



*agriculture*

IMPACT  
FACTOR  
**3.6**

CITESCORE  
**6.3**

Article

---

# Application of Machine Learning Algorithms for On-Farm Monitoring and Prediction of Broilers' Live Weight: A Quantitative Study Based on Body Weight Data

---



Peng Lyu, Jeongik Min and Juwhan Song



<https://doi.org/10.3390/agriculture13122193>

## Article

# Application of Machine Learning Algorithms for On-Farm Monitoring and Prediction of Broilers' Live Weight: A Quantitative Study Based on Body Weight Data

Peng Lyu <sup>1,2</sup>, Jeongik Min <sup>1,2</sup> and Juwhan Song <sup>1,2,\*</sup>

<sup>1</sup> Graduate School of Artificial Intelligence, Jeonju University, Jeonju-si 55069, Republic of Korea; lyupaif@jj.ac.kr (P.L.); minji@jj.ac.kr (J.M.)

<sup>2</sup> Artificial Intelligence Research Center, Jeonju University, Jeonju-si 55069, Republic of Korea

\* Correspondence: jwsong@jj.ac.kr

**Abstract:** A non-invasive automatic broiler weight estimation and prediction method based on a machine learning algorithm was developed to address the issue of high labor costs and stress responses caused by the traditional broiler weighing method in large-scale broiler production. Machine learning algorithms are a data-driven strategy that enables computer systems to make predictions and judgments based on patterns and regularities that they have learned. To estimate the current weight of individual live broilers on farms, machine learning algorithms such as the Gaussian mixture model, Isolation Forest, and Ordering Points To Identify the Clustering Structure (OPTICS) are used to filter and extract data features using a two-stage clustering and noise reduction process. Real-time weight prediction was also achieved by combining polynomial fitting and the gray models and adjusting the model parameters based on prediction accuracy feedback. The symmetric mean absolute percentage error (SMAPE) value is a metric that is commonly used to evaluate the predictive performance of a model by comparing the degree of error between the model's predicted value on the day of slaughter and the true value measured manually, and the results of the experiments on 111 datasets showed that 7.21% were less than or equal to 0.03, 28.83% were less than or equal to 0.1 and greater than 0.03, and 31.53% were less than or equal to 0.2 and greater than 0.1. This method can be used as a prediction scheme for broiler weight monitoring in a large-scale rearing environment, considering the cost of implementation and the accuracy of estimation.



**Citation:** Lyu, P.; Min, J.; Song, J. Application of Machine Learning Algorithms for On-Farm Monitoring and Prediction of Broilers' Live Weight: A Quantitative Study Based on Body Weight Data. *Agriculture* **2023**, *13*, 2193. <https://doi.org/10.3390/agriculture13122193>

Received: 10 October 2023  
Revised: 2 November 2023  
Accepted: 21 November 2023  
Published: 23 November 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** machine learning; multilevel clustering; broiler weight estimation; weight prediction; GMM; OPTICS

## 1. Introduction

Broilers' body weight is an important indicator of their health, and effective broiler body weight monitoring is a problem that must be solved in the process of large-scale broiler farming. Meanwhile, large poultry companies typically cover a vertical chain from breeding farms, chicken farms, slaughterhouses, and distributors to form a complete broiler supply chain in order to maximize profits. Due to the lack of a younger generation of workers willing to work in the broiler industry, companies in Korea contract with hundreds of farms to meet market demand for broilers with acceptable specifications, and to meet the production capacity of their slaughterhouses. Because of the large price difference between qualified and substandard broilers, agribusinesses must increase the qualification rate, i.e., broiler production within a specified weight and size range, to avoid potential profit losses. Farmers, on the other hand, can earn additional incentive gains based on the number of broilers that meet the weight standards. As a result, on-farm real-time monitoring of live broiler weights and slaughter time control are critical for revenue management.

Broiler body weight estimation and monitoring have been extensively studied and can be categorized into four main directions:

- i. Traditional growth curves and growth models based on fitting mathematical functions. For example, Topal et al. [1] studied the fitting and prediction of avian weight–age relationships and compared the goodness of fit between MMF, Weibull, logistic, Gompertz, and von Bertalanffy models. Moharrery et al. [2] proposed a methodology to study and predict the growth characteristics of commercial broilers and indigenous chickens using a nonlinear function, and they used several statistical methods to evaluate the fit of the function and differences in growth parameters. Rizzi et al. [3] investigated growth patterns and sex differences in poultry meat production by comparing different models, such as linear, logistic, Gompertz, and Richards models, and the effect of fit analysis revealed a flexible growth function. Mouffok et al. [4], for the Cobb500 strain of meat birds, found that the Gompertz model was more accurate in estimating body weights in the early stages when comparing and evaluating the fit and predictive effects of different models.
- ii. Live weight estimation methods based on digital image processing. For example, Wet et al. [5] proposed a method to analyze images of broilers using commercial software and established a nonlinear regression equation to estimate the body weight of broilers by statistically analyzing the nonlinear relationship between their surface area, girth, and body weight, which was found to be less accurate compared to image analysis of pigs. Chedad et al. [6] proposed a method to estimate the body weight of chickens by image analysis, and the results of the study showed that the results of automated weighing systems tend to underestimate the actual body weight of chickens at the end stage of the growth period. Bazlur et al. [7] proposed a method to develop a linear equation for estimating the body weight of broilers by analyzing the digital images of their body surface area, validated using a random sample of 100 broilers, and the highest error between the manually measured weight and estimated weight was 16.47%, while the lowest error was 0.04%. Mortensen et al. [8] proposed a method for predicting the weight of broilers based on a 3D camera and image processing algorithms, where the average relative error between the predicted and true weight on the test dataset was 7.8%, and as the density of chickens increased, the absolute error of prediction became larger in the later stages of breeding. Amraei et al. [9] proposed a research method that includes the use of machine vision techniques to extract features related to body weight and the use of artificial neural network algorithms for predicting body weight, with prediction errors mainly centered on less than 50 g.
- iii. Body weight monitoring methods based on audio analysis. For example, Aydin et al. [10] conducted a study to determine the feed intake of chickens by detecting the birds' pecks and comparing them with feed intake measured by a weighing system. They discovered a linear correlation between the number of pecks and feed intake, with 93% of pecks being accurately identified. Fontana et al. [11] developed a tool that can automatically detect the growth status of broiler chickens at varying ages based on the frequency of calls emitted by the chickens. The results of the statistical analysis showed a significant correlation between the age and weight of the chickens and the maximum power frequency (PF) emitted in their calls. Fontana et al. [12] applied SAS 9.3 software programs, including PROC TTEST, PROC CORR, and PROC REG, to perform regression analyses and statistical tests. Statistical and regression analyses indicated a notable correlation between the sound frequency, age, and body weight of broilers. Fontana et al. [13] conducted a study on the use of sound analysis to predict the body weight of broilers and found a considerable correlation between age and body weight. Incidentally, they established that frequency analyses of chickens' crowing may be disrupted by filters and ambient noise during the final stages of broiler growth. The study revealed that filters and environmental noise during the final stages of broiler growth may interfere with the frequency analysis of chicken calls. Abdel-Kafy et al. [14] utilized statistical analysis software and regression modeling to predict the body weight of turkeys by recording their

- vocalizations and corresponding body weights. The results demonstrated a decrease in the frequency of vocalizations with age.
- iv. Direct predictive modeling based on other sensor data (nutritional intake, ventilation, temperature, humidity, etc.) or weight data. For example, Johansena et al. [15] proposed a research methodology to predict broiler weight utilizing a dynamic neural network model. The model was trained using an LM optimization algorithm, with input variables selected based on mutual information. Additionally, kernel density estimation was employed to estimate the joint probability density function. The system achieved an average root-mean-square error of prediction of 66.8 g. Lee et al. [16] developed an automated chicken weighing system composed of weighing scales and workstations. The weighing scale was built using an aluminum plate and a 5 kg load cell, and weight data were transmitted wirelessly to the workstation via a transmission module. The workstation collects data every 15 s and compares the average weight per day with a reference value to monitor the growth and development of the chickens. Weihong Ma et al. [17] introduced an effective method for extracting values using dynamic weighing. Their approach involves an improved amplitude-limited filtering algorithm and a BP neural network model to analyze data such as age, daily weight gain, average speed, and preprocessed weight values. The weighing error was reduced from 6% to less than 3% through a data-driven framework proposed by Chunyao Wang et al. [18] This framework employs Gaussian mixture modeling, self-sampling, and weighted averaging techniques to enhance the accuracy of monitoring and predicting live chicken weights. Birzniece et al. [19] suggested utilizing a long short-term memory (LSTM) artificial neural network for broiler weight prediction, based on environmental factors including temperature, gas concentration, humidity, broiler weight, and feed consumption.

A detailed enumeration and analysis of the methods, evaluation indicators, and main contributions that have appeared in previous research is provided in Table 1 below.

As shown in Table 1, in the traditional field of growth pattern and function fitting, researchers primarily rely on mainstream nonlinear regression to estimate different growth curve models for broiler chickens (e.g., Gompertz, Richards, and logistic), in order to describe and predict growth trends. However, this method has some drawbacks. It can be challenging to find a function that is applicable to diverse broiler breeds or breeding cycles, resulting in fitted functions that are often only suitable for characterizing the growth pattern of a specific breed of broiler chickens at a certain stage. When evaluating models, more emphasis is typically placed on goodness of fit, such as the coefficient of determination ( $R^2$ ), rather than the accuracy of single-point or overall live weight predictions. Although some studies report a mean absolute percentage error (MAPE) value of 4% for prediction errors, these values were measured on small datasets consisting of no more than 100 chickens.

**Table 1.** Summary of previous studies' methodology, scale, findings, and conclusions.

Related Research	Research Approach	Author	Methods	Specifications	Evaluation Metrics and Error	Main Conclusions
Growth patterns and function fitting		M. Topall, S.C. Bolukbasi et al. [1]	Nonlinear regression to estimate the parameters of different growth curve models in broiler chickens; 96 broiler chickens were weighed weekly from birth to 42 days of age.	Compared the predictive performance of different growth curve models, namely, Gompertz, logistic, Bertalanffy, Morgan–Mercer–Flodin (MMF), and Weibull.	The top performing MMF, Weibull, and Gompertz models had MAPE values of 0.0381, 0.0391, and 0.0385, respectively.	MMF, Weibull, and Gompertz models performed better compared to logistic and von Bertalanffy models.
		A. Moharrery and M. Mirzaei et al. [2]	Used nonlinear functions to describe the growth characteristics of commercial broiler and native chickens. Statistical procedures were used to analyze the data.	Compared the fitting performance of different equations: logistic, Gompertz, Richards, and Weibull.	The most successful Richards equation exhibited adjusted determination coefficient ( $R^2$ ) values of 99.51% for the commercial strain and 99.12% for native chickens.	The Richards function provided the best fit. Commercial birds grew faster and reached a higher final weight than native chickens.
		C. Rizzi, B. Contiero, and M. Cassandro et al. [3]	Compared the growth of an Italian commercial hybrid (Berlanda) and a local Italian breed called Padovana (in two color varieties) and their crosses. The birds were reared from 1 day to 180 days of age in an environmentally controlled breeder house.	Male and female chicks from five different genotypes were used to compare growth patterns using linear, logistic, Gompertz, and Richards growth models.	The most successful Gompertz and Richards growth models exhibit adjusted determination coefficient ( $R^2$ ) values of 99.51% for the commercial strain and 99.12% for native chickens.	The growth rates of the studied genotypes were lower than those of the commercial hybrids. The Gompertz and Richards growth models gave better estimates of weight parameters than the logistic model.
		Charef Eddine Mouffok, Semara L, and Farida Belkasmi et al. [4]	Describes the retrospective analysis of 50 broiler chicks and their division into three weight classes: light, middle, and heavy. Highlights the use of goodness-of-fit criteria to evaluate the accuracy of the models.	Discusses the comparison of six mathematical models: Gompertz, Richards, logistic, Weibull, von Bertalanffy, and exponential.	The total determination coefficient ( $R^2$ ) values of the Gompertz, logistic, von Bertalanffy, and WLS models at the three weight classes were all 0.954.	Concludes that the Gompertz model is the most suitable for describing the growth curve up to four weeks of age, while the logistic, von Bertalanffy, and WLS models accurately describe the growth curve after one month of age.
Image processing and computer vision		Lourens de Wet, Erik Vranken, Jean-Marie Aerts, and Daniel Berckmans et al. [5]	Use of commercial software to analyze captured images and determine body size based on surface area and peripheral pixel count. Nonlinear regression analysis was used to determine the relationship between body weight and image characteristics.	Used digital image processing techniques to estimate the live weight of broiler chickens and discussed the challenges and limitations of image analysis, such as variations in lighting, animal movement, occlusion, and background clutter.	The body weight of the chickens was estimated with an average relative error of about 11% from the image surface and 16% from the image periphery.	Suggested the possibility of using image sequences for behavioral characterization and real-time observation systems.

Table 1. Cont.

Related Research	Research Approach	Author	Methods	Specifications	Evaluation Metrics and Error	Main Conclusions
		Erik Vranken, Jean-Marie Aerts, and Daniel Berckmans et al. [6]	Individual broiler chickens' weights are determined by analyzing their two-dimensional surface area (top view) under optimal lighting, with the integration of the Otsu adaptive threshold algorithm enhancing the system's ability to distinguish between individual birds and the background.	Study conducted using three broiler houses to compare automatically recorded body weights with manually recorded body weights on days 36 and 42.	Automatically obtained body weights were comparable to manual recordings on day 36 in both House I ( $1445 \pm 15$ g vs. $1477 \pm 20$ g) and House II ( $2124 \pm 27$ g vs. $2142 \pm 20$ g), but on day 42 the automatic system underestimated the weights for House I ( $1839 \pm 30$ g vs. $2140 \pm 25$ g) and House II ( $2430 \pm 18$ g vs. $2555 \pm 35$ g).	The notion that bigger animals utilized the weighing equipment less frequently was confirmed. For example, statistical disparities in the area of broilers on and near the weighing system might be proven during week 5.
Image processing and computer vision		Md. Bazlur, R. Mollah, Md. A. Hasan, Md. A. Salam, and Md. A. Ali et al. [7]	Digital image analysis using raster image analysis software (IDRISI 32) to capture digital images of broilers. Determined broilers' body surface area from the images and developed a linear equation to estimate broilers' weight based on surface area pixels.	Reared 100 Arbor Acres broiler chicks under standard rearing conditions and captured 1200 digital images from 20 randomly selected broilers during the 7–42 day growing period for analysis.	Relative error in weight estimation of broiler chickens by image analysis: 0.04% to 16.47%. The overall value of the determination coefficient ( $R^2$ ) of the final linear equation was 0.999.	Although the precision of live weight estimation depends on many factors, the presented data indicate that the development of a practical imaging system for weighing broiler is feasible.
		Anders Krogh Mortensen, Pavel Lisouski, Peter Ahrendt et al. [8]	Use of a low-cost 3D camera (Kinect) with its own infrared light source and image processing algorithm, along with a range-based watershed algorithm for segmentation, extraction of weight descriptors, and weight prediction using a Bayesian artificial neural network, and the comparison and evaluation of four other models for weight prediction.	Using a commercial broiler house of 48,000 broilers (Ross 308) during the last 20 days of the breeding period and a test set of 83 broilers, manually annotated images, and a traditional platform scale for reference weights to explore the different 1D, 2D and 3D features for weight prediction.	A relative mean error of 7.8% was achieved on a separate test set, and the range of absolute errors was 20–100 g in the first half of the period and 50–250 g in the last half. Larger errors were observed at the end of the rearing period as the broiler density increased.	The system shows promise as a non-intrusive, robust solution for weighing broilers in commercial production environments, with potential for additional applications.

Table 1. Cont.

Related Research	Research Approach	Author	Methods	Specifications	Evaluation Metrics and Error	Main Conclusions
Image processing and computer vision		S. Amraei, S. A. Mehdizadeh and S. Salari et al. [9]	Machine vision technology was used to extract six features from the captured image features: area, perimeter, convex area, major axis length, minor axis length, and eccentricity. Then, a variety of artificial neural network techniques were used (gradient descent algorithm, scaling conjugate gradient algorithm, Levenberg–Marquardt algorithm, Bayesian fitting algorithm, etc.) to predict the live weight of broilers.	Live weight estimation was performed on 30 one-day-old broiler chickens reared for 42 days and imaged twice a day.	Bayesian regression with an $R^2$ value of 0.98 was the best network for predicting broiler weight, and its root-mean-square error (RMSE) value was 82.37.	The five remaining features (area, perimeter, convex area, major and minor axis length) showed a strong relationship with body weight. Computer vision systems provide a valuable alternative to manual weighing.
Audio processing and correlation exploration		A. Aydin, C. Bahr, S. Viazzi, V. Exadaktylos, J. Buyse, D. Berckmans et al. [10]	Sound recordings using a microphone attached to the feeding pen. Each hen was deprived of food for four hours before the experiment. Feed intake was automatically recorded using a weighing system, and feed wastage was manually collected and weighed. The results of the algorithm were compared with reference feed intake values obtained by weighing and video observation.	Twelve individual 28-day-old male broiler chickens (Ross-308) were used. Pecking sounds were recorded for 15 min during each trial. A total of 36 trials were performed, with three laboratory trials per broiler.	The algorithm correctly identified 93% of pecking sounds but had a false positive rate of 7%. The coefficient of determination ( $R^2$ ) between the number of pecks and the feed uptake was 0.995. The coefficient of determination ( $R^2$ ) between the feed intake and the number of pecks (pecking frequency) was 0.985.	Provided a non-invasive and automated method to measure feed intake in broiler chickens. The real-time data provided by this algorithm have potential applications in the study of broilers' feeding behavior and welfare.

Table 1. Cont.

Related Research	Research Approach	Author	Methods	Specifications	Evaluation Metrics and Error	Main Conclusions
		Ilaria Fontana, Emanuela Tullo, Alberto Peña Fernández and Daniel Berckmans et al. [11]	Precision livestock farming (PLF) technologies were used to monitor the welfare and health status of broilers. Sound recordings were made at regular intervals throughout the broiler production cycle, and the sound data were compared with the weight of the birds, which was automatically measured using a 'step-on scale' placed on the floor of the broiler house. Regression analysis, <i>t</i> -tests, and correlation analysis were used to evaluate the relationship between PF, age, and bird weight.	Data collection on two farms: one in the UK and one in the Netherlands. Only 18 h of sound recordings from specific days was used for analysis, together with the bird weights collected on those days. Fast Fourier transform (FFT) was used for frequency analysis.	A <i>t</i> -test of the expected values and the observed ones resulted in a <i>p</i> -value of 0.8807. The correlation between the weight and age of the broilers was highly positive (0.97, <i>p</i> -value < 0.001). The correlation between the PF of the sounds and the age of the broilers was −0.96. The correlation between the PF and the weight was −0.92 ( <i>p</i> -value < 0.001).	As broilers grow older and gain weight, the PF of their vocalizations decreases. Individual vocalizations and the whole audio file can be considered equivalent for PF analysis. These results have potential for the development of a weight prediction algorithm based on sound analysis in broiler production.
Audio processing and correlation exploration		Ilaria Fontana, Emanuela Tullo, Andy Butterworth and Marcella Guarino et al. [12]	Used precision livestock farming (PLF) techniques to combine audio and video information for automated monitoring of broiler chickens. PROC TTEST, PROC CORR, and PROC REG were used to compare the peak frequencies (PFs) recorded in two trials, and to examine the correlation between the PF, age, and weight of chickens. The GLM procedure was used to estimate the effects influencing PF. The LSMEANS analysis was used to analyze the variation in PF with the age of the broilers.	Sound recordings were made at regular intervals throughout the life of the birds over a period of 38 days. A total of 55 h and 20 min of recordings was collected, and 600 birds were weighed; a total of 600 sounds (50 sounds per day), randomly selected from 12 days of recordings, were manually labeled and analyzed.	There was a strong positive correlation (0.97, <i>p</i> -value < 0.001) between the weight and age of the broilers, and a strong negative correlation (0.95, <i>p</i> -value < 0.001) between the PF of the sounds and the age of the broilers. There was a significant negative correlation (0.80; <i>p</i> -value < 0.001) between the frequency of the vocalizations and the weight of the broilers.	A strong positive correlation was observed between weight and age. PF showed significant negative correlations with both age and weight. PF could be used as an early warning or continuous monitoring system to assess the health and status of broiler chickens.

Table 1. Cont.

Related Research	Research Approach	Author	Methods	Specifications	Evaluation Metrics and Error	Main Conclusions
Audio processing and correlation exploration		Ilaria Fontana, Emanuela Tullo, Lenn Carpentier, Dries Berckmans, Andy Butterworth, Erik Vranken, Tomas Norton, Daniel Berckmans, and Marcella Guarino et al. [13]	Sound analysis techniques were used to measure and analyze the frequency of broiler vocalizations. Peak frequencies (PFs) of vocalizations were manually processed to remove outliers and exclude sounds collected during dark periods with background noise. A polynomial regression model was estimated using PFs and weight data collected for each production cycle to predict weight.	Sounds and body weight were continuously recorded throughout the cycles. The chosen frequency interval was 1100 Hz to 3700 Hz. A final dataset was created by merging data from rounds 1 to 8. Data from laps 1 to 5 were used to predict weight.	The correlation coefficient between the expected and observed weights was high and positive ( $R^2 = 0.96$ , $p$ value $\leq 0.001$ ). The regression model between the expected and observed weights was highly significant ( $R^2 = 0.93$ , $p$ value $\leq 0.001$ ).	The identified model for predicting weight as a function of peak frequency confirmed that birds' weight could be predicted by frequency analysis of sounds emitted at the farm level.
		El-Sayed M. Abdel-Kafy, Samya E. Ibraheim, Alberto Finzi, Sabbah F. Youssef, Fatma M. Behiry, and Giorgio Provolo et al. [14]	SAS 9.3 software was used to estimate the relationships between age, weight, and PF. Regression models were developed to predict weight and PF from age and weight from PF. Pooled data were analyzed using ANOVA to test for differences between age, weight, and PF variables. Regression models based on pooled data were used to predict weight and PFs.	Four trials were conducted in Egypt to record sounds and weights of turkeys during an 11-day growth period. A total of 2200 sounds were manually analyzed and labeled using peak frequency (PF).	The correlation coefficient between turkeys' weight and age was high and positive ( $R^2 = 0.96$ , $p < 0.0001$ ). The correlation coefficient between the PF of turkey vocalizations and their age was high and negative ( $R^2 = 0.97$ , $p < 0.001$ ). The correlation coefficient between the PF of the vocalizations and the weight of the turkeys was high and negative ( $R^2 = 0.97$ , $p < 0.001$ ). The RMSE values during calibration and validation differed by 5.1%.	Audio monitoring provides a non-contact method of monitoring turkeys' growth, eliminating manual handling. Potentially useful for farmers to automate turkey growth monitoring.

Table 1. Cont.

Related Research	Research Approach	Author	Methods	Specifications	Evaluation Metrics and Error	Main Conclusions
Direct live weight prediction for broilers using other sensors		Simon V. Johansen, Jan D. Bendtsen, Martin R.-Jensen, and Jesper Mogensen et al. [15]	Dynamic neural network (DNN) model with input variable selection (IVS). Mutual information-based input variable selection using kernel density estimation (KDE). DNN model training using the Levenberg–Marquardt optimization algorithm and cross-validation. Used ensemble predictions generated by training multiple sub-models with different initial weights. Calculation of prediction evaluation metrics such as root-mean-square error (RMSE).	Data collected from a broiler house over a period of 3 years and 4 months, consisting of 29 batches. Inputs to the model: environmental variables (i.e., temperature, humidity, light intensity, ventilation demand, heating demand). Outputs: broiler behavior indicators (i.e., weight, feed consumption, water consumption).	Mean forecasting RMSE of 66.8 g.	The dynamic influence of environmental conditions on broiler growth was found to be significant, showing the potential usefulness of the method in industry and as a basis for future research on broiler production optimization.
		C.C. Lee, A.H. Adom, M.A. Markom and E.S.M.M. Tan et al. [16]	The automated chicken weighing system uses a wireless sensor network (WSN) for data transmission. The system includes rugged aluminum plate scales and 5 kg load cells. The weight data are transmitted to a workstation via a wireless transceiver module.	The scales were designed to accurately measure the weight of the chickens. They were made of aluminum plates and equipped with 5 kg load cells to ensure robustness and accuracy, and they were placed inside the three pens for 38 days to record and monitor the development of the chickens from day to day.	From day 1 to day 12, the hens' development rate was as planned. However, beginning on day 13, the development rate of the hens was 3.38% to 12.21% slower than projected, causing the animals to take 40–42 days to reach 1.8 kg.	The successful development of an automated chicken weighing system using a wireless sensor network (WSN) capable of collecting real-time weight data from broiler chickens.

Table 1. Cont.

Research Approach	Author	Methods	Specifications	Evaluation Metrics and Error	Main Conclusions
Related Research	Weihong Ma, Qifeng Li, Jiawei Li, Luyu Ding and Qinyang Yu et al. [17]	A sliding window technique was used to create a low-pass filter with continually updated samples. Valid data were extracted via dynamic weighing, which entails analyzing weight values, applying amplitude-limiting filtering, recording stable values, and comparing differences within an error range. Finally, the modeling and implementation of BP (backpropagation) neural networks using age (days), daily weight increase, average speed, and preprocessed weight as input variables.	Thirteen groups of Beijing fatty chickens ranging in weight from 500 to 1800 g were tested. Data were obtained from 200 of the 2000 sets. Three distinct weighing methods were used: no filtering algorithm or BP neural networks, just an enhanced amplitude-limiting filtering algorithm, and a hybrid of an improved amplitude-limiting filtering algorithm and BP neural networks.	The hybrid method decreased error from 6% (no filtering algorithm or BP neural networks) to less than 3%.	The hybrid technique of the increased amplitude-limiting filter algorithm and BP neural networks performed better in terms of minimizing error in weighing broiler chicks. It outperformed the other methods examined by lowering the error to less than 3%.
Direct live weight prediction for broilers using other sensors	Chun-Yao Wang, Ying-Jen Chen, and Chen-Fu Chien et al. [18]	Proposed a data-driven framework for weight monitoring and prediction in the broiler industry. The weight monitoring module estimates live broiler body weight using a Gaussian mixture model (GMM). It employs a bootstrap resampling approach to decrease cluster noise, identify outliers, and compute the cluster-weighted mean as a single individual average weight. To model broiler growth and anticipate future weight, the weight prediction module employs mathematical growth functions, notably the Gompertz function. The cumulative mean absolute percentage error (Cu-MAPE) is used as a model fit indicator.	Empirical studies were carried out in six broiler farms to validate the proposed approach. For each batch, the estimated value was compared with manual weighing on four reference days, i.e., day 14, day 21, day 28, and the day of delivery.	For batch 165–1, the error on day 28 was greater than 8%; for batch 165–2, the error on day 21 was 7.55%, while that on day 28 was 11.2%; for batch 165–4, the error on day 14 was the highest of all, at about 16.34%; for batch 165–9, the error on day 14 was 5.2%, while that on day 21 was 7.75%. As regards the error on the day of delivery, several batches had an error of less than 3%, with some of them even less than 1%.	The proposed data-driven framework for weight monitoring and prediction in the broiler industry was successfully implemented and validated. The study demonstrated the practicality of the approach and highlighted the potential for Industry 3.5 solutions in the agricultural sector.

Table 1. Cont.

Related Research	Research Approach	Author	Methods	Specifications	Evaluation Metrics and Error	Main Conclusions
Direct live weight prediction for broilers using other sensors		Ilze Birzniece, Signe Balina, Ilze Andersone, Andris Kikans, and Agris Nikitenko et al. [19]	Data from numerous production cycles, including environmental indicators and poultry growth data, were collected for the study. Data were preprocessed to account for mistakes and missing values. To enhance the sample size for broiler weight data, augmentation procedures were applied. The forecasting model's accuracy was assessed by comparing it to the Gompertz model. Long short-term memory (LSTM) artificial neural networks were used to build the forecasting model.	Data on broiler raising were obtained for three independent rooms that had sensor systems installed. In each room, mixed-sex flocks of Ross 308 AP95 broilers were housed. The dataset contained manually gathered data, such as breeders' notes on the number of broilers, age, feeding, and weight, along with sensor data such as temperature, humidity, carbon dioxide (CO <sub>2</sub> ) levels, ammonium (NH <sub>3</sub> ) levels, and so on.	The results obtained were evaluated at three different durations of the prediction step—1, 3, and 6 days. The highest accuracy was found for the 3-day forecast horizon, with an RMSE value of 0.295.	The LSTM model can effectively predict broiler weight in poultry production, considering relevant environmental factors. The research and development work demonstrates the potential of machine learning techniques to improve production quality and profitability in the poultry industry.

In the widely popular fields of image processing and computer vision (referenced in Table 1), researchers commonly use commercial image processing software or artificial neural networks to establish the relationship between broilers' area in images and their actual weight. These methods often involve the fundamental task of distinguishing the broiler from the background, frequently using the Otsu method. However, this technique is highly sensitive to noise and may face challenges in accurately segmenting images that lack distinct bimodal peaks or exhibit multimodal characteristics. Furthermore, conducting research in this domain requires the use of expensive camera equipment, usually incorporating infrared night-vision capabilities for capturing scenes in low-light conditions, as well as high pixel resolution for enhanced calculation accuracy in obtaining high-definition images. In some studies, 3D cameras are even employed to capture spatial information. Additionally, both image processing and training neural networks are time-consuming tasks, and the former also requires complex data annotation work. Finally, the evaluation of these models involves calculations performed on small test datasets consisting of less than 100 chickens. For example, the best-performing Bayesian neural network model achieved a root-mean-square error (RMSE) value of 82.37 when tested on a dataset comprising 30 broiler chickens.

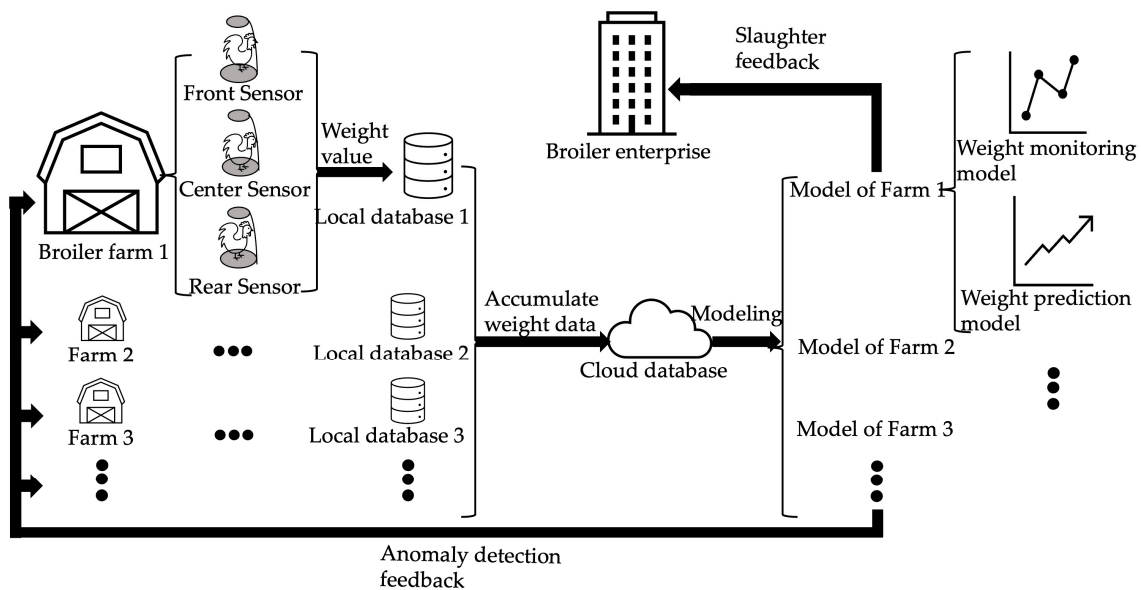
As indicated in Table 1, in the increasingly popular field of audio processing and correlation exploration, researchers commonly utilize audio processing software in conjunction with multiple regression models to estimate the relationship between broilers' age, weight, and peak frequency (PF). The first challenge that they face is removing background noise. Furthermore, expensive recording equipment is necessary for this type of research, requiring high-resolution recording capabilities to capture more detail and improve audio quality. These devices often incorporate noise suppression technology to minimize the interference of background noise. Data preprocessing and annotation work are also time-consuming and labor-intensive, similar to video processing. Finally, when conducting model evaluation, most studies focus on correlation studies, examining the  $p$ -values between age, weight, and PF, with little attention given to direct prediction accuracy measurement.

As observed in Table 1, in the limited domain of direct live weight prediction for broilers using other sensors, researchers have attempted to use alternative types of sensors to directly capture data on factors influencing body weight (such as temperature, carbon dioxide concentration, and ammonia concentration). They established neural networks or regression models to directly estimate broiler chickens' weight. The first step for researchers is to reduce data noise through data preprocessing, filtering out the real weight data of individual broiler chickens. It should be noted that the resampling technique mentioned in previous studies originally served as a method of expanding the dataset, not a noise reduction algorithm. The sensor equipment required for this method is relatively simple, and model calculations on the same scale are faster compared to indirect data processing methods like images and sounds. Finally, when conducting model evaluation, although the best prediction accuracy error can be as low as 3%, these results are based on a limited number of datasets and small samples.

Based on the previous research, we believe that the live weight monitoring and prediction of broilers in large-scale breeding industries has not really been studied, because the actual data are not for tens or even hundreds of thousands of chickens in breeding. Moreover, directly using the weight sensor to calculate the live weight of broiler chickens is undoubtedly the simplest and most economical method, and the construction of the Internet of Things system and model calculation are relatively simple. The only difficulty is extracting individual body weight data from highly contaminated data. Our research is based on large datasets and combined with the most popular machine learning methods in current artificial intelligence to provide advanced solutions to the above challenges.

The solution consists of two stages: non-invasive automatic real-time monitoring, and prediction of broiler weight. The first stage uses a multi-clustering machine learning algorithm to continuously eliminate outliers to obtain the average weight of a single individual. In the second stage, short- and long-term prediction models are flexibly established accord-

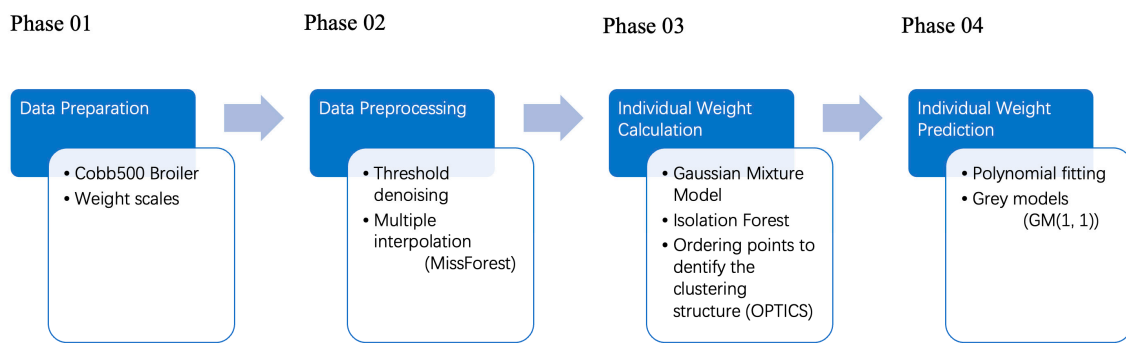
ing to the feeding cycle to improve the prediction accuracy. An experimental study using 113 large datasets provided by Harim's Cooperative Farms, the largest and leading broiler company in South Korea, demonstrated the practical feasibility of this approach. Figure 1 shows the general process framework, from data collection to model building, used in this study. It includes a hardware IoT system and a software algorithm system. The IoT system consists of three parts: an electronic scale fixed on the farm floor, a database that stores the local measurement data of the electronic scale, and a cloud database that aggregates all of the datasets. The software part constructs two types of models for monitoring and prediction based on the weight dataset (in terms of farms and feeding cycles) from the cloud database.



**Figure 1.** Proposed system architecture.

## 2. Materials and Methods

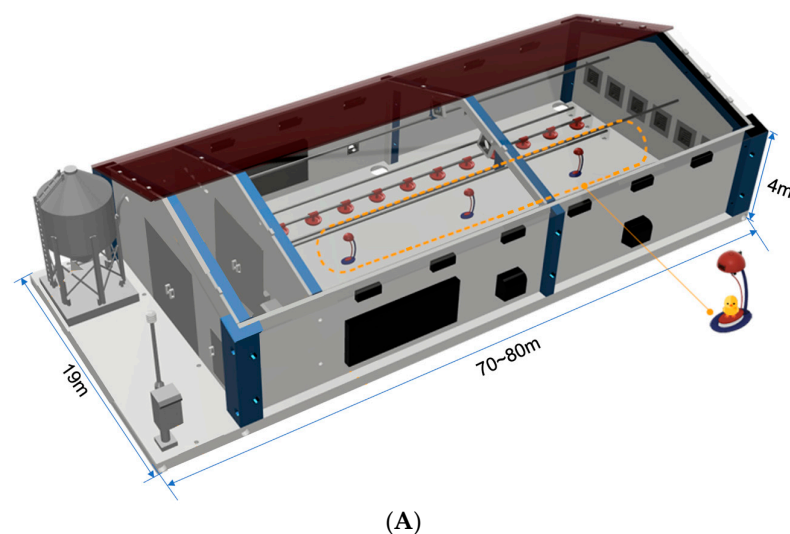
The objective of this study was to explore an automated and cost-effective method, based on machine learning algorithms, that uses weight sensors as the sole source of data collection to determine the live weight of individual broilers. To this end, several machine learning clustering algorithms were introduced to develop a methodology for processing and analyzing the data in three stages, allowing the average weight of an individual chicken to be derived from noisy and outlier data. Machine learning itself is a collection of algorithms that automatically learn and extract patterns and features from large amounts of data. It can be divided into supervised learning, which includes labels (i.e., positive solutions), unsupervised learning, which does not include labels, and reinforcement learning, which simulates maximizing cumulative utility, with unsupervised learning algorithms often used in data analysis and mining. The overall processing flow is shown in Figure 2 (Phases 2–4): In the first phase, all obvious outliers in the original data, including negative and 0 values, are replaced by data within the range of valid values. Here, we used the multiple interpolation machine learning algorithm. In subsequent stages, the entire dataset is divided into a number of monitoring units at fixed time intervals, and then the mean body weights of the individuals in each unit block are calculated using multiple clustering, anomaly detection, and other machine learning algorithms. In the third stage, the mean weight for the next time interval is predicted using the compiled list of individual weights.



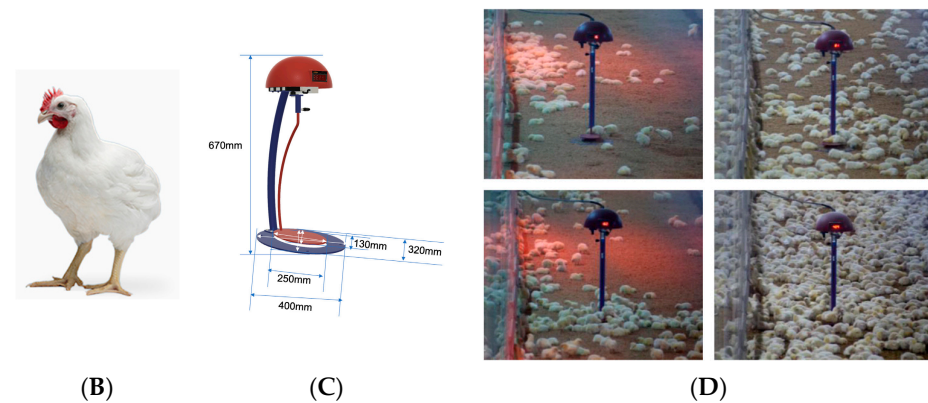
**Figure 2.** Research phases and methods.

### 2.1. Broiler Sample and Live Weight Data Collection

This study was conducted at 36 farms. A regular supply of Cobb 500 chicks was provided by the firm, and Cobb-Vantress, Inc. of Siloam Springs, AR, USA, was the output company. The breed's minimum theoretical weight was 42 g (0 days of age), and the maximum weight was 4641 g (56 days of age). KOKOFARM electronic scales (model: KOKOFARM, provider: EMOTION Co., Ltd. of Jeonju, Korea, output company: CAS Co., Ltd. of Yangju, Korea) were used for precise measurements. We established and consistently calibrated monitoring systems at the front, middle, and rear locations of each farm before introducing a new group of breeders. Each scale was equipped with six types of sensors, including weight sensors, with a maximum range of 1500 g, a divisional value of 1 g, and a measurement error range of 5%. The breeding cycle on the farms lasts for 28 to 31 days, during which 30,000 broilers are raised in a single batch. The theoretical body weights of the broilers ranged from 42 to 2094 g based on the actual breeding cycle, and within this range the sample of broilers was representative. The sensors for body weight produced one reading per second, which was sent to the data server for storage in a CSV file format every minute of accumulation. The raw data, consisting of time and scale data, were obtained by filtering redundant sensor data and synchronizing timestamps with three weight values. The program was developed using Python 3.9, with core algorithmic libraries such as scikit-learn 1.1.2, SciPy 1.9.1, missingpy 0.2.0, and others being employed. All experiments were conducted on the server, which was equipped with an Intel (R) Xeon (R) Gold 5218R CPU operating at 2.10 GHz, an NVIDIA A100 80 GB PCIe, 251 GB RAM, and the Ubuntu Linux operating system. The experimental farm layout, broilers, and electronic scales are shown in Figure 3 below.



**Figure 3.** Cont.



**Figure 3.** Laboratory space, equipment, and materials: (A) Scheme of the length, width, and height of the experimental farm and location of the three electronic scales; (B) Cobb 500 broiler; (C) KOKO-FARM electronic scale; (D) farm environment during feeding operations.

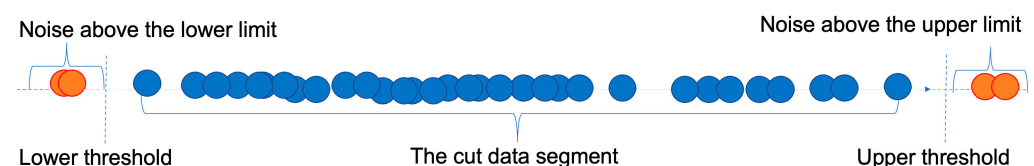
## 2.2. Data Preparation and Preprocessing

The raw data were time-series data with one time column (in seconds) and three columns of body weight values (in grams), i.e., three body weight measurements at the same time corresponding to different farm locations. Due to the unpredictable individual and group behavior of chickens, as well as the uncontrollable contingencies of external environmental factors such as earthquakes, strong winds, hardware short-circuits, and so on, it was inevitable that the collected data would be filled with negative numbers, zeros, and other noisy data, such as 1.5 chicken data versus 2.5 chicken data or even 3.5 chicken data. When the weight of a chicken was accurately identified, the numbers were mixed with many outliers due to the sensitivity of the sensors. Therefore, these noise and outliers needed to be removed before analyzing the data.

### 2.2.1. Critical-Value-Based Data Processing for Noise Reduction

We can define thresholds based on observation or life experience to eliminate data that appear invalid. For example, setting the threshold to a very small positive number, such as 1, and deleting values less than that will filter out all negative integers and zeros. Similarly, a higher number can be used to remove noisy data that exceed the upper limit of the observed value. The advantages of this method are that it is fast, effective, and has a low computational overhead. The disadvantage is that it requires the user to have sufficient a priori knowledge about the range of values of the data and the labeling standards.

After processing, the cut data segment, as shown in Figure 4, contains the true weight data. This is particularly useful when there is a lot of noise outside the threshold.



**Figure 4.** Principle of critical-value noise reduction.

### 2.2.2. MissForest-Based Approach to Null-Filling

Noises are not removed immediately to maintain the timeliness of the data but are replaced with null values and then replenished periodically. This method attempts to maximize the value of the data. The MissForest [20] approach, known for its robustness, is used to impute noisy data, especially in cases where missing data are large and dispersed. It is a modern nonparametric approach based on random forests, a widely used nonlinear modeling tool. Its advantages include its ability to capture interactions and nonlinear characteristics within data variables, along with its flexibility to handle a wide range of data

formats, including mixed types with numerical classifications. Missing value imputation involves filling in all missing data with the median or mean for continuous values, training a random forest model on the entire dataset, predicting missing values, and iterating until convergence. For the set of continuous variables  $N$ , the convergence equation is defined as follows:

$$\Delta_N = \frac{\sum_{j \in N} (X_{new}^{imp} - X_{old}^{imp})^2}{\sum_{j \in N} (X_{new}^{imp})^2}, \quad (1)$$

where  $X_{new}^{imp}$  denotes the interpolation matrix and  $X_{old}^{imp}$  is the previous one.

Following the two phases of processing in the first stage, we obtained input data that could be used to begin the second stage of data analysis.

### 2.3. Time Interval Segmentation and Average Weight Calculation

To achieve real-time monitoring of the average body weight of individual chickens while reducing the computational overhead, the data from the preprocessing stage were split by temporal frequency. Direct clustering or finding the mean was difficult and imprecise due to the continuity of the data. Multiple clustering techniques and anomaly detection methods were used with data blocks as processing units for continuous cleaning and screening.

#### 2.3.1. Segmentation and Gaussian-Mixture-Model-Based Data Modeling

Data segmentation using time–frequency equivalent intervals is straightforward and, therefore, is not provided individually. We can divide the amount of data collected in real time into days or hours (in this case, 3 h). Then we can move on to clustering.

Because the dataset included various possibilities, like multiple chickens on the platform, or just a wing touch, we employed a Gaussian mixture model (GMM) [21] for segmenting the data to decompose them into individual Gaussian distribution components. As per the central limit theorem, if a random variable is composed of many tiny and independent elements, the variable is considered to follow a Gaussian distribution. Consequently, numerous incidental variables can be described roughly by uni- or multivariate Gaussian distributions. Specific weights merge multiple Gaussian models into a solitary model, and distinct data points possess diverse likelihoods of belonging to all Gaussian models, creating the GMM. Based on the time–weight value pairs, a two-dimensional GMM was utilized, as presented in Equations (2) and (3):

$$P(X) = \sum_{k=1}^K p_k N(X | \mu_k, \sigma_k), \quad (2)$$

$$N(X | \theta) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x - \mu_k)^T \Sigma^{-1}(x - \mu_k)\right), \quad (3)$$

where  $p_k$  is the weight of the  $k$ th component,  $\sum_{k=1}^K p_k = 1$ , and  $N(X | \theta)$  is the probability density function of the  $k$ th component. The data dimension is denoted by  $d$  (2 in this study),  $X$  and  $\mu$  are both  $d$ -dimensional data represented by a matrix with one row and  $d$  columns, and  $\Sigma$  is a non-heterogeneous square covariance matrix with  $d$  rows and  $d$  columns.

The expectation maximization (EM) method was utilized to estimate parameters, and upon completion of the iterations, the raw dense data were divided into clusters, each representing a unique parameter model of the initial data distribution.

#### 2.3.2. Isolation-Forest-Based Outlier Sieving

After breaking down the initial data into smaller subsets and patterns, it is necessary to determine the representative values of each subset by calculating averages. Centralization is crucial to prevent excessive variation in each set. To exclude outliers from each cluster and enable the calculation of a more general and robust average, the Isolation Forest

method [22] was employed. Isolation is the action of separating samples with characteristics from other data. These characteristics have two distinct meanings: firstly, the samples have a minimal size compared to the overall data, and secondly, the values of the samples differ significantly from the surrounding sample data. The Isolation Forest algorithm is a rapid and efficient anomaly detection algorithm. It calculates the necessary number of hyperplanes to isolate an instance by dividing the hyperplane and then assesses the anomaly of said instance. Owing to the miscellaneous and sporadic characteristics of the anomalous image components, they are typically detected close to the base node in the constructed tree structure, while the background is more likely to be identified at the deeper end of the tree, which, in turn, leads to a smaller depth of the anomaly in the isolated tree. Consequently, if image elements in an Isolation Forest made up of multiple standalone trees possess short path lengths, they are considered to be abnormal.

### 2.3.3. Average- and OPTICS-Based Multiple Clustering

The individual Gaussian components excluding outliers were averaged to convert the mixed data into discrete feature points. These representatives possessed both true and noisy values. The density-based soft clustering algorithm Ordering Points To Identify the Clustering Structure (OPTICS) [23] was then used to re-cluster the above discrete values into clusters of varying densities. The average weight of a single chicken within the current time was determined by selecting the value with the highest weight from the cluster with the most objects. This decision was based on our assumption that clusters with the most objects meeting the required density are more likely to contain accurate values, as the actions that reliably trigger the sensors during data collection produce the most noise, whereas temporary, random triggers produce less noise. Valid and accurate data typically arise when the sensor is in a stable condition (e.g., triggered continuously). OPTICS is a density-based clustering algorithm that can identify clusters of any shape and detect anomalies in the data. It is not sensitive to initial parameters and can find noisy points effectively. Additionally, OPTICS provides a result in the form of a set of sequences of points in all possible classification cases, making it versatile for various data scenarios. Multiple iterations are possible depending on the size of the data.

We determined average broiler body weights using algorithmic model analysis based on dual clustering combining GMM and OPTICS. Farmers can monitor body weights in real time based on these data, which can be used for predictive modeling as they accumulate.

## 2.4. Adaptive Forecasting Combining Multinomial Regression with Gray Models

Once enough real-time weight measurements have been collected, a new model can be created to forecast weight at a future point in time. Body weight data obtained during the monitoring stage in hours are sampled at a fixed frequency for daily data. The application of polynomial functions for modeling biological growth curves has been a longstanding area of research; however, this method typically necessitates additional data to attain a high level of precision. To address the gaps in prediction during data accumulation, we employed a gray model, which demands fewer data and provides good accuracy.

### 2.4.1. Multinomial-Regression-Based Medium- and Long-Term Forecasting

Polynomial regression, a subset of multiple linear regression, estimates the relationship as an  $n$ th-order polynomial [24]. To determine the best function match for the data, the widely used least squares method (LSM) [25], a mathematical optimization technique, minimizes the sum of squares of errors (also known as residuals). That is, a function is selected from the set of  $n$ th-order polynomial functions with strong predictive ability for both known and unknown data. Frequently, fitting a mathematical function is used to model animals' growth characteristics. Based on prior research, the Gompertz function was chosen to offer a polynomial fit to the invested data for up-to-date weight estimates as of the anticipated delivery date. Technical abbreviations will be explained at their first use. Equation (4) below provides the function, where  $a$  signifies the maximum limit,

b signifies the displacement along the x-axis (shifting the graph to the left or right), c signifies the growth rate (scaling along the y-axis), and e is Euler's number ( $e = 2.71828\dots$ ). However, as chickens' growth is influenced by environmental conditions, health status, and farmers' feeding strategies, the initial fitted curve may not accurately predict the average weight of future broilers. Therefore, to enhance the computational accuracy, this study incorporated a cumulative symmetric mean absolute percentage error (SMAPE) feedback model adjustment mechanism. SMAPE is a symmetry-based measure of the percentage error between anticipated and actual values that is particularly useful for evaluating predictive model performance in regression situations. SMAPE is assessed according to Equation (5), where  $A_t$  denotes the measured value and  $P_t$  denotes the projected value. The absolute deviation between  $A_t$  and  $P_t$  is divided by half the amount of the absolute values of  $A_t$  and  $P_t$  (the measured and projected values, respectively). Then, the calculated value for each fitted point t is divided by the number of fitted points n.

$$f(t) = ae^{-be^{-ct}}, \quad (4)$$

$$SMAPE = \frac{100\%}{n} \sum_{t=1}^n \frac{|P_t - A_t|}{\frac{(|P_t| + |A_t|)}{2}}, \quad (5)$$

The chicken growth cycle usually spans 28 to 32 days, consisting of a starter phase, growing phase, and terminal phase. The growth function cannot be fitted until the chickens have completed the growing period, due to inadequate data in the beginning. This study began by fitting a mathematical curve to the weight monitoring data from day 1 to day 14, which included the starter and grower periods. The deviation between the projected value according to the mathematical curve and the measured value from the weight monitoring was computed utilizing SMAPE for each period following the initial model fitting. Whenever the cumulative SMAPE of the vested growth function surpasses a specific threshold, the program adaptively refits a new growth function.

#### 2.4.2. Gray-Model-Based Short-Term Forecasting

To compensate for the late starting point of polynomial fitting prediction models (generally, two weeks of data are required to achieve high accuracy), and because the time series was only dedicated to data fitting and not law discovery, a gray model with more flexible application was introduced (generally, only four data points are needed to be sufficient). Gray models (GMs) are fuzzy long-term descriptions of the development patterns of things created by building gray differential prediction models with limited information [26]. They are generally used to determine the degree of dissimilarity of development trends among system factors, i.e., correlation analysis and gray generation of the original data to find the pattern of system changes and generate a data series with strong regularity, and then establish the corresponding differential equation model to predict the future development trends of things. This study's prediction method is known as the GM (1, 1) model, which is a type of gray prediction model [27]. Where G denotes gray, M denotes model, the first 1 in parentheses denotes that the differential equation is first-order, and the second 1 denotes that the equation has only one variable. Because each fit uses four new data points from the neighborhood, there is no feedback adjustment.

Combining the above methods, we achieved automatic real-time monitoring and prediction of the average body weight of a single chicken.

### 3. Results and Discussion

#### 3.1. Results of Data Pre-Processing

Typically, a CSV file from the server stores data from a farm during a breeding cycle, with a size of about 50 MB and 70 columns. The first 10 columns contain time (in minutes), ammonia concentration, and so on, while the last 60 columns are broiler weight data, one second at a time, with the number of rows varying according to the time duration. The

example data have 133,346 rows. The time columns and body weight data were filtered with Pandas and then classified by electronic scales 1, 2, and 3, and finally the timestamps in minutes were matched to each body weight value in seconds, as shown in Figure 5 below. The weight data extraction code, the timestamp matching code, and the data analysis code were combined to create an end-to-end automatic computational model.

Electronic scale No.		Temperature		Humidity		CO <sub>2</sub> concentration		Ammonia concentration		Dust concentration		w <sub>01</sub> ~ w <sub>60</sub> : Weight data								
A	B	C	D	E	F	G	H	I	J	K	L	M	BL	BM	BN	BO	BP	BQ	BR	
id	farmID	dongID	jeoulID	getTime	temp	humid	co	nh	dust	w01	w02	w03	w54	w55	w56	w57	w58	w59	w60	
2	KF00210101	KF0021	1	1 2021-12-10 09:24:23	27.9	76	5808	0	-100	-100	-100	-100	0	0	0	0	0	0	0	0
3	KF00210101	KF0021	1	1 2021-12-10 09:25:23	27.9	70.3	6371	0	-100	0	0	0	0	0	0	0	0	0	0	0
4	KF00210101	KF0021	1	1 2021-12-10 09:26:23	28	67.4	6377	0	-100	0	0	0	0	0	0	0	0	0	0	0
5	KF00210102	KF0021	1	2 2021-12-10 09:26:23	28	75.2	6113	0	-100	-100	-100	-100	0	0	0	0	0	0	0	0
6	KF00210101	KF0021	1	1 2021-12-10 09:27:23	28	65.3	6144	0	-100	0	0	0	0	0	0	0	0	0	0	0
7	KF00210102	KF0021	1	2 2021-12-10 09:27:23	28.1	69.8	6489	0	-100	0	0	0	0	0	0	0	0	0	0	0
8	KF00210101	KF0021	1	1 2021-12-10 09:28:23	28	64.5	6119	0	-100	0	0	0	0	0	0	0	0	0	0	0
9	KF00210102	KF0021	1	2 2021-12-10 09:28:23	28	67	6495	0	-100	0	0	0	0	0	0	0	0	0	0	0
10	KF00210103	KF0021	1	3 2021-12-10 09:28:23	27.7	72	6087	0	-100	-100	-100	-100	0	0	0	0	0	0	0	0
11	KF00210101	KF0021	1	1 2021-12-10 09:29:23	28	62.8	6105	0	-100	0	0	0	0	0	0	0	0	0	0	0
12	KF00210102	KF0021	1	2 2021-12-10 09:29:23	28	65	6491	0	-100	0	0	0	0	0	0	0	0	0	0	0
13	KF00210103	KF0021	1	3 2021-12-10 09:29:23	27.5	67.6	6003	0	-100	0	0	0	0	0	0	0	0	0	0	0
14	KF00210101	KF0021	1	1 2021-12-10 09:30:23	28	62	6072	0	-100	0	0	0	0	0	0	0	0	0	0	0
15	KF00210102	KF0021	1	2 2021-12-10 09:30:23	27.9	63.6	6426	0	-100	0	0	0	0	0	0	0	0	0	0	0
16	KF00210103	KF0021	1	3 2021-12-10 09:30:23	27.5	65.5	5948	0	-100	0	0	0	0	0	0	0	0	0	0	0
17	KF00210101	KF0021	1	1 2021-12-10 09:31:23	28	61.1	6071	0	-100	0	0	0	0	0	0	0	0	0	0	0
18	KF00210102	KF0021	1	2 2021-12-10 09:31:23	27.8	62.4	6401	0	-100	0	0	0	0	0	0	0	0	0	0	0
19	KF00210103	KF0021	1	3 2021-12-10 09:31:23	27.4	63.6	5886	0	-100	0	0	0	0	0	0	0	0	0	0	0
20	KF00210101	KF0021	1	1 2021-12-10 09:32:23	27.6	60.8	6075	0	-100	0	0	0	0	0	0	0	0	0	0	0
21	KF00210102	KF0021	1	2 2021-12-10 09:32:23	27.8	61.8	6418	0	-100	0	0	0	0	0	0	0	0	0	0	0
22	KF00210103	KF0021	1	3 2021-12-10 09:32:23	27.4	62.8	5976	0	-100	0	0	0	0	0	0	0	0	0	0	0
23	KF00210101	KF0021	1	1 2021-12-10 09:33:23	27.9	60.1	6059	0	-100	0	0	0	0	0	0	0	0	0	0	0
24	KF00210102	KF0021	1	2 2021-12-10 09:33:23	27.9	61	6425	0	-100	0	0	0	0	0	0	0	0	0	0	0
25	KF00210103	KF0021	1	3 2021-12-10 09:33:23	27.6	62	5926	0	-100	0	0	0	0	0	0	0	0	0	0	0
26	KF00210101	KF0021	1	1 2021-12-10 09:34:23	27.9	59.6	6011	0	-100	0	0	0	0	0	0	0	0	0	0	0
27	KF00210102	KF0021	1	2 2021-12-10 09:34:23	27.9	60.6	6432	0	-100	0	0	0	0	0	0	0	0	0	0	0
28	KF00210103	KF0021	1	3 2021-12-10 09:34:23	27.6	61.3	5964	0	-100	0	0	0	0	0	0	0	0	0	0	0
29	KF00210101	KF0021	1	1 2021-12-10 09:35:23	27.9	59.3	5998	0	-100	0	0	0	0	0	0	0	0	0	0	0
30	KF00210102	KF0021	1	2 2021-12-10 09:35:23	28	59.7	6423	0	-100	0	0	0	0	0	0	0	0	0	0	0
31	KF00210103	KF0021	1	3 2021-12-10 09:35:23	27.6	60.6	5959	0	-100	0	0	0	0	0	0	0	0	0	0	0
32	KF00210101	KF0021	1	1 2021-12-10 09:36:23	28	59.4	6005	0	-100	0	0	0	0	0	0	0	0	0	0	0
33	KF00210102	KF0021	1	2 2021-12-10 09:36:23	28	59.6	6417	0	-100	0	0	0	0	0	0	0	0	0	0	0
34	KF00210103	KF0021	1	3 2021-12-10 09:36:23	27.6	59.8	5955	0	-100	0	0	0	0	0	0	0	0	0	0	0

133336	KF00210102	KF0021	1	2 2022-01-10 06:12:50	26.5	59.4	4698	0	-100	758	747	754	1038	1043	1041	1021	1022	1069	958
133337	KF00210101	KF0021	1	1 2022-01-10 06:13:50	28.6	56.3	4505	0	-100	1515	1467	1405	1359	1353	1377	1362	1359	1371	1355
133338	KF00210102	KF0021	1	2 2022-01-10 06:13:50	27.3	60.2	4793	0	-100	854	856	846	775	700	691	694	750	700	698
133339	KF00210101	KF0021	1	1 2022-01-10 06:14:50	28	52.6	3744	0	-100	1364	1367	1355	1360	1360	1360	1359	1359	1363	1367
133340	KF00210102	KF0021	1	2 2022-01-10 06:14:50	26.5	54.2	3998	0	-100	696	737	743	1879	1713	1464	1258	1236	1226	1313
133341	KF00210101	KF0021	1	1 2022-01-10 06:15:50	28.4	55.5	3905	0	-100	1353	1356	1363	1362	1363	1361	1342	1369	1352	1347
133342	KF00210102	KF0021	1	2 2022-01-10 06:15:50	26.5	56	4229	0	-100	1235	1073	1010	578	584	572	586	592	588	570
133343	KF00210101	KF0021	1	1 2022-01-10 06:16:50	29.5	57	4290	0	-100	1364	1358	1358	267	382	25	0	-100	563	349
133344	KF00210101	KF0021	1	1 2022-01-10 06:17:50	28.8	54.2	4342	0	-100	1230	461	1023	-100	-100	-100	0	-100	0	0
133345	KF00210101	KF0021	1	1 2022-01-10 06:18:50	27	54.1	3687	0	-100	0	0	-100	224	235	263	277	0	0	0
133346	KF00210101	KF0021	1	1 2022-01-10 06:19:50	27.3	56.7	4098	0	-100	2	0	0	1273	1272	1272	1269	1269	1272	1272

(A)

	getTime	j01	j02	j03
0	2021-12-10 09:24:23	0	0	0
1	2021-12-10 09:24:24	0	0	0
2	2021-12-10 09:24:25	0	0	0
3	2021-12-10 09:24:26	0	0	0
4	2021-12-10 09:24:27	0	0	0
...	...	...	...	...
2667355	2022-01-10 06:20:45	1272.0	-100	1355.0
2667356	2022-01-10 06:20:46	1269.0	1341.0	1343.0
2667357	2022-01-10 06:20:47	1269.0	1347.0	1355.0
2667358	2022-01-10 06:20:48	1272.0	0	0
2667359	2022-01-10 06:20:49	1272.0	0	1272.0

(B)

	getTime	j01	j02	j03		getTime	j01	j02	j03
0	2021-12-10 09:24:23	NaN	NaN	NaN	0	2021-12-10 09:24:23	56	81	33
1	2021-12-10 09:24:24	NaN	NaN	NaN	1	2021-12-10 09:24:24	56	81	33
2	2021-12-10 09:24:25	NaN	NaN	NaN	2	2021-12-10 09:24:25	56	81	33
3	2021-12-10 09:24:26	NaN	NaN	NaN	3	2021-12-10 09:24:26	56	81	33
4	2021-12-10 09:24:27	NaN	NaN	NaN	4	2021-12-10 09:24:27	56	81	33
...	...	...	...	...	...	...	...	...	...
2667355	2022-01-10 06:20:45	1272.0	NaN	NaN	2667355	2022-01-10 06:20:45	1272	1058	356
2667356	2022-01-10 06:20:46	1269.0	NaN	NaN	2667356	2022-01-10 06:20:46	1269	1058	356
2667357	2022-01-10 06:20:47	1269.0	NaN	NaN	2667357	2022-01-10 06:20:47	1269	1058	356
2667358	2022-01-10 06:20:48	1272.0	NaN	NaN	2667358	2022-01-10 06:20:48	1272	1058	356
2667359	2022-01-10 06:20:49	1272.0	NaN	NaN	2667359	2022-01-10 06:20:49	1272	1058	356

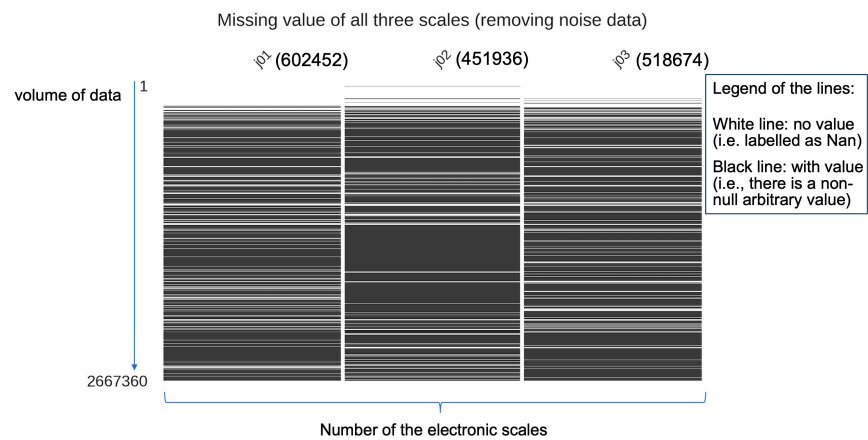
Figure 5. Sensor data vs. raw data: (A) Sensor data CSV file downloaded from the server containing data from 5 additional sensors and weight sensor data; (B) raw data used as preprocessing inputs.

The raw data consisting of timestamps (in seconds) and weight values (in grams) after the preprocessing are shown in Figure 6 below.

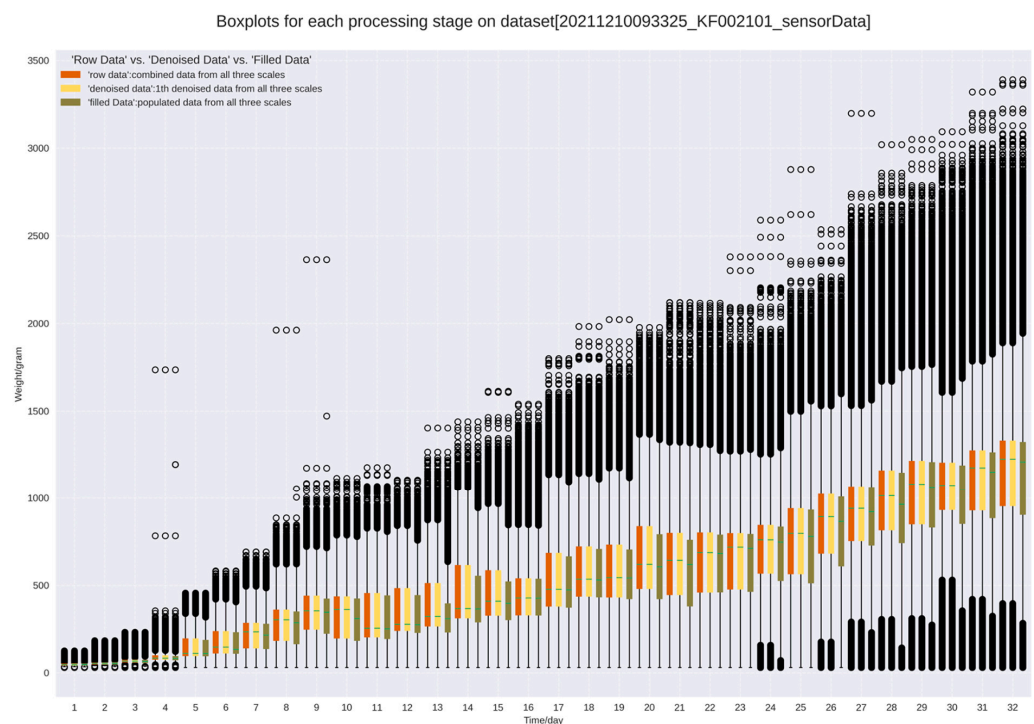
Figure 6. Raw data after preprocessing: (A) Noise reduction results based on critical values; (B) using MissForest to fill all null values.

### 3.2. Analysis of Data Preprocessing

From previous research, we know that electronic scales collect a large number of zeros when the broiler does not walk up to the scale, or when there is no touching at all. And when the scale is touched, various other noise values are generated. Thus, after removing the noise values, there is a large number of missing values, almost 1/5 of the data shown in Figure 7 below. Since the weight of the missing values is usually large and destroys the real-time performance, they have to be filled in. We used MissForest, which has excellent performance, to multiply impute the table of numbers rich in missing values, and we used boxplots to observe five statistics, including the minimum, first quartile (Q1), median, third quartile (Q3), and maximum of all data before and after the populations, and found that the overall distribution of the data did not change significantly, as shown in Figure 8. The need for padding and the effectiveness of the algorithm were demonstrated.



**Figure 7.** Visualization of the missing values of the three scales' data: j01 represents the first scale's data, with a missing data volume of 602,452 lines (seconds); j02 represents the second scale, and j03 represents the third; the full data volume is 2,667,360 lines.



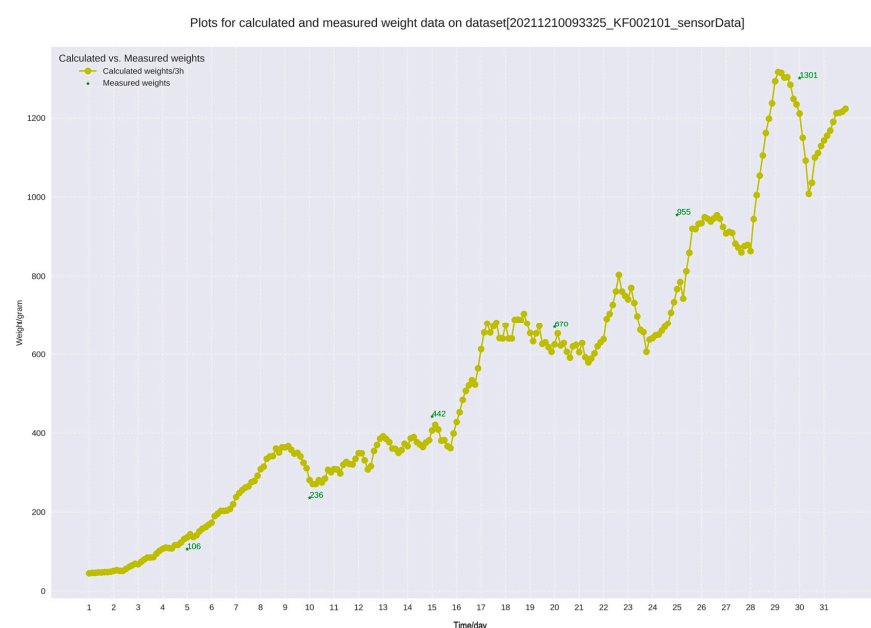
**Figure 8.** Visualization of raw data, data based on critical-value noise reduction, and populated data: The 1st denoised data represent the critical-value noise reduction operation.

### 3.3. Results of the Calculation of the Average Live Weight of Individual Broilers

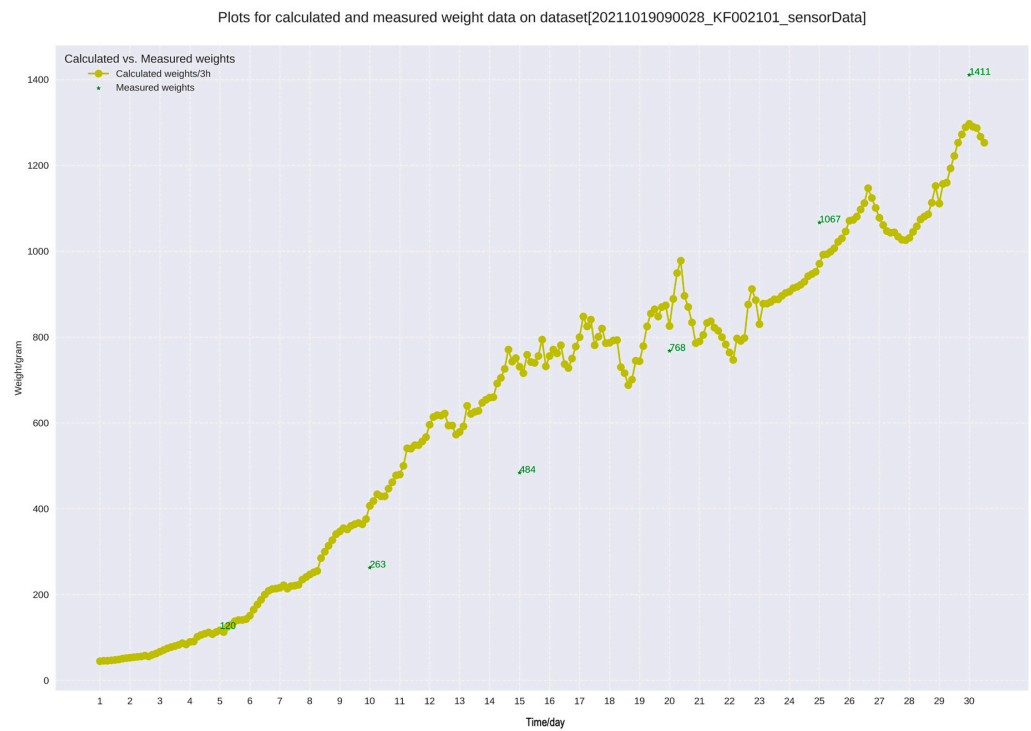
It is difficult to weigh each chicken during a feeding cycle, especially when almost every farm has thousands of chickens. The traditional method is to weigh a random number of chickens in different areas of the farm each week to represent the broiler growth on the sampled farm on that day. However, such a small sample measurement is usually biased and inaccurate. Either way, the truck collects the entire batch of broilers at once on the day of delivery (usually between days 28 and 32) and then calculates the gross weight of all of the broilers. In this study, a checkpoint was set every five days to validate the proposed framework, with days 5, 10, 15, 20, 25, and 30 serving as reference days to calculate differences and compare results, as shown in Figures 9 and 10. However, due to the high cost of the inputs, only two such datasets were available, in addition to 111 datasets containing only the delivery day measurement data, which were used for validation as shown in Figure 11. Although all of the broilers were measured at one time in this case, the reference value of the data on the delivery date was not high, due to human interference during transportation. In particular, in order to avoid fines, staff usually take unqualified broilers away at will. Nevertheless, the validation results of this method on 111 datasets showed that 67.57% of the datasets had calculation errors below 0.2, of which 40 farms had SMAPE values below 0.1 and 7 farms had SMAPE values below 0.03. This demonstrates the practical feasibility of this method.

### 3.4. Analysis of the Calculation of the Average Live Weight of Individual Broilers

As our experimental data included both time and weight values, we used a two-dimensional Gaussian mixture model to perform a first-round clustering on the data after the above filling. Obviously, it makes sense to discuss the time values of these rapidly growing and changing broilers. The number of hyperparametric components  $k$  of the Gaussian mixture model was determined by calculating the Bayesian information criterion (BIC) (here, the BIC values were calculated exhaustively for the first 20 components under four covariance types: differential, spherical, joint, and full, as shown in Figure 12 below). Observing the decreasing trend of the BIC and the final calculation results, we found that it is reasonable to set the  $k$  value to 20, which not only reduces the computational overhead of the system but also ensures a high accuracy rate. And the cluttered data were decomposed into twenty or fewer components, as shown in Figure 13 below.

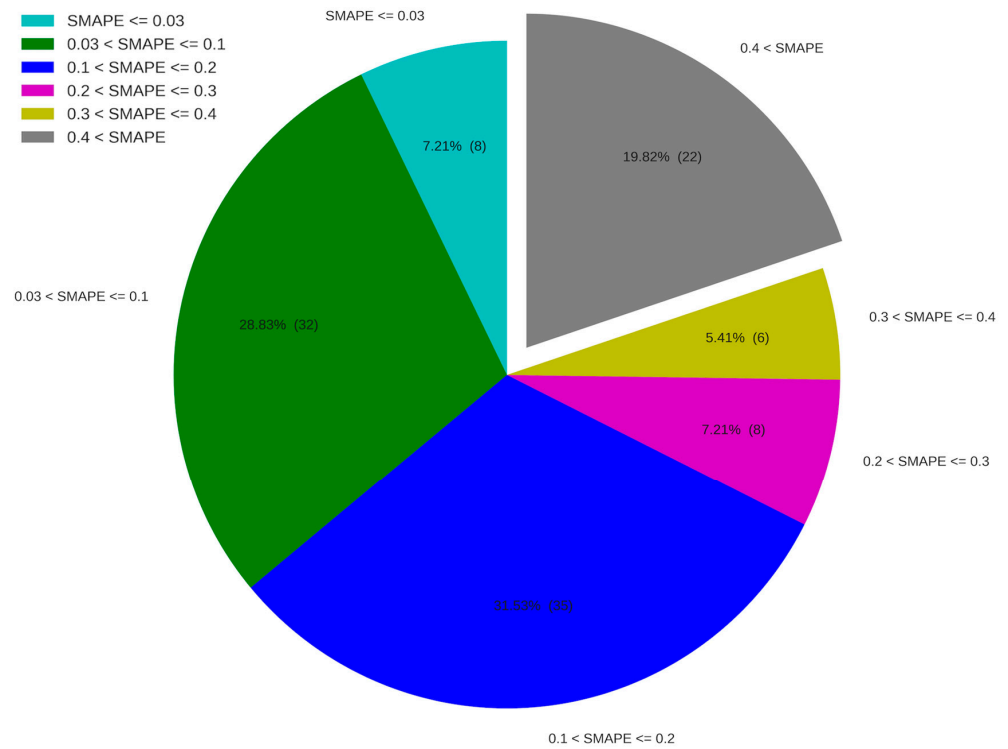


**Figure 9.** Comparison of the measured values (1) at 5-day intervals with the results calculated by the algorithm.

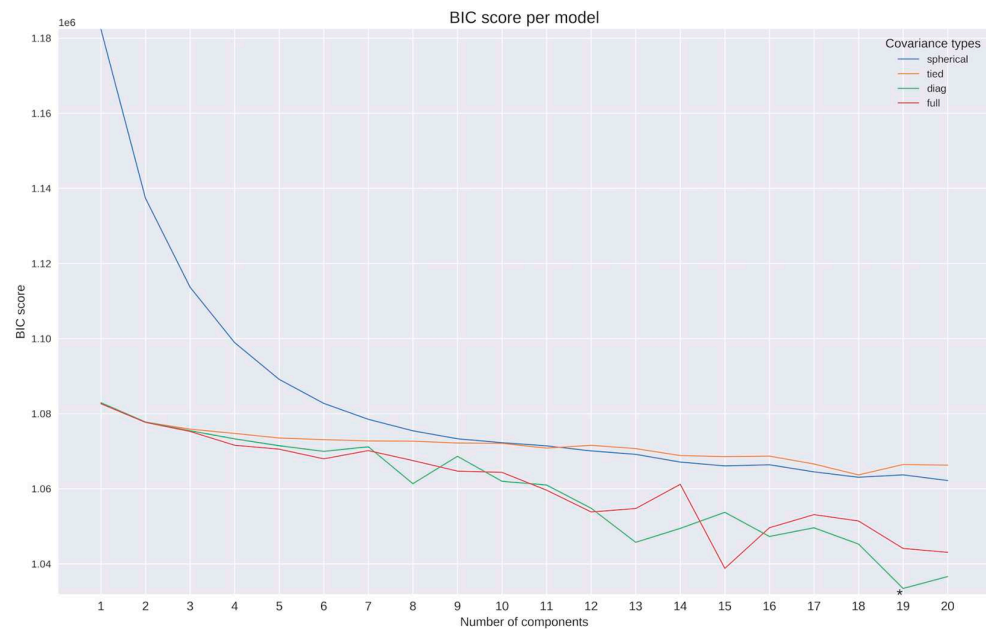


**Figure 10.** Comparison of the measured values (2) at 5-day intervals with the results calculated by the algorithm.

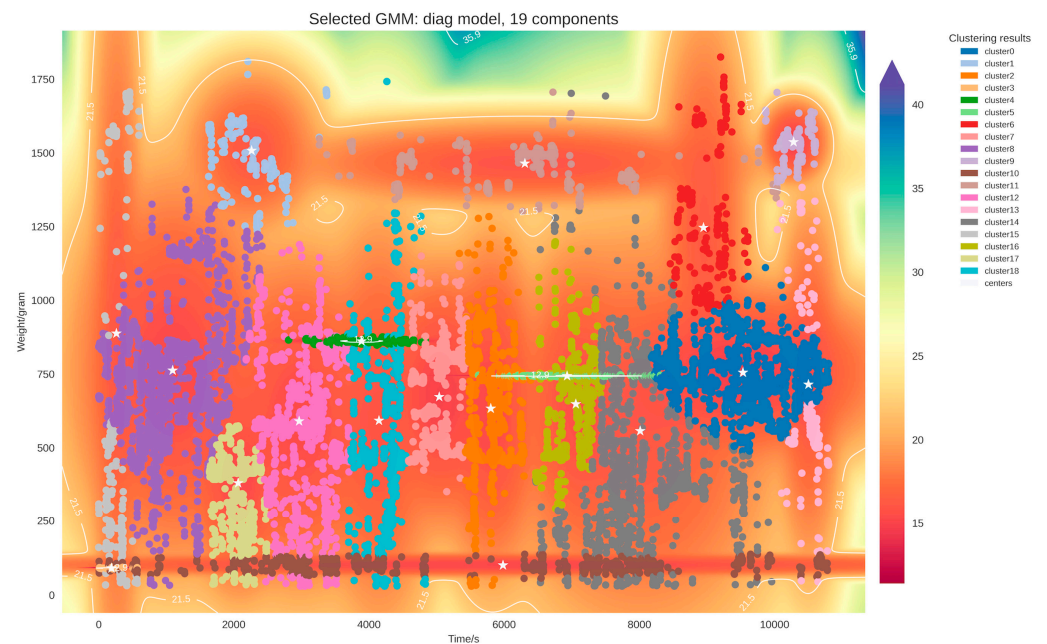
A pie chart of the errors between the results calculated by the algorithm and the manually measured values



**Figure 11.** Comparison of the delivery-day measurements of the other 111 datasets with the results calculated by the algorithm.



**Figure 12.** BIC values for different variance types; \* marks the position of the minimum value.



**Figure 13.** Two-dimensional Gaussian mixture model clustering results. Here, the number of components corresponding to the BIC minimum was set to 19. The different colors represent different clusters, and the white pentagrams are the means for each cluster.

A powerful Isolation Forest was introduced to make the data of the components more centralized. Outliers, in general, were fewer than regular observations and differ in terms of values (they were far from regular observations in the feature space), as shown in Figure 14 below. The average of each filtered component was then calculated as the representative value of the overall data, as shown by the white five-pointed stars in Figure 13 above. However, these representative values were not equalized. Instead, weights were calculated for each value based on the ratio of the total number of values in its corresponding cluster to the total number of values in all clusters.

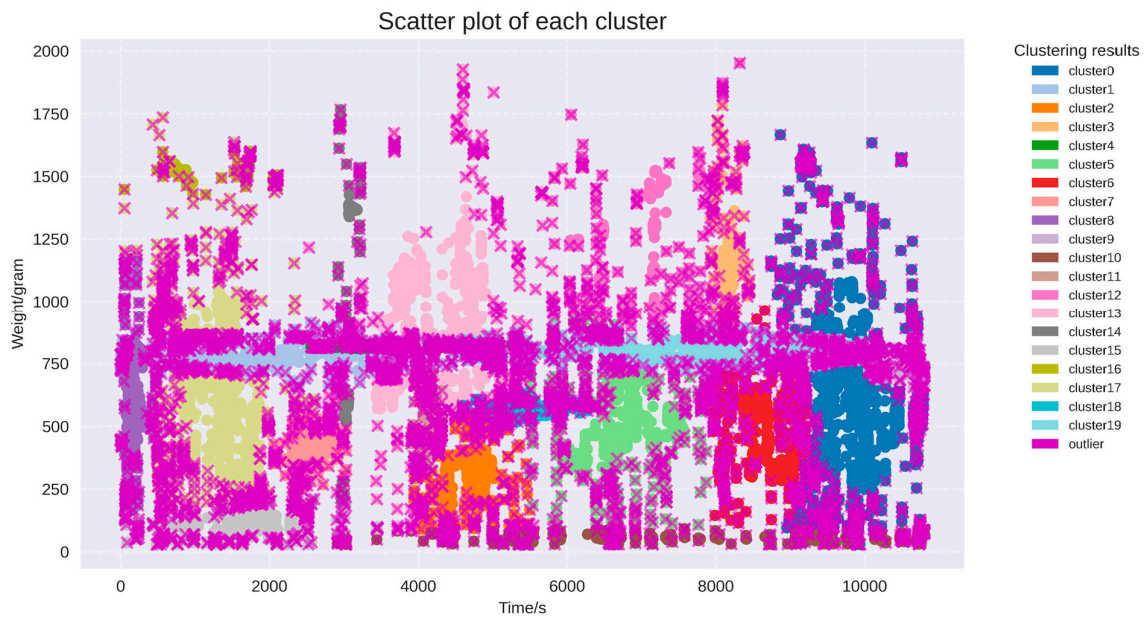


Figure 14. Outliers filtered by Isolation Forest; the purple Xs are outliers.

To screen the target values, we performed quadratic clustering of the above feature values based on the effective density (e.g., the electronic scale was restored from the perturbed state to the stationary state to obtain the effective measurement data). The range of feature data was significantly reduced after two iterations of the OPTICS process. Then, as shown in Figure 15 below, the datum with the highest weight among the feature values was selected as the target weight value.

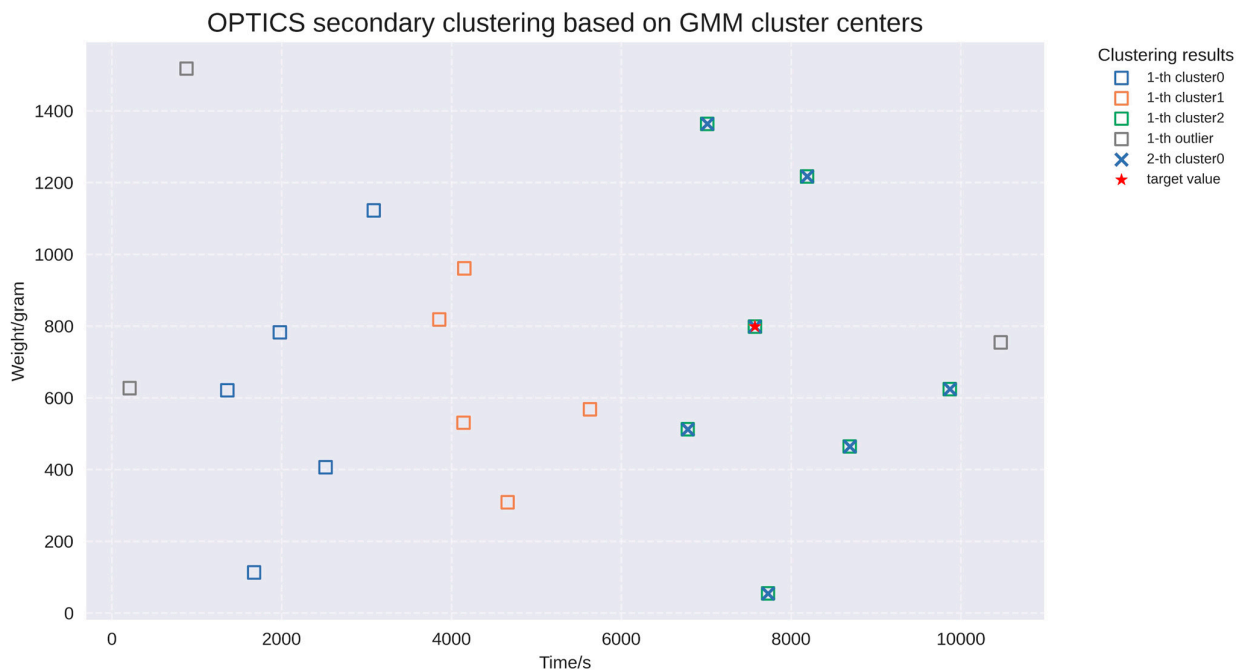
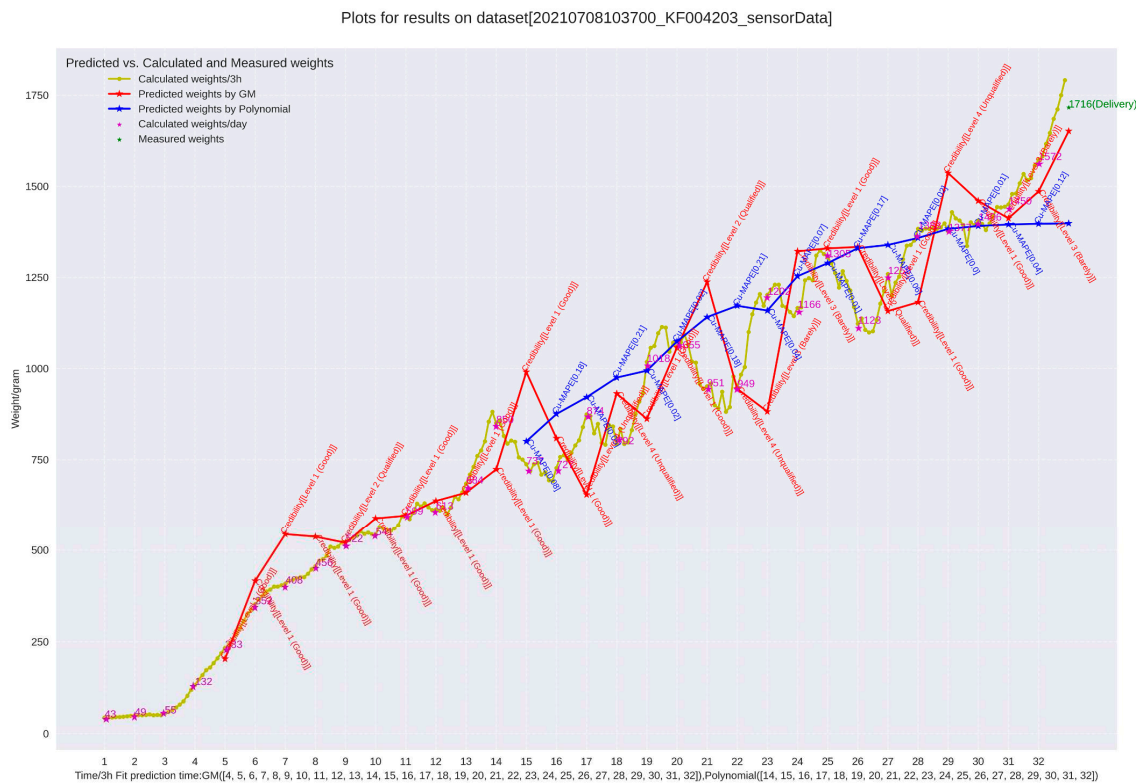


Figure 15. Secondary clustering based on OPTICS. The boxes indicate the first iteration, the X indicates the second iteration, and the red pentagram indicates the final weight.

### 3.5. Results of the Prediction of the Average Live Weight of Individual Broilers

We labeled both the medium- and long-term prediction results using multinomial regression and the short-term prediction results based on the gray model (1, 1) in real time on the same image. As shown in Figure 16 below, the medium- and long-term prediction accuracy was still evaluated by the SMAPE values, while the gray model was evaluated by four prediction accuracy classes based on the C-value (a posteriori difference ratio value) and the *p*-value (probability of small error), as follows: good ( $p \geq 0.95, C \leq 0.35$ ), qualified ( $0.80 \leq p < 0.95, 0.35 < C \leq 0.50$ ), barely qualified ( $0.70 \leq p < 0.80, 0.50 < C \leq 0.65$ ), and unqualified ( $p < 0.70, C > 0.65$ ).

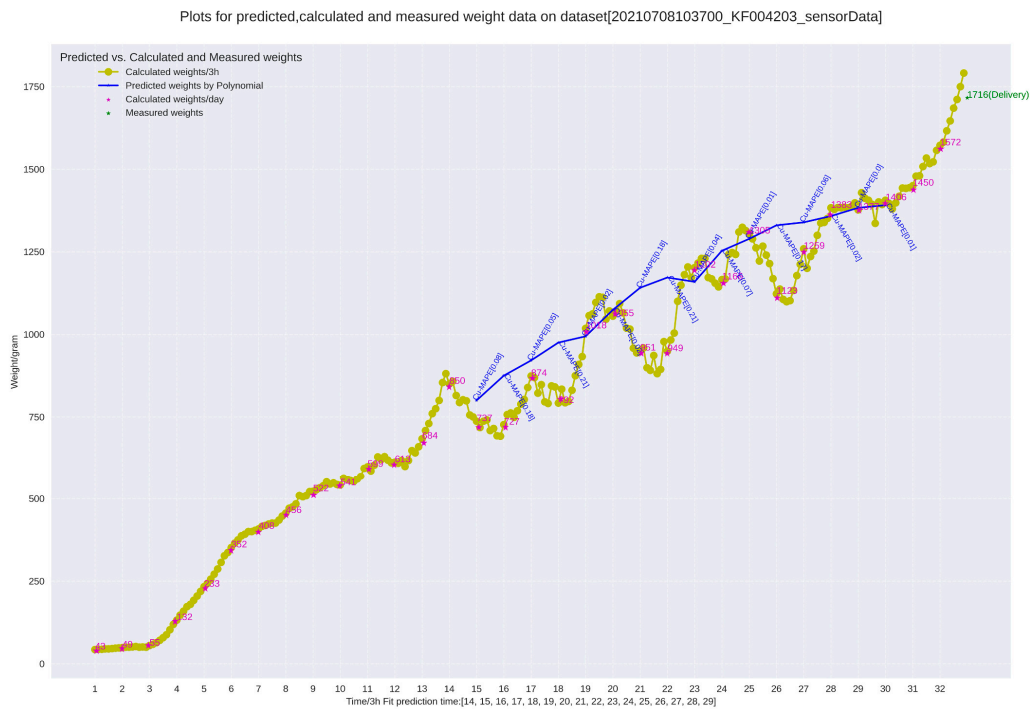


**Figure 16.** Monitoring and prediction results. Here, the yellow line is the individual broiler body weight value calculated by the body weight monitoring model, the purple pentagram (with the purple text) is the representative body weight value for each day, the green pentagram is the true mean body weight value measured manually on the day of slaughter, the blue line is the polynomial regression prediction value, and the red line is the GM (1, 1) regression prediction value.

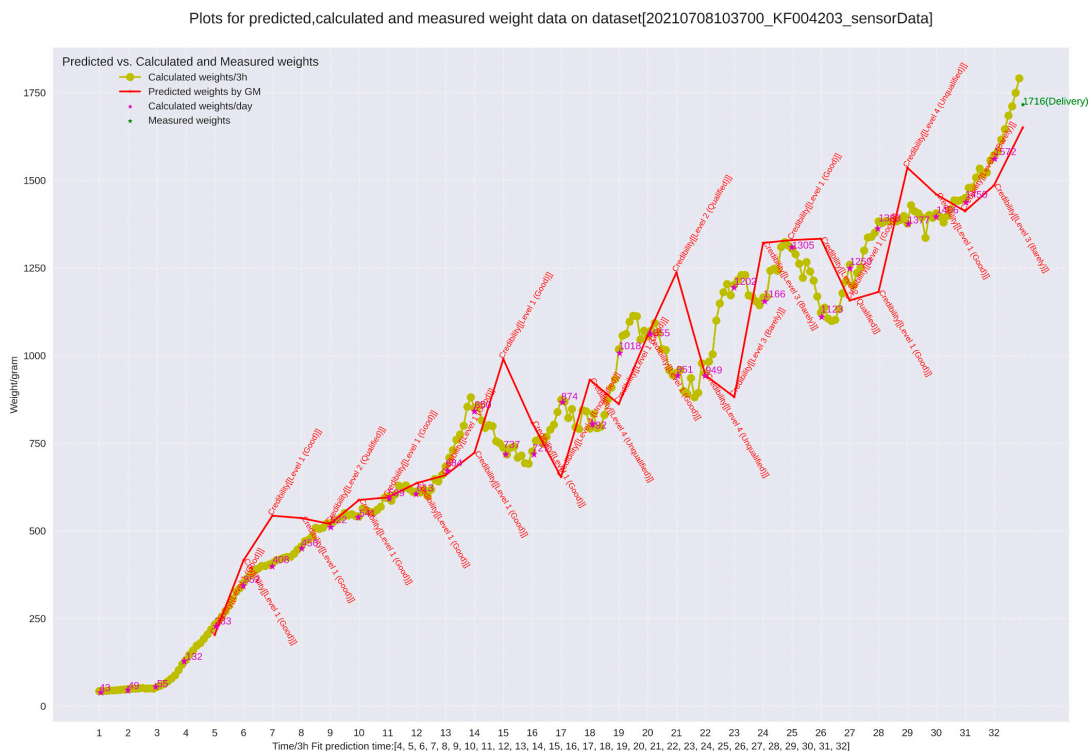
### 3.6. Analysis of the Prediction of the Average Live Weight of Individual Broilers

In the polynomial fit, the predicted values for day 15 were first fitted using the calculated weights for the first 14 days of age, each predicted value was compared to the actual calculated value for the next day, and the SMAPE value was calculated, and so on until the day of delivery. As shown in Figure 17 below, if the cumulative SMAPE value exceeded the set threshold, the model was refitted using all data prior to the current day of age.

In the GM (1, 1) model, each predicted value is the result of modeling predictions based on the weight values of the previous 4 days of age. In addition, each prediction is labeled with a confidence level for the breeder’s reference, as shown in Figure 18 below.



**Figure 17.** Body weight predicted by a polynomial regression model with a feedback regulation mechanism.



**Figure 18.** Short-term forecasting based on the GM (1, 1) model.

#### 4. Conclusions

Despite the fact that broiler rearing has become highly scientific and standardized, broiler weight management has long been largely dependent on farmers' experience and expertise. In this study, artificial intelligence machine learning algorithms were used to develop cost-effective solutions to improve the accuracy of broiler weight monitoring and

prediction based on a simple electronic scale device to support intelligent production decisions, thereby saving labor and resources, improving production, and increasing income.

The innovation of this study is the use of machine learning algorithms based on Gaussian mixture models, Isolation Forest, and OPTICS, which simulate the natural process of weighing and obtaining reliable readings using electronic scales in real life through a two-stage clustering and noise reduction process. In addition, a polynomial fitting model and a gray model were combined to achieve real-time weight prediction over long periods of time. Instead of using video cameras and image processing algorithms, we have pioneered a new scheme for calculating and predicting the average body weight of an individual broiler at low cost, with program code that can be run by an ordinary computer and graphs that can be understood by the average person, using only a simple electronic scale and data analysis algorithms.

Future research could include more factors that can be collected on modern broiler farms to efficiently and effectively estimate broiler weights in real time and derive broiler weight distributions to adjust decisions related to broiler feeding, health improvement, harvesting, and chicken meat production planning for overall resource management and income optimization. Secondly, this system could also be used for the weight estimation of other farm animals.

**Author Contributions:** Conceptualization, J.M. and J.S.; Data Curation, J.S.; Methodology, P.L.; Software, J.S.; Supervision, J.M. and J.S.; Validation, J.M.; Visualization, P.L.; Writing—Original Draft, P.L.; Writing—Review and Editing, J.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data described in this study are accessible from the corresponding author upon request. The data is not publicly available since it refers to farm operation and management information, and because confidentiality agreements with commercial farms have been signed, it is not disclosed until expressly asked.

**Acknowledgments:** We sincerely thank EMOTION Co. for providing equipment and site information, as well as Sunwoo Ko for his methodological guidance and consistent support.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Topal, M.; Bolukbasi, Ş.C. Comparison of Nonlinear Growth Curve Models in Broiler Chickens. *J. Appl. Anim. Res.* **2008**, *34*, 149–152. [[CrossRef](#)]
2. Moharrery, A.; Mirzaei, M. Growth Characteristics of Commercial Broiler and Native Chickens as Predicted by Different Growth Functions. *J. Anim. Feed Sci.* **2014**, *23*, 82–89. [[CrossRef](#)]
3. Rizzi, C.; Contiero, B.; Cassandro, M. Growth Patterns of Italian Local Chicken Populations. *Poult. Sci.* **2013**, *92*, 2226–2235. [[CrossRef](#)] [[PubMed](#)]
4. Mouffok, C.E.; Semara, L.; Belkasmi, F. Comparison of Some Nonlinear Functions for Describing Broiler Growth Curves of Cobb500 Strain. *Poult. Sci. J.* **2019**, *7*, 51–61.
5. De Wet, L.; Vranken, E.; Chedad, A.; Aerts, J.-M.; Ceunen, J.; Berckmans, D. Computer-Assisted Image Analysis to Quantify Daily Growth Rates of Broiler Chickens. *Br. Poult. Sci.* **2003**, *44*, 524–532. [[CrossRef](#)]
6. Chedad, A.; Aerts, J.-M.; Vranken, E.; Lippens, M.; Zoons, J.; Berckmans, D. Do Heavy Broiler Chickens Visit Automatic Weighing Systems Less than Lighter Birds? *Br. Poult. Sci.* **2003**, *44*, 663–668. [[CrossRef](#)] [[PubMed](#)]
7. Mollah, M.B.R.; Hasan, M.A.; Salam, M.A.; Ali, M.A. Digital Image Analysis to Estimate the Live Weight of Broiler. *Comput. Electron. Agric.* **2010**, *72*, 48–52. [[CrossRef](#)]
8. Mortensen, A.K.; Lisouski, P.; Ahrendt, P. Weight Prediction of Broiler Chickens Using 3D Computer Vision. *Comput. Electron. Agric.* **2016**, *123*, 319–326. [[CrossRef](#)]
9. Amraei, S.; Abdanan Mehdizadeh, S.; Salari, S. Broiler Weight Estimation Based on Machine Vision and Artificial Neural Network. *Br. Poult. Sci.* **2017**, *58*, 200–205. [[CrossRef](#)]

10. Aydin, A.; Bahr, C.; Viazzi, S.; Exadaktylos, V.; Buyse, J.; Berckmans, D. A Novel Method to Automatically Measure the Feed Intake of Broiler Chickens by Sound Technology. *Comput. Electron. Agric.* **2014**, *101*, 17–23. [[CrossRef](#)]
11. Fontana, I.; Tullo, E.; Peña Fernández, A.; Berckmans, D.; Koenders, E.; Vranken, E.; Mckinstry, J.; Butterworth, A.; Guarino, M. Frequency Analysis of Vocalisation in Relation to Growth in Broiler Chicken. *Precis. Livest. Farming* **2015**, *15*, 174–182.
12. Fontana, I.; Tullo, E.; Butterworth, A.; Guarino, M. An Innovative Approach to Predict the Growth in Intensive Poultry Farming. *Comput. Electron. Agric.* **2015**, *119*, 178–183. [[CrossRef](#)]
13. Fontana, I.; Tullo, E.; Carpentier, L.; Berckmans, D.; Butterworth, A.; Vranken, E.; Norton, T.; Berckmans, D.; Guarino, M. Sound Analysis to Model Weight of Broiler Chickens. *Poult. Sci.* **2017**, *96*, 3938–3943. [[CrossRef](#)] [[PubMed](#)]
14. Abdel-Kafy, E.-S.M.; Ibraheim, S.E.; Finzi, A.; Youssef, S.F.; Behiry, F.M.; Provolò, G. Sound Analysis to Predict the Growth of Turkeys. *Animals* **2020**, *10*, 866. [[CrossRef](#)]
15. Johansen, S.V.; Bendtsen, J.D.; Jensen, M.R.; Mogensen, J. Broiler Weight Forecasting Using Dynamic Neural Network Models with Input Variable Selection. *Comput. Electron. Agric.* **2019**, *159*, 97–109. [[CrossRef](#)]
16. Lee, C.C.; Adom, A.H.; Markom, M.A.; Tan, E.S.M.M. Automated Chicken Weighing System Using Wireless Sensor Network for Poultry Farmers. *IOP Conf. Ser. Mater. Sci. Eng.* **2019**, *557*, 012017. [[CrossRef](#)]
17. Ma, W.; Li, Q.; Li, J.; Ding, L.; Yu, Q. A Method for Weighing Broiler Chickens Using Improved Amplitude-Limiting Filtering Algorithm and BP Neural Networks. *Inf. Process. Agric.* **2021**, *8*, 299–309. [[CrossRef](#)]
18. Wang, C.-Y.; Chen, Y.-J.; Chien, C.-F. Industry 3.5 to Empower Smart Production for Poultry Farming and an Empirical Study for Broiler Live Weight Prediction. *Comput. Ind. Eng.* **2021**, *151*, 106931. [[CrossRef](#)]
19. Birzniece, I.; Andersone, I.; Nikitenko, A.; Balina, S.; Kikans, A. Time Series Forecast Model Application for Broiler Weight Prediction Using Environmental Factors. In Proceedings of the 2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), Maldives, 16–18 November 2022; pp. 1–7.
20. Stekhoven, D.J.; Bühlmann, P. MissForest—Nonparametric Missing Value Imputation for Mixed-Type Data. *Bioinformatics* **2012**, *28*, 112–118. [[CrossRef](#)]
21. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum Likelihood from Incomplete Data Via the EM Algorithm. *J. R. Stat. Soc. Ser. B Methodol.* **1977**, *39*, 1–22. [[CrossRef](#)]
22. Liu, F.T.; Ting, K.M.; Zhou, Z.-H. Isolation Forest. In Proceedings of the 2008 Eighth IEEE International Conference on Data Mining, Pisa, Italy, 15–19 December 2008; pp. 413–422.
23. Ankerst, M.; Breunig, M.M.; Kriegel, H.-P. OPTICS: Ordering Points to Identify the Clustering Structure. *ACM Sigmod Rec.* **1999**, *28*, 49–60. [[CrossRef](#)]
24. Stigler, S.M. Gergonne’s 1815 Paper on the Design and Analysis of Polynomial Regression Experiments. *Hist. Math.* **1974**, *1*, 431–439. [[CrossRef](#)]
25. Golub, G.H.; Loan, C.F.V. An Analysis of the Total Least Squares Problem. *SIAM J. Numer. Anal.* **1980**, *17*, 883–893. [[CrossRef](#)]
26. Deng, J. Control Problems of Grey System, Systems and Control Letters. *Syst. Control Lett.* **1982**, *1*, 288–294.
27. Dong, X.; Zhang, S. Impact Factor Dynamic Forecasting Model for Management Science Journals Based on Grey System Theory. *Open J. Soc. Sci.* **2015**, *3*, 22–25. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.