Computers and Electronics in Agriculture 232 (2025) 110019

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Weighing finishing pigs in motion: A walk-over scale for accurate weight estimation

François Decarie [a],*, Charles Grant [b], Gabriel Dallago [c]

[a] Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, Canada
[b] Department of Agribusiness and Agricultural Economics, University of Manitoba, Winnipeg, Canada
[c] Department of Animal Science, University of Manitoba, Winnipeg, Canada

ARTICLE INFO

Keywords:
Precision livestock farming
Machine learning
Dynamic weighing
Swine production
Animal welfare
Feature engineering

ABSTRACT

Accurate and efficient weight estimation of pigs is crucial for optimizing production, ensuring animal welfare, and making informed decisions in swine farming. Despite technological advancements, obtaining precise individual pig weights remains challenging due to the dynamic nature of pig movement and the stress induced by traditional weighing methods, highlighting the need for innovative, non-invasive solutions. This study presents an automated walk-over scale system that leverages high-frequency load cell data, feature engineering, and machine learning techniques to estimate pig weights in motion, addressing the limitations of traditional weighing methods. The system's effectiveness was validated in a real-world setting with 50 pigs across 944 walk-throughs, achieving a Root Mean Square Error (RMSE) of 2.87 kg and a Mean Absolute Percentage Error (MAPE) of 2.65% on a 20% pig-wise holdout validation set, demonstrating its potential as a practical solution for non-invasive, accurate weight monitoring in commercial pig farming operations.

1. Introduction

The global pig meat market has seen dynamic trends in recent years, significantly shaped by economic factors and evolving consumer preferences. According to the latest Agricultural Outlook by the (OECD-FAO, 2023), global pig meat production in 2032 is projected to reach 129 million tonnes. This growth is primarily driven by increased production in China, Brazil, Vietnam, and the United States, while notable declines are expected in the European Union and other regions due to high production costs and ongoing challenges such as African Swine Fever (OECD-FAO, 2023). Economically, the pig meat market remains vital, contributing significantly to the global meat trade. In 2022, pig meat was the world's 136th most traded product, with a total trade value of $34 billion (OECD-FAO, 2023). Major exporters include Spain, the United States, Germany, Canada, and Denmark, while key importers are China, Japan, Italy, Mexico, and South Korea. This trade landscape reflects the broader economic value and competitive dynamics within the pig meat market (OECD-FAO, 2023). With the global demand for pork rising, particularly in the Americas, Asia and Europe (OECD-FAO, 2023), efficient and accurate pig growth measurement is essential to meet market demands and ensure farm sustainability (See, 2024). Accurate weight estimation of finisher pigs is fundamental for the effective management of commercial finishing environments, impacting economic outcomes, animal welfare, and overall farm profitability. Traditional weighing methods, which often involve confinement and handling, can induce stress, negatively affecting pigs' health and welfare. Animal Scientist Temple Grandin (1997) emphasized that reducing handling stress is critical for maintaining animal welfare, as stress during weighing can lead to behavioral distress, physiological disruptions, and poor meat quality (Grandin, 1997). Grandin (1988) also noted that pigs appear to "enjoy" walking through alleyways during gentle driving, where they show excitement and enthusiasm rather than fear, reinforcing the benefits of stress-reducing techniques. Automated walk-over weighing systems provide a non-invasive alternative that minimizes handling stress while enabling precise weight monitoring, aligning with advancements in precision livestock farming.

The ability to accurately determine pig weights plays a crucial role in various aspects of swine production, including feed management, health interventions, and optimal market delivery timing. Accurate weights are essential for transactions involving finishing pig shipments, determining optimal feeding regimes, and assigning shipping dates to maximize revenue according to scoring grids provided by processing plants (Pomar and Remus, 2019). Furthermore, precise weight monitoring enhances feed efficiency, informs health management decisions, and improves overall productivity, collectively supporting both

* Corresponding author.
E-mail address: Decarief@myumanitoba.ca (F. Decarie).

https://doi.org/10.1016/j.compag.2025.110019
Received 8 September 2024; Received in revised form 30 December 2024; Accepted 20 January 2025
Available online 6 February 2025

animal welfare and profitability (Gómez et al., 2021). Incorporating stress-reducing techniques, such as alleyway movement and automated systems, not only improves animal handling but also contributes to calmer, healthier pigs and better meat quality outcomes.

Traditional weighing methods, which typically involve confining pigs in crates or on static scales, present several challenges. These methods are labor-intensive, time-consuming, and can induce stress in the animals, potentially affecting their well-being and growth performance (Liu et al., 2023b). The stress associated with handling and confinement during weighing can lead to inaccurate measurements and may have negative impacts on animal welfare and productivity (Conte et al., 2014). Early attempts to automate pig weighing faced significant hurdles. Smith and Turner (1974) explored the use of electronic circuits to enhance conventional scales but found these systems unreliable due to the dynamic movements of pigs. Slader and Gregory (1988) developed a system for monitoring feed intake using radio frequency transponders but encountered issues with mechanical malfunctions and stress induced by single-file feeding. While innovative for their time, these historical methods failed to provide practical solutions for dynamic, high-throughput weighing environments.

More recent research has focused on developing non-invasive and automated weighing systems for pigs. Kashiha et al. (2014) proposed an image analysis method for automatic weight estimation of individual pigs in group-housing systems, achieving a mean absolute percentage error of 3.2%. Condotta et al. (2018) demonstrated the potential of depth sensors for mass estimation of growing and finishing pigs, achieving an R^2 of 0.988 for weight prediction. Shi et al. (2016) developed a method combining 2D and 3D cameras for pig weight estimation in group-housing environments while accounting for various pig postures; their approach achieved a mean relative error of 2.57%. Most recently, Paudel et al. (2024) developed an approach using 3D point cloud data and deep learning to estimate finishing pig weights. Their PointNet-based method processed data from freely moving pigs in a holding pen, achieving an overall R^2 of 0.94 and RMSE of 6.87 kg.

These advancements in automated pig weighing systems represent significant progress toward overcoming the limitations of traditional methods. However, significant challenges remain in developing systems that are accurate, robust, and practical for large-scale commercial operations. While the academic focus has been on refining technologies for precision and accuracy, the market has simultaneously introduced commercial solutions based on load cells and sensor-integrated systems, designed for weight monitoring, automatic feeding, and sorting of pigs. Numerous commercial systems are already available in Europe, with Gómez et al. (2021) highlighting market-ready tools in their supplementary materials. These systems boast advanced automation and user-friendly integration into existing farm management practices, reflecting a high technology readiness level. Nevertheless, despite their market availability, independent validation studies of these commercial systems are scarce. It remains unclear how reliably these technologies perform under diverse and dynamic farming conditions, especially regarding factors such as pig behavior, environmental variation, and calibration drift over time. The lack of rigorous, peer-reviewed evaluations creates a gap between commercial claims and scientific validation.

The primary challenge lies in obtaining a validated dataset with accurate weights that is sufficiently large and diverse (L'Heureux et al., 2017). Such a dataset needs to include a wide range of pig sizes as well as various farm conditions and animal movements on the scale (Wang et al., 2024). It must account for the dynamic nature of walk-through weighing, where pigs are in motion and may exhibit different gaits, speeds, and movements, such as jumping or running across the scale. Collecting such comprehensive data is labor-intensive and time-consuming, often requiring coordination between researchers and farm operators. Despite these difficulties, a high-quality, diverse dataset is crucial for developing and validating accurate weight estimation models that can perform reliably in real-world commercial settings (Gómez et al., 2021).

While advancements in automated weighing systems have progressed significantly for other livestock species, such as dairy cows and grazing steers, the challenges in pig production remain distinct. Walk-over weighing systems (WOW), for instance, are widely integrated with milking robots to monitor cow body weights (Alawneh et al., 2011; Mardhati et al., 2021). However, dairy producers primarily focus on milk production and body condition scores (BCS) rather than body weight alone. In contrast, for pig producers, precise body weight measurements are critical for optimizing feed efficiency, growth monitoring, shipment scheduling, and ultimately maximizing revenue. Pig behavior—such as erratic movements—poses unique challenges to automated weighing systems (Vranken and Berckmans, 2017).

Modern livestock weighing methods pose significant operational challenges. Typically, animals must be restrained and weighed on static scales—often placed as gates between feeding stations and holding pens—requiring each animal to be weighed individually. This approach demands extensive pig training, increases waiting times for feed, and elevates animal stress, leading to welfare concerns. While advanced machine learning techniques show promise, they still rely on precise reference weights—commonly gathered via static load cells—which is highly labor-intensive. The adoption rates of continuous weighing technologies, remain limited due to several interconnected challenges. One primary issue is the unavailability of reliable methods to verify the precision of these systems under diverse commercial farming conditions. Without standardized validation protocols, potential users face uncertainty regarding the reliability of weight predictions, which hinders widespread implementation (Banhazi et al., 2022). Furthermore, anecdotal evidence suggests that technical problems, such as suboptimal variable farm management practices, exacerbate the difficulty of achieving consistent accuracy in weight estimation (Banhazi et al., 2012) These challenges, combined with the lack of proper communication about the expected precision and limitations of these tools, contribute to a general mistrust among farm managers (Banhazi et al., 2022). This mistrust can result in low adoption rates, despite the potential economic and operational benefits that such technologies can provide when implemented correctly. Addressing these gaps through rigorous precision studies and enhanced user engagement is crucial for the broader adoption of continuous weighing technologies (Banhazi et al., 2022). This underscores the need for reliable weight measurement systems that can function effectively under actual farm conditions.

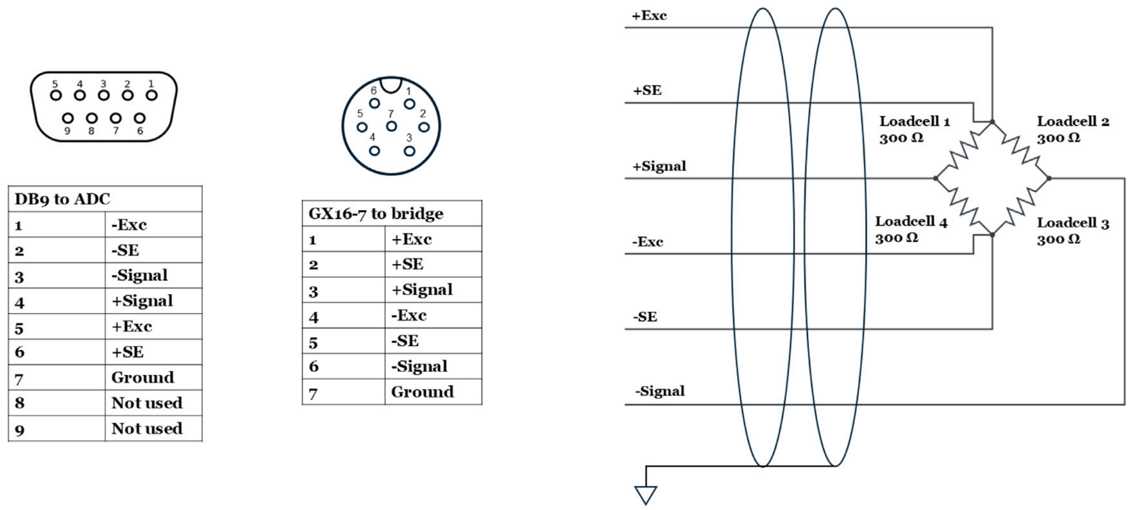
Our research aimed to develop a parallelizable weighing station prototype for finishing pigs that removes the need for restraint or extensive pig training. By integrating high-frequency load cell data, signal processing, and machine learning algorithms, we capture continuous weight measurements while pigs move freely through the station. We hypothesize that a pig's mass can be accurately correlated to its dynamic weight across various walking speeds. The high sampling rate adapts to diverse movement patterns, and the machine learning models handle non-linear dynamics and complex interactions.

Our goal is to streamline the weighing process and increase operational throughput, making large-scale implementation more feasible. Furthermore, we aim to address concerns regarding expected precision by transparently presenting the methodology and validation procedures. By clearly defining the limitations and strengths of our approach, we hope to enhance trust in continuous weighing technologies and provide a reproducible framework for evaluating precision under commercial conditions.

2. Materials and methods

2.1. Animals, apparatus, and data collection

The study was conducted at the Glenlea Research Station at the University of Manitoba. All experimental procedures were approved by the Animal Care Committee of the Research Ethics Board (Protocol



(a) Connections to the scale head. **(b)** Wheatstone bridge configuration.

Fig. 1. Scale setup showing (a) connections to the scale head and (b) Wheatstone bridge configuration.

Table 1
Characteristics of the pigs in the study.

Item	Group 1	Group 2
Total Pig Count	25	25
Female Count	7	14
Male Count	18	11
Average Birth Weight	1.84 kg	2.01 kg
Standard Deviation	± 0.30	± 0.29
Birth Date Range	24–28 November 2023	6–8 January 2024
Mortality	0	2
Initial Scaling Date	2 February 2024	6 March 2024
Final Scaling Date	22 April 2024	26 May 2024

Reference Number: F23-023, AC11851) at the University of Manitoba, Fort Garry campus. A total of 50 pigs of mixed sex, fed *ad libitum*, divided into two groups, were used in the study, as shown in Table 1. These animals were of a grower hybrid breed from Landrace-Yorkshire dames and DanBred Duroc sires and were clinically healthy at the start of the experiment.

The 50 pigs were assigned to two groups of 25 pigs each based on date of birth. Two scaling trials were conducted, whereby each pig was scaled once on a stable scale and once on a dynamic scale. In total, the study conducted a combined 944 scaling events on both the stable and dynamic scales.

The scales employed in this study were industrial floor models (Model WB244243, Global Industrial) with a maximum capacity of 454.55 kg and a manufacturer's accuracy of 0.23 kg. Each scale measured 1.48 meters in length and 0.51 meters in width and featured a tread plate surface for slip prevention with a factory-calibrated scale head. The scales were modified to add bars on the lateral edges, and additional tread plate treatment was applied to provide further slip prevention.

One scale was designated for obtaining stable weight measurements, and the other scale was designated for obtaining dynamic weight measurements. Each scale was equipped with four 300-ohm load cells arranged in a Wheatstone bridge configuration, connected via a 7-wire cable to the scale head, as shown in Fig. 1.

For the dynamic scale, the manufacturer's scale head unit, limited to a 1 Hz output rate, was replaced with a Digital Strain Gauge Converter to USB (DSCusb) from Mantracourt. This device was chosen for its high speed, accuracy, and USB connectivity, which facilitated efficient data transfer.

Custom CAT5e cables were fabricated to connect the scales' Wheatstone bridges via GX14-7 Aircraft connectors to the DSCusb's DB-9 connector. The ADC was configured with a nominal sensitivity of 2.5 mV/V and linearized. All optional filtering functions were disabled in the device configuration. The cables were tested, and both the stable and dynamic scales were calibrated prior to each trial using known 20 kg weights, ensuring linear responses across the expected range of pig weights.

The data collection process began with creating a sample label that included the pig's ear tag ID, the walk-through ID, and a timestamp to ensure traceability and record keeping. Each pig was then guided toward an aluminum-railing hallway onto the stable scale one at a time. A thin 3 mm transparent polycarbonate panel hung with plastic chains from the ceiling was used to discourage the pig from moving forward once standing on the stable scale. The operator visually confirmed that all four of the pig's hooves were in contact with the scale surface, indicating a stable stance. The panel served as a non-obtrusive obstacle that successfully prevented the pigs from progressing beyond the stable scale until the stable weight was measured and recorded. The stable weight was recorded when the reading did not fluctuate for two seconds. Data was collected using an Intel NUC with an i7 processor for real-time data capture.

Fig. 2 shows the layout and dimensions of the experimental apparatus and Fig. 3 shows a photograph of the setup.

Once the stable weight was recorded, the dynamic scale was activated automatically, and the operator manually lifted the polycarbonate panel, granting the pig unimpeded access to proceed onto the dynamic scale at its own pace. The pig's movement across the dynamic scale was recorded at a high sampling rate of 200 Hz using custom-developed Python software that interfaced with the scales through serial ports. The software continuously monitored the weight readings from the dynamic scale and implemented an online rolling average calculation over 20 samples to debounce the output signal.

A 2 kg threshold was applied to the rolling average to detect the pig's presence on the scale and initiate data recording. The high-frequency load cells, operating at 200 Hz, are highly sensitive and subject to minor fluctuations even when no pigs are on the scale. Setting the threshold at 0 kg would trigger the presence detection prematurely, resulting in false positives when no pigs are present. Conversely, setting the threshold too high, such as 20 kg, would risk cutting off parts of the waveform when only the first or last hoof is on the platform, leading to incomplete or inaccurate weight recordings. By establishing a 2 kg threshold, we ensure robust presence detection that eliminates both false positives and false negatives, allowing the recording process to start and end precisely when the pig fully enters and exits the scale.

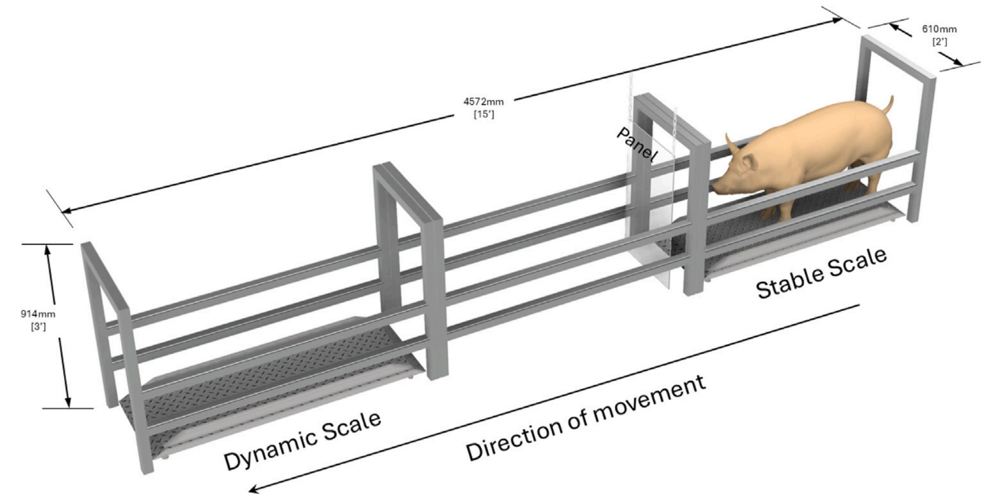


Fig. 2. Technical drawing of the experimental apparatus.

Fig. 3. Photograph of the experimental setup.

This threshold guarantees that the complete walk-through waveform is captured accurately, isolating the pig's weight data from background noise while preserving the integrity of the measurement.

Hesitation of a pig's progress across the dynamic scale occurred when the pig engaged the scale but subsequently backed away from the scale. The software incorporated a hesitation detection algorithm

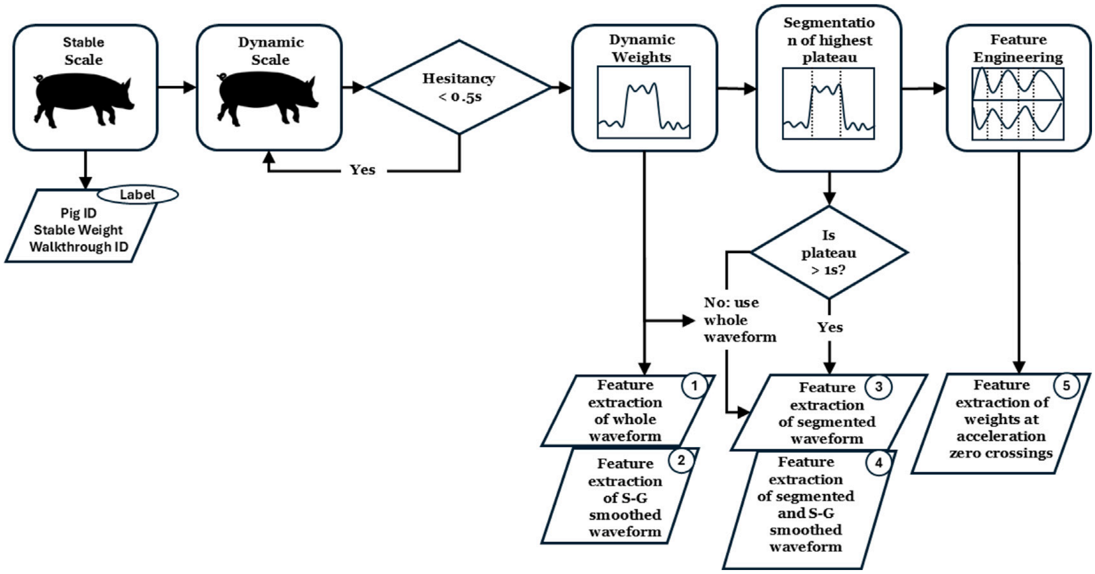


Fig. 4. Data flow diagram illustrating the process from pig walk-through to feature extraction.

to disregard the data collected during hesitation within the dynamic weight data readings. This algorithm monitored the duration of the pig's walk-through on the dynamic scale using the same rolling average calculation and 2 kg threshold, ensuring that only waveform recordings with more than 100 samples or 0.5 s, representing the pig's full weight in motion across the scale, were captured and logged. Data recording for each pig was terminated when the rolling average values from the dynamic scale fell below the 2 kg threshold, indicating that the pig had exited the scale.

Before each trial, the scales were calibrated and fully re-linearized from the tare value to 160 kg in 20 kg increments. This procedure ensured accurate measurements by compensating for temporal drift, temperature variations, electrical fluctuations, and any shifts in load cell contacts. The scale's design also addresses practical concerns such as slip prevention, durability, and ease of maintenance. Specifically, the Wheatstone bridge circuitry was protected by silicone caulking, allowing it to withstand high-pressure washing.

Data collection was conducted twice per week throughout the growth cycle of the pigs. In alignment with the National Research Council's range of the grower-finisher phase, which ranges from 20 to 140 kg (National Research Council, 2012), this study collected data from pigs starting at 20 kg and extending up to 150 kg. This wider range, exceeding typical commercial operation weights by 10 kg, was adopted to ensure a comprehensive representation of the entire grower-finisher stage.

2.2. Data pre-processing

Each walk through waveform was used to generate five sets of extracted features Fig. 4 as well as the stable weight label. The features were extracted using the raw waveforms but also after applying the Savitzky–Golay (S–G) filter, segmenting the waveform to keep the highest plateau and obtaining zero-acceleration moment weights.

The S–G filter is a digital smoothing polynomial filter that increases the precision of data by fitting successive subsets of adjacent data points with a low-degree polynomial (Savitzky and Golay, 1964). To determine optimal parameters for the S–G filter, a static weight of 20 kg was placed on the dynamic scale, and the window length was progressively increased until the filtered loadcell response stabilized. Given the 200 Hz sampling rate, the Nyquist frequency is 100 Hz. This indicates the highest frequency that can be accurately represented in the data without aliasing. With this in mind, and to ensure a balance between noise reduction and preservation of meaningful weight fluctuations, a 3rd order S–G filter with a window length of 51 samples was selected (Schafer, 2011). This configuration provided the best compromise between smoothing out noise and retaining the signal fidelity necessary for accurate weight estimation.

The application of the S–G filter was not limited to the weight data alone. The same filtering approach was extended to other derived measurements, including the calculated speed and the acceleration of the pig on the dynamic scale (see Figs. 5 and 6).

In our dynamic scale system setup, a typical walk-through weight begins when the rolling average threshold of 2 kg is exceeded. Initially, the weight readings increase as the two front hooves apply pressure on the scale. This is followed by a plateau where all four hooves are on the scale, representing the pig's full weight. Subsequently, the weight decreases as the pig continues walking, with the two rear hooves and eventually half of the pig's weight leaving the scale until the reading drops below the threshold. Accurately identifying and segmenting each loadcell waveform at key transition points was essential for the subsequent feature extraction and weight prediction methods to function effectively. Binary segmentation, a changepoint detection technique for time series data (Fryzlewicz, 2014), was employed to identify and isolate individual pig walk-throughs within the continuous weight data recorded by the dynamic scale. This method was implemented using the Ruptures Python library (Truong et al., 2018) with a least-squares model, allowing for the precise identification of weight fluctuations corresponding to pig entries and exits on the scale. By identifying individual pig walk-through events within the continuous weight data, we were able to isolate the portion of the signal most relevant for weight estimation. This segmentation method focused on the stable weight readings when the pig was fully on the scale, minimizing the influence of transitional weight changes as the animal entered and exited. This approach enhanced the accuracy of weight calculations by prioritizing data from the most representative periods of the weighing process.

2.3. Feature engineering

When analyzing the motion of a pig on a dynamic scale, we considered the forces acting upon it. In the vertical axis, the pig is subject to two primary forces: the constant force of gravity (F_g), determined by its mass (m) and the acceleration due to gravity (g), expressed as $F_g = mg$, and the variable force due to the pig's self-motion (F_p), which

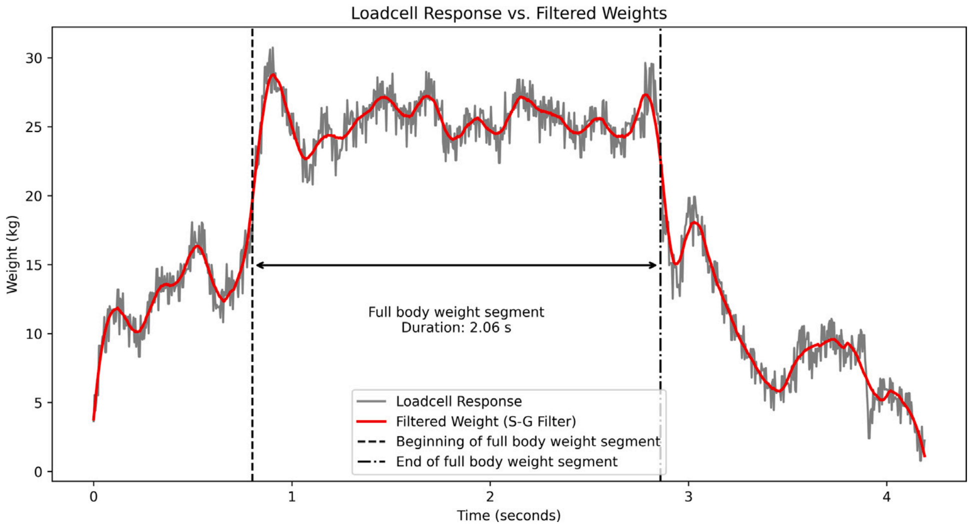


Fig. 5. Comparison of the raw loadcell response (gray) and the smoothed signal (red) using a 3rd-order Savitzky–Golay filter. The vertical dotted lines indicate the beginning and end of the stable weight segment, detected using binary segmentation (Ruptures Python library, least-squares model) with two breakpoints. The horizontal arrow indicates the segment's duration. This approach reduces high-frequency noise while preserving the overall trends, enabling processing of the full dynamic waveform.

is a function of the pig's vertical acceleration (a) and its mass (m), expressed as $F_p = ma$.

While the gravitational force remains constant, the vertical component of the self-motion force fluctuates with the pig's movement and requires estimation to accurately determine the pig's weight. The output of the load cell (LC) is the sum of these two forces:

$$LC \propto F_g + F_p = mg + ma. \tag{1}$$

To isolate the gravitational force and obtain an accurate weight measurement, a method was developed to identify the precise moments when the pig's vertical acceleration transitions through zero acceleration ($a = 0$). These zero-crossing points signified a temporary state of equilibrium in the vertical motion of the pig, where the forces acting on the pig in the vertical direction were momentarily balanced. It is important to note that this equilibrium state did not necessarily indicate that the pig was at rest; rather, it represented a point of inflection in the pig's vertical trajectory, where it transitioned from an upward to a downward movement or vice versa. At these points, the load cell output simplified to:

$$LC_{a=0} \propto F_g = mg. \tag{2}$$

We identified moments when the pig's vertical acceleration was zero and recorded the load cell output at these points. These readings, which primarily reflect the pig's true mass under gravity with minimal influence from movement, formed the basis for Series 5 in Fig. 4. Fig. 7 provides a visual representation of this process, illustrating the identification of zero-crossing points in the smoothed acceleration data and the corresponding load cell values at those specific moments.

The identification of zero-acceleration moments required careful signal processing to account for load cell response characteristics. We first smoothed the raw acceleration data to reduce noise, then analyzed the smoothed data to identify zero-crossing points. At these points, we recorded the corresponding load cell output. This approach helped minimize the impact of load cell response delays and potential S–G filter bias (Piskorowski and Barcinski, 2008).

2.4. Feature extraction

Feature extraction was conducted on the dynamic weight time series and the zero-acceleration moment weight series using a combination of

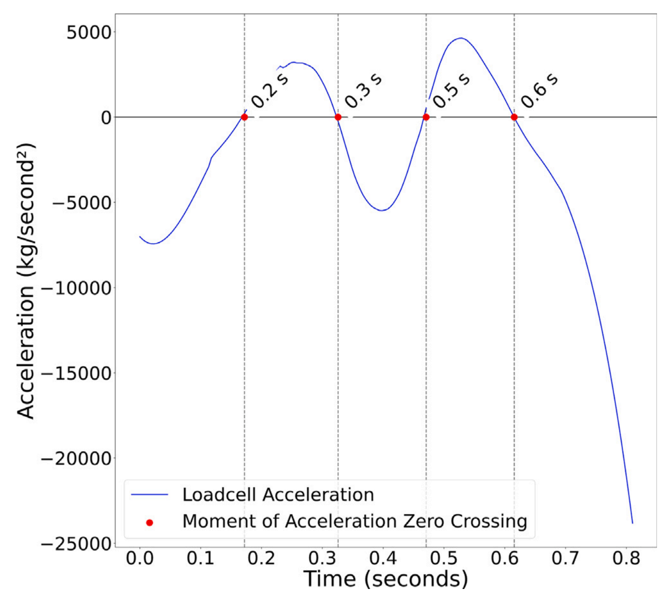


Fig. 6. Loadcell acceleration data showing moments of zero-crossing (red points) where vertical acceleration equals zero. These moments mark key inflection points in the pig's motion on the dynamic scale.

automated techniques and targeted statistical measures, respectively. For the primary time series data, which included both the whole and segmented weight measurements in their raw and smoothed forms, we employed the Python Time Series Feature Extraction Library (TSFEL) developed by Barandas et al. (2020). The overall feature extraction workflow is illustrated in Fig. 4, where dynamic weights are processed at multiple levels to extract a wide range of features.

The TSFEL library initially generated 1556 features across several domains. Statistical measures (e.g., mean, standard deviation) described the central tendencies and variability of the weight data. Temporal features (e.g., autocorrelation) captured time-dependent patterns associated with the pig's movement. Spectral features, such as dominant frequency, were derived to characterize cyclical patterns in the dynamic weight data. Fractal features, including the Hurst exponent, were computed to quantify long-term dependencies within the time series. As depicted in Fig. 4 (Steps 1–4), features were extracted

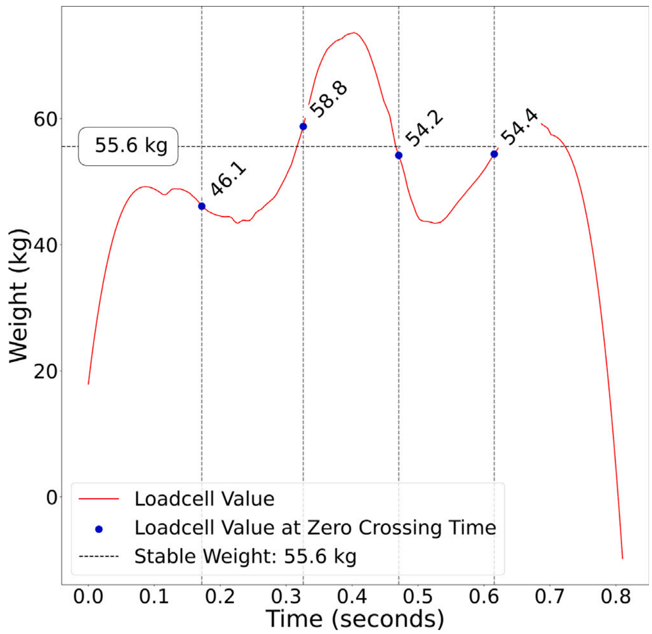


Fig. 7. Weight data corresponding to zero-acceleration crossings (blue points) plotted alongside the stable weight (dotted line). This visualization highlights the transition moments during the pig's walk-through.

from the whole waveform and from the segmented stable plateaus, including both raw and Savitzky–Golay (S–G) smoothed forms.

For the zero-acceleration moment weights series, feature extraction focused exclusively on 13 additional statistical features, including the mean, median, standard deviation, skewness, kurtosis, interquartile range (IQR), and percentiles. Unlike the dynamic weight time series, temporal and spectral features were not applicable to this dataset, as it consists of discrete weight measurements captured at specific equilibrium points (Step 5 in Fig. 4).

2.5. Feature reduction and normalization

The combined 1,569 features (1,556 from TSFEL and 13 from zero-acceleration moment weights) underwent a systematic reduction process to ensure consistency across all pig walk-through events. As shown in Fig. 4, dynamic weight data inherently contain variability due to pig movement, requiring robust preprocessing.

First, 528 features were removed because they were not consistently generated across all events, primarily affecting higher-order Fast Fourier Transform (FFT) and wavelet coefficients, which were sensitive to differences in signal quality or duration. Next, 287 features containing 'NaN' values were excluded. These 'NaN' values resulted from mathematical underflow or overflow conditions during computation (Goldberg, 1991). Rather than introducing uncertainty through imputation, these features were entirely removed to preserve data integrity.

After these preprocessing steps, the feature set was reduced to 754 features. To ensure that all remaining features contributed equally to the model's learning process, Z-score normalization was applied. Each feature was standardized by subtracting its mean and dividing by its standard deviation, transforming the values to have zero mean and unit variance (Jayalakshmi and Santhakumaran, 2011).

The final identification and selection of the top 15 most informative features, determined using the ReliefF algorithm, is described in detail in the next section. Although the full set of extracted features included statistical, temporal, spectral, and fractal components, only the most relevant subset contributed to the final model.

Table 2
Model performance metrics.

Metric	Formula[a]
Mean Absolute Error	$\frac{1}{n}\sum_{i=1}^{n}\lvert y_i-\hat{y}_i\rvert$
Mean Squared Error	$\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$
Root Mean Squared Error	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}$
R-squared	$1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}$
Root Mean Squared Log Error	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\log(y_i+1)-\log(\hat{y}_i+1))^2}$
Mean Absolute Percentage Error	$\frac{1}{n}\sum_{i=1}^{n}\left\lvert\frac{y_i-\hat{y}_i}{y_i}\right\rvert\times 100$

[a] Where y_i is the actual value, $\hat{y}_i$ is the predicted value, n is the number of observations, and $\bar{y}$ is the mean of the actual values.

2.6. Model development process

The model development process consisted of three main phases: model type selection, feature reduction, and hyperparameter tuning (Fig. 8), which is aligned with best practices in machine learning model development for complex regression tasks (Kuhn and Johnson, 2013). The data comprised 944 individual pig walk-throughs and 754 features, and was split pig-wise using an 80:20 ratio to create the training and validation sets, respectively.

The first step in our analysis involved an exploratory search to identify the most suitable model architecture for the pig weight estimation task. This process was conducted using the entire preprocessed dataset, including all 754 features derived from the dynamic scale measurements, with PyCaret's autoML (Ali, 2020). We evaluated a diverse range of regression models, including linear models, tree-based ensembles, and boosting algorithms, similar to recent studies on livestock weight estimation (Kashiha et al., 2014). The performance of each model was assessed using 10-fold cross-validation (Kohavi, 1995).

To identify the most informative subset of features, we employed the ReliefF algorithm, which is effective in handling complex, non-linear relationships between predictors and target variables (Urbanowicz et al., 2018). The ReliefF algorithm assigns a score to each feature, reflecting its relevance to the target variable. This scores ranges between 0 and 1, where higher scores indicate features that are more strongly associated with accurate predictions. ReliefF operates by iteratively comparing feature values for similar and dissimilar instances in the dataset, determining how well each feature can distinguish between them. This approach is particularly useful in our context, as it accounts for potential feature interactions, which are prevalent given the variability in pig movement on the scale (Kononenko, 1994).

Once the model type and features were selected, hyperparameter optimization was performed using PyCaret's tuning capabilities. We employed a systematic grid search across 1000 iterations, exploring a wide parameter space to identify the optimal configuration. This process was crucial for maximizing model performance while maintaining generalizability (Bergstra et al., 2012).

2.7. Performance metrics

In assessing the model's performance, we explored various metrics, each offering a unique perspective on the accuracy of the predictions. These metrics included Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), R-squared (R^2), Root Mean Squared Log Error (RMSLE), and Mean Absolute Percentage Error (MAPE). The formulas for these metrics are presented in Table 2. [1]

[1] Where y_i is the actual value, $\hat{y}_i$ is the predicted value, n is the number of observations, and $\bar{y}$ is the mean of the actual values.

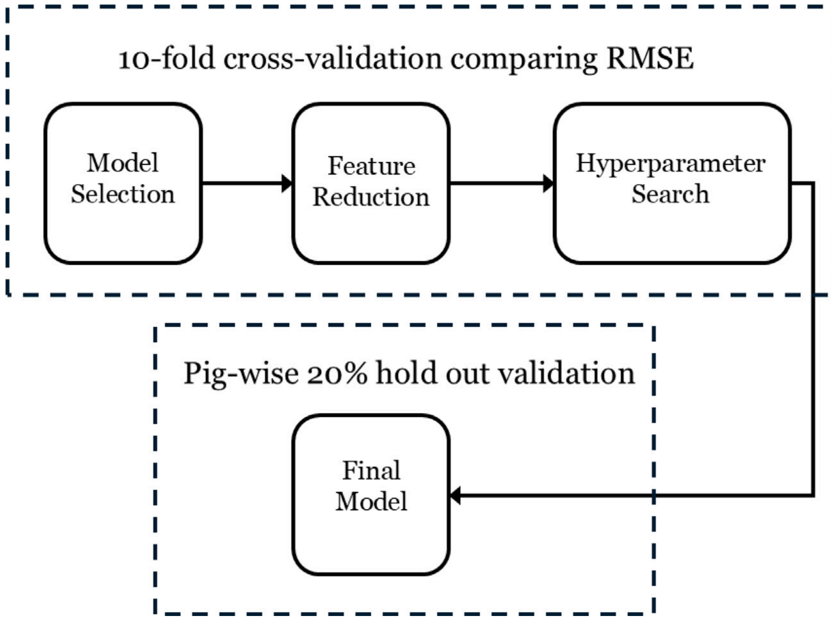


Fig. 8. Flow diagram illustrating the model development process, from initial data preprocessing through to final model validation.

3. Results and discussion

3.1. Analysis of dynamic scale measurements

In our study, we collected data from 944 individual pig walk-throughs on a dynamic weighing scale, revealing a diverse range of movement patterns that posed challenges for developing a universal weight estimation method. The loadcell waveforms in Fig. 9 illustrate the complexity of pig movements on the scale. For example, rapid movements produced short, intense spikes in the weight readings, while hesitation resulted in prolonged partial weight readings before full body weight was registered. These movement patterns are consistent with findings from other studies on pig behavior during weighing procedures (Kashiha et al., 2014; Liu et al., 2023a). For instance, Yu et al. (2021) noted a similar variability in pig behavior described as image blurring during automated weighing of unrestrained pigs.

Traditional approaches to dynamic weighing often discard data from the entry and exit portions of the scale, as well as readings affected by sudden movements like running or jumping (Kashiha et al., 2014). However, our study uniquely retained this diverse data and employed different techniques to correlate dynamic weight patterns with stable weights. We implemented a data processing approach including the application of a 3rd order Savitzky–Golay filter to smooth the loadcell and acceleration readings. This filter was particularly suitable for our application as it effectively reduced high-frequency noise while preserving the underlying trends and patterns in the pig's movement (Schafer, 2011).

An important observation from our analysis was the variability of zero-acceleration moment weights during a pig's movement across the scale. These zero-crossings, which signify equilibrium points in the pig's motion, occurred at irregular intervals due to differences in individual movement patterns, such as walking, trotting, or jumping. This variability posed challenges in accurately identifying and extracting these critical moments. Additionally, the inherent response time of the load cell introduced minor delays between weight changes and recorded outputs, potentially affecting the precision of these measurements.

To mitigate these challenges, the use of the Savitzky–Golay filter proved essential in reducing noise while preserving the integrity of the signal. By smoothing the raw acceleration data and pinpointing zero-crossing events with enhanced accuracy, we minimized the effects of noise and potential bias introduced by the filtering process. This approach allowed us to reliably capture key moments in the pigs' movement, reflecting the combined force of gravity and motion, and provided a robust basis for weight estimation. The findings underscore the importance of advanced signal processing techniques in overcoming the challenges posed by dynamic, real-world conditions, ensuring reliable weight predictions even under highly variable movement patterns.

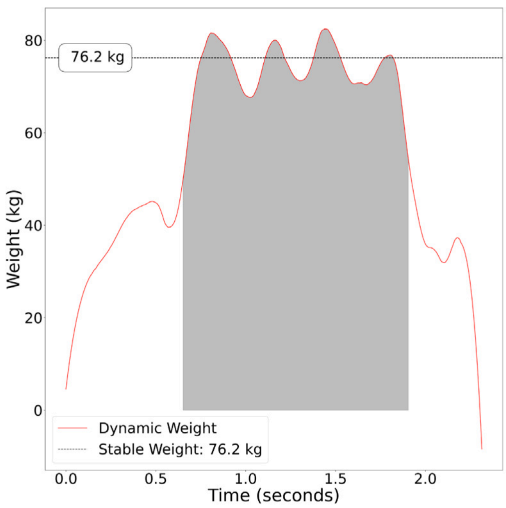
3.2. Model type selection

We assessed and compared 18 regression models types using the initial set of 754 features, encompassing a diverse range of approaches, including linear models, tree-based ensembles, and boosting algorithms. This approach aligns with methodologies employed in recent studies on livestock weight estimation (Kashiha et al., 2014).

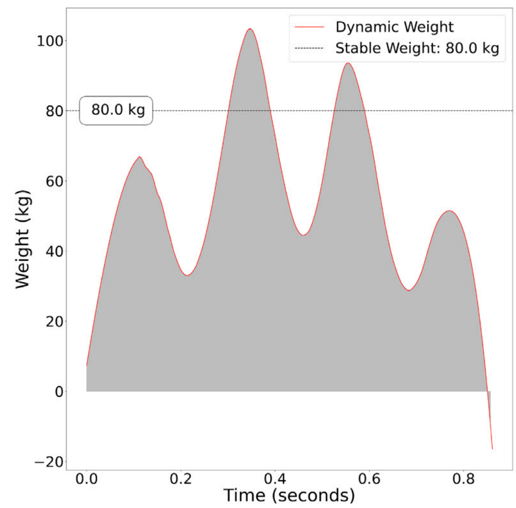
The Gradient Boosting Regressor (GBR) consistently outperformed other models in our study, achieving the lowest RMSE of 3.55 kg (see Table 3). While models like the Extra Trees Regressor demonstrated better performance when evaluated solely using the MAE, RMSE was chosen as the primary metric for model selection due to the financial implications associated with large prediction errors in this problem set.

Specifically, in the context of maximizing pig shipment revenue and adhering to scoring grids, small errors in weight estimation may have minimal financial impact, but larger errors can lead to significant revenue losses by failing to meet weight thresholds or incorrectly classifying pigs into the wrong scoring categories. RMSE, which penalizes larger errors more heavily than MAE, aligns with the need to minimize the likelihood of extreme deviations that could result in substantial financial losses. By prioritizing RMSE as the selection criterion, we ensure that the model optimizes predictions in a way that reduces the overall risk of such costly misclassifications. This preference for RMSE over MAE reflects the practical requirements of the application, where robust accuracy is critical for economic outcomes.

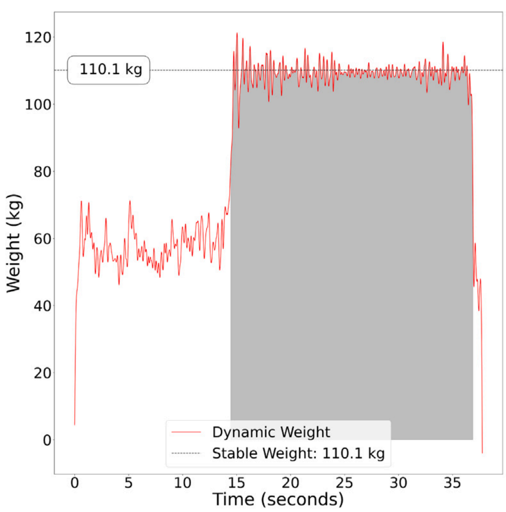
This performance of GBR aligns with findings from other complex regression tasks in agricultural settings. For instance, Morota et al. (2018) highlighted the effectiveness of ensemble methods like gradient boosting in handling complex, high-dimensional data in precision animal agriculture. The effectiveness of GBR can be attributed to its ability to capture non-linear relationships and handle interactions between features, which is particularly relevant given the complex nature of pig movements on the dynamic scale (Bentéjac et al., 2021). This capability



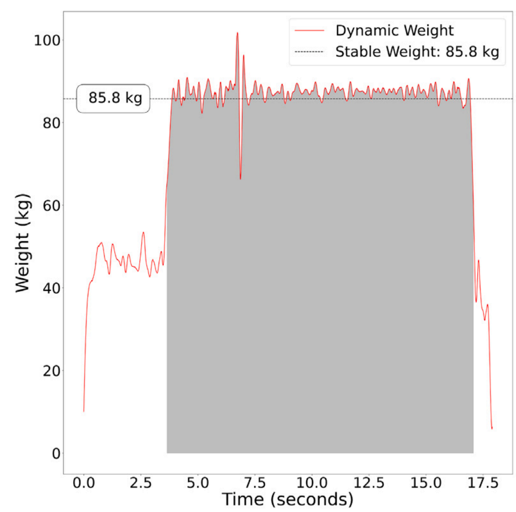
(a) Continuous forward movement.



(b) Running or trotting.



(c) Half weight from front hooves followed by full weight and lingering.



(d) Abrupt movements and jumping.

Fig. 9. Examples of movements observed on the dynamic scale: (a) continuous forward movement, (b) running or trotting, (c) half weight from front hooves followed by full weight and lingering, and (d) abrupt movements and jumping.

is especially valuable in animal science applications, where multiple interacting factors influence outcomes. Based on these results, the GBR was selected as the base model for further optimization in subsequent stages of the model development process.

To determine the optimal number of features, we plotted the RMSE of the Gradient Boosting Regressor against the number of features selected by ReliefF as seen in Fig. 10.

The elbow method was used to select the number of features for the final model. This technique involves plotting the model's error metric (e.g., RMSE) against the number of features and identifying the point where the rate of error reduction diminishes—commonly referred to as the "elbow". In our analysis, the first inflection point was observed near 15 features, beyond which additional features contributed less improvements to model performance. This guided our decision to select the top 15 features, balancing computational efficiency and predictive accuracy.

Table 4 presents the selected features along with their ReliefF scores, which quantify the importance of each feature for weight prediction. The scores range from 0 to 1, with higher values indicating greater relevance. For instance, the Maximum Zero-Acceleration Moment Weight feature received a score of 0.823, signifying its high predictive value, while the Median Segmented Weights feature scored 0.670, indicating relatively lower but still significant importance.

The selected features encompass a diverse range of measurements. Notably, features derived from the zero-acceleration moment weights were among the selected features. This aligns with our hypothesis that these points, where the pig's vertical acceleration momentarily reaches zero, provide particularly informative snapshots of the animal's true weight.

By reducing the number of input features from 754 to 15, we have streamlined the model's structure and improved interpretability. This feature reduction likely simplifies certain aspects of the computational pipeline (e.g., data storage, processing steps, and model complexity), potentially lowering the computational overhead associated with running the model. In principle, a more parsimonious model could require fewer processing resources, which might ease integration into existing farm management systems, especially those with limited computational infrastructure.

However, it is important to emphasize that we have not conducted a formal analysis to quantify how much computational cost or run-time is actually saved by using fewer features, nor have we directly assessed the potential savings on equipment or maintenance costs that may arise

Table 3
The performance metrics for the top-performing models using 10-fold cross-validation.

Model	Mean Absolute Error	Mean Squared Error	Root Mean Squared Error	R-Squared Error	Root Mean Squared Log Error	Mean Absolute Percentage Error
Gradient Boosting Regressor	2.24	13.94	3.55	0.99	0.04	0.03
Light Gradient Boosting Machine	2.23	14.05	3.56	0.99	0.04	0.03
Extreme Gradient Boosting	2.38	15.22	3.71	0.98	0.05	0.03
Lasso Regression	2.43	15.28	3.76	0.98	0.05	0.04
Lasso Least Angle Regression	2.43	15.36	3.77	0.98	0.05	0.04
Extra Trees Regressor	2.13	15.67	3.71	0.98	0.05	0.03
CatBoost Regressor	2.49	18.32	4.03	0.98	0.05	0.04
AdaBoost Regressor	3.06	19.25	4.27	0.98	0.07	0.05
Elastic Net	2.65	20.10	4.21	0.98	0.06	0.04
Decision Tree Regressor	3.11	23.47	4.70	0.97	0.06	0.04
Orthogonal Matching Pursuit	2.98	32.29	5.14	0.97	0.06	0.04
Huber Regressor	2.93	46.67	5.71	0.95	0.09	0.04
Bayesian Ridge	3.01	49.18	5.90	0.95	0.08	0.04
Passive Aggressive Regressor	3.42	54.81	6.39	0.94	0.09	0.05
Ridge Regression	3.47	72.03	7.17	0.92	0.10	0.05
K Neighbors Regressor	8.16	126.15	11.07	0.86	0.16	0.12
Dummy Regressor	25.95	947.85	30.69	−0.03	0.47	0.48
Linear Regression	22.35	2357.61	44.06	−1.59	0.45	0.32

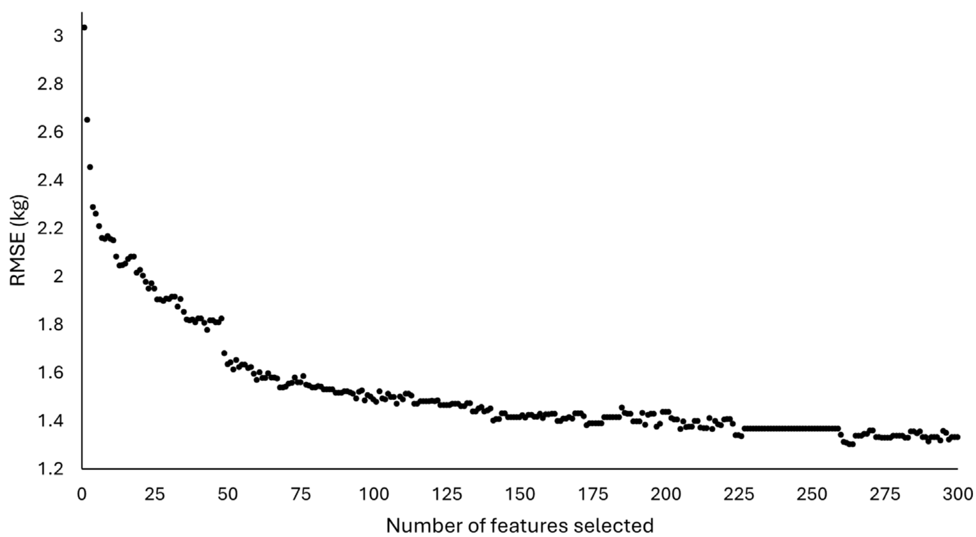


Fig. 10. Root Mean Squared Error (RMSE) of the Gradient Boosting Regressor plotted against the number of features selected by the ReliefF algorithm. The elbow method was applied with 10-fold cross-validation to reduce number of features.

Table 4
Features selected using the ReliefF algorithm. The ReliefF score ranges from 0 to 1, with higher values indicating greater importance in predicting the target variable. The features are ranked by their ReliefF scores in descending order of relevance.

Rank	Feature	Series	ReliefF score
1	Maximum	Zero-Acceleration Moment Weights	0.823
2	Empirical Cumulative Distribution Function (ECDF) Percentile 1	Segmented Smoothed Weights	0.813
3	Empirical Cumulative Distribution Function (ECDF) Percentile 1	Segmented Weights	0.807
4	75th Percentile	Zero-Acceleration Moment Weights	0.802
5	Empirical Cumulative Distribution Function (ECDF) Percentile 1	Weights	0.759
6	Root Mean Square	Segmented Weights	0.746
7	Root Mean Square	Segmented Smoothed Weights	0.743
8	Mean	Segmented Smoothed Weights	0.704
9	Mean	Segmented Weights	0.703
10	Mean	Zero-Acceleration Moment Weights	0.700
11	Max	Segmented Smoothed Weights	0.699
12	Median	Zero-Acceleration Moment Weights	0.692
13	Root Mean Square	Weights	0.677
14	Median	Segmented Smoothed Weights	0.674
15	Median	Segmented Weights	0.670

Table 5
Model tuning results.

Hyperparameter	Value
Number of trees	120
Learning rate	0.05
Maximum tree depth	11
Minimum samples to split a node	2
Minimum samples per leaf node	5
Feature selection criterion	Friedman MSE
Minimum impurity decrease	0.02
Subsample ratio	0.4

Table 6
Model performance metrics.

Metric	Score
Mean Absolute Error (MAE)	1.79 kg
Mean Squared Error (MSE)	8.26 kg^2
Root Mean Squared Error (RMSE)	2.87 kg
R^2	0.9912
Root Mean Squared Logarithmic Error (RMSLE)	0.0385
Mean Absolute Percentage Error (MAPE)	2.65%

from a simpler model structure. For instance, we do not provide metrics comparing the model's inference time or resource consumption when using the full feature set versus the reduced set. Without these data, it remains speculative to assert that our approach would be substantially more cost-effective than using a larger feature set.

Moreover, a deeper understanding of the economic implications of this feature reduction would require additional studies. For example, one could simulate scenarios comparing the running time and server maintenance costs of the full feature set model against the reduced feature set model. Such simulations could also consider the economic impact of the trade-off between a slightly less accurate model and the financial consequences of predicting pig body weight with greater uncertainty—e.g., evaluating the difference between a 1 kg versus a 2 kg error in Body Weight (BW) estimation in terms of feed allocation, shipment timing, and ultimate farm revenue (Pomar and Remus, 2019).

3.3. Hyperparameter tuning

The final optimized model configuration included the following key parameters in Table 5. This configuration balanced model complexity with generalization capability. The relatively low learning rate (0.05) and moderate number of estimators (120) helped prevent overfitting, while the maximum depth of 11 allowed the model to capture complex interactions in the data (Friedman, 2001).

3.4. Model performance

The optimized GBR model demonstrated adequate performance in predicting pig weights from dynamic scale data. The final model achieved an RMSE of 2.87 kg and an R^2 of 0.99 in the unseen 20% holdout pig-wise validation set, stratified by pig ID, as shown in Table 6 and Fig. 11.

The performance of the GBR can be attributed to its ability to capture complex, non-linear relationships in the data. As described by Friedman (2001), gradient boosting constructs additive regression models by sequentially fitting simple base learners to current pseudo-residuals. In our application to pig weight estimation, this allows the model to progressively refine its predictions, effectively learning the intricate patterns of weight distribution associated with various pig movements across the scale. This sequential error-correcting process enables the GBR to adapt to the nuanced relationships between the extracted features and the actual pig weights. While the manufacturer's listed accuracy for our industrial floor scales was 0.23 kg, our model achieved a Mean Absolute Error (MAE) of 1.79 kg across the full range of pig weights (20–150 kg). This difference between theoretical static accuracy and dynamic measurement error is expected and comparable to other studies using load cells for livestock weighing. For example, Alawneh et al. (Alawneh et al., 2011) reported measurement errors of 1.5–2.0 kg for dairy cattle walk-over weighing systems despite using scales with sub-kilogram static accuracy. Similarly, Parsons et al. (Parsons et al., 2023) found errors of 1.3–1.8 kg in their cattle weighing system.

Importantly, our observed error remains well within acceptable limits for commercial pig production, where weight estimates within ±2–3 kg are considered sufficient for management decisions like sorting and shipping (Stygar et al., 2018). The consistency of our error rates across multiple trials suggests that while we may not achieve the theoretical precision of static weighing, our system provides reliable and reproducible measurements suitable for both production settings and research applications.

Our accuracy demonstrates noteworthy improvements over existing automated pig weighing systems reported in recent literature. When examining comparable studies, Liu et al. (2023b)'s vision-based system achieved an RMSE of 12.45 kg (5.36%), while Wang et al. (2008)'s machine vision-based walk-through system reported a relative error of approximately 3%. In comparison, our system achieved an RMSE of 2.87 kg, indicating enhanced precision despite operating under more challenging conditions and across a broader weight range. This improvement becomes particularly significant when considering other benchmark studies. For instance, Hou et al. (2024) evaluated a system for growing-finishing pigs that achieved a MAE of 3.48 kg and an RMSE of 4.43 kg for animals weighing between 60 to 120 kg. Additionally, an assessment of the PigVision camera system by Farmer (2022) revealed varying accuracy levels depending on pig size, with lower accuracy (87.7%) for lighter pigs around 32.7 kg compared to higher accuracy (97.6%) for heavier pigs weighing approximately 117.5 kg. These comparative results provide valuable context for evaluating our system's performance and highlight its robust capability in measuring pig body weights across diverse conditions.

The model's performance can be attributed to its use of a diverse set of features. While features derived from zero-acceleration moment weights proved particularly informative, they represent only 4 out of the 15 features used in the final model. These zero-acceleration features correspond to moments when the pig's vertical acceleration is momentarily zero, providing crucial information about the animal's true weight with minimal influence from dynamic movement.

However, the model's effectiveness also stems from its incorporation of other key features. These include statistical measures from segmented weight data, such as root mean square and mean values, as well as features derived from the entire weight time series. This combination of features allows the model to capture both the static weight information and the dynamic aspects of the pig's movement across the scale.

3.5. Analysis of residuals

Examination of the model's residuals reveals a symmetrical distribution centered around zero, with no apparent systematic bias, as seen in Fig. 12(a). This suggests that the model is not consistently over- or under-predicting weights across the range of measurements.

The lack of clear patterns in a scatter plot of residuals against predicted values (Fig. 12(b)) further supports the model's consistent performance across the weight range. This consistency is a crucial factor for practical application in diverse farm settings.

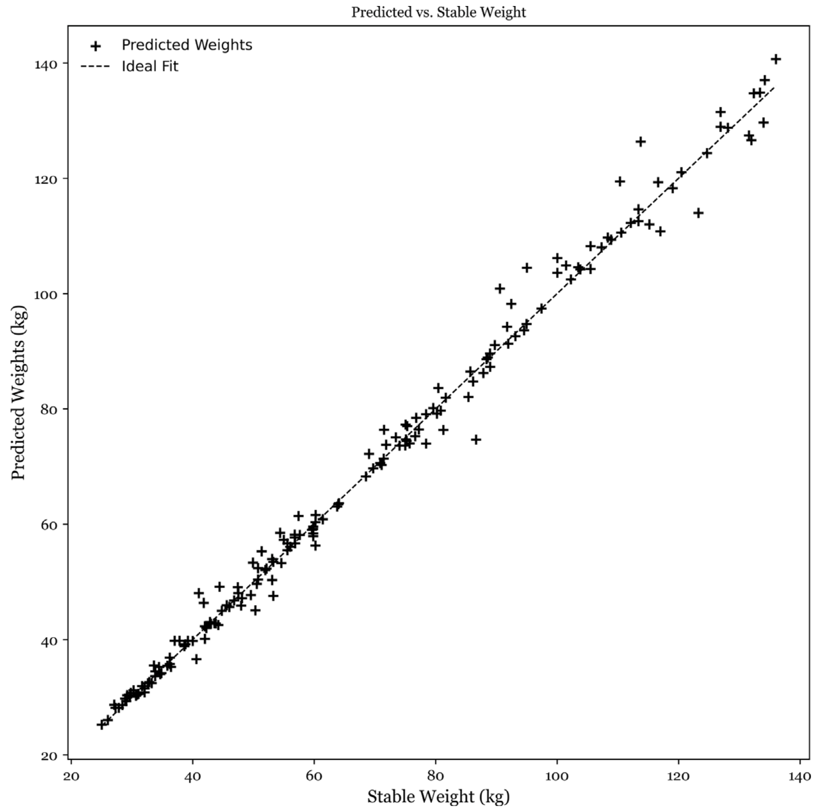


Fig. 11. Performance of the GBR model on the validation set: predicted versus actual weights.

4. Key contributions

Multiple automated walk-over weighing systems have been developed for livestock species such as sheep, beef, and dairy cattle, focusing primarily on animals in grazing or semi-controlled environments. While these systems have proven effective, our study addresses the unique challenges posed by pig production systems, where animals are housed in confined environments and exhibit erratic and unpredictable movement patterns. Prior work on automated pig weighing has often relied on vision-based methods or setups that involve some level of static positioning or specific camera configurations (Son et al., 2022; Mattina et al., 2024; Riekert et al., 2020). In contrast, our approach is tailored to the dynamic, free-motion conditions of finishing pigs in barn environments, providing a methodology that does not depend on constrained animal behavior or controlled conditions.

The novelty of our contribution lies in demonstrating that a walk-over scale system, when combined with high-frequency load cell data, feature engineering—including adaptive segmentation techniques and zero-acceleration moment features—and a systematic hyperparameter search, can achieve precise weight estimation in motion. This approach extends beyond existing solutions that operate effectively only under simpler conditions, involve limited methodological transparency, or rely on costly proprietary equipment(Gómez et al., 2021). Furthermore, our system serves as more than a standalone tool—it is a foundational enabler for future research.

Our methodology provides accurate weight measurements dynamically captured at speed, forming a robust data pipeline capable of generating high-quality labeled instances. By anchoring these datasets with consistent, validated weight labels, the system offers a reliable foundation for developing and refining precision livestock farming tools. These tools can drive advancements in feed planning, health monitoring, and shipment scheduling. By addressing the limitations of traditional systems, this methodology represents a scalable, transparent framework for improving livestock systems and practices. It complements existing approaches while enabling novel applications in integrated farming systems, establishing a foundation for future innovations in non-invasive weight estimation and efficient livestock management.

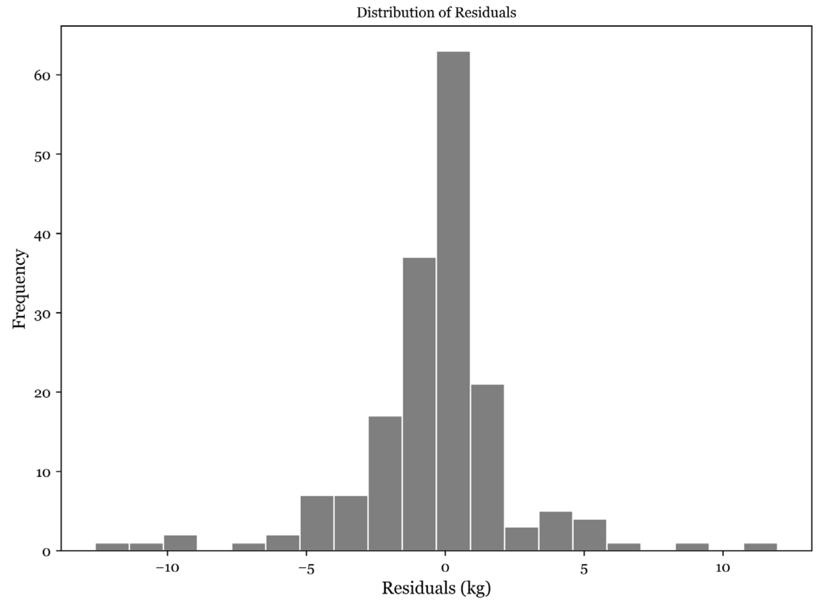
4.1. Future applications and research

Our current system is a prototype aimed at improving weight monitoring processes in commercial finishing environments. Existing systems, such as automated sorting technologies using static, constrained scales (Stygar et al., 2018), have effectively demonstrated the value of frequent body weight (BW) monitoring and the importance of accounting for diurnal weight variation. However, these systems rely on pigs passing through fixed, controlled weighing stations, which limits their ability to capture dynamic and natural weight fluctuations during free movement.

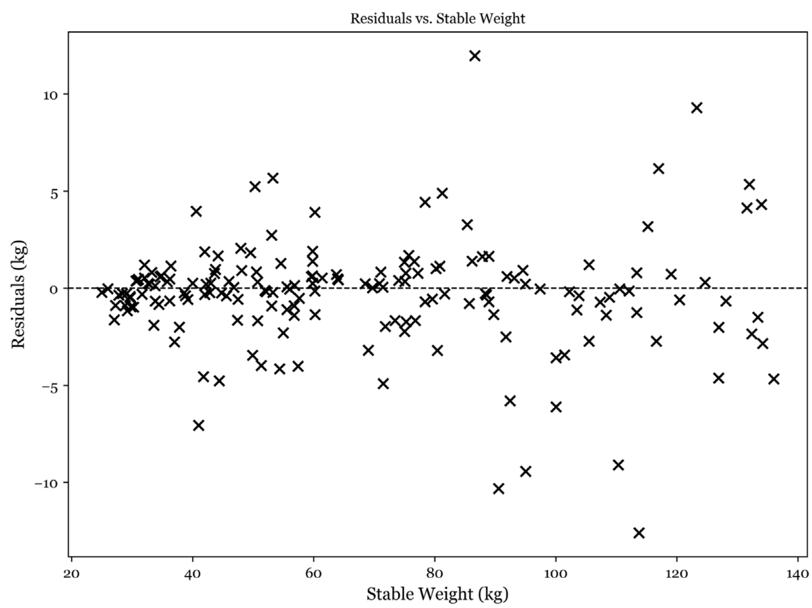
Our dynamic walk-through scale addresses these constraints by enabling continuous, unconstrained monitoring of pigs' body weights during natural movements. This system eliminates the need for static sorting scales, captures a broader range of weight dynamics—including diverse motion patterns—and maintains precision through advanced filtering and segmentation techniques. The system could be integrated with RFID readers for pig identification and could generate high-speed weight measurements with unique IDs. This capability, if used in parallel, could accelerate the development of labeled datasets, which are essential for validating and improving future advancements in precision livestock farming technologies.

Beyond computational efficiency and accuracy, real-world adoption also depends on equipment costs, installation, and maintenance, factors that were reported as barriers for smaller operations implementing Precision Livestock Farming (PLF) tools (Karunathilake et al., 2023).

While simplifying the model may be a step toward making the system more manageable and potentially more economical, we currently lack empirical evidence to substantiate any claims of reduced



(a) Histogram of residuals showing symmetrical distribution around zero.



(b) Scatter plot of residuals against predicted values showing no systematic bias.

Fig. 12. Analysis of residuals from the final Gradient Boosting Regressor model.

real-world operating costs or broad-scale efficiency improvements. Further research—including computational benchmarks, economic simulations, and assessments of technology readiness—would be necessary to draw meaningful conclusions about the economic benefits of feature reduction in our system.

During data collection trials, there was a higher frequency of measurements for smaller pigs compared to larger ones, not due to behavioral differences but because the time commitment required by researchers decreased toward the end of trials. This imbalance highlights the need for more consistent sampling strategies in future studies to enhance model generalizability. Additionally, variance in body weight measurements increased over time, particularly for larger pigs, as shown in Fig. 12 and consistent with findings from Gómez et al. (2021). Addressing this natural variability may require the collection of more balanced and extensive datasets to ensure even representation across growth phases. Additionally, techniques such as Synthetic Minority Over-sampling Technique (SMOTE) could be explored as complementary approaches to enhance model robustness and accuracy when real-world data collection is constrained.

By providing these datasets, our system lays the groundwork for predictive models that refine growth projections and improve feed planning and market timing (Taylor et al., 2023). Furthermore, integrating these predictive capabilities with decision-support tools aids in determining shipment schedules, resource allocation, and market strategies while addressing practical constraints such as transport costs and on-farm logistics (Zong and Guan, 2024). Importantly, this method reduces the need for frequent, labor-intensive weighing events, minimizing stress on animals and streamlining farm operations (Kapun et al., 2023).

Although commercial adoption faces economic challenges such as the high costs of sensor installations and pig identification systems prioritizing cost-effective, scalable solutions that integrate seamlessly with existing farm infrastructures is essential. Moreover, the practical placement of the scale in high-traffic areas, such as alleys leading to feeding or sorting zones, ensures better use, while its design accommodates frequent cleaning and maintenance to support reliable operation during trials.

Finally, future research should explore integration with computer vision techniques, such as image segmentation and regression-based video analysis, enabling simultaneous monitoring of multiple pigs. This advancement significantly increases throughput and enhances the system's feasibility for large-scale commercial use, while maintaining its role as a research-driven foundation for further technological development.

5. Conclusion

This study developed and validated an automated walk-over scale system for estimating pig weights in motion. The system used high-frequency load cell data and machine learning techniques to address the challenges associated with dynamic weighing of pigs. Through a combination of signal processing, feature engineering, and model optimization, we achieved a RMSE of 2.87 kg, R^2 of 0.99, and a MAPE of 2.65% on a holdout validation set.

The Gradient Boosting Regressor model was selected as the best algorithm, demonstrating robust performance across a range of pig weights and movement patterns. Feature selection using the ReliefF algorithm identified 15 key features, composed of both zero-acceleration moment weights and segmented weight features.

While the current system is limited to weighing one pig at a time, it represents a step toward more efficient and less stressful weight monitoring in swine production. Future work should focus on expanding the system's capabilities to handle multiple pigs simultaneously and integrating it with broader farm management systems.

CRediT authorship contribution statement

François Decarie: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Charles Grant:** Supervision, Project administration, Investigation. **Gabriel Dallago:** Writing – review & editing, Supervision, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors thank the staff at Pioneer Meat for sharing their expertise. We also acknowledge the invaluable assistance of the Glenlea Research Station staff, including Jason Bourcier, Don Chaput, Archimedes Isit, and Sam Ariyib, in facilitating data collection and providing care to our pigs.

Data availability

The authors do not have permission to share data.

References

Alawneh, J.I., Stevenson, M.A., Williamson, N.B., Lopez-Villalobos, N., Otley, T., 2011. Automatic recording of daily walkover liveweight of dairy cattle at pasture in the first 100 days in milk. J. Dairy Sci. 94, 4431–4440. http://dx.doi.org/10.3168/jds.2010-4002.

Ali, M., 2020. PyCaret: An open source, low-code machine learning library in Python.

Banhazi, T., Dunn, M., Banhazi, A., 2022. Weight-detect™: On-farm evaluation of the precision of image analysis based weight prediction system. In: Banhazi, T., Halas, V., Maroto-Molina, F. (Eds.), Practical Precision Livestock Farming. Wageningen Academic Publishers, pp. 29–38. http://dx.doi.org/10.3920/978-90-8686-934-3_1.

Banhazi, T., Lehr, H., Black, J., Crabtree, H., Schofield, C., Tscharke, M., Berckmans, D., 2012. Precision livestock farming: An international review of scientific and commercial aspects. Int. J. Agric. Biol. Eng. 5, 1–9. http://dx.doi.org/10.3965/j.ijabe.20120503.00?.

Barandas, M., Folgado, D., Fernandes, L., Santos, S., Abreu, M., Bota, P., Liu, H., Schultz, T., Gamboa, H., 2020. TSFEL: Time series feature extraction library. SoftwareX 11, 100456. http://dx.doi.org/10.1016/j.softx.2020.100456, URL https://www.sciencedirect.com/science/article/pii/S2352711020300017.

Bentéjac, C., Csörgő, A., Martínez-Muñoz, G., 2021. A comparative analysis of gradient boosting algorithms. Artif. Intell. Rev. 54, 1937–1967. http://dx.doi.org/10.1007/s10462-020-09896-5.

Bergstra, J., Ca, J.B., Ca, Y.B., 2012. Random search for hyper-parameter optimization yoshua bengio. J. Mach. Learn. Res. 13, 281–305, URL http://scikit-learn.sourceforge.net.

Condotta, I.C., Brown-Brandl, T.M., Silva-Miranda, K.O., Stinn, J.P., 2018. Evaluation of a depth sensor for mass estimation of growing and finishing pigs. Biosyst. Eng. 173, 11–18. http://dx.doi.org/10.1016/J.BIOSYSTEMSENG.2018.03.002.

Conte, S., Bergeron, R., Gonyou, H., Brown, J., Rioja-Lang, F.C., Connor, L., Devillers, N., 2014. Measure and characterization of lameness in gestating sows using force plate, kinematic, and accelerometer methods1. J. Anim. Sci. 92, 5693–5703. http://dx.doi.org/10.2527/jas.2014-7865.

Farmer, N.H., 2022. Commercial camera solutions for weighing finishing pigs automatically. Natl. Hog Farmer URL https://www.nationalhogfarmer.com/hog-health/commercial-camera-solutions-for-weighing-finishing-pigs-automatically.

Friedman, J.H., 2001. 999 Reitz lecture greedy function approximation: A gradient boosting machine 1. Ann. Statist. 29, 1189–1232.

Fryzlewicz, P., 2014. Wild binary segmentation for multiple change-point detection. Ann. Stat. 42, 2243–2281. http://dx.doi.org/10.1214/14-AOS1245.

Goldberg, D., 1991. What every computer scientist should know about floating-point arithmetic. ACM Comput. Surv. 23 (1), 5–48, Available online at the ACM Digital Library URL: https://doi.org/10.1145/103162.103163.

Gómez, Y., Stygar, A.H., Boumans, I.J.M.M., Bokkers, E.A.M., Pedersen, L.J., Niemi, J.K., Pastell, M., Manteca, X., Llonch, P., 2021. A systematic review on validated precision livestock farming technologies for pig production and its potential to assess animal welfare. Front. Vet. Sci. 8, 660565. http://dx.doi.org/10.3389/fvets.2021.660565, URL: https://doi.org/10.3389/fvets.2021.660565.

Grandin, T., 1988. Environmental enrichment for confinement pigs. In: Proceedings of the Livestock Conservation Institute Annual Meeting. Livestock Conservation Institute, Kansas City, Missouri, pp. 119–123, Available at: https://www.grandin.com/references/LCIhand.html.

Grandin, T., 1997. Assessment of stress during handling and transport. J. Anim. Sci. 75, 249–257. http://dx.doi.org/10.2527/1997.751249x.

Hou, G., Li, R., Tian, M., Ding, J., Zhang, X., Yang, B., Chen, C., Huang, R., Yin, Y., 2024. Improving efficiency: Automatic intelligent weighing system as a replacement for manual pig weighing. Animals 14 (11), 1614, URL https://www.mdpi.com/2076-2615/14/11/1614.

Jayalakshmi, T., Santhakumaran, A., 2011. Statistical normalization and back propagation for classification. Int. J. Comput. Theory Eng. 89–93. http://dx.doi.org/10.7763/IJCTE.2011.V3.288.

Kapun, A., Adrion, F., Gallmann, E., 2023. Evaluating the activity of pigs with radio-frequency identification and virtual walking distances. Animals 13, http://dx.doi.org/10.3390/ani13193112.

Karunathilake, E.M., Le, A.T., Heo, S., Chung, Y.S., Mansoor, S., 2023. The path to smart farming: Innovations and opportunities in precision agriculture. Agric. (Switzerland) 13, http://dx.doi.org/10.3390/agriculture13081593.

Kashiha, M., Bahr, C., Ott, S., Moons, C.P., Niewold, T.A., Ödberg, F.O., Berckmans, D., 2014. Automatic weight estimation of individual pigs using image analysis. Comput. Electron. Agric. 107, 38–44. http://dx.doi.org/10.1016/J.COMPAG.2014.06.003.

Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. URL http://roboticsStanfordedu/"ronnyk.

Kononenko, I., 1994. Estimating attributes: Analysis and extensions of RELIEF. pp. 171–182. http://dx.doi.org/10.1007/3-540-57868-4_57.

Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer Link, URL https://link.springer.com/book/10.1007/978-1-4614-6849-3.

L'Heureux, A., Grolinger, K., Elyamany, H.F., Capretz, M.A.M., 2017. Machine learning with big data: Challenges and approaches. IEEE Access 5, 7776–7797. http://dx.doi.org/10.1109/ACCESS.2017.2696365, URL https://ieeexplore.ieee.org/document/7906512.

Liu, G., Guo, H., Ruchay, A., Pezzuolo, A., 2023a. Recent advancements in precision livestock farming. Agric. (Switzerland) 13, http://dx.doi.org/10.3390/agriculture13091652.

Liu, Z., Hua, J., Xue, H., Tian, H., Chen, Y., Liu, H., 2023b. Body weight estimation for pigs based on 3D hybrid filter and convolutional neural network. Sensors 23, http://dx.doi.org/10.3390/s23187730.

Mardhati, M., González, L.A., Thomson, P.C., Clark, C.E., García, S.C., 2021. Short-term liveweight changes of dairy cows measured by stationary and walk-over weighing scales. J. Dairy Sci. 104, 8202–8213. http://dx.doi.org/10.3168/jds.2020-19912.

Mattina, M., Benzinou, A., Nasreddine, K., Richard, F., 2024. An efficient center-based method for real-time pig posture recognition and tracking. Appl. Intell. 54, 5183–5196. http://dx.doi.org/10.1007/s10489-024-05439-5, URL https://link.springer.com/article/10.1007/s10489-024-05439-5.

Morota, G., Ventura, R.V., Silva, F.F., Koyama, M., Fernando, S.C., 2018. Big data analytics and precision animal agriculture symposium: Machine learning and data mining advance predictive big data analysis in precision animal agriculture. J. Anim. Sci. 96, 1540–1550. http://dx.doi.org/10.1093/jas/sky014.

National Research Council, 2012. Nutrient Requirements of Swine: Eleventh Revised Edition. The National Academies Press, Washington, DC, http://dx.doi.org/10.17226/13298, URL: https://doi.org/10.17226/13298.

OECD-FAO, 2023. OECD-FAO agricultural outlook 2023–2032. URL: https://doi.org/10.1787/08801ab7-en.

Parsons, I., Boudreau, M.R., Karisch, B.B., Stone, A.E., Smith, D.R., 2023. Automated walk-over-weighing system: methods to track daily body mass and growth in grazing steers. J. Anim. Sci. 101 (Suppl. 1), 85.

Paudel, S., de Sousa, R.V., Sharma, S.R., Brown-Brandl, T., 2024. Deep learning models to predict finishing pig weight using point clouds. Animals 14, http://dx.doi.org/10.3390/ani14010031.

Piskorowski, J., Barcinski, T., 2008. Dynamic compensation of load cell response: A time-varying approach. Mech. Syst. Signal Process. 22, 1694–1704. http://dx.doi.org/10.1016/j.ymssp.2008.01.001.

Pomar, C., Remus, A., 2019. Precision pig feeding: A breakthrough toward sustainability. Anim. Front. 9, 52–59. http://dx.doi.org/10.1093/af/vfz006.

Riekert, M., Klein, A., Adrion, F., Hoffmann, C., Gallmann, E., 2020. Automatically detecting pig position and posture by 2D camera imaging and deep learning. Comput. Electron. Agric. 174, 105391. http://dx.doi.org/10.1016/j.compag.2020.105391, URL https://www.academia.edu/99958836/Automatically_detecting_pig_position_and_posture_by_2D_camera_imaging_and_deep_learning.

Savitzky, A., Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures.. Anal. Chem. 36, 1627–1639. http://dx.doi.org/10.1021/ac60214a047.

Schafer, R.W., 2011. What is a savitzky-golay filter? IEEE Signal Process. Mag. 28, 111–117. http://dx.doi.org/10.1109/MSP.2011.941097.

See, M.T., 2024. — invited review — Current status and future trends for pork production in the United States of America and Canada. Anim. Biosci. 37, 775–785. http://dx.doi.org/10.5713/ab.24.0055.

Shi, C., Teng, G., Li, Z., 2016. An approach of pig weight estimation using binocular stereo system based on LabVIEW. Comput. Electron. Agric. 129, 37–43. http://dx.doi.org/10.1016/j.compag.2016.08.012.

Slader, R.W., Gregory, A.M.S., 1988. An automatic feeding and weighing system for ad libitum fed pigs. Comput. Electron. Agric. 3, 157–170.

Smith, R., Turner, M., 1974. Electronic aids for use in fastock weighing. J. Agric. Eng. Res. 19, 299–311, URL https://www.sciencedirect.com/science/article/pii/0021863474900699.

Son, S., Ahn, H., Baek, H., Yu, S., Suh, Y., Lee, S., Chung, Y., Park, D., 2022. StaticPigDet: Accuracy improvement of static camera-based pig monitoring using background and facility information. Sensors 22 (21), 8315. http://dx.doi.org/10.3390/s22218315, URL https://www.mdpi.com/1424-8220/22/21/8315.

Stygar, A.H., Dolecheck, K.A., Kristensen, A.R., 2018. Analyses of body weight patterns in growing pigs: A new view on body weight in pigs for frequent monitoring. Animal 12, 295–302. http://dx.doi.org/10.1017/S1751731117001690.

Taylor, C., Guy, J., Bacardit, J., 2023. Estimating individual-level pig growth trajectories from group-level weight time series using machine learning. Comput. Electron. Agric. 208, http://dx.doi.org/10.1016/j.compag.2023.107790.

Truong, C., Oudre, L., Vayatis, N., 2018. Selective review of offline change point detection methods. Signal Process. 153, 278–299. http://dx.doi.org/10.1016/j.sigpro.2019.107299.

Urbanowicz, R.J., Meeker, M., Cava, W.L., Olson, R.S., Moore, J.H., 2018. Relief-based feature selection: Introduction and review. J. Biomed. Inform. 85, 189–203. http://dx.doi.org/10.1016/j.jbi.2018.07.014.

Vranken, E., Berckmans, D., 2017. Precision livestock farming for pigs. Anim. Front. 7, 32–37. http://dx.doi.org/10.2527/af.2017.0106.

Wang, Z., Li, Q., Yu, Q., Qian, W., Gao, R., Wang, R., Wu, T., Li, X., 2024. A review of visual estimation research on live pig weight. Sensors 24 (21), 7093. http://dx.doi.org/10.3390/s24217093, URL https://www.mdpi.com/1424-8220/24/21/7093.

Wang, Y., Yang, W., Winter, P., Walker, L., 2008. Walk-through weighing of pigs using machine vision and an artificial neural network. Biosyst. Eng. 100, 117–125. http://dx.doi.org/10.1016/j.biosystemseng.2007.08.008.

Yu, H., Lee, K., Morota, G., 2021. Forecasting dynamic body weight of nonrestrained pigs from images using an RGB-D sensor camera. Transl. Anim. Sci. 5, http://dx.doi.org/10.1093/tas/txab006.

Zong, Z., Guan, Y., 2024. AI-driven intelligent data analytics and predictive analysis in industry 4.0: Transforming knowledge, innovation, and efficiency. J. Knowl. Econ. http://dx.doi.org/10.1007/s13132-024-02001-z.