Article:

Tashim, Najeebah Az-Zahra, Lim, Tiong Hoo, Zariful, Muhammad WafiqAbd et al. (1 more author) (2025) Regression-based artificial intelligence length and weight estimation for sustainable prawn aquaculture. Smart Agricultural Technology. 101089. ISSN 2772-3755

https://doi.org/10.1016/j.atech.2025.101089

Smart Agricultural Technology 12 (2025) 101089

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

Regression-based artificial intelligence length and weight estimation for sustainable prawn aquaculture

Najeebah Az-Zahra Tashim [a], Tiong Hoo Lim [a], Wafiq Zariful [a], Pengcheng Liu [b],*

[a] Electrical and Electronic Engineering, Universiti Teknologi Brunei, Jalan Tungku Link, Gadong, Bandar Seri Begawan, BE1410, Brunei Muara, Brunei Darussalam
[b] Computer Science, University of York, Deramore Lane, York, YO10 5GH, United Kingdom

ARTICLE INFO

Keywords:
Prawn weigh prediction
Aquaculture
YoloV8
Linear regression
Non-linear regression
Length-weight
Segmentation
Deep learning
SDG1

ABSTRACT

The need for sustainable aquaculture practices has become very important to ensure sufficient production in addressing the increasing global demand for seafood. In this context, accurately assessing the size and weight of prawns is pivotal for efficient farming and resource utilization, allowing farmers to make informed decisions and productions. The integration of advanced AI algorithms into aquaculture practices holds great promise for fostering sustainability, thereby enhancing the overall productivity and resilience of prawn farming in the face of growing global challenges. This paper compares different length-weight regression techniques to estimate the weight of prawns and proposed a novel Regression-based Artificial Intelligence Biomass Estimation (RAIBE) systems for prawn aquaculture. RAIBE leverages deep learning and regression models to estimate the weight from images captured from a mobile device. The proposed methodology employs YOLOv8 with Segmentation for precise prawn identification. A unique biomarker is applied to estimate the length information. Subsequently, a polynomial based regression model is selected to correlate prawn length with actual weights, utilising comprehensive datasets collected under real-world farm conditions. As many different regression approaches have been proposed for the length-weight relationship, four commonly used approaches have been analysed. Results from extensive statistical analysis revealed that the modified polynomial regression with correction factor provides the best weight prediction. The integration of these techniques has equipped farmers with a reliable tool for predicting prawn weight during the sampling process, thereby minimizing stress on the prawns, and optimizing the segregation process.

1. Introduction

Sustainable aquaculture practices are crucial in meeting the escalating global demand for seafood. Ensuring optimal production, especially considering the influence of local breeding environments, depends on precise sampling and accurate size and weight measurement. Ye et al. emphasised the considerable spatial variations in oceanographic conditions across the Middle East's coastal waters, underscoring their potential impact on localised fish stocks [1]. Traditional weight assessment methods, involving labour-intensive physical measurements, are time-consuming. Conventional prawn sampling techniques, such as baited traps, visual census methods, and hand netting, often proven challenging and inaccurate [2].

In response to these limitations, alternative technological approaches like precision aquaculture offer promising solutions. Precision aquaculture leverages advanced Artificial Intelligence (AI) techniques, Internet of Things (IoT), and data-driven approaches to transition from traditional practices to knowledge-based methodologies, ensuring profitability, sustainability, and environmental protection. The utilisation of IoT sensors and edge devices, such as the Raspberry Pi, enables in-situ processing, facilitating immediate decision-making with data transmission to the farmer over the network. Machine learning algorithms and image processing have demonstrated efficacy in disease detection, predicting the growth and survivable of farmed fish, as well as forecasting pond water quality [3–5]. However, despite its proven potential, the practical implementation and deployment of these AI methods on actual aquaculture farms has been constrained by computational complexity, including challenges in edge-processing capabilities and prediction accuracy.

To bridge this gap between AI potential and real-world application, particularly in biomass and size estimation, it is essential to explore methods that are computationally efficient yet practically useful for deployment on resource-constraint edge devices. While many AI systems

* Corresponding author.
E-mail address: pengcheng.liu@york.ac.uk (P. Liu).

https://doi.org/10.1016/j.atech.2025.101089
Received 24 March 2025; Received in revised form 22 May 2025; Accepted 8 June 2025
Available online 13 June 2025

rely on complex deep learning models for high-level tasks such as disease classification, simpler statistical approaches like regression analysis are better suited for continuous variables such as weight and length estimation, which require lower processing power and are easier to deploy at the farm level. This highlights a shift from AI for classification problems to AI-supported regression models for numerical prediction, tailored to on-site aquaculture needs.

To overcome these challenges, regression analysis methods, known for their low processing requirements, have gained popularity in various domains, including weather forecasting, earthquake prediction, and stock market analysis. Additionally, they have been applied for length and weight inference in fishes and prawns [6], utilising formulas such as the growth model based on the length-weight relationship, $W = aL^b$ [7] in linear regression. Nevertheless, the relationship between length and weight can be inherently non-linear, thus the reliance on linear regression models may not yield accurate results.

The primary contribution of this paper proposes a Regression-based Artificial Intelligence Biomass Estimation (RAIBE) model, which integrates image processing and deep learning algorithms with a non-linear regression-based model to estimate prawn weight using images from mobile devices. The proposed approach employed YOLOv8 for prawn detection and segmentation to enhance size estimation accuracy. Each prawn is assigned a unique biomarker, serving as a reference for length and weight information. The secondary contribution of the paper provides a comprehensive comparison between different regression models using actual farm data to identify the most accurate model for prawn weight estimation. The results of the paper also highlight that the commonly used length-weight formula ($W = aL^b$) does not consistently predict the optimal weight predictions for prawns and instead incurred additional processing.

The subsequent sections of the paper are organized as follows: Section 2 provides an overview of related works in the research area, Section 3 presents the proposed architecture to predict the weight of the prawn, and Section 4 details the experimental approaches to determine and compare different linear and non-linear regression model to predict the weight. Section 5 compares and discusses various regression models results. Finally, Section 6 concludes the research, summarising the finding and recommended regression model for weight prediction in prawn.

2. Related works

There have been many research works published on weight predictions based on regression models for aquatic species. Nahavandi et al. [7] studied the length-weight relationship of giant tiger shrimp (*Penaeus monodon*) cultured in artificial seawater. They have analysed the length-weight relationship by using the formula $W = aL^b$ based on linear regression. Results revealed an isometric growth of *P. monodon* at b = 2.94, indicating proportional growth in length and weight. When the length-weight exponent b = 3.0, the body form remains constant in proportion to the length and the fish develops isometrically, resulting in an ideal shape. This study suggested the potential of using artificial water for culturing *P. monodon*. Silva et al. [8], also conducted a similar approach in identifying the length-weight relationship of cage-farmed Nile tilapia. Likewise, the b coefficient indicates isometric fish growth at b = 3.06 at $R^2 = 0.9952$.

A similar study by Waiho et al. [9], investigated the length-weight relationship of a mud spiny lobster *Panulirus polyphagus* by using a linear regression approach. The study noted a significant strong positive correlation between body length and weight of *P. polyphagus* at $R^2 > 0.94$, indicating robust growth patterns in relation to body size. They have also engaged in a sex-specific allometry study, revealing a positive allometry among males, while females displayed an opposite trend-negative allometry. These specifics are often related to reproductive function and feed resource allocation. As such, females are more likely to allocate more resources into egg production compared to males for reproduction purposes [10].

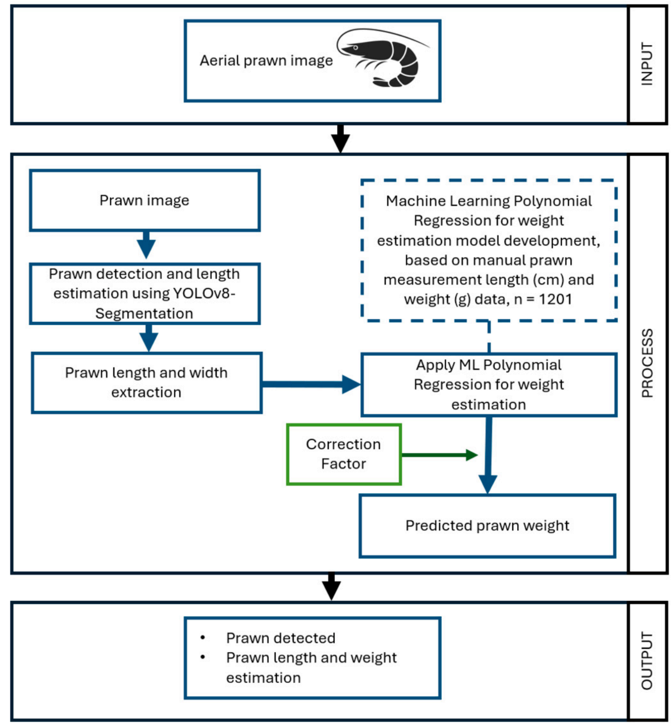

Fig. 1. Proposed Regression-based Artificial Intelligence Biomass Estimation (RAIBE) system.

Alternatively, Clain et al. [11], studied the length-weight relationship of South-Eastern Australian population of large-head hairtail (*Trichiurus lepturu*) by employing non-linear regression approach. The power regression best described the relationship between length and weight of the samples, with females revealing positive allometry where weight increases faster with length compared to males. This suggests potential differences of strategies in the allocation of feed resources.

3. Regression based AI Biomass Estimation model

The aquaculture industry depends on accurately estimating product bio-mass to ensure sustainable practices and enhanced profitability. Regression-based AI models can provide a powerful tool to allow biomass and weight predictions even based on singular parameters such as body length. In this section, we present our proposed RAIBE Model shown in Fig. 1 to estimate the biomass of the prawn in terms of length and weight.

All images in the dataset of *Macrobrachium rosenbergii* prawns were annotated, excluding the legs (chelipeds and pereiopods) to concentrate on the body segments essential for length estimation.

The system first estimates the prawn's length, which serves as the input for a regression model to predict its weight. To accurately measure length, prawn detection is required within the captured images. This detection is performed using the YOLOv8 Convolutional Neural Network (CNN) trained on a dataset of prawn images to identify and localize prawns accurately. Additionally, the YOLOv8-Segmentation model is applied to segment each detected prawn, further improving the detection precision.

Following segmentation, an oriented bounding box is generated based on the minimum area rectangle that aligns with the prawn's shape. This bounding box adapts to various angles and orientations of the prawn, enabling the systems to detect the prawn and estimate the size. The goal of the detection is to detect the length of the prawn from the tip of the head (rostrum) to the tip of the telson (tail), while the width will not be considered in this research as there was not enough data collected to consider trying out multiple linear regression with the carapace width of the prawn.

To calculate the length of the prawn, a known biomarker reference point based on ArUco marker is employed. ArUco marker is a type of fiducial marker used for scaling in object detection and estimation in computer vision [12]. The system measures the perimeter of the ArUco Marker in pixels and converts to find the pixel-to-cm ratio. This is used to calculate the length of the prawn by multiplying the pixel-to-cm ratio with the length of the bounding box of the prawn, in which the longer side of the oriented bounding box represents as the length of the prawn. Since the bounding box is rotated at an angle, a distance formula is applied between two corners of the bounding box to calculate the length of the prawn through the oriented bounding boxes,

Based on the calculated length value, the weight can be estimated using the regression model. In existing literature, most common model utilises the growth equation $W = aL^b$, where W represents weight, L as the length and a and b represent coefficients determining the relationship between length and weight [13]. Traditional linear regression identifies a and b coefficients as y-intercept and slope directly from the equation: $y = bx + a$ [8]. While power regression provides a direct estimation of coefficients a and b with its equation $y = ax^b$ [6]. For our proposed systems, we will evaluate different regression models to identify the best R^2 values and compare the accuracy of the predicted and actual weight in the next section.

4. Evaluation approaches

4.1. Stage one: data collection

The data collection for model training and validation was completely derived from actual farm conditions to ensure model applicability and reliability. For the development of the prawn detection and measurement model, a comprehensive aerial image data set was collected and processed. Specifically, the training data set comprised 8370 aerial images, and validation involved an additional 711 images, all captured from farm conditions and used in training the YOLOv8 segmentation model for detecting and segmenting prawns. A separate and independent dataset, also collected from real farm environments, included 1201 manual measurements of *M. rosenbergii* prawn lengths and weights across 3 grow-out ponds over a 7 month period, provided by ODE Aquaculture and Agriculture Company Sdn Bhd, Brunei. This dataset was instrumental in developing robust linear and non-linear regression models, including power, exponential, and polynomial regressions, for predicting prawn weight based on length. Finally, to rigorously assess the accuracy of the model, an independent set of 397 manually measured prawn data points was used, allowing comparison between the predicted and actual weights using standard accuracy metrics (RMSE, MAE and R^2).

4.2. Stage two: prawn detection

The prawn detection system stage was carried out by retrieving aerial images of the prawn. This system utilises YOLOv8-Segmentation to create set of masks that outlines the shape of the prawns. The Open CV library utilises these segmentation results to create and calculate the minimum area of a rectangle to create an oriented bounding box that follows the shape of the prawn. This allows prawn detection and size estimation at any angle or rotation of the prawn. The YOLOv8-Segmentation model, trained on the largest YOLOv8 backbone (YOLOv8x-seg), was used for this purpose. The model was trained on a dataset containing 8370 training images and 711 validation images, with an image size set to 800, for 500 epochs. Using YOLOv8, the system calculates the minimum area of the rectangle to estimate the prawn's length. The different detection annotations are as displayed in Fig. 2. A biomarker based on ArUco marker, with a side length of 4 cm and a perimeter of 16 cm, is used to determine the pixel-to-cm ratio. This ratio is then applied to the

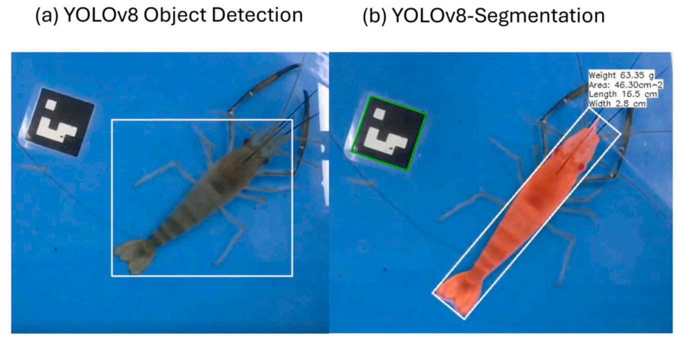

Fig. 2. Prawn detection comparison between (a) YOLOv8 object detection and (b) YOLOv8-Segmentation model, with ArUco marker reference.

length of the bounding box drawn from the segmentation to calculate the prawn's length.

4.3. Stage three: prawn length-weight model

This stage involves taking manual measurements of the prawn length and weight for model development. Exactly 1201 readings of *M. rosenbergii* length (cm), measured from the rostrum to the telson tail tip and weight (g). This dataset is used to develop regression models. Both linear and non-linear regressions, including power, exponential and polynomial regression were considered.

Linear regression stands as the direct and most widely used statistical technique for modelling relationships between variables. This method assumes a straight-line relationship between an independent variable and a dependent variable [14]. In the context of seafood weight prediction, length often serves as the independent variable, while weight is the dependent variable. By fitting a straight line through the data points, the model estimates a weight value for any given length. This allows for quick and straightforward weight predictions based on easily obtainable length measurements. Power regression, a non-linear regression model, offers a valuable tool for weight prediction in various seafood species. This approach assumes a power-law relationship between length as the independent variable and weight as the dependent variable, which are often observed in biological development patterns [15]. In contrast with exponential regression model, it assumes an exponential increase in weight with increasing length, which are potentially suitable for rapid development phases. Finally, polynomial regression allows the capturing of a more complex relationship between length and weight by including higher-order terms [16]. This flexibility is particularly beneficial, suggesting a suitable approach for modelling crustacean growth that often exhibits curvature in length-weight relationships.

4.4. Stage four: model testing

The evaluation stage was then carried out, testing the model for the weight estimation of 397 prawn data. To evaluate the accuracy of the trained model's prediction against the observed data, the models was assessed through root mean squared error (RMSE, Equation (1)), mean absolute error (MAE, Equation (2)) and R^2 (Equation (3)). The model with the lowest RMSE and MAE, combined with the highest R^2 value, is considered to be the most accurate [17]. The equations are as follows [18]:

$$RMSE = \sqrt{\frac{\sum (y_o - y_p)^2}{n}} \tag{1}$$

$$MAE = (|(yo - yp)|)/n \tag{2}$$

$$R^2 = 1 - \frac{\sum (y_o - y_p)^2}{\sum (y_o - y_m)^2} \tag{3}$$

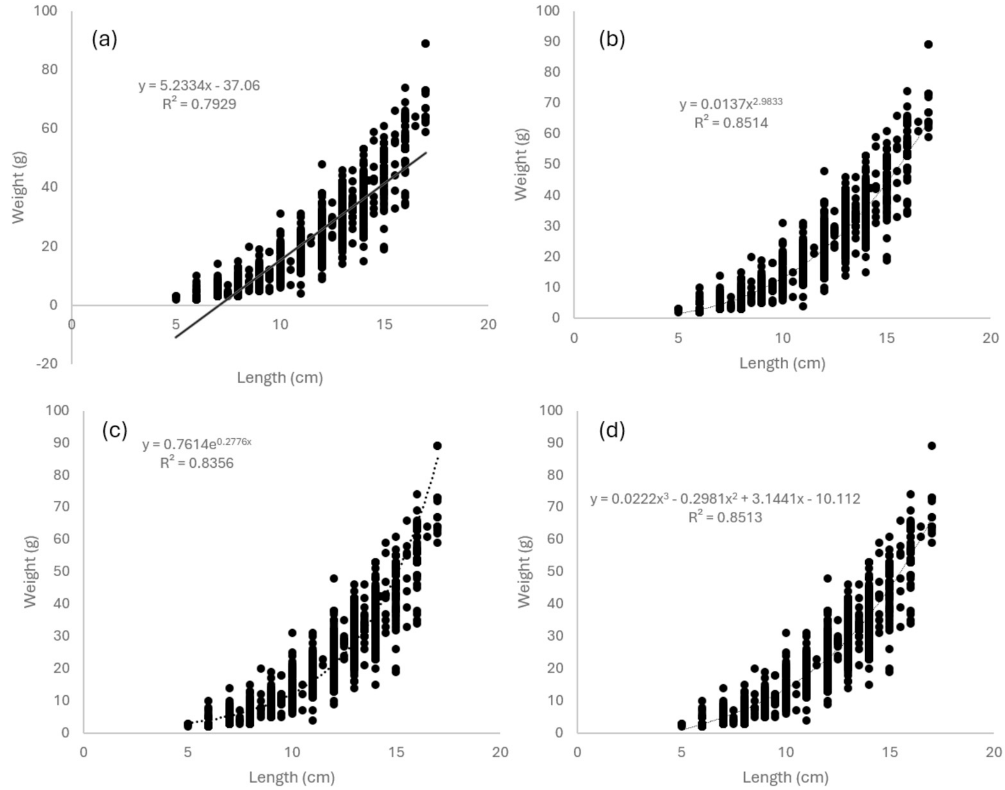

Fig. 3. Length-weight relationship of *M. rosenbergii* based on (a) Linear regression, (b) Power regression, (c) Exponential regression and (d) Polynomial regression to establish the best-fit model for weight prediction based on prawn length.

where, yo represents the observed value, yp as the predicted value, ym as the mean value of y and n as the number of observations.

Following the selection of the best-performing regression model, a further post-processing step involving a correction factor (CF) was introduced to further enhance the accuracy of the model's predictions. Regression models can exhibit subtle systematic biases when applied to validation or real-world datasets due to factors such as minor differences between training and application environments, or limitations in capturing complex non-linearities across the entire data range [19,20]. Thus, the CF was introduced to mitigate such systematic deviations and improve the overall predictive accuracy, particularly in terms of aligning the mean predicted weight with the mean observed weight in the target population. The CF value was then used to calculate the new corrected weight values (CW) for individual prawns (Equation (5))

$$CF = mo/mp \tag{4}$$

$$CW = CF \times pw \tag{5}$$

where, mo represents the mean observed weight, mp as the mean predicted weight and pw as the predicted weight.

This correction based on mean ratio is a practical approach applied in predictive modelling to calibrate predictions and reduce overall systematic bias observed in validation data. This serves as an additional refinement step to further improve the model's performance characteristics in real-world application.

5. Experimental results

The proposed RAIBE system integrates image processing and regression analysis to estimate prawn weight. The initial step of the RAIBE system utilises YOLOv8-Segmentation model to detect individual prawns in a single frame and with the use of ArUco marker as the reference point, the length of *Macrobrachium rosenbergii* can be determined. The YOLOv8-Segmentation model allows detection of the prawn with more accurate length estimation in any angle of the or rotation, based on the bounding box and segmentation result, when compared to YOLOv8 object detection (Fig. 2). This was confirmed in our previous study where YOLOv8-Segmentation recorded the least MAE and percentage error of 0.93 cm and 5.79% [21]. This estimated length information from the image processing step is then passed as input into the selected regression model to predict the prawn's weight.

The regression weight prediction for *M. rosenbergii* was initially developed using a dataset of 1201 training data points based on the actual farm data from the manually measured prawn length and weight values. The objective was to develop a reliable predictive model that could accurately determine the weight of prawns based on length measurements. To achieve this, we applied various machine learning regression techniques, specifically linear, power, exponential, and polynomial regression models (Fig. 3). The developed weight prediction model for *M. rosenbergii* was validated by using a blind dataset (n = 397), which was not used in the model development. The observed lengths of *M. rosenbergii* individuals ranged from 6.0 cm to 15.0 cm, with corresponding weights between 3.0 g and 44.0 g. The mean observed weight in a separate validation dataset was 21.5 g. This range of data provides the model with a comprehensive foundation for training and validation, to ensure that the models could effectively generalise to prawns of all sizes.

In evaluating the performance of our models, we utilised a combination of metrics to assess accuracy, goodness-of-fit, and the significance of results. Two key accuracy metrics were Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). RMSE calculates the average magnitude of the error, taking the square root of the average of the squared differences between predicted and observed values. Lower RMSE indicates a better fit, as smaller errors contribute less to the final value. MAE, on the other hand, calculates the average of the absolute differences between predicted and observed values. It provides a more comprehensive understanding of the average prediction error, with lower values signifying better performance.

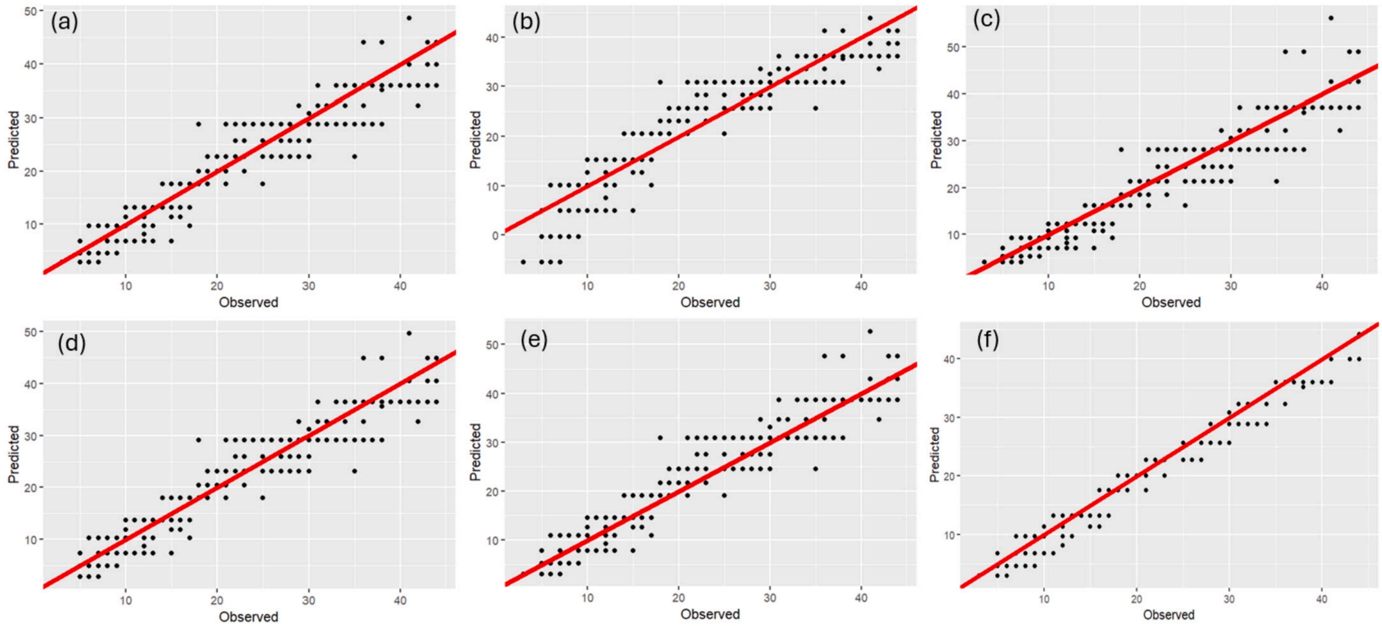

Fig. 4. Predicted prawn weights using different regression models: (a) Power (b) Linear (c) Exponential (d) Polynomial, (e) Modified Polynomial regression and (f) Growth model, against observed weight values. The line represents the ideal fit.

Table 1
Comparison of mean, RMSE, MAE, R^2, Pearson's Correlation test (r) and T-test, between observed and predicted weights of prawns based on the growth model ($W = aL^b$) and regression models; Linear, Power, Exponential (Exp), Polynomial (Poly), Modified Polynomial with correction factor (Poly + CF).

Model	Mean (g)	RMSE	MAE	R^2	r	T-test
Actual weight	21.5					
$W = aL^b$	20.0	2.47	2.09	0.95	0.98	< 0.05
Linear reg.	20.9	3.69	2.87	0.90	0.93	< 0.05
Power reg.	20.0	2.47	2.06	0.96	0.95	< 0.05
Exp reg.	19.5	2.87	2.34	0.96	0.94	< 0.05
Poly reg.	20.4	2.26	1.92	0.96	0.98	< 0.05
Poly reg. + CF	21.5	2.04	1.77	0.96	0.98	> 0.05

To assess model fit, we utilised the coefficient of determination (R^2). R^2 represents a statistical measure that quantifies the proportion of variance in the response variable explained by the model's independent variables. It ranges from 0 to 1, with higher values indicating the best fit, suggesting better model performance (Fig. 4).

Furthermore, we employed paired t-test analysis to evaluate the statistical significance of the difference between observed and predicted weight values. This test allows us to identify if the model's predictions are systematically different from the actual observations. This information is crucial for understanding the model's generalization and potential biases. Finally, we implemented Pearson's correlation coefficient (r) to quantify the strength and direction of the linear association between the observed and predicted weight values. This coefficient ranges from -1 to 1, with values closer to 1 indicating a strong positive linear relationship, signifying that the model's predictions tend to increase as the observed values increase.

5.1. Comparative analysis of prawn length-weight models

Among the models tested (Table 1), polynomial regression demonstrated promising initial results. It achieved an average predicted value of 20.4 g, with a low RMSE and MAE of 2.26 g and 1.92 g, respectively, while having a high R^2 value of 0.96. This was further corroborated by its high Pearson correlation coefficient of 0.98 between observed and predicted weights, suggesting a positive association.

However, the introduction of a correction factor (CF) within the polynomial regression yielded more compelling results. The modified polynomial model achieved a comparable mean value of predicted weight to the observed weight of 21.5 g. This model also produced the lowest RMSE value of 2.04 g and MAE value of 1.77 g, while maintaining a high R^2 of 0.96 and Pearson coefficient (r) of 0.98. These collectively demonstrate the modified polynomial regression model's ability to capture the intricate relationship between prawn length and weight, with consistently producing most accurate weight predictions with minimal deviations from the actual measurements. The t-test analysis further confirmed the model's reliability by indicating that there is no significant difference between the observed and predicted weights, thus signifying that the model's predictions were statistically indistinguishable from the observed measurements.

Alternatively, the growth model with the equation $W = 0.0138L^{2.9801}$, linear, power, and exponential regression models were less effective as each of these models had resulted in RMSE and MAE value of greater than 2.00 g. This translates to greater prediction errors and a much larger deviation from the actual prawn weights. Although these models showed similarly high significant positive R^2 and Pearson's correlation values ranging between 0.90 - 0.98, the t-test revealed significant differences between the observed and predicted weight values, highlighting their lower accuracy and reliability. Linear regression had demonstrated good accuracy in various seafood species [22]. However, it fails to capture the complexities that are inherent in most seafood growth patterns. As highlighted by Gumus et al. [23], the inability to account for non-linear growth dynamics consistently leads to under-performance compared to other models.

For power and exponential regressions, although effective in predicting weight for certain fish species [11] are less suitable for prawns due to their limited applicability across a wide weight range. This restricted scope also raises concerns about generalisation to different prawn varieties. These models struggle to accommodate the variations in growth patterns observed across diverse prawn populations.

5.2. Enhanced accuracy of the polynomial models through correction factor (CF) integration for overall effectiveness of the RAIBE system

The superior performance of the polynomial models, especially the modified polynomial model, can be attributed to their ability to capture non-linear relationships in the data. Prawn growth does not comply with a simple linear progression. It is largely influenced by a number of biological and environmental factors, such as genetics, nutrition, water quality, and temperature [24]. These factors can cause the growth rate to fluctuate at different stages of a prawn's life-cycle. Polynomial models are more adept at accommodating these complexities. By incorporating higher-order terms, this model creates a more intricate curve that closely mirrors the observed non-linear growth patterns in prawns [25]. This enhanced flexibility allows the model to provide a more accurate fit to the data, resulting in lower prediction errors and a closer correspondence between predicted and actual prawn weights. Furthermore, the introduction of the CF in the modified polynomial regression model further refines its ability to capture the variables of prawn growth. This CF aids in addressing the systematic bias that may have skewed the initial model's predictions [26]. By incorporating this factor, the modified polynomial model achieves even lower RMSE and MAE values, while maintaining a high R^2 and Pearson's correlation. This improved performance underscores the effectiveness of the CF in fine-tuning the model and enhancing its overall accuracy.

5.3. Future considerations

Although this study presents a valuable approach for non-invasive weight prediction in *M. rosenbergii*, several limitations need to be addressed. One significant limitation is the consideration of additional factors such as water quality, sex, and diet. The incorporation of these factors may provide more comprehensive predictions, as they largely influence the prawn growth and development. Water quality parameters including pH, dissolved oxygen and temperature can influence the health and growth rate of prawns [27]. Often, with sub-optimal water conditions lead to stress and reduced growth, which the current model does not account for.

The sex of the prawns may also influence the growth rates, with males and females exhibiting different growth patterns [28]. Incorporating these data into the model could further improve prediction accuracy by capturing these biological differences. Additionally, diet is a critical factor affecting prawn growth. Variations in nutrient intake, feed composition, and feeding frequency can lead to differences in weight gain that are not reflected in a model based solely on length measurements.

Moreover, the *Macrobrachium* prawn species exhibit diverse morphotypes including blue-clawed, orange-clawed, and smaller morphs. These are also known to display complex growth patterns that are size-dependent based on their social hierarchy, influencing individual development [29]. While we acknowledge these complex biological factors, the current study did not categorise the prawns based on their specific claw morphotype. Therefore, future studies will aim to address this limitation by categorising prawns based on their distinct morphotypes and life stages to investigate their specific impact on weight estimation and further refine the predictive models for increased accuracy and potentially improving the applicability of the model.

6. Conclusion

This concludes that in the current approach, while the growth equation of $W = aL^b$ has been generally accepted as an effective approach in estimating the length-weight relationship, it may not be suitable in certain aquatic animals. The non-linear regression models should also be taken into consideration as it was observed with a similar outcome as the growth equation. In this case, the modified polynomial regression is the best-suited model in our proposed RAIBE for determining the length-weight relationship and as a weight prediction model for *M. rosenbergii*.

By investigating the intricacies of length-weight relationships, it can pave ways to optimise prawn aquaculture and promote sustainability practices in the aquaculture industry. For future research, intensive sampling in terms of great numbers, sizes and sexes of prawns may assist in revealing further insights into the length-weight relationships.

CRediT authorship contribution statement

Najeebah Az-Zahra Tashim: Writing – original draft, Validation, Software, Methodology, Formal analysis, Data curation. **Tiong Hoo Lim:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Funding acquisition, Conceptualization. **Wafiq Zariful:** Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Data curation. **Pengcheng Liu:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Investigation, Funding acquisition, Conceptualization.

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.

Acknowledgement

This work is financially supported by ICT Virtual Organisation of ASEAN Institutes and NICT (ASEAN IVO, http://www.nict.go.jp/en/index.html), and the Science and Technology Research program (fund vote No. 1117-001) under CREATES. This work is supported in part by the Engineering and Physical Sciences Research Council (EPSRC) under Grant number EP/Y000773/1. This project is also an output from the collaboration with O.D.E. Aquaculture and Agriculture Farm, and the Department of Fisheries, Ministry of Primary Resources and Tourism, Brunei Darussalam, with the research support of the Center for Research on AgriFood Science and Technology (CrAFT), Universiti Teknologi Brunei.

Data availability

Data will be made available on request.

References

[1] Y. Ye, J. Bishop, N. Fetta, E. Abdulqader, J. Al-Mohammadi, A. Alsaffar, S. Almatar, Spatial variation in growth of the green tiger prawn (Penaeus semisulcatus) along the coastal waters of Kuwait, eastern Saudi Arabia, Bahrain, and Qatar, ICES J. Mar. Sci. 60 (4) (2003) 806–817.

[2] P.K. Ng, Collecting and processing freshwater shrimps and crabs, J. Crustac. Biol. 37 (1) (2017) 115–122.

[3] R. Garcia, R. Prados, J. Quintana, A. Tempelaar, N. Gracias, S. Rosen, H. Vågstøl, K. Løvall, Automatic segmentation of fish using deep learning with application to fish size measurement, ICES J. Mar. Sci. 77 (4) (2020) 1354–1366.

[4] D. Voskakis, A. Makris, N. Papandroulakis, Deep learning based fish length estimation. An application for the Mediterranean aquaculture, in: OCEANS 2021: San Diego–Porto, IEEE, 2021, pp. 1–5.

[5] J.J. Dabrowski, A. Rahman, D.E. Pagendam, A. George, Enforcing mean reversion in state space models for prawn pond water quality forecasting, Comput. Electron. Agric. 168 (2020).

[6] H. Richter, C. Luckstadt, U. Focken, K. Becker, An improved procedure to assess fish condition on the basis of length-weight relationships, Arch. Fish. Mar. Res. 48 (3) (2000) 255–264.

[7] R. Nahavandi, S.N. Amin, S. Zakaria, M.N. Shamsudin, Growth and length-weight relationship of Penaeus monodon (Fabricius) cultured in artificial sea water, Res. J. Fish. Hydrobiol. 5 (1) (2010) 52–55.

[8] T.S.d.C. Silva, L.D. d. Santos, L.C.R. d. Silva, M. Michelato, V.R.B. Furuya, W.M. Furuya, Length-weight relationship and prediction equations of body composition for growing-finishing cage-farmed Nile tilapia, Rev. Bras. Zootec. 44 (2015) 133–137.

[9] K. Waiho, H. Fazhan, A.C. Shu-Chien, M.H. Abualreesh, H. Ma, M. Syahnon, G. Azmie, N.J. Razman, M. Ikhwanuddin, Size distribution, length-weight relationship, and size at morphometric maturity of the mud spiny lobster Panulirus polyphagus (Herbst, 1793) in the Johor strait, Front. Mar. Sci. 8 (2021).

[10] A. Al Nahdi, C. Garcia de Leaniz, A.J. King, Spatio-temporal variation in length-weight relationships and condition of the ribbonfish Trichiurus lepturus (Linnaeus, 1758): implications for fisheries management, PLoS ONE 11 (8) (2016).

[11] C. Clain, J. Stewart, A. Fowler, S. Diamond, Age, growth and length-to-weight relationship of largehead hairtail (Trichiurus lepturus) in south-eastern Australia suggest a distinct population, Aquacult. Fish. (2023).

[12] R. Berral-Soler, R. Muñoz-Salinas, R. Medina-Carnicer, M.J. Marín-Jiménez, Deeparuco++: improved detection of square fiducial markers in challenging lighting conditions, Image Vis. Comput. (2024).

[13] R. Froese, Cube law, condition factor and weight–length relationships: history, meta-analysis and recommendations, J. Appl. Ichthyol. 22 (4) (2006) 241–253.

[14] R.P.A. Carrion, H. de la Torre, S. Estrada, An extreme value regression approach to the prediction of shrimp size distribution, J. Mod. Appl. Stat. Methods 21 (2022).

[15] A. Dikou, Weight–length relationship in fish populations reflects environmental regulation on growth, Hydrobiologia 850 (2) (2023) 335–346.

[16] F. Akhter, H.R. Siddiquei, M.E.E. Alahi, S.C. Mukhopadhyay, Recent advancement of the sensors for monitoring the water quality parameters in smart fisheries farming, Computers 10 (3) (2021) 26.

[17] G. De Luca, M. Gallo, The use of artificial neural networks for extending road traffic monitoring data spatially: an application to the neighbourhoods of Benevento, Transp. Res. Proc. 45 (2020) 635–642.

[18] A. Setiawan, H. Hadiyanto, C.E. Widodo, Shrimp body weight estimation in aquaculture ponds using morphometric features based on underwater image analysis and machine learning approach, Rev. Intell. Artif. 36 (6) (2022).

[19] H. Lee, S. Chen, Systematic bias of machine learning regression models and correction, IEEE Trans. Pattern Anal. Mach. Intell. 47 (6) (2025) 4974–4983, https://doi.org/10.1109/TPAMI.2025.3552368.

[20] B.M. Sanderson, On the estimation of systematic error in regression-based predictions of climate sensitivity, Clim. Change 118 (2013) 757–770.

[21] M.W.A. Zariful, N.A.-Z. Tashim, T.H. Lim, H. Fakhrurroja, S. Chuprat, S.A. Putra, P. Liu, A.M. Basri, Comparison of biocode based machine learning and segmentation model for automated prawn size prediction for real prawn farm, in: 2024 Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), 2024, pp. 581–586.

[22] X. Yu, Y. Wang, J. Liu, J. Wang, D. An, Y. Wei, Non-contact weight estimation system for fish based on instance segmentation, Expert Syst. Appl. 210 (2022).

[23] B. Gümüş, E. Gümüş, M.O. Balaban, Image analysis to determine length-weight and area-weight relationships, and color differences in scaled carp and mirror carp grown in fiberglass and concrete tanks, Turk. J. Fish. Aquat. Sci. 23 (1) (2022).

[24] J. Wei, L. Tian, Y. Wang, L. Yu, X. Zhu, Effects of salinity, photoperiod, and light spectrum on larval survival, growth, and related enzyme activities in the giant freshwater prawn, Macrobrachium rosenbergii, Aquaculture 530 (2021).

[25] D. Maulud, A.M. Abdulazeez, A review on linear regression comprehensive in machine learning, J. Appl. Sci. Technol. Trends 1 (2) (2020) 140–147.

[26] F. Wang, D. Tian, On deep learning-based bias correction and downscaling of multiple climate models simulations, Clim. Dyn. 59 (11) (2022) 3451–3468.

[27] C. Leigh, B. Stewart-Koster, N. Van Sang, V.B. Xoan, N.T.N. Tinh, J. Sammut, M.A. Burford, et al., Rice-shrimp ecosystems in the Mekong Delta: linking water quality, shrimp and their natural food sources, Sci. Total Environ. 739 (2020).

[28] A. Indarjo, G. Salim, C.D. Nugraeni, I.M. Sari, M. Zein, L.Y. Prakoso, G.A. Daengs, The growth model composition of giant prawns (Macrobrachium rosenbergii) in Muara Tepian Sembakung, Nunukan, IOP Conf. Ser. Earth Environ. Sci. 564 (1) (2020).

[29] S. Ibrahim, Z. Zhong, X. Lan, J. Luo, Q. Tang, Z. Xia, S. Yi, G. Yang, Morphological diversity of different male morphotypes of giant freshwater prawn Macrobrachium rosenbergii (De Man, 1879), Aquacult. J. 3 (2) (2023) 133–148, https://doi.org/10.3390/aquacj3020012, https://www.mdpi.com/2673-9496/3/2/12.