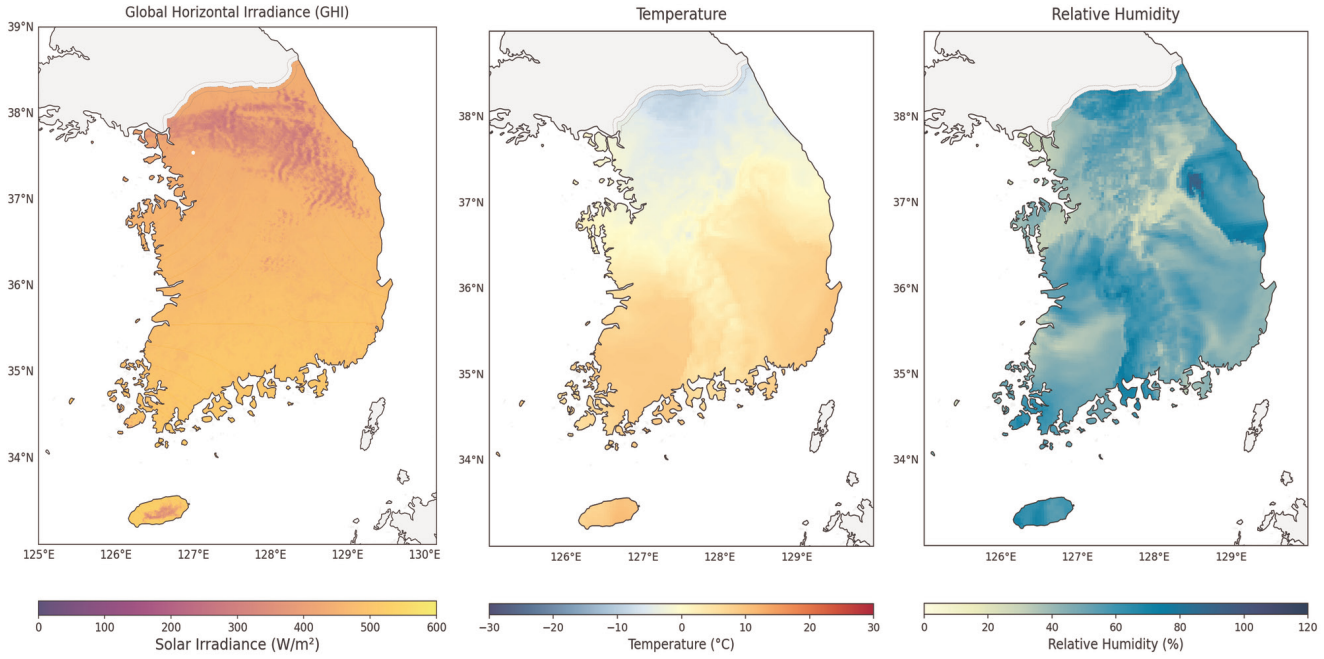

Figure 4. Spatial maps of solar irradiance, temperature, and humidity in South Korea based on GK2A satellite observations on 1 July 2025 at 12:00 KST.

formed moderately well, with RMSE values between 0.50 and 1.12°C and MAE values from 0.38 to 0.75°C, outperforming RF. RF recorded the highest errors, particularly in broiler farms, with RMSE values up to 2.52°C and MAE values reaching 1.78°C, indicating difficulty in capturing nonlinear temperature dynamics. Overall, these results confirm that XGBoost provides the most accurate and reliable temperature predictions across both swine and broiler farms.

Figure 5 presents scatterplots comparing estimated and measured indoor temperatures across all farms using the XGBoost model, which was selected for visualization because it achieved the highest accuracy among the tested machine learning approaches. The prediction points closely follow the one-to-one reference line in both swine and broiler farms, confirming strong agreement between estimated and observed values. The high coefficients of determination (R^2), ranging from 0.989 at the Bethel swine farm to 0.997 at the Koryeo broiler farm, demonstrate that XGBoost explains nearly all temperature variability in the test datasets. Histograms of predicted and measured temperatures also show similar distribution patterns, indicating that the model captures both central tendencies and the full range of temperature variations. These results confirming that XGBoost provides the most reliable and stable performance across different farm types and operational conditions.

Relative humidity

Table 4 summarizes the indoor relative humidity estimation results for the swine and broiler farms. XGBoost again provided the strongest performance, with RMSE values ranging from 1.56% to 1.95% and MAE between 0.83% and 1.32%. MBE values remained close to zero, indicating that the model did not systematically overor underestimate humidity. Relative errors were small, with rRMSE values between 2.50% and 2.91% and MAPE values below 2.0% in most farms, confirming that XGBoost maintained high accuracy under varying humidity conditions. LightGBM showed slightly higher errors, with RMSE values between 1.94% to 2.39% and MAE values between 1.38 and 1.65%, though performance remained consistent. CatBoost achieved moderate accuracy, with RMSE values ranging from 2.14% to 2.47% and MAE from 1.15% to 1.55%. RF performed the weakest, particularly in broiler farms, with RMSE values up to 5.31% and MAE values reaching 3.47%, reflecting difficulty in capturing nonlinear humidity dynamics. Overall, the results confirm that XGBoost consistently provides the most accurate humidity predictions across all farms, aligning with its superior performance in temperature estimation.

Figure 6 presents scatterplots of the XGBoost model's performance in estimating indoor relative humidity across the farms. Strong correlations between measured and estimated values were observed, with R^2 ranging from 0.964 at Bethel to 0.989 at Handon. The slightly higher R^2 at Handon suggested that the model captured humidity dynamics most accurately there, likely due to lower variability or better alignment of input features. In contrast, the marginally lower R^2 at Bethel indicated slightly higher residual errors, possibly arising from farm-specific factors or measurement noise. Overall, the consistently high R^2 values across all sites confirm the model's robustness, while small differences reflect local variations in environmental conditions or data characteristics.

Temperature-humidity index

Table 5 presents the THI estimation results across all farms. XGBoost consistently provided the most accurate predictions, achieving the lowest rRMSE and MAE values. For Bethel,

Table 3. Comparison of indoor temperature estimation performance across machine learning models.

	RMSE (°C)	MAE (°C)	MBE (°C)	rRMSE (%)	MAPE (%)	rMBE (%)
XGBoost						
Bethel	0.311	0.226	-0.006	1.218	0.873	-0.024
Handon	0.342	0.254	0.002	1.429	1.117	0.007
Jangjiuk	0.643	0.407	-0.005	2.658	3.588	-0.022
Koryeo	0.397	0.233	0.006	1.756	1.349	0.027
LightGBM						
Bethel	0.367	0.269	-0.006	1.435	1.042	-0.025
Handon	0.41	0.304	0.002	1.713	1.343	0.009
Jangjiuk	0.739	0.481	-0.001	3.052	4.697	-0.003
Koryeo	0.479	0.287	0.009	2.117	1.649	0.039
RF						
Bethel	0.853	0.651	-0.008	3.34	2.534	-0.031
Handon	0.853	0.651	-0.008	3.34	2.534	-0.031
Jangjiuk	2.519	1.784	0.03	10.403	22.24	0.125
Koryeo	1.8	1.288	0.02	7.954	9.047	0.09
CatBoost						
Bethel	0.503	0.38	-0.008	1.969	1.477	-0.03
Handon	0.607	0.462	0.01	2.536	2.045	0.04
Jangjiuk	1.125	0.755	0.001	4.645	7.106	0.005
Koryeo	0.75	0.493	0.016	3.316	2.844	0.069

XGBoost yielded an rRMSE of 0.623% and an MAE of 0.340, while Handon showed slightly higher errors with an rRMSE of 0.693% and an MAE of 0.379. Jangjiuk and Koryeo display somewhat larger deviations, with rRMSE values of 1.332% and 0.827%, respectively; however, these remained the lowest among all evaluated models. LightGBM demonstrated marginally higher errors, with rRMSE ranging from 0.732% to 1.542%, indicating reasonable performance yet consistently lower precision compared to XGBoost. RF exhibited substantially larger errors, particularly for Jangjiuk and Koryeo, where rRMSE values reached 5.466% and 3.886%, respectively, reflecting reduced reliability in THI estimation. CatBoost showed intermediate performance, with rRMSE values between 0.983% and 2.361%, highlighting lower accuracy than XGBoost while outperforming RF. Across all farms, XGBoost not only achieved the lowest rRMSE but also maintained the smallest MAE and MAPE values, demonstrating the robustness of its predictions. Performance differences among models were most pronounced in farms with higher variability, such as Jangjiuk, where cumulative errors in less precise models notably affected THI estimates.

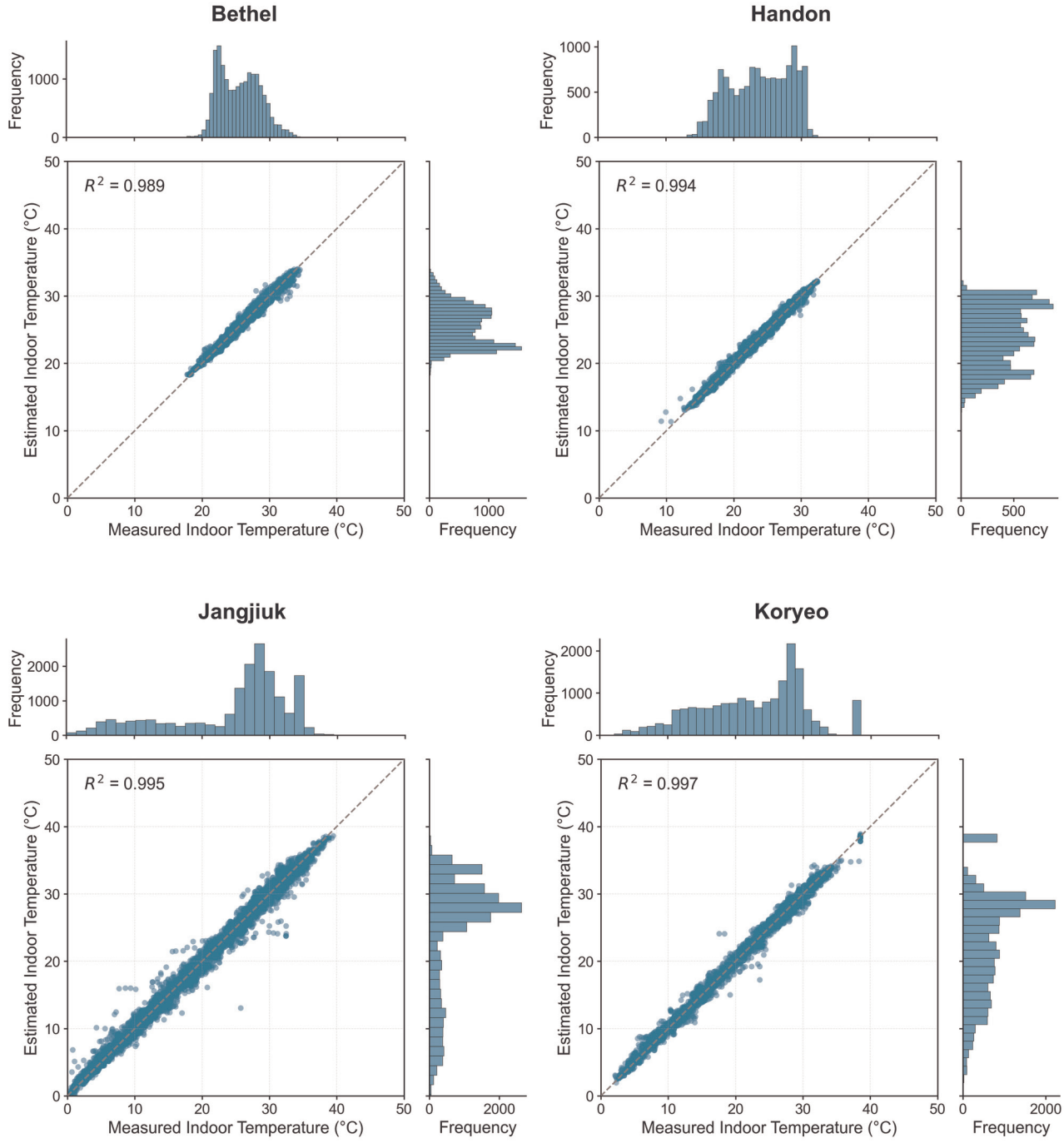

Figure 5. Scatterplots of estimated *vs* measured indoor temperature using the XGBoost model for swine and broiler farms.

Table 4. Comparison of indoor humidity estimation performance across machine learning models.

	RMSE (%)	MAE %)	MBE (%)	rRMSE (%)	MAPE (%)	rMBE (%)
XGBoost						
Bethel	1.71	1.201	-0.019	2.5	1.81	-0.027
Handon	1.506	1.067	0.023	2.148	1.672	0.033
Jangjiuk	1.95	1.299	0.022	2.591	1.815	0.030
Koryeo	1.42	0.832	0.019	2.014	1.220	0.027
LightGBM						
Bethel	1.944	1.375	-0.02	2.843	2.073	-0.029
Handon	1.783	1.265	0.036	2.543	1.993	0.051
Jangjiuk	2.277	1.543	0.021	3.026	2.150	0.028
Koryeo	1.685	1.003	0.017	2.39	1.479	0.024
RF						
Bethel	4.000	3.025	-0.052	5.848	4.645	-0.075
Handon	4.000	3.025	-0.052	5.848	4.645	-0.075
Jangjiuk	5.255	4.049	0.010	6.982	5.720	0.013
Koryeo	5.311	3.447	-0.034	7.533	5.263	-0.048
CatBoost						
Bethel	2.416	1.743	-0.028	3.533	2.628	-0.041
Handon	2.416	1.749	0.027	3.447	2.758	0.039
Jangjiuk	3.04	2.138	0.017	4.038	2.991	0.022
Koryeo	2.468	1.553	0.012	3.501	2.321	0.017

Table 5. Comparison of temperature-humidity index (THI) estimation performance across machine learning models, calculated using the THI equation based on temperature and humidity predictions.

	RMSE (%)	MAE (%)	MBE (%)	rRMSE (%)	MAPE (%)	rMBE (%)
XGBoost						
Bethel	0.466	0.340	-0.008	0.623	0.450	-0.011
Handon	0.505	0.379	0.007	0.693	0.526	0.010
Jangjiuk	0.967	0.615	0.008	1.332	0.949	0.011
Koryeo	0.582	0.346	0.015	0.827	0.520	0.022
LightGBM						
Bethel	0.547	0.403	-0.008	0.732	0.535	-0.011
Handon	0.598	0.450	0.011	0.820	0.626	0.015
Jangjiuk	1.119	0.725	0.020	1.542	1.131	0.027
Koryeo	0.708	0.425	0.021	1.006	0.638	0.030
RF						
Bethel	1.275	0.973	-0.015	1.705	1.292	-0.021
Handon	1.275	0.973	-0.015	1.705	1.292	-0.021
Jangjiuk	3.966	2.700	0.259	5.466	4.542	0.357
Koryeo	2.736	1.953	0.041	3.886	3.093	0.058
CatBoost						
Bethel	0.735	0.555	-0.009	0.983	0.737	-0.012
Handon	0.881	0.675	0.023	1.209	0.942	0.031
Jangjiuk	1.713	1.136	0.039	2.361	1.797	0.054
Koryeo	1.096	0.719	0.037	1.557	1.083	0.052

Figure 7 illustrates the comparison between THI values derived from indoor temperature and humidity measurements and the THI estimates produced by the XGBoost model using outdoor meteorological inputs. Across all four farms, the predicted THI values showed strong correspondence with the reference values, with points closely clustered around the 1:1 line. Model performance was consistently high, yielding R^2 values of 0.991 for Bethel, 0.996 for Handon, 0.995 for Jangjiuk, and 0.997 for Koryeo. These high coefficients of determination indicated that more than 99% of the variability in indoor-derived THI was explained by the model predictions. Although slight deviations from the reference line were observed, the dispersion was minimal, demonstrating that the model accurately captured indoor THI dynamics based solely on outdoor meteorological conditions. Overall, these results highlight the strong predictive capability and generalizability of the XGBoost model across diverse farm environments.

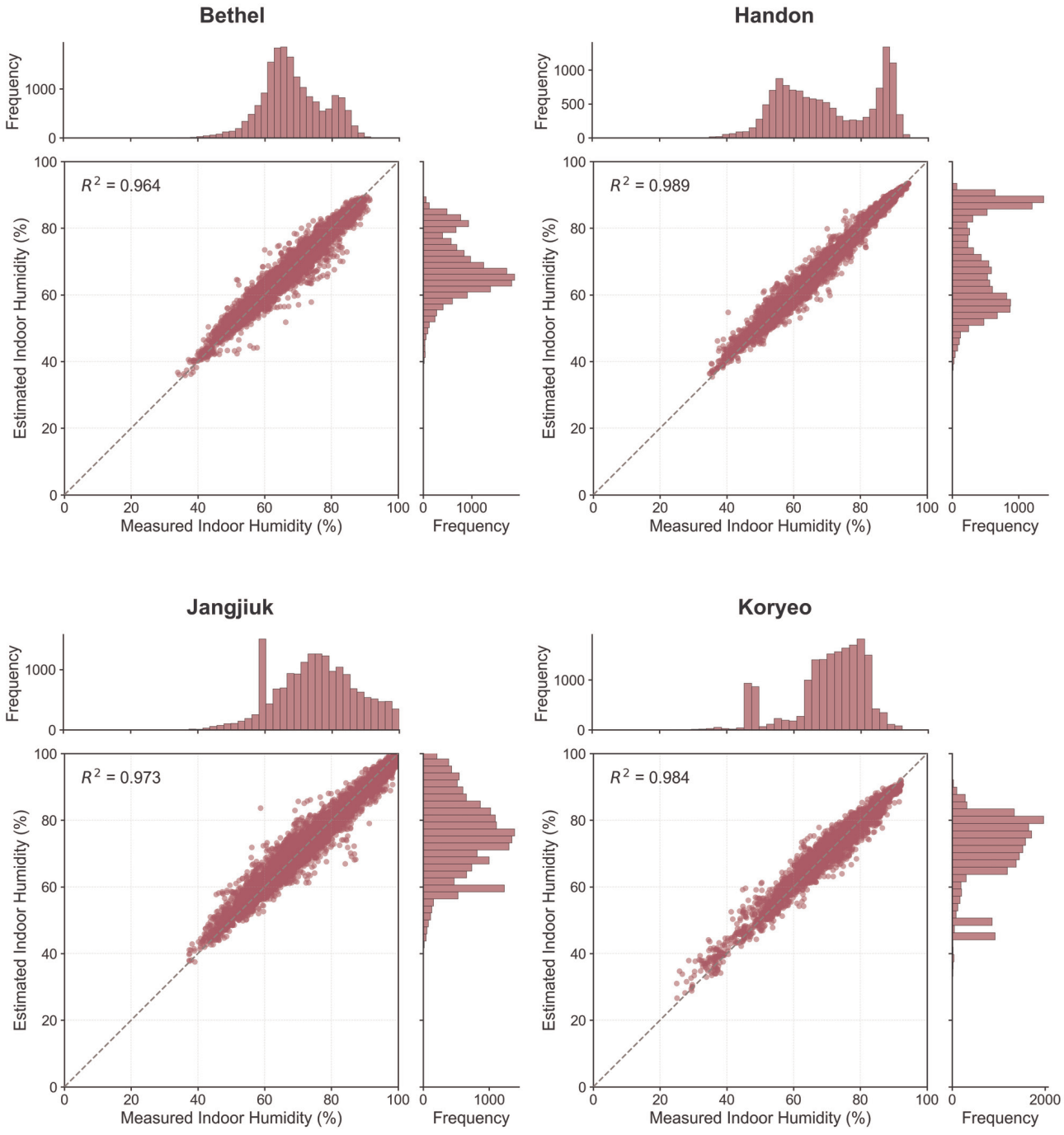

Figure 6. Scatterplots of estimated *vs* measured indoor humidity using the XGBoost model for swine and broiler farms.

Discussions

Temperature

The SHAP analysis presented in Figure 8 highlights the relative importance and directional impact of individual features on the XGBoost model's predictions of indoor temperature across the four farm sites (Bethel, Handon, Jangiuk, and Koryeo). Across all sites, the model demonstrated a consistent hierarchy of predictors: satellite-derived temperature and temporal indicators (time of day, day of week, day of month) emerged as dominant drivers, with mean SHAP magnitudes substantially exceeding those of secondary features. Notably, TL estimates derived from PVLib, which characterize atmospheric clarity and aerosol optical depth, ranked as important secondary features at several sites, reflecting their role in modulating solar radiation transmission to ground level and consequently influencing indoor thermal conditions. Solar irradiance metrics (SZA, SAA, EI), satellite-derived relative humidity, and other environmental variables exhibited lower but non-negligible importance, indicating that the model effectively

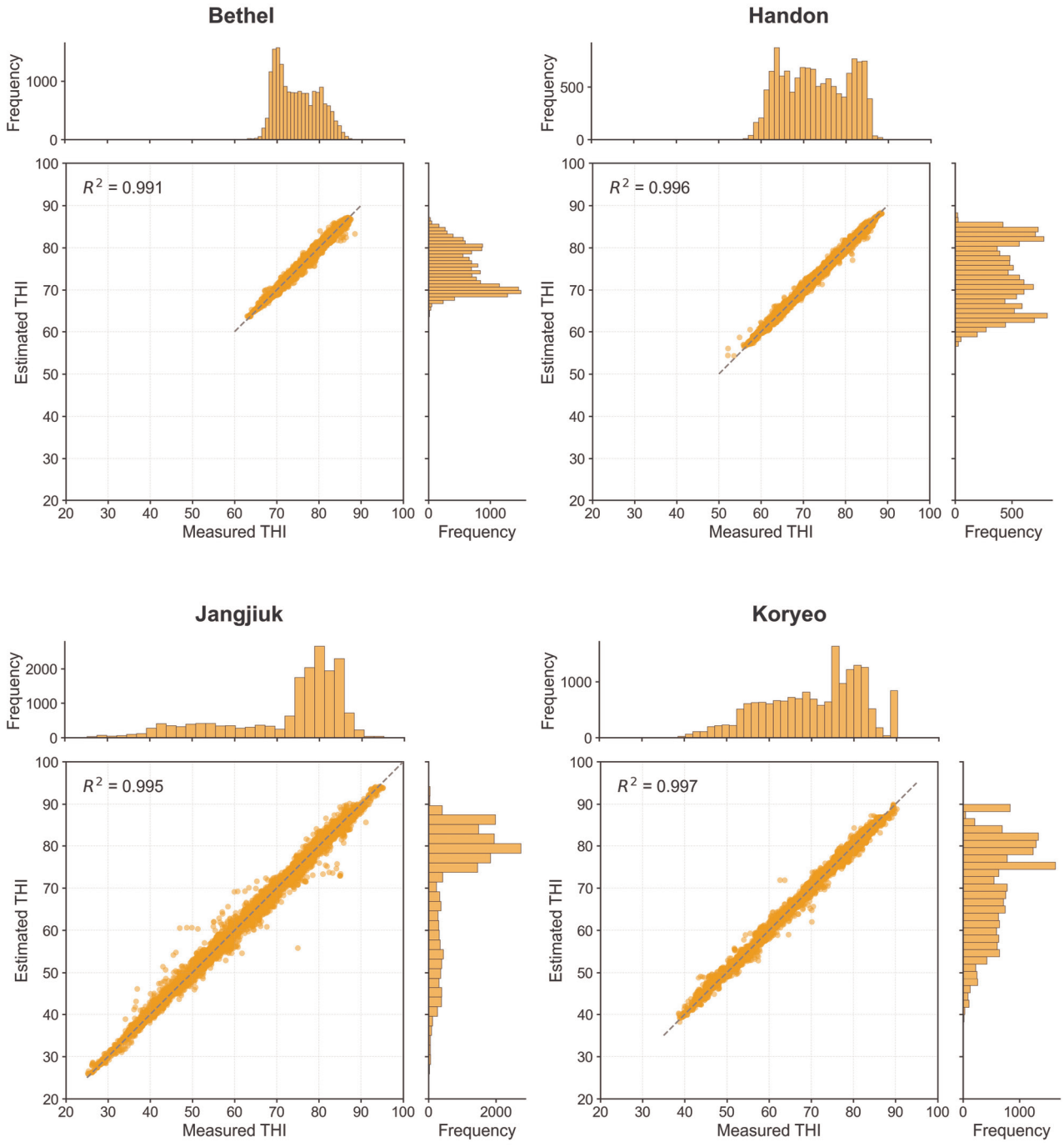

Figure 7. Scatterplots of estimated *vs* measured indoor humidity using the XGBoost model for swine and broiler farms.

integrated multiple atmospheric and temporal data streams to capture complex, non-linear relationships between satellite observations, atmospheric properties, and indoor temperatures.

The wide distribution of SHAP values across features demonstrated the model's adaptive responsiveness to varying input magnitudes, validating its applicability as a practical tool for accurate thermal environment estimation in commercial broiler production systems.

Relative humidity

The SHAP analysis for indoor relative humidity estimation (Figure 9) revealed notable differences in feature importance compared to temperature prediction across the four farm sites. Satellite-derived relative humidity consistently ranked as the most influential predictor, reflecting the strong correspondence between atmospheric moisture and indoor conditions.

However, the hierarchy of secondary features varied across sites. At Bethel and Handon, TL estimates from PVlib and satellite-derived temperature emerged as important secondary drivers. In contrast, at Jangiuk and Koryeo, temporal variables (day of month, hour, day of week) and radiative parameters (EI, CI, and SE) exhibited greater relative influence. This site-specific heterogeneity indicates that indoor relative humidity is modulated by distinct combinations of atmospheric moisture transport, turbidity effects, and localized temporal and solar geometry patterns associated with ventilation management, occupancy schedules, and seasonal variations in solar forcing.

The broader distribution of SHAP values and wider range of feature contributions for humidity estimation compared to temperature prediction suggested a more complex, multi-factorial control of indoor moisture conditions. This required the model to integrate diverse atmospheric, radiative, and operational variables to achieve accurate estimation across different farm configurations and climatic contexts.

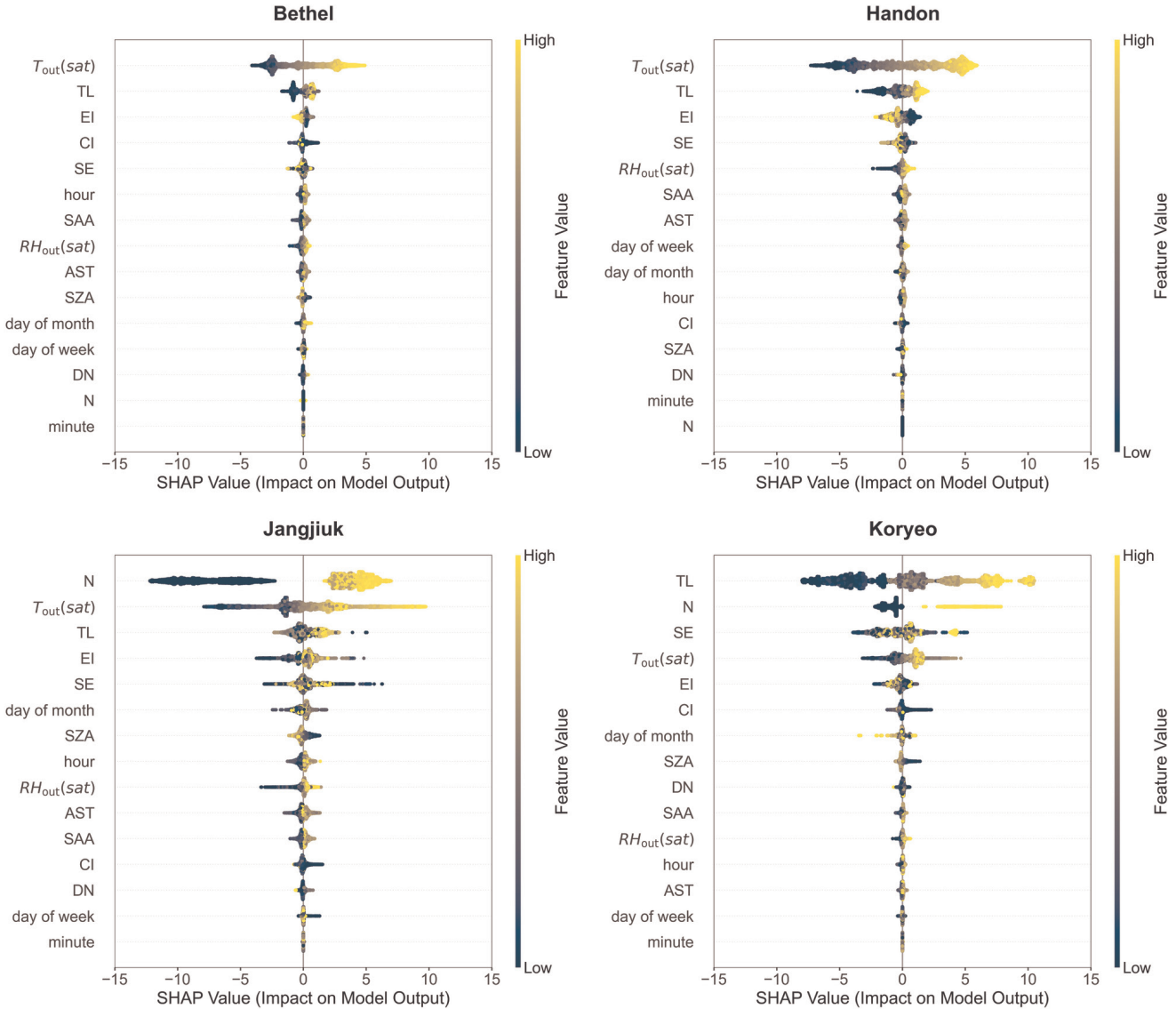

Figure 8. SHAP feature importance for XGBoost-based indoor temperature estimation across four farms.

Temperature-humidity index

Figure 10a presents results for swine farms at Bethel and Handon. At Bethel, the model accurately estimated the dominant "caution" (~35%) and "warning" (~42%) categories, with deviations below 2%, indicating reliable performance under moderate to high thermal stress conditions. At Handon, where "optimal" conditions accounted for ~32% of observations, the model achieved similar accuracy for the dominant category but slightly underestimated the "caution" category by 2–3%. For both swine facilities, minor underestimations were observed in the "danger" and "lethal" categories, which represented smaller proportions of the overall distribution. The broiler farm sites exhibited comparable estimation performance (Figure 10b). At Jangjuk, the model accurately replicated the "optimal" category (~27%) and achieved high precision in estimating "warning" (~33%) and "danger" (~37%) categories, with deviations within 1-2%. At Koryeo, which exhibited the most favorable thermal profile with "optimal" conditions comprising 41% of observations, the model demonstrated excellent model accuracy for this predominant category.

The close agreement between observed and estimated THI distributions validates the XGBoost model's practical utility for livestock management. Accurate estimation of the frequency and distribution of thermal stress categories enables several critical applications: real-time thermal stress monitoring and early warning systems, strategic infrastructure and cooling capacity planning, precision feed management adjusted for anticipated thermal conditions, and production performance forecasting. The slight tendency to underestimate extreme categories ("danger" and "lethal") suggests that model outputs should be interpreted conservatively for critical management decisions, with safety margins incorporated into response protocols. Overall, mean absolute errors below 2% points for major THI categories across diverse production environments confirmed the model's reliability as a practical tool for thermal

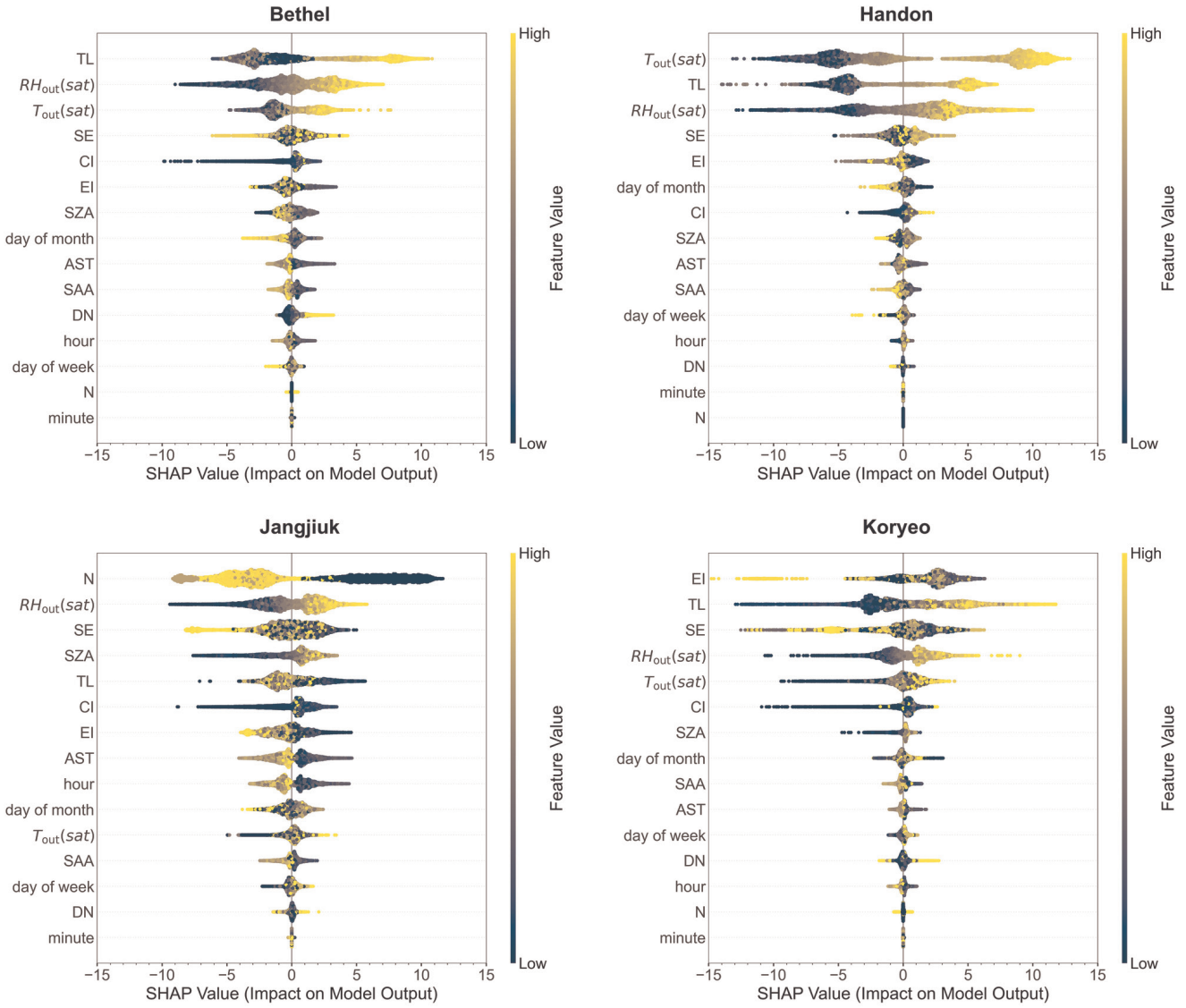

Figure 9. SHAP feature importance for XGBoost-based indoor relative humidity estimation across four farms.

environment assessment and decision support in modern livestock production systems.

Limitations and future directions

Although the satellite-based framework demonstrated strong potential for estimating indoor temperature and humidity using GK2A data, several limitations should be noted. The evaluation was conducted using observations from only two swine farms and two broiler farms, restricting the generalizability of the results. Consequently, farm-specific models were employed rather than a single unified model applicable across different production environments. Furthermore, the spatial resolution of the satellite data, which ranges from approximately 1 to 2 km depending on the product, is still much coarser than the scale of individual livestock buildings. Consequently, the satellite-derived variables should be interpreted as indicators of the surrounding outdoor atmospheric conditions rather than direct representations of farm-level microenvironments. Additionally, local land-cover heterogeneity, including nearby buildings, paved surfaces, or surrounding urban areas, may affect the representativeness of the satellite data and introduce uncertainty into the estimation of indoor climate conditions. Further validation using farms located in diverse landscape settings is therefore necessary to better quantify these influences and assess the generalizability of the proposed framework.

In addition, the model did not include detailed information on farm management practices such as ventilation settings, fan operation, or cooling system use. The only internal input variable available was the number of animals housed, which may not fully represent the factors influencing indoor microclimate conditions.

Future work should expand the dataset to include a wider range of farms, geographic regions, and seasonal conditions to support the development of a more transferable model. Incorporating management and structural characteristics, including ventilation rate, cooling strategies, building materials, and stocking density, would likely enhance prediction accuracy and robustness. Furthermore, integrating satellite-derived data with real-time in-house sensor networks could enable more adaptive and farm-independent modeling approaches. These improvements would contribute to a scalable and reliable framework for estimating indoor environmental conditions and supporting precision livestock management.

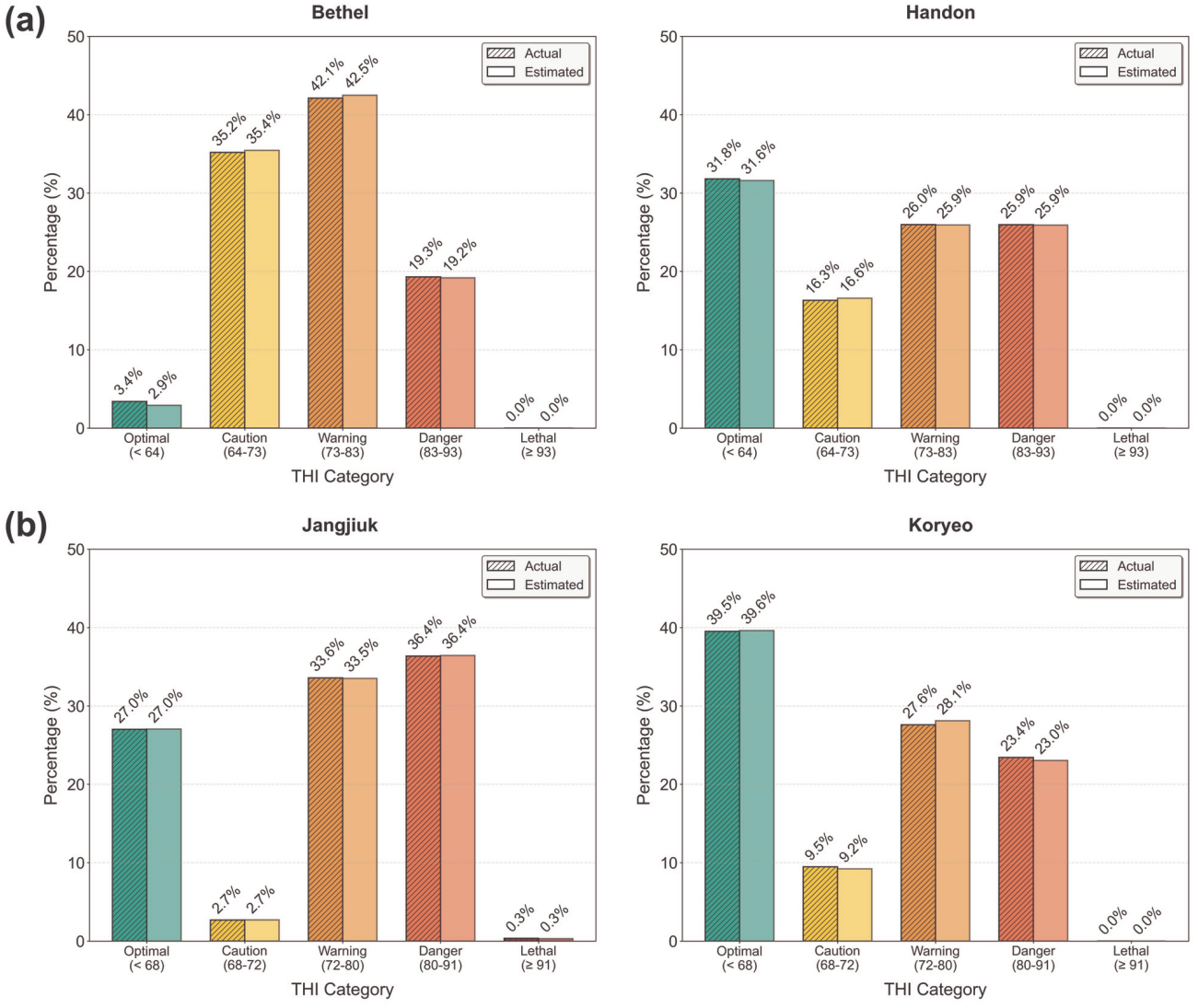

Figure 10. Observed *vs* estimated temperature-humidity index (THI) category distributions for swine farms at Bethel and Handon (**a**) and broiler farms at Jangiuk and Koryeo (**b**).

Conclusions

This study successfully developed and validated a satellite-based framework for estimating indoor heat stress in livestock facilities using GK2A satellite data and machine learning models. GK2A products showed reasonable accuracy when validated against ground measurements, with RMSE values of 58.06 W/m^2 for solar irradiance, 2.78°C for temperature, and 17.88% for relative humidity. An important advantage of GK2A geostationary satellite observations is their high temporal continuity and regional coverage, which enable frequent updates and capture rapid changes in outdoor weather conditions relevant to heat stress assessment. These satellite-derived meteorological variables served as reliable inputs for the machine learning models to estimate indoor environmental conditions using limited site-specific information. Among the four machine learning algorithms evaluated (XGBoost, LightGBM, RF, and CatBoost), XGBoost consistently demonstrated superior performance across all farms and parameters. For indoor temperature estimation, XGBoost achieved rRMSE values ranging from 1.218% to 2.658%, substantially outperforming RF (3.340-10.403%) and showing better accuracy than LightGBM and CatBoost. For humidity estimation, XGBoost exhibited the lowest errors with rRMSE values between 2.014% and 2.591%. When converted to THI, the primary indicator of heat stress, XGBoost achieved remarkable accuracy with rRMSE of 0.623-0.693% for swine farms (Bethel and Handon) and 0.827-1.332% for broiler farms (Jangjiuk and Koryeo). The THI category distributions (Figure 10) further confirm the model's capability to capture thermal stress levels accurately, with estimated and observed percentages showing close agreement across "optimal," "caution," "warning," and "danger" categories. The proposed framework provides a practical and scalable approach for enhancing continuous heat stress surveillance across livestock facilities. By integrating high-temporal-resolution geostationary satellite observations with machine learning algorithms, the framework can complement conventional farm monitoring systems by extending coverage to facilities where dense sensor deployment is limited, impractical, or economically constrained. In this regard, the approach is particularly valuable as an early warning and decision-support tool for identifying periods of elevated indoor heat stress risk, while also enabling broader regional-scale assessment across multiple farms. Nevertheless, this framework is not intended to replace direct on-farm monitoring or environmental control infrastructure. Effective real-time operation of ventilation, cooling, and other mitigation measures still depends on local sensors and automated control systems that directly capture and respond to indoor environmental conditions. Rather, the main utility of the proposed approach lies in its ability to supplement existing monitoring strategies and provide meaningful indoor heat stress estimates under data-limited conditions. Accurate estimation of indoor THI may therefore support farm managers in improving heat stress preparedness, optimizing ventilation and cooling interventions, and strengthening animal welfare management during hot and humid periods. In addition, the framework has practical significance because it remains applicable even when detailed building metadata are unavailable, a common constraint in commercial livestock production environments. Future research should incorporate more comprehensive structural and operational descriptors, including floor area, orientation, insulation properties, and quantitative ventilation performance, and should further investigate integration with real-time alert and environmental control systems to support proactive heat stress management in intensive livestock production.

References

Al-Hadhrami LM, 2013. Comprehensive review of cooling and heating degree days characteristics over Kingdom of Saudi Arabia. Renew Sustain Energy Rev 27:305-314.

Blanes-Vidal V, Guijarro E, Balasch S, Torres AG, 2008. Application of computational fluid dynamics to the prediction of airflow in a mechanically ventilated commercial poultry building. Biosyst Eng 100:105-116.

Chapman NH, Chlingaryan A, Thomson PC, Lomax S, Islam MA, Doughty AK, Clark CEF, 2023. A deep learning model to forecast cattle heat stress. Comput Electron Agric 211:107932.

Chen T, Guestrin C, 2016. XGBoost: A scalable tree boosting system. Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., San Francisco. Available from: https://www.kdd.org/kdd2016/papers/files/rfp0697-chenAemb.pdf

Deng S, Li Z, Wei Y, Wang Y, Li B, Zheng W, 2024. Assessing temperature distribution inside commercial stacked cage broiler houses in winter. Animals (Basel) 14:2638.

do Amaral Vercellino R, Nääs, I. de A., Moura, D.J. de, 2025. Tracking heat stress in broilers: a thermographic analysis of anatomical sensitivity across growth stages. Animals (Basel) 15:2233.

Georgiades P, Economou T, Proestos Y, Araya J, Lelieveld J, Neira M, 2025. Global projections of heat stress at high temporal resolution using machine learning. Earth Syst Sci Data 17:1153-1171.

Hammer A, Heinemann D, Hoyer C, Kuhlemann R, Lorenz E, Müller R, Beyer HG, 2003. Solar energy assessment using remote sensing technologies. Remote Sens Environ 86:423-432.

Herbut P, Angrecka S, Walczak J, 2018. Environmental parameters to assessing of heat stress in dairy cattle—a review. Int J Biometeorol 62:2089-2097.

Jang JC, Sohn EH, Park KH, 2022. Estimating hourly surface solar irradiance from GK2A/AMI Data using machine learning approach around Korea. Remote Sens 14:1840.

Kim HR, Seong P, Seol KH, Park JE, Kim H, Park W, et al., 2025. Effects of heat stress on growth performance, physiological responses, and carcass traits in broilers. J Therm Biol 127:103994.

Lee IB, Bitog JPP, Hong SW, Seo IH, Kwon KS, Bartzanas T, Kacira M, 2013. The past, present and future of CFD for agro-environmental applications. Comput Electron Agric 93:168-183.

Lin S, Gao W, Chen Y, Yao W, Fang Z, Guo C, Lu Z, 2026. Advancing thermal and humidity comfort assessment in hot-humid climates in summer through machine learning. Build Environ 288:113947.

Liu C, Cao Y, Luo Z, Liu Y, Reynolds CK, Humphries D, et al., 2025. Heat stress monitoring, modelling, and mitigation in a dairy cattle building in reading, UK: Impacts of current and projected heatwaves. Build Environ 279:113046.

Liu F, Zhao W, Le HH, Cottrell JJ, Green MP, Leury BJ, et al., 2022. Review: What have we learned about the effects of heat stress on the pig industry? Animal 16:100349.

Liu L, Ren M, Ren K, Jin Y, Yan M, 2020. Heat stress impacts on broiler performance: a systematic review and meta-analysis. Poult Sci 99:6205-6211.

Lundberg SM, Lee SI, 2017. A unified approach to interpreting model predictions. Proc. 31st Conf. Neural Information Processing Systems (NIPS2017), Long Beach. Available from: https://proceedings.neurips.cc/paper_files/paper/2017/file/8a2

0a8621978632d76c43dfd28b67767-Paper.pdf
M'Hamdi N, Darej C, Attia K, El Akram Znaidi I, Khattab R, Djelailia H, et al., 2021. Modelling THI effects on milk production and lactation curve parameters of Holstein dairy cows. J Therm Biol 99:102917.
Machado MR, Karray S, De Sousa IT, 2019. LightGBM: An effective decision tree gradient boosting method to predict customer loyalty in the finance industry. Proc. 14th Int. Conf. Comput. Sci. Educ. (ICCSE), Toronto; pp. 1111–1116.
Mueller R, Behrendt T, Hammer A, Kemper A, 2012. A new algorithm for the satellite-based retrieval of solar surface irradiance in spectral bands. Remote Sens 4:622-647.
Niu K, Zhong J, Hu X, 2024. Impacts of climate change-induced heat stress on pig productivity in China. Sci Total Environ 908:168215.
Park T, Lee J, Oh H, Yun WJ, Lee KW, 2025. Optimizing indoor farm monitoring efficiency using UAV: yield estimation in a GNSS-denied cherry tomato greenhouse. arXiv:2505.00995.
Patrick B, Johnson T, Kanjo E, 2024. Internetless low-cost sensing system for real-time livestock monitoring. IEEE Sensors Lett 8:7500504.
Perez R, Ineichen P, Moore K, Kmiecik M, Chain C, George R, Vignola F, 2002. A new operational model for satellite-derived irradiances: Description and validation. Sol Energy 73:307-317.
Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A, 2018. Catboost: Unbiased boosting with categorical features. Proc. 32nd Conf. Neural Information Processing Systems (NIPS2018), Montreal. Available from: https://proceedings.neurips.cc/paper_files/paper/2018/file/14491b756b3a51daac41c24863285549-Paper.pdf
Rajagukguk RA, Ramadhan RAA, Lee HJ, 2020. A review on deep learning models for forecasting time series data of solar Irradiance and photovoltaic power. Energies 13:6623.
Soci C, Hersbach H, Simmons A, Poli P, Bell B, Berrisford P, et al., 2024. The ERA5 global reanalysis from 1940 to 2022. Q J R Meteorol Soc 150:4014–4048.
Stein JS, Holmgren WF, Forbess J, Hansen CW, 2016. PVLIB: Open source photovoltaic performance modeling functions for Matlab and Python. Proc. IEEE 43rd Photovoltaic Specialists Conference (PVSC), Portland. pp. 3425-3430.
Uea-Anuwong T, Lam YK, Yau D, Li CM, Xin C, Paudel S, et al., 2025. Impact of temperature-humidity index on heat stress behavior and welfare of battery-caged yellow-feathered native broilers during summertime in Hong Kong: An observational study. J Therm Biol 132:104224.
Wan Q, Smigaj M, Brede B, Kooistra L, 2024. Optimizing UAV-based uncooled thermal cameras in field conditions for precision agriculture. Int J Appl Earth Obs Geoinf 134:104184.
Yan G, Shi Z, Cui B, Li H, 2022. Developing a new thermal comfort prediction model and web-based application for heat stress assessment in dairy cows. Biosyst Eng 214:72-89.
Yang D, Gu Y, Mayer MJ, Gueymard CA, Wang W, Kleissl J, et al., 2024. Regime-dependent 1-min irradiance separation model with climatology clustering. Renew Sustain Energy Rev 189:113992.
Yu P, Zhao T, Shi J, Ran Y, Jia L, Ji D, Xue H, 2022. Global spatiotemporally continuous MODIS land surface temperature dataset. Sci Data 9:143.

Received: 12 January; Accepted: 11 March 2026.

Contributions: Rial Arifin Rajagukguk, software, methodology, investigation, formal analysis, writing – original draft. Jinseon Park, writing – review & editing. Heejou Kim, writing – review & editing Chae-rin Lee, writing – review & editing Ji-yeon Park, writing – review & editing Se-yeon Lee, contribution to manuscript writing and editing. Taehwan Ha, writing – review & editing. Se-woon Hong, software, methodology, investigation, formal analysis, writing – review & editing.

Conflict of interest: the authors declare no competing interests, and all authors confirm accuracy.

Availability of data and materials: the datasets generated and/or analyzed during the current study are not publicly available due to confidentiality agreements with the collaborating farms, but are available upon reasonable request from the corresponding author.

Funding: this research was supported by the Cooperative Research Program for Agricultural Science and Technology Development, National Institute of Agricultural Sciences, Rural Development Administration, South Korea (grant number: RS-2025-00398327).

Publisher's note: all claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article or claim that may be made by its manufacturer is not guaranteed or endorsed by the publisher.