



# Application of information theory-based decision support system for high precision modeling of the length-weight relationship (LWR) for five marine shrimps from the northwestern Bay of Bengal

Gyanaranjan Dash<sup>a,\*</sup>, Swatipriyanka Sen<sup>a</sup>, Rajesh Kumar Pradhan<sup>a</sup>, Shubhadeep Ghosh<sup>b</sup>, Josileen Jose<sup>c</sup>

<sup>a</sup> Puri Field Centre of ICAR-Central Marine Fisheries Research Institute, Puri 752 002, Odisha, India

<sup>b</sup> Visakhapatnam Regional Station of ICAR-Central Marine Fisheries Research Institute, Visakhapatnam 530 003, Andhra Pradesh, India

<sup>c</sup> ICAR-Central Marine Fisheries Research Institute, Head Quarter, Cochin 682 018, Kerala, India

## ARTICLE INFO

### Article history:

Received 10 May 2022

Received in revised form 4 July 2023

Accepted 31 July 2023

Available online 16 August 2023

### Keywords:

AIC

Likelihood

Akaike weight

Length-weight relationship

Model selection

## ABSTRACT

The study was conducted to develop an information theory-based decision support system to understand the variance distribution structure of the data so that a proper modeling approach could be implemented to explain the relationship between the body length and weight of shrimps. Based on biological reasoning, initially, a log-normal multiplicative error structure was assumed and therefore, a log-linearized model was applied. Secondly, the support for normal additive error structure was assessed by fitting a weighted nonlinear model with a power variance structure (wNLM) to address the heteroscedasticity in shrimp weight. The likelihood support for the error structures was ascertained by comparing the  $AIC_c$  of the two competing models. As the general cut-off criterion ( $\Delta AIC_c > 2.0$ ) did not give conclusive evidence from the scrutiny of the probability density diagnostic plot of the residuals, an alternative model scaling criteria, i.e., Akaike weight (Aw) of 0.9 was used for model selection. The corresponding  $\Delta AIC_c$  cut-off score of 4.2 was estimated by regressing the  $\Delta AIC_c$  score of the competing model against the Aw scores of the best model. The competing models with  $\Delta AIC_c > 4.2$  were rejected and the alternate models with  $Aw \geq 0.9$  were selected for modeling the length-weight relationship. Both the models were observed to be well founded, as narrow differences in the root mean squared error (RMSE) were observed. A lower RMSE was almost always observed from wNLM despite a higher  $\Delta AIC_c$  score, which indicates that RMSE may not be efficient in detecting the model overfitting issue. Contrary to popular belief, only 26.7% of the datasets exhibited a log-normal error structure, whereas, a normal error structure was evident in 33.3% of the datasets. Interestingly, 40.0% of the datasets showed data ambivalence ( $\Delta AIC_c < 4.2$ ) and therefore, an Akaike weighted model averaging was performed to reduce model uncertainty for the accurate estimation of model parameters and their confidence intervals.

© 2023 Elsevier B.V. All rights reserved.

## 1. Introduction

The collection of information pertaining to the biomass of aquatic animals directly from field observation is quite cumbersome, expensive and time-consuming. However, such information can be easily derived by establishing a relationship between body length and weight of animals. Unlike body weight, the length of the aquatic animal can be collected much more easily and quickly from the field and is also widely available in developing fisheries across the world. Due to the high dependence of conventional fisheries stock assessment studies on readily

available length-based data, an accurate modeling framework is essential for converting these length-based data to weight data so that the uncertainty during the prediction of stock biomass can be minimized. As the information on this biomass is further used in ecological modeling and quantification of various resource management reference points, inappropriate modeling of the relationship between body length and weight (LWR) could have severe implications for the uncertainty associated with the final model output. The information on the length-weight relationship (LWR) has several applications in fishery biology studies, such as estimation of stock biomass for their sustainable utilization, various ecological and biomass modeling studies and also to understand the spatio-temporal as well as ontogenetic variations in the overall well-being of animals (Ricker, 1968; Pauly, 1984;

\* Corresponding author.

E-mail address: [gyanranjandashcmfri@gmail.com](mailto:gyanranjandashcmfri@gmail.com) (G. Dash).

Petrakis and Stergiou, 1995; Froese, 2006; Froese et al., 2011; Türker et al., 2018).

Accurate estimation of the LWR is of paramount importance, especially when it is used as an input parameter for population biomass estimation, because wrongly predicted modeling parameters could disproportionately amplify the uncertainty and errors during the raising process from sample level to population or stock level. Though often neglected, the variance distribution structure of data plays a critical role in the selection of an appropriate modeling approach while establishing the relationship between length and weight. If the variance structure is assumed to be normal and additive, then an ordinary nonlinear model should be used for the correct estimation of modeling parameters (Xiao et al., 2011). However, the ordinary nonlinear model does not account for the heteroscedasticity in the variance structure of the body weight data that gradually increases as the animal becomes larger. Such a gradual inequality in the variance of fish weight with an increase in body length can be addressed by a weighted nonlinear modeling approach with an appropriate weighting function to address the unequal variance structure (De Giosa and Czerniejewski, 2016). On the other hand, if the variance is suspected of having a log-normal multiplicative error structure, it is appropriate to apply a log-linearized (transformed) model (Xiao et al., 2011). The log transformation not only makes the variance distribution Gaussian (normal) but also stabilizes the inequality in the variance structure.

It is, therefore, very important to ascertain the variance distribution structure before an appropriate modeling approach can be followed. This can be achieved by a statistical approach that analyzes the likelihood of data being generated from an underlying normal distribution with an additive variance structure or an underlying log-normal distribution with a multiplicative error structure (Xiao et al., 2011). Consequently, commonly used information theory (IT-based) indices such as the Akaike information criterion (AIC) or the Bayesian information criterion (BIC) can be used to evaluate the best modeling approach that is in congruence with the assumed probability distribution structure of residuals generated from the data at hand (Burnham and Anderson, 2002). An IT-based criterion such as AIC basically assumes that though real-world data is generated from random variable(s) with an unknown probability distribution, the true probability distribution (true model) of the data can be derived. AIC calculates the information that would be lost in the form of a distance (technically known as the Kullback–Leibler or K–L discrepancy) between the predicted probability distribution of the fitted model and that of the true probability distribution of real-world data. It also assumes that the K–L discrepancy will get smaller as the accuracy of the fitted model gets better, which will result in a smaller ‘information loss’ between the true probability distribution of the real-world data and the predicted probability distribution of the fitted model. As a result, the best-fitted model will be the one that approximates the true probability distribution of the data in a way that will result in the least amount of information loss (Burnham and Anderson, 2001; Banks and Joyner, 2017).

However, it must be noted that the absolute value of these IT-based indices (i.e., AIC and BIC) does not provide any information about the performance of the model. It is the comparison of these IT-based indices values in a relative term that provides the necessary insight for drawing the multimodel inference from a set of models, where the model with the lowest information loss is selected as the best model. The AIC values of competing models can be compared in terms of their relative values, where the difference in the absolute AIC value (i.e.,  $\Delta\text{AIC}$ ) of the competing model compared to the best model (with the lowest AIC) gives information about the support for the competing model

based on specific threshold values. As a general rule of thumb, the following three cutoff criteria are used for acceptance or rejection of the competing model: (1) If  $\Delta\text{AIC} = 0\text{--}2$ , then both the competing model and the best model are accepted due to substantial empirical support for both models; (2) If  $\Delta\text{AIC} = 4\text{--}7$ , then the competing model might be considered for rejection due to the lack of adequate empirical support in its favor and (3) If  $\Delta\text{AIC} > 10$ , then the competing model should be straight away rejected due to the complete lack of empirical support in its favor (Burnham and Anderson, 2002). Nevertheless, there is considerable ambiguity regarding the treatment (acceptance or rejection) of the competing model if the  $\Delta\text{AIC}$  falls within the doubtful zone that primarily ranges from 2 to 4 and to some extent even up to 7.

Though the application of IT-based indices such as AIC and BIC for evaluation of the model performance to draw multimodel inference is very popular for ecological hypothesis testing (Burnham and Anderson, 2002; Johnson and Omland, 2004; Richards, 2005; Thorup et al., 2006; Burnham et al., 2011; Symonds and Moussalli, 2011; Brewer et al., 2016), the indices have not been much publicized as a decision support system in fishery biology investigations. According to a recent review conducted by Flinn and Midway (2021), only about 28% of the studies conducted for the growth modeling of fish have used some form of IT-based indices. Nevertheless, there are few noteworthy studies where IT-based indices have been used in fishery biological investigation, mainly while modeling the growth of fish (Imai et al., 2002; Porch et al., 2002; Zhu et al., 2009; Panhwar et al., 2010; Diniz et al., 2012) and also to some extent while conducting other biological investigations such as modeling the relationship between stock and recruitment of fish (Hiramatsu et al., 1994; Wang and Liu, 2006) and even the relationship between body length and weight of fish (De Giosa and Czerniejewski, 2016).

The present study was envisaged to develop a decision support system using an IT-based criterion to establish the relationship between the body length and weight of five commercially important shrimps. This was done by giving adequate consideration to the likelihood support for the variance distribution structure in order to reduce bias and improve the accuracy of modeling parameters and their confidence intervals.

## 2. Materials and methods

### 2.1. Data collection

The data on the length and weight of the selected five commercially important shrimps, viz. *Penaeus monodon* ( $n = 480$ , Male to Female ratio (M:F) = 1:1.04), *Penaeus semisulcatus* ( $n = 420$ , M:F = 1:0.75), *Parapenaeopsis stylifera* ( $n = 480$ , M:F = 1:7.42), *Parapenaeopsis sculptilis* ( $n = 480$ , M:F = 1:10.71) and *Solenocera crassicornis* ( $n = 480$ ; M:F = 1:5.40) were collected from the most important fish landing centre (FLC), i.e., Digha Mohana FLC (West Bengal) of the northeast coast of India, at monthly intervals from July 2018 to March 2021. The shrimps were primarily caught by the mechanized crafts using non-selective trawl nets having a cod end mesh size of 25 mm from a depth range of 15–40 m operating between the latitudinal range of  $21^{\circ}28'N$  to  $21^{\circ}35'N$  along the northwestern portion of the Bay of Bengal (Fig. 1). All the specimens were washed and stored immediately in the field using ice and brought to the laboratory of the Digha Research Centre of ICAR-CMFRI (West Bengal, India), where their identity was further confirmed following the standard identification protocols (Fischer and Bianchi, 1984). The sexes of the shrimps (i.e., male or female) were identified based on the presence or absence of petasma between the first pair of pleopods and the total body length (TL), i.e., the straight line distance between

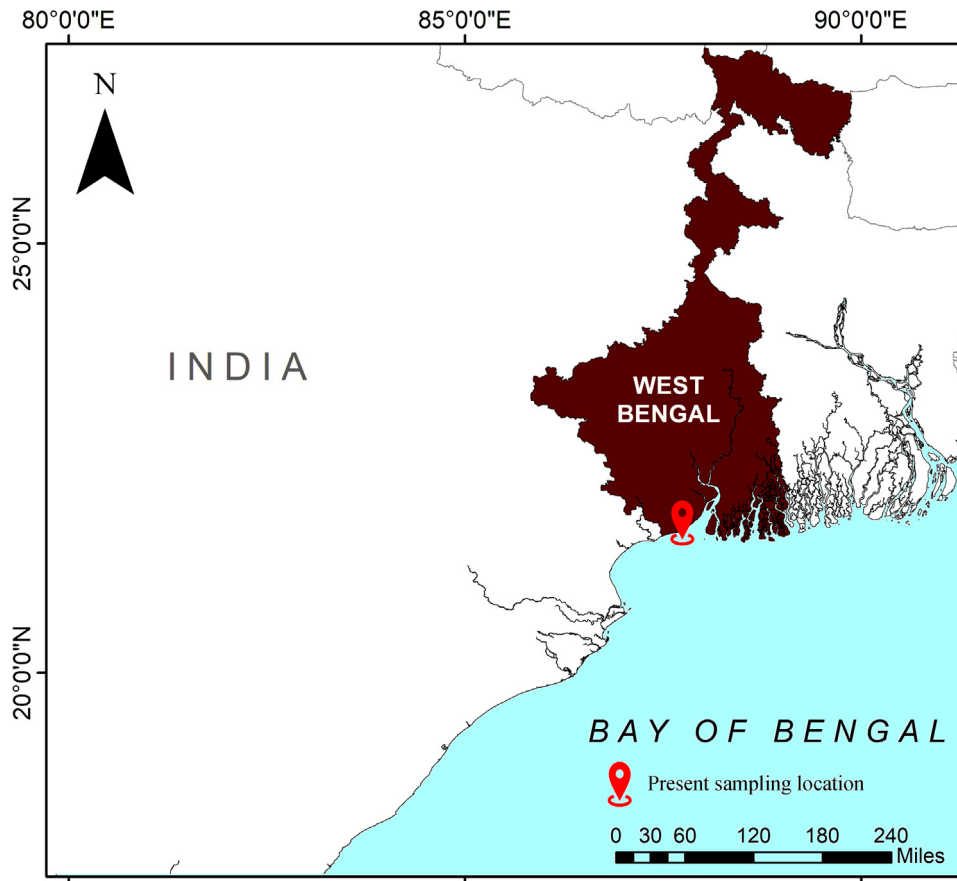


Fig. 1. Fishing location from which the shrimp specimens for the present study have been collected.

the tip of the rostrum and the tip of the telson to the nearest ‘0.1 cm’ was recorded for all the specimens. The body weight was recorded to the nearest ‘0.1 g’. The shrimps were dissected in the laboratory for routine biological investigation after length and weight measurements were taken.

2.2. Determination of the error structure of data

The variance distribution structure was ascertained by conducting a likelihood analysis and comparing the derived information theory (IT-based) indicators. Based on biological reasoning, first, the data were assumed to have been generated from a log-normal distribution with multiplicative error structure (LNM) (Eq. (1)) and therefore, a log-log transformed linearized modeling approach (LM) (Eq. (2)) was followed. The modeling parameters (a, b, and  $\sigma^2$ ) were derived for the LM and were back-transformed to the original scale. The likelihood score ( $L_{\lognorm}$ ) of data assuming a log-normal probability distribution was derived in the original untransformed scale using Eq. (3) (Banks and Joyner, 2017).

$$\text{LNM, } \log(W) = a \times L^b \times \exp(\varepsilon), \varepsilon \sim \text{Normal}(0, \sigma^2) \tag{1}$$

$$\text{LM, } \log(W) = \log(a) + b \times \log(L) + \varepsilon, \varepsilon \sim \text{Normal}(0, \sigma^2) \tag{2}$$

$$L_{\lognorm} = \prod_{i=1}^n \left[ \frac{1}{\sqrt{2\pi\sigma_{LM}^2}} \exp\left( \frac{-\left(\log(y_i) - \log(a_{LM}x_i^{b_{LM}})\right)^2}{2\sigma_{LM}^2} \right) \right] \tag{3}$$

Secondly, the data were assumed to have been generated from a normal distribution with additive error structure (NLM)

(Eq. (4)). However, assuming heterogeneous variance distribution structure in fish weight data, a weighted nonlinear modeling approach with power variance structure, i.e.,  $\text{var}(\varepsilon) = \sigma^2[\text{fitted weight}]^{2d}$  (Eq. (5)) was followed to account for the heteroscedasticity in the data. The modeling parameters (a, b, and  $\sigma^2$ ) were derived for the wNLM and the likelihood score ( $L_{\text{norm}}$ ) of the data assuming a normal probability distribution was derived using Eq. (6) (Banks and Joyner, 2017).

$$\text{NLM, } W = a \times L^b + \varepsilon, \varepsilon \sim \text{Normal}(0, \sigma^2) \tag{4}$$

$$\text{wNLM, } W = a \times L^b + \varepsilon, \varepsilon \sim \text{Normal}(0, \sigma^2[\text{fitted weight}]^{2\Delta}) \tag{5}$$

$$L_{\text{norm}} = \prod_{i=1}^n \left[ \frac{1}{\sqrt{2\pi w_i^2 \sigma_{wNLM}^2}} \exp\left( \frac{-\left(y_i - a_{wNLM} x_i^{b_{wNLM}}\right)^2}{2 w_i^2 \sigma_{wNLM}^2} \right) \right] \tag{6}$$

The linear modeling (LM) and weighted nonlinear modeling (wNLM) were performed using the linear regression ‘lm’ routine and a generalized nonlinear regression ‘gnls’ routine, respectively, in the R statistical software package. The statistical support for the error structure was determined using an IT-based indicator, i.e., the Akaike information criterion (AIC). However, due to the small sample size in the case of males of certain shrimps, a sample size-corrected AIC (second-order information criterion), i.e.,  $AIC_c$  (Sugiura, 1978; Hurvich and Tsai, 1995), was used to compare the likelihood support for the appropriate error structure. This small sample unbiased AIC (i.e.,  $AIC_c$ ) value for log-normal and normal probability distribution assumptions was derived using the corresponding likelihoods ( $L_i$ ) of the two distribution models,

i.e.,  $L_{\log\text{norm}}$  and  $L_{\text{norm}}$  respectively, in the following equation suggested by Burnham and Anderson (2002).

$$\text{AIC}_{c-i} = 2k - 2 \log(L_i) + \frac{2k(k+1)}{n-k-1} \quad (7)$$

where,  $\text{AIC}_{c-i}$  represents the  $\text{AIC}_c$  score for the corresponding models (i.e., LM and NLM).  $k$  is the number of parameters in each of the models and  $n$  is the sample size. This  $\text{AIC}_c$  was further used to derive  $\Delta\text{AIC}_c$  using the following equation suggested by Burnham and Anderson (2002).

$$\Delta\text{AIC}_{c-i} = \text{AIC}_{c-i} - \min(\text{AIC}_{c-\text{norm}}, \text{AIC}_{c-\log\text{norm}}) \quad (8)$$

where,  $\Delta\text{AIC}_{c-i}$  represents the  $\Delta\text{AIC}_c$  score for the corresponding models (i.e., LM and NLM). This  $\Delta\text{AIC}_c$  score was further used to calculate the Akaike weight ( $Aw_i$ ) for the corresponding models (i.e., LM and NLM) using the following equation suggested by Burnham and Anderson (2002).

$$Aw_i = \frac{\exp(-\frac{1}{2}\Delta\text{AIC}_{c-i})}{\sum_{r=1}^R (-\frac{1}{2}\Delta\text{AIC}_{c-k})} \quad (9)$$

where, the numerator represents the relative likelihood of the corresponding model and the denominator represents the sum of the relative likelihoods of the competing models under consideration. Therefore, the value of  $Aw_i$  ranges from 0 to 1.

The ratio of the Akaike weight ( $Aw$ ) of the best model to the other competing model in a pair of models under consideration was used to derive the evidence ratio (ER) in favor of the best model using the following equation suggested by Burnham and Anderson (2002).

$$\text{ER} = \frac{Aw_i}{Aw_j} = \frac{\exp(-\frac{1}{2}\Delta\text{AIC}_{c-i})}{\exp(-\frac{1}{2}\Delta\text{AIC}_{c-j})} \quad (10)$$

where,  $Aw_i$  is the Akaike weight of the best among the competing models (with the lowest  $\text{AIC}_c$ ) and  $Aw_j$  is the Akaike weight of the other competing model. ER can also be calculated from the relative likelihood of the corresponding competing models.

### 2.3. Decision support for modeling approach

Initially, an  $\Delta\text{AIC}_c$  cutoff of higher than 2 was used to ascertain the appropriate error structure for the data, which was cross-validated using probability density diagnostic plots for the residuals generated from both competing models. As no apparent distinguishable probability density distribution of residuals could be visualized in certain cases at  $\Delta\text{AIC}_c > 2.0$  and the corresponding  $W$  ( $Aw$ ) of 0.88 and ER of 7.5, a further higher  $W$  ( $Aw$ ) of 0.9 and ER of 9.0 was used for the rejection of the competing model. The corresponding  $\Delta\text{AIC}_c$  score for the model rejection was determined by regressing the  $\Delta\text{AIC}_c$  of the competing model against the Akaike weight ( $Aw$ ) of the best model (Fig. 3 and Eq. (11)).

$$\Delta\text{AIC}_c = a - \left( \frac{1}{b} \times \text{LN}(1 - W) \right) \quad (11)$$

where,  $a = -0.53342$  and  $b = 0.48834$  are regression parameters. The cut-off  $\Delta\text{AIC}_c$  for model rejection, corresponding to an  $Aw$  of 0.9 and an ER of 9.0 was estimated to be 4.2. The following three

criteria were used for the selection of an appropriate modeling approach.

Criterion-1:  $\Delta\text{AIC}_c > 4.2$  with  $W_{\text{LM}} \geq 0.9$ : Better support for log-normal distribution with multiplicative error structure

Criterion-2:  $\Delta\text{AIC}_c > 4.2$  with  $W_{\text{wNLM}} \geq 0.9$ : Better support for normal distribution with additive error structure

Criterion-3:  $\Delta\text{AIC}_c < 4.2$  with  $W_{\text{LM}}$  or  $W_{\text{wNLM}} < 0.9$ : Equal support for both the log-normal distribution with multiplicative error structure and the normal distribution with additive error structure.

Under criterion-1, log-transformed linearized modeling (LM) approach was used for modeling. On the other hand, if criterion-2 was fulfilled, then weighted nonlinear modeling (wNLM) was used. However, under criterion-3, an AIC weighted model averaging (AwMA) approach was followed, assuming inconclusive support for both the probability density structures of error. A full model averaging following the recommendations of Lukacs et al. (2009) and Xiao et al. (2011) was used to derive the model parameters. The model parameters ( $\beta$ ) i.e.,  $a$  and  $b$ , were estimated using the following equation.

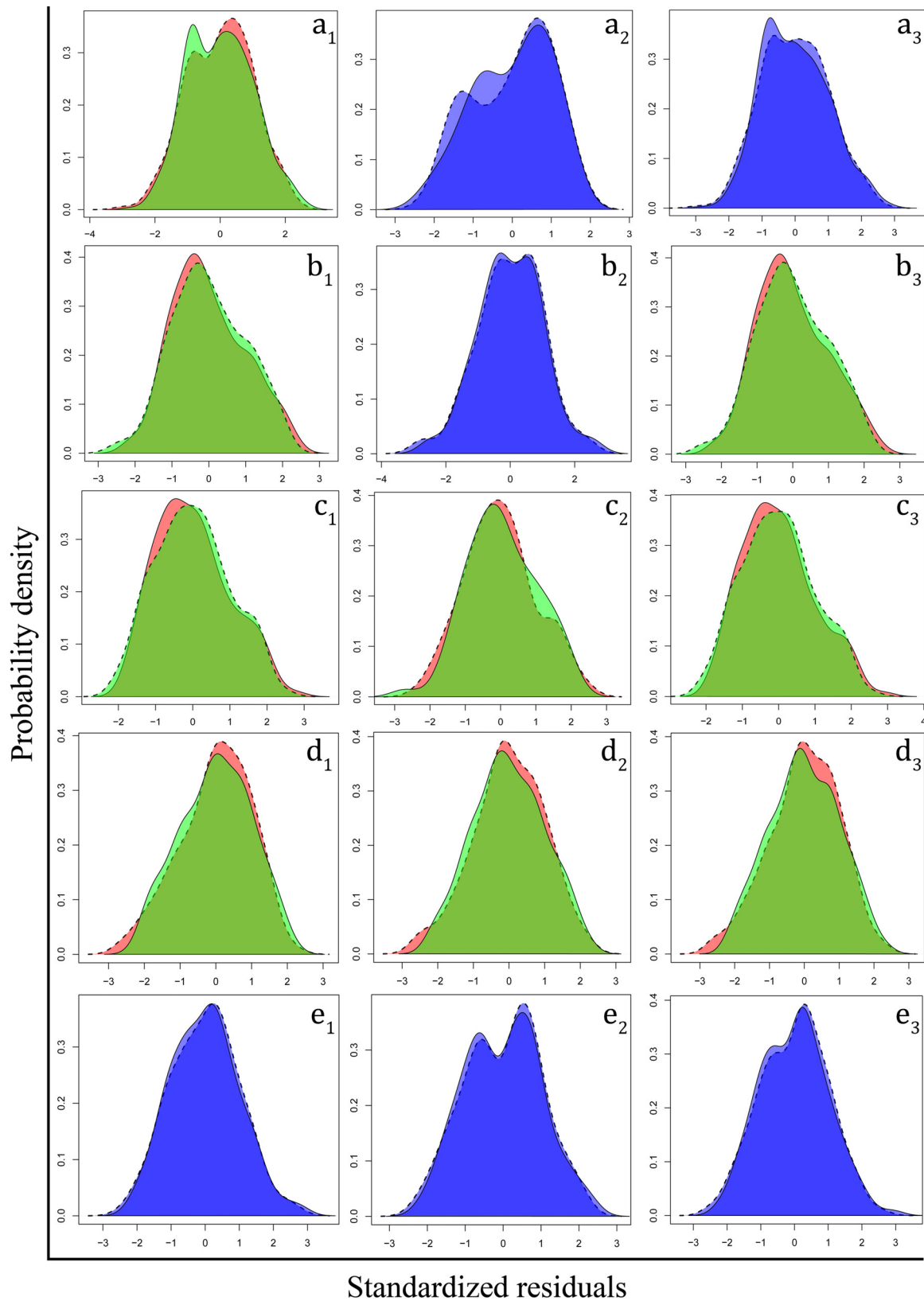
$$\beta = \sum_{i=1}^R W_i (\text{AIC}) \times \beta_i \quad (12)$$

Subsequently, the CIs for  $a$  and  $b$  for the AwMA were generated by bootstrapping datasets using the 'boot' routine in the R statistical software package. The models performance was further cross-checked using the conventional goodness of fit (GOF) indicator, i.e., root mean squared error (RMSE).

### 3. Results

The information-theoretic criterion ( $\text{AIC}_c$ ) and its derivatives for the evaluation of support for two competing models with log-normal multiplicative error (LM) and normal additive error structure (wNLM) are summarized in Table 1. An initial scrutiny of the probability density diagnostic plot of residuals (Fig. 2) generated from two modeling approaches does not provide conclusive evidence for the two competing models at a  $\Delta\text{AIC}_c$  of more than 2.0 or even at an Akaike weight ( $Aw$ ) threshold of 0.88. Therefore, a higher  $Aw$  of 0.9 and an evidence ratio of 9.0 were used for the rejection of the competing model. The Akaike weights ( $Aw$ ) of the best models as predictor variables were regressed against the  $\Delta\text{AIC}_c$  scores of the competing models as response variables to derive the  $\Delta\text{AIC}_c$  cut-off score of 4.2 for the rejection of the competing model (Fig. 3).

Using this derived  $\Delta\text{AIC}_c$  cut-off, a higher likelihood support for the log-normal multiplicative error structure of the data was clearly evident in only 4 out of 15 datasets (26.7% of overall data), i.e., *P. sculptilis* (sex pooled and female datasets), *S. crassicornis* (sex pooled and female datasets), suggesting the applicability of the log-linearized model (LM) as the best modeling approach. On the other hand, 5 datasets (33.3% of overall data), i.e., *P. stylifera* (female dataset), *S. crassicornis* (male dataset) and *P. monodon* (sex pooled, male and female datasets) showed higher likelihood support for normal additive error structure, suggesting the weighted nonlinear model (wNLM) as the best modeling approach. In maximum datasets, i.e., 6 out of 15 (40% of overall data), i.e., *P. stylifera* (sex pooled and male datasets), *P. sculptilis* (male dataset) and *P. semisulcatus* (sex pooled, male and female datasets), equal support for two of the above error distribution structures was observed, indicating an ambivalent error structure for most of the observations. For these six datasets, an AIC weighted model averaging (AwMA) approach was used for the accurate estimation of modeling parameters and their confidence



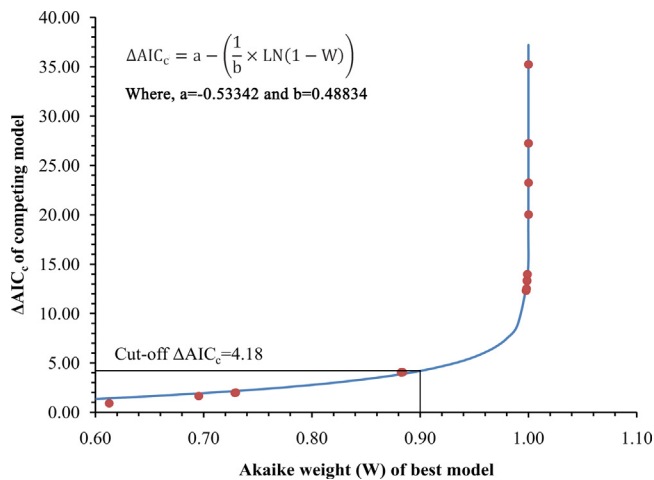
**Fig. 2.** Probability density diagnostic plot for validating the support for lognormal multiplicative error (LM) and normal additive error structure (wNLM). The alphabets in the figure represent shrimp species, i.e., a: *Parapenaopsis stylifera*; b: *Parapenaopsis sculptilis*, c: *Solenocera crassicornis*; d: *Penaeus monodon* and e: *Penaeus semisulcatus*; the numerical subscripts represent dataset category, i.e., 1: female dataset; 2: male dataset and 3: sex-pooled dataset. The green and red probability plots with dotted lines indicate higher and lesser supports, respectively, for the log-normal distribution of residuals. The green and red probability plots with solid lines indicate higher and lesser supports, respectively, for the normal distribution of residuals. The blue probability plots indicate equal support for both the log-normal and normal distributions of residuals.

**Table 1**

Information-theoretic criterion (AIC<sub>c</sub>) and its derivatives for evaluation of support for two competing models with lognormal multiplicative error (LM) and normal additive error structures (wNLM).

Species	Sex	AIC <sub>c-lognorm</sub>	AIC <sub>c-norm</sub>	ΔAIC <sub>c</sub>	exp (-1/2 ΔAIC <sub>c-lognorm</sub> )	exp (-1/2 ΔAIC <sub>c-norm</sub> )	Aw <sub>lognorm</sub>	Aw <sub>norm</sub>	ER	Best modeling approach
<i>P. stylifera</i>	P	1192.02	1187.99	4.03	0.13	1.00	0.12	0.88	7.51	AwMA
	M	47.62	45.54	2.08	0.35	1.00	0.26	0.74	2.83	AwMA
	F	1093.45	1080.93	12.51	0.00	1.00	0.00	1.00	521.80	wNLM
<i>P. sculptilis</i>	P	2332.80	2346.12	13.32	1.00	0.00	1.00	0.00	778.91	LM
	M	153.10	155.08	1.98	1.00	0.37	0.73	0.27	2.70	AwMA
	F	2169.65	2183.02	13.37	1.00	0.00	1.00	0.00	800.31	LM
<i>S. crassicornis</i>	P	1329.52	1349.53	20.01	1.00	0.00	1.00	0.00	22088.12	LM
	M	115.19	91.94	23.25	0.00	1.00	0.00	1.00	111983.83	wNLM
	F	1213.14	1225.46	12.32	1.00	0.00	1.00	0.00	473.65	LM
<i>P. monodon</i>	P	3259.97	3224.73	35.24	0.00	1.00	0.00	1.00	44884578.07	wNLM
	M	1614.60	1587.37	27.23	0.00	1.00	0.00	1.00	817651.54	wNLM
	F	1647.37	1633.40	13.97	0.00	1.00	0.00	1.00	1079.61	wNLM
<i>P. semisulcatus</i>	P	2563.37	2564.28	0.92	1.00	0.63	0.61	0.39	1.58	AwMA
	M	1365.30	1366.95	1.65	1.00	0.44	0.70	0.30	2.28	AwMA
	F	1169.29	1173.35	4.06	1.00	0.13	0.88	0.12	7.61	AwMA

P: sex-pooled data; M: male data; F: female data; AIC<sub>c-lognorm</sub>: sample size corrected AIC for LM; AIC<sub>c-norm</sub>: sample size corrected AIC for wNLM; ΔAIC<sub>c</sub>: AIC<sub>c</sub> difference of the competing model from best model; exp (-1/2ΔAIC<sub>c-lognorm</sub>): relative likelihood of LM; exp (-1/2ΔAIC<sub>c-norm</sub>): relative likelihood of wNLM; Aw<sub>lognorm</sub>: Akaike weight for lognormal distribution of error; Aw<sub>norm</sub>: Akaike weight for normal distribution of error; ER: Evidence ratio.



**Fig. 3.** Regression of the Akaike weight (Aw) of the best model against the ΔAIC<sub>c</sub> score of the competing model for deriving the cut-off ΔAIC<sub>c</sub> for the model rejection.

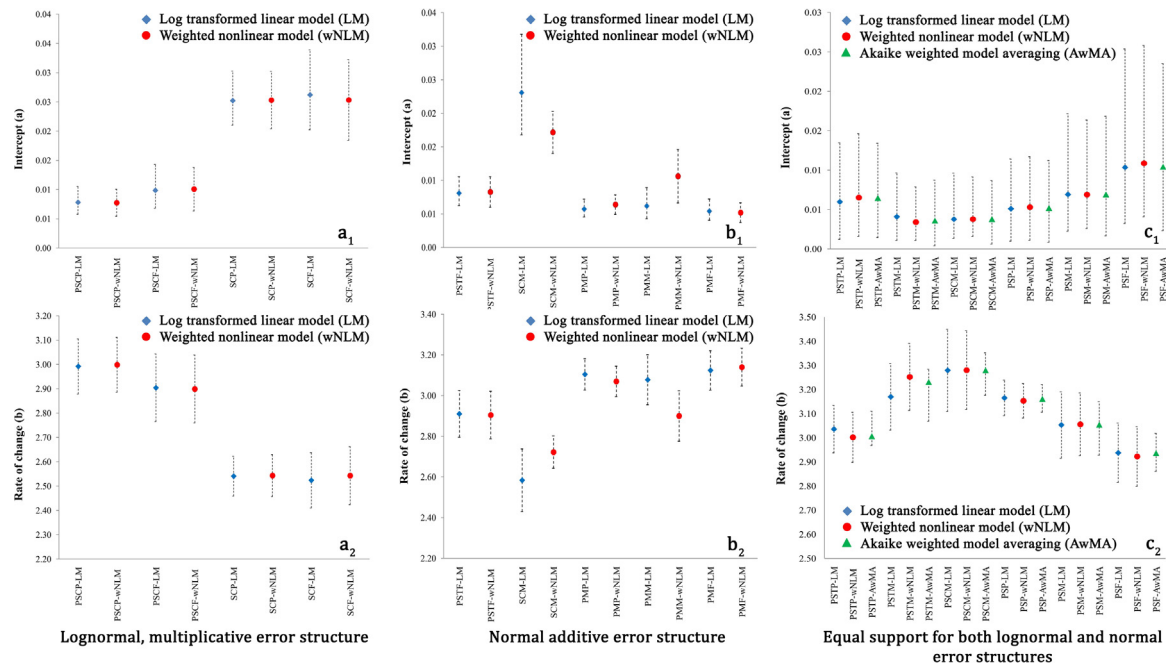
intervals. The modeling parameters and their confidence intervals obtained from LM, wNLM and AwMA are summarized in Table 2. The comparison of inappropriate modeling approaches such as the application of wNLM when errors are log-normally distributed with multiplicative error structures and the application of LM when the errors are normally distributed with additive error structures is shown in Fig. 4a<sub>1</sub>-a<sub>2</sub> and Fig. 4b<sub>1</sub>-b<sub>2</sub>, respectively. The applicability of AwMA in reducing the model bias when both the error structures (log-normal and normal) are ambivalent is also shown in Fig. 4c<sub>1</sub>-c<sub>2</sub>. The root mean squared error (RMSE) compared to check the soundness of the competing models revealed narrow differences among the models and therefore, the competing models could be concluded to be well-founded. Nevertheless, lower RMSE was almost always observed in the wNLM compared to the LM and AwMA, except for the *Parapanaeopsis sculptilis* (sex pooled) dataset.

#### 4. Discussion

The choice of an appropriate modeling approach is mainly influenced by the variance distribution structure (Xiao et al.,

2011). As the fish grows in three dimensions, its weight grows nonlinearly as a power function of its length, and therefore, it is biologically reasonable to assume a log-normal multiplicative error structure for the variance of the data. Furthermore, as the weight of bigger fish varies more than that of smaller fish, the relationship between fish length and weight does not follow homogeneity (heteroscedasticity) in terms of variance structure (De Giosa and Czerniejewski, 2016). Therefore, it makes sense why the log-linearized models (LM) are so popular for deriving LWRs in fishery biology studies. By log transformation, the log-normal multiplicative error structure with heteroscedasticity (Eq. (1)) could be briefly normalized (Eq. (2)) with an additive and near homogeneous error structure and therefore, this model is very common in the biological literature (Ogle, 2015). However, real-world data may not always follow this popular and biologically sound assumption (Xiao et al., 2011). There is a possibility that the data at hand may have come from a normal distribution with an additive error structure (Eq. (4)) due to inherent limitations in the sampling process or due to the influence of some unknown biological process that is yet to be explained. Under such circumstances, it is better to use a nonlinear modeling approach (Xiao et al., 2011). As the ordinary nonlinear model (NLM) is better at characterizing data with normal, additive, homoscedastic error structure, an improvised version of the nonlinear modeling approach, i.e., weighted nonlinear modeling (wNLM) has been used in the present study to address the heteroscedasticity in error structure in fish weight data. The application of an appropriate weight function (here, power variance structure) in the model (Eq. (5)) in the presence of heteroscedasticity could effectively reduce the influence of data points with high noise (variance or error) and thus makes the model less biased. Such a weighted nonlinear modeling approach (wNLM) has previously been proven to outperform conventional log-linearized modeling approach (De Giosa and Czerniejewski, 2016; Dash et al., 2022).

Since the data collected from nature could have come from either a log-normal distribution with a multiplicative error structure or a normal distribution with an additive error structure, the appropriate approach for modeling is to first ascertain the variance (error) structure so that the modeling parameters and their confidence intervals can be estimated accurately. It has been found that IT-based indices such as AIC and BIC can be successfully used to ascertain the likelihood support for the log-normal multiplicative variance structure or the normal additive



**Fig. 4.** Comparison of modeling parameters (intercept (a); and rate of change (b)) and their confidence intervals derived from different modeling approaches (LM vs. wNLM vs. AwMA) under different error structures. The alphabets in the figure represent different error structures, where Fig. 4a<sub>1</sub> and Fig. 4a<sub>2</sub> represent the prevalence of lognormal multiplicative error structure; Fig. 4b<sub>1</sub> and Fig. 4b<sub>2</sub> represent the prevalence of normal additive error structure; and Fig. 4c<sub>1</sub> and Fig. 4c<sub>2</sub> represent the prevalence of support for both the error structures in the data; The abbreviations in X-axis represent the species-level data categories and applied modeling approaches as follows: Fig. 4a<sub>1</sub> and 4a<sub>2</sub>: PSCP: *Parapenaeopsis sculptilis* sex-pooled data; PSCF: *Parapenaeopsis sculptilis* female data; SCPS: *Solenocera crassicornis* sex-pooled data; SCF: *Solenocera crassicornis* female data; Fig. 4b<sub>1</sub> and 4b<sub>2</sub>: PSTF: *Parapenaeopsis stylifera* female data; SCM: *Solenocera crassicornis* male data; PMP: *Penaeus monodon* sex-pooled data; PMM: *Penaeus monodon* male data; PMF: *Penaeus monodon* female data; Fig. 4c<sub>1</sub> and 4c<sub>2</sub>: PSTP: *Parapenaeopsis stylifera* sex-pooled data; PSTM: *Parapenaeopsis stylifera* male data; PSCM: *Parapenaeopsis sculptilis* male data; PSP: *Penaeus semisulcatus* sex-pooled data; PSM: *Penaeus semisulcatus* female data; PSF: *Penaeus semisulcatus* female data; -LM: linear modeling approach with log transformed data; -wNLM: weighted nonlinear modeling approach with untransformed data; -AwMA: Akaike weighted model averaging approach.

variance structure (Xiao et al., 2011). From the brief introduction about IT-based indices, it is apparent that such indices help to draw multimodel inferences by comparing the information loss under a likelihood framework from several competing models. They determine the accuracy (maximum likelihood) of model parameters while maintaining the principle of model parsimony by penalizing for the inclusion of extra parameters in the models and thus reducing the effect of model over-fitting bias (Burnham and Anderson, 2002). In the present study, the small-sample equivalent of AIC, i.e.,  $AIC_c$  (Eq. (7)) has been used as the IT-based criterion for model selection due to small sample sizes (the ratio between the sample size ( $n$ ) and the number of modeling parameters ( $k$ ), i.e.,  $n/k$  is less than 40) in certain data categories such as *Parapenaeopsis stylifera* (males), *Parapenaeopsis sculptilis* (males) and *Solenocera crassicornis* (males) following the recommendations of Sugiura (1978) and Hurvich and Tsai (1995). In the present study,  $AIC_c$  has been used instead of AIC and BIC as it has been observed that  $AIC_c$  provides a stronger penalty compared to AIC and BIC for small and very small sample sizes, respectively (Brewer et al., 2016).

In the present study, the likelihoods have been calculated from the least squares estimated parameters using Eq. (3) for the LM-derived parameters and Eq. (6) for the wNLM-derived parameters for the subsequent calculation of  $AIC_c$ , which are valid as per the recommendations of Burnham and Anderson (2002). Such switching from a least squares framework to a likelihood-based framework is possible as the maximum likelihood estimates (MLE) for model parameters under the Gaussian distributed error assumption are the same as the ordinary least squares estimates (OLS) or weighted least squares estimates (WLS) (Banks and Joyner, 2017). Furthermore, it must be noted that though an operational normal and constant error structure has been assumed in log-linearized LM to derive the model

parameters, it basically represents an underlying log-normal multiplicative (heterogeneous) error structure in an untransformed scale and therefore, the likelihood of LM has been estimated in the untransformed scale in order to make the response variables from both the models (i.e., LM and wNLM) in an identical scale for the comparison, which is an important criterion for drawing multimodel inference under an information-theoretic paradigm (Burnham and Anderson, 2002).

As the absolute value of  $AIC_c$  is meaningless, it cannot elucidate if the model is good or poor in an absolute sense. Rather, it is the relative difference in the  $AIC_c$  value (i.e.,  $\Delta AIC_c$ , Eq. (8)) of the competing model compared to the best model (with the lowest  $AIC_c$ ), that gives vital information regarding the performance of the competing model. Though a  $\Delta AIC_c$  score-based criterion has been proposed by Burnham and Anderson (2002) for the treatment of competing models, there is still considerable ambiguity regarding the treatment (acceptance or rejection) of competing models if the  $\Delta AIC_c$  falls within the doubtful zone that primarily ranges from 2 to 4 and to some extent even up to 7. In the present study, the  $\Delta AIC_c$  values for the competing model were observed to be approximately 4.0 in the case of *Parapenaeopsis stylifera* (sex-pooled dataset) and *Penaeus semisulcatus* (female dataset). Though, from the  $\Delta AIC_c$  score, it appears that NLM and LM were the best modeling approaches for *Parapenaeopsis stylifera* (sex-pooled dataset) and *Penaeus semisulcatus* (female dataset) respectively, a revalidation attempt using density plots of the residuals does not produce any apparent perceivable difference in the probability densities of the residuals from the competing models (Fig. 2a<sub>3</sub> and 2e<sub>1</sub>). This could be due to the inherent biological limitations in the unbiased random sampling process that affect the variance distribution structure of the datasets. Large individuals in a population exhibit a higher variance in body

**Table 2**

Modeling parameters, their confidence intervals and modeling performance indicators obtained from different modeling approaches (LM, wNLM and AwMA) followed to establish relationship between length and weight of five commercially important marine shrimps of West Bengal, India.

Species	Sex	Model	a	95%CI of a	b	95% CI of b	SD	d	DF	Log Lik.	AIC <sub>c</sub>	RMSE
<i>P. stylifera</i>	P	LM	0.0060	0.0048–0.0075	3.0358	2.9373–3.1342	0.8447	NA	478	–592.99	1192.02	0.8452
	P	wNLM	0.0065	0.0049–0.0081	3.0019	2.8980–3.1058	0.8441	0.50	478	–589.95	1187.99	0.8433
	P	AwMA*	<b>0.0064</b>	<b>0.0050–0.0070</b>	<b>3.0059</b>	<b>2.9678–3.1104</b>	<b>0.8442</b>	NA	<b>478</b>	<b>–590.31</b>	<b>1188.46</b>	<b>0.8433</b>
<i>P. stylifera</i>	M	LM	0.0041	0.0030–0.0056	3.1696	3.0315–3.3077	0.3634	NA	55	–20.59	47.62	0.3610
	M	wNLM	0.0034	0.0023–0.0045	3.2521	3.1135–3.3907	0.3492	0.49	55	–18.39	45.54	0.3461
	M	AwMA*	<b>0.0036</b>	<b>0.0032–0.0051</b>	<b>3.2305</b>	<b>3.0677–3.2844</b>	<b>0.3509</b>	NA	<b>55</b>	<b>–18.96</b>	<b>46.09</b>	<b>0.3480</b>
<i>P. stylifera</i>	F	LM	0.0081	0.0062–0.0106	2.9101	2.7954–3.0247	0.8713	NA	421	–543.70	1093.45	0.8720
	F	wNLM*	<b>0.0083</b>	<b>0.0060–0.0105</b>	<b>2.9039</b>	<b>2.7870–3.0209</b>	<b>0.8713</b>	<b>0.44</b>	<b>421</b>	<b>–536.42</b>	<b>1080.93</b>	<b>0.8702</b>
<i>P. sculptilis</i>	P	LM*	<b>0.0078</b>	<b>0.0058–0.0106</b>	<b>2.9917</b>	<b>2.8785–3.1050</b>	<b>2.9439</b>	NA	<b>478</b>	<b>–1163.38</b>	<b>2332.80</b>	<b>2.9451</b>
	P	wNLM	0.0077	0.0054–0.0101	2.9983	2.8857–3.1109	2.9486	1.03	478	–1169.02	2346.12	2.9456
<i>P. sculptilis</i>	M	LM	0.0038	0.0024–0.0058	3.2791	3.1086–3.4497	1.6244	NA	39	–73.22	153.10	1.6056
	M	wNLM	0.0038	0.0022–0.0054	3.2800	3.1169–3.4431	1.6247	1.02	39	–72.99	155.08	1.6048
	M	AwMA*	<b>0.0038</b>	<b>0.0031–0.0049</b>	<b>3.2794</b>	<b>3.1757–3.3520</b>	<b>1.6244</b>	NA	<b>39</b>	<b>–73.16</b>	<b>153.63</b>	<b>1.6050</b>
<i>P. sculptilis</i>	F	LM*	<b>0.0099</b>	<b>0.0068–0.0143</b>	<b>2.9039</b>	<b>2.7649–3.0428</b>	<b>3.0064</b>	NA	<b>437</b>	<b>–1081.80</b>	<b>2169.65</b>	<b>3.0081</b>
	F	wNLM	0.0101	0.0063–0.0138	2.8989	2.7594–3.0385	3.0075	0.95	437	–1087.46	2183.02	3.0041
<i>S. crassicornis</i>	P	LM*	<b>0.0252</b>	<b>0.0210–0.0303</b>	<b>2.5406</b>	<b>2.4586–2.6226</b>	<b>1.0411</b>	NA	<b>478</b>	<b>–661.74</b>	<b>1329.52</b>	<b>1.0424</b>
	P	wNLM	0.0253	0.0204–0.0302	2.5430	2.4568–2.6292	1.0398	0.81	478	–670.72	1349.53	1.0387
<i>S. crassicornis</i>	M	LM	0.0231	0.0168–0.0317	2.5833	2.4290–2.7376	0.4526	NA	73	–54.42	115.19	0.4508
	M	wNLM*	<b>0.0172</b>	<b>0.0140–0.0203</b>	<b>2.7217</b>	<b>2.6423–2.8012</b>	<b>0.4418</b>	<b>0.68</b>	<b>73</b>	<b>–41.68</b>	<b>91.94</b>	<b>0.4406</b>
<i>S. crassicornis</i>	F	LM*	<b>0.0262</b>	<b>0.0202–0.0339</b>	<b>2.5237</b>	<b>2.4099–2.6375</b>	<b>1.1172</b>	NA	<b>403</b>	<b>–603.54</b>	<b>1213.14</b>	<b>1.1186</b>
	F	wNLM	0.0253	0.0184–0.0322	2.5427	2.4237–2.6617	1.1149	0.61	403	–608.68	1225.46	1.1135
<i>P. monodon</i>	P	LM	0.0057	0.0045–0.0072	3.1042	3.0275–3.1809	7.1869	NA	478	–1626.96	3259.97	7.1889
	P	wNLM*	<b>0.0064</b>	<b>0.0049–0.0078</b>	<b>3.0695</b>	<b>2.9946–3.1443</b>	<b>7.1817</b>	<b>0.55</b>	<b>478</b>	<b>–1608.32</b>	<b>3224.73</b>	<b>7.1742</b>
<i>P. monodon</i>	M	LM	0.0062	0.0043–0.0089	3.0776	2.9545–3.2006	7.2425	NA	233	–804.25	1614.60	7.2316
	M	wNLM*	<b>0.0106</b>	<b>0.0066–0.0146</b>	<b>2.8996</b>	<b>2.7757–3.0236</b>	<b>7.0573</b>	<b>0.31</b>	<b>233</b>	<b>–789.60</b>	<b>1587.37</b>	<b>7.0423</b>
<i>P. monodon</i>	F	LM	0.0054	0.0040–0.0072	3.1238	3.0270–3.2206	7.0388	NA	243	–820.64	1647.37	7.0395
	F	wNLM*	<b>0.0052</b>	<b>0.0037–0.0066</b>	<b>3.1394</b>	<b>3.0466–3.2323</b>	<b>7.0238</b>	<b>0.62</b>	<b>243</b>	<b>–812.62</b>	<b>1633.40</b>	<b>7.0095</b>
<i>P. semisulcatus</i>	P	LM	0.0051	0.0041–0.0063	3.1650	3.0914–3.2386	5.4613	NA	418	–1278.66	2563.37	5.4588
	P	wNLM	0.0053	0.0042–0.0064	3.1532	3.0819–3.2245	5.4597	0.83	418	–1278.09	2564.28	5.4533
	P	AwMA*	<b>0.0052</b>	<b>0.0043–0.0060</b>	<b>3.1604</b>	<b>3.1061–3.2202</b>	<b>5.4609</b>	NA	<b>418</b>	<b>–1278.44</b>	<b>2563.72</b>	<b>5.4555</b>
<i>P. semisulcatus</i>	M	LM	0.0069	0.0047–0.0103	3.0529	2.9153–3.1905	4.2912	NA	238	–679.60	1365.30	4.2876
	M	wNLM	0.0069	0.0043–0.0095	3.0555	2.9250–3.1859	4.2836	0.83	238	–679.39	1366.95	4.2747
	M	AwMA*	<b>0.0069</b>	<b>0.0053–0.0099</b>	<b>3.0537</b>	<b>2.9278–3.1497</b>	<b>4.2888</b>	NA	<b>238</b>	<b>–679.53</b>	<b>1365.80</b>	<b>4.2825</b>
<i>P. semisulcatus</i>	F	LM	0.0104	0.0071–0.0150	2.9374	2.8143–3.0605	6.3632	NA	178	–581.58	1169.29	6.3495
	F	wNLM	0.0109	0.0068–0.0149	2.9225	2.7990–3.0461	6.3617	0.82	178	–582.56	1173.35	6.3440
	F	AwMA*	<b>0.0104</b>	<b>0.0081–0.0131</b>	<b>2.9357</b>	<b>2.8602–3.0186</b>	<b>6.3631</b>	NA	<b>178</b>	<b>–581.69</b>	<b>1169.77</b>	<b>6.3482</b>

P: Sex pooled data; M: Male data; F: Female data; a and b: Model parameters; SD: Standard deviation of the residuals; d: Power used in the power variance structure of wNLM; DF: Degrees of freedom; CI: Confidence interval; LM: log-transformed linearized model; wNLM: Weighted nonlinear model; AwMA: Akaike weighted model averaging; NA: Not available; Log Lik.: Log Likelihood score of the model; AIC<sub>c</sub>: sample size corrected Akaike Information Criterion for the model; RMSE: Root mean squared error; Models in bold and star marked are the appropriate modeling approaches for the data at hand.

weight compared to their smaller counterparts, but are scarcely represented in a sample due to their decreased abundance in a population. Such a limitation might have constrained these real-world datasets in such a manner that they are producing ambiguous variance distribution structures that equally support both the log-normal residual assumption of the LM and the normal residual assumption of the wNLM, leading to data ambivalence even at a  $\Delta AIC_c$  score of 4. This suggests that the competing model should not be rejected simply because their  $\Delta AIC_c$  score is merely above 2; instead, they should also be cross-validated with appropriate tools, such as a residual diagnostic plot, to assess the fulfillment of their assumed variance distribution structures. Therefore, in the present study, some other scaling methods, such as Akaike weights (Aw, Eq. (9)) and evidence ratios (ER, Eq. (10)) have been used to derive additional insights for drawing holistic conclusions about the competing models. As the Aw is derived from the relative likelihood of the competing model, it represents the weight of evidence (also considered analogous to the probability) in favor of the competing model being the best model. In the present study, instead of a  $\Delta AIC_c$  score above 2.0, an Aw score of 0.9 for the best model has been used as the minimum cutoff score for the rejection of the competing model. This was further corroborated by the inspection of the residual probability

density diagnostic plot, as no apparent perceivable evidence for the model superiority could be established even at an Aw score of 0.88 corresponding to a  $\Delta AIC_c$  value above 2.0 (i.e., 4.03–4.06) (Fig. 2a<sub>3</sub> and 2e<sub>1</sub>, Table 1). As the Aw is a value between 0 and 1, a high score of 0.9 for the best model eventually indicates a very low score of 0.1 for the competing model and thus the ratio values (Eq. (10)), i.e., 9.0, which is also known as evidence ratios (ER), can also be used as the minimum cutoff value of ER for the rejection of the competing model. ER, being the ratio of the weight of evidence of the best model (lowest AIC<sub>c</sub>) to that of the competing model (higher AIC<sub>c</sub>), represents the number of times (here, 9 times) the model is likely to be the best model compared to the competing model in terms of Kullback–Leibler discrepancy (Wagenmakers and Farrell, 2004). As the wAIC<sub>c</sub> (Aw) score of the best model increases exponentially in relation to the  $\Delta AIC_c$  of the competing model, an exponential nonlinear regression (Fig. 3) has been performed in the present study to estimate the cut-off value of  $\Delta AIC_c$  using Eq. (11). The  $\Delta AIC_c$  cut-off value of 4.2 has been used for the model treatment, above which the competing models have been rejected in favor of the best model. This is found to be in congruence with the general observation of Burnham et al. (2011), in which it has been recommended that the plausible hypotheses are identified by a narrow  $\Delta AIC_c$



value that probably ranges from 4 to 7. However, under situations where the likelihoods of competing models are equally good at approximating the information in the dataset (i.e.,  $Aw < 0.9$  or  $ER < 9.0$  or  $\Delta AIC_c < 4.2$ ), robust inference has been derived by model averaging (Burnham and Anderson, 2002). Model averaging helps to reduce the impact of model misspecification induced by data ambivalence while accounting for model uncertainty by giving due consideration to the likelihood support for both types of error structures (Burnham et al., 2011).

Furthermore, in the present study, the  $AIC_c$  has also been used in association with the root mean squared error (RMSE), which is a conventional goodness of fit (GOF) indicator, to assess the prediction quality of the competing model and cross-check if the model is well-founded just like the best-performing model. However, in the present study, this conventional GOF indicator has almost always shown that wNLM is the best modeling approach despite the higher  $AIC_c$  score in certain cases (Table 2), which indicates the limitations of RMSE in accurately predicting the performance of the model. Contrary to the  $AIC_c$ , RMSE is biased towards model overfitting and therefore, it cannot reveal if the model is overfitting the observations as it does not penalize the model for the inclusion of extra variable(s) or parameter(s) even though they are biologically not significant. RMSE as a fit (accuracy) performance indicator is usually influenced by outliers and tends to select the model that maximizes fit without giving consideration to model complexity (Liemohn et al., 2021).

## 5. Conclusions

In the present study, it has been illustrated how the variance (error) distribution structure of the data could be ascertained in order to select the appropriate modeling approach for reducing uncertainty while establishing the relationship between the length and weight data of shrimps. A likelihood-based information theoretic criterion such as  $AIC_c$  should be used in congruence with the probability density of residuals through a residual diagnostic plot for the problem-specific formulation of cutoff criteria for the evaluation of competing models. The study reveals that an improvised IT-based cutoff criterion, i.e., a  $\Delta AIC_c$  score of 4.2 and above (corresponding to  $Aw \geq 0.9$  and  $ER \geq 9.0$ ) could be used as a decision support threshold to ascertain the variance distribution structure of data so that the best modeling approach could be followed to increase the accuracy of the model prediction. Though a log-linearized model (LM) is commonly used in fishery biology studies to establish the relationship between the body length and weight of aquatic animals with a biologically sound assumption in favor of a log-normal multiplicative error structure, the present study revealed that such an assumption is not always valid. Contrary to this popular belief, the lowest number of datasets in the present study exhibited a log-normal error structure. Furthermore, the highest number of datasets exhibited equal support for both the log-normal and normal error structures. Therefore, in such a data ambivalence situation, the study also recommends the use of an Akaike weighted model averaging approach to reduce model uncertainty.

## CRedit authorship contribution statement

**Gyanaranjan Dash:** Developed the concept of the paper and formal analysis, Created the initial draft and structure, Preparation of manuscript. **Swatipriyanka Sen:** Participated in writing and validation, Preparation of manuscript. **Rajesh Kumar Pradhan:** Helped in data curation, Preparation of manuscript. **Shubhadeep Ghosh:** Overall supervised the manuscript preparation process, Preparation of manuscript. **Josileen Jose:** Participated in editing of the original manuscript, Preparation of manuscript.

## 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.

## Data availability

Data will be made available on request

## Acknowledgments

The authors are thankful to Dr. A. Gopalakrishnan, the Director, ICAR-Central Marine Fisheries Research Institute, Kochi for providing the infrastructure and facilities to carry out the work. The authors are also indebted to the Indian Council of Agricultural Research, New Delhi, India for funding the research work.

## References

- Banks, H.T., Joyner, M.L., 2017. AIC under the framework of least squares estimation. *Appl. Math. Lett.* 74, 33–45. <http://dx.doi.org/10.1016/j.aml.2017.05.005>.
- Brewer, M.J., Butler, A., Cooksley, S.L., Freckleton, R., 2016. The relative performance of AIC, AICC and BIC in the presence of unobserved heterogeneity. *Methods Ecol. Evol.* 7, 679–692. <http://dx.doi.org/10.1111/2041-210X.12541>.
- Burnham, K.P., Anderson, D.R., 2001. Kullback–Leibler information as a basis for strong inference in ecological studies. *Wildlife Res.* 28 (2), 111–119. <http://dx.doi.org/10.1071/WR99107>.
- Burnham, K.P., Anderson, D.R., 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, second ed. Springer-Verlag, New York, New York, USA. <http://dx.doi.org/10.1007/b97636>.
- Burnham, K.P., Anderson, D.R., Huyvaert, K.P., 2011. AIC model selection and multimodel inference in behavioral ecology: Some background, observations, and comparisons. *Behav. Ecol. Sociobiol.* 65, 23–35. <http://dx.doi.org/10.1007/s00265-010-1029-6>.
- Dash, G., Sen, S., Pradhan, R.K., Ghosh, S., Josileen, J., Jayasankar, J., 2022. Modeling framework for establishing the power law between length and weight of fishes and a meta-analysis for validation of LWRs for six commercially important marine fishes from the northwestern Bay of Bengal. *Fish. Res.* 257, 106496. <http://dx.doi.org/10.1016/j.fishres.2022.106496>, 1–13.
- De Giosa, M., Czerniejewski, P., 2016. A generalized, nonlinear regression approach to the length-weight relationship of European perch (*Perca fluviatilis* L.) from the polish coast of the southern Baltic Sea. *Arch. Pol. Fish.* 24, 169–175. <http://dx.doi.org/10.1515/aopf-2016-0014>.
- Diniz, C.A.R., Louzada-Neto, F., Morita, L.H.M., 2012. The multiplicative heteroscedastic Von Bertalanffy model. *Braz. J. Probab. Stat.* 26 (1), 71–81. <http://dx.doi.org/10.1214/10-BJPS120>.
- Fischer, W., Bianchi, G., 1984. *FAO Species identification sheets for fishery purposes, Western Indian ocean, Fishing Area 51, V. Food and agriculture organization of the United Nations (FAO), Rome, Italy*. <https://www.fao.org/3/ad468e/ad468e00.htm>.
- Flinn, S.A., Midway, S.R., 2021. Trends in growth modeling in fisheries science. *Fishes* 6, 1–18. <http://dx.doi.org/10.3390/fishes6010001>.
- Froese, R., 2006. Cube law, condition factor and weight-length relationships: History, meta-analysis and recommendations. *J. Appl. Ichthyol.* 22, 241–253. <http://dx.doi.org/10.1111/j.1439-0426.2006.00805.x>.
- Froese, R., Tsikliras, A.C., Stergiou, K.I., 2011. Editorial note on weight length relations of fishes. *Acta Ichthyol. Piscat.* 41, 261–263. <http://dx.doi.org/10.3750/AIP2011.41.4.01>.
- Hiramatsu, K., Matsumiya, Y., Kitada, S., 1994. Introduction of suitable stock-recruitment relationship by a comparison of statistical models. *Fish. Sci.* 60 (4), 411–414. <http://dx.doi.org/10.2331/FISHSCI.60.411>.
- Hurvich, C.M., Tsai, C.-L., 1995. Model selection for extended quasi-likelihood models in small samples. *Biometrics* 51, 1077–1084.
- Imai, C., Sakai, H., Katsura, K., Honto, W., Y. Hida., 2002. Growth model for the endangered cyprinid fish *Tribolodon nakamurai* based on otolith analyses. *Fish. Sci.* 2002 68, 843–848. <http://dx.doi.org/10.1046/j.1444-2906.2002.00501.x>.
- Johnson, J.B., Omland, K.S., 2004. Model selection in ecology and evolution. *Trends Ecol. Evol.* 19, 101–108. <http://dx.doi.org/10.1016/j.tree.2003.10.013>.
- Liemohn, M.W., Shane, A.D., Azari, A.R., Petersen, A.K., Swiger, B.M., Mukhopadhyay, A., 2021. RMSE is not enough: Guidelines to robust data-model comparisons for magnetospheric physics. *J. Atmos. Sol.-Terr. Phys.* 218, 105624. <http://dx.doi.org/10.1016/j.jastp.2021.105624>.

- Lukacs, P.M., Burnham, K.P., Anderson, D.R., 2009. Model selection bias and Freedman's paradox. *Ann. Inst. Stat. Math.* 62, 117–125. <http://dx.doi.org/10.1007/s10463-009-0234-4>.
- Ogle, D.H., 2015. Introductory fisheries analysis with R-Chapman and Hall/CRC. Boca Raton, USA, p. 337. <http://dx.doi.org/10.1201/9781315371986>.
- Panhwar, S.K., Liu, Q., Khan, F., 2010. Selecting the best growth model for fish using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). *J. Fish. Soc. Taiwan* 37 (3), 183–190. <http://dx.doi.org/10.29822/JFST.201009.0003>.
- Pauly, D., 1984. Fish population dynamics in tropical waters: A manual for use with programmable calculators. In: *ICLARM Stud. Rev.* 8, Manila, Philippines, p. 325. <https://hdl.handle.net/20.500.12348/3445>.
- Petrakis, G., Stergiou, K.I., 1995. Weight-length relationships for 33 fish species in greek waters. *Fish. Res.* 21, 465–469. [http://dx.doi.org/10.1016/0165-7836\(94\)00294-7](http://dx.doi.org/10.1016/0165-7836(94)00294-7).
- Porch, C.E., Wilson, C.A., Nieland, D.L., 2002. A new growth model for red drum (*Sciaenops ocellatus*) that accommodates seasonal and ontogenetic changes in growth rates. *NOAA Fish* 100, 149–152. <https://spo.nmfs.noaa.gov/content/new-growth-model-red-drum-sciaenops-ocellatus-accommodates-seasonal-and-ontogenetic-changes>.
- Richards, S.A., 2005. Testing ecological theory using the information-theoretic approach: Examples and cautionary results. *Ecol.* 86 (10), 2805–2814. <http://dx.doi.org/10.1890/05-0074>.
- Ricker, W.E., 1968. *Methods for Assessment of Fish Production in Freshwaters*. Blackwell Scientific Publications, Oxford, UK, p. 313.
- Sugiura, N., 1978. Further analysts of the data by Akaike's information criterion and the finite corrections. *Commun. Stat. - Theory Methods* 7 (1), 13–26. <http://dx.doi.org/10.1080/03610927808827599>.
- Symonds, M.R.E., Moussalli, A., 2011. A brief guide to model selection, multi-model inference and model averaging in behavioural ecology using Akaike's information criterion. *Behav. Ecol. Sociobiol.* 65, 13–21. <http://dx.doi.org/10.1007/s00265-010-1037-6>.
- Thorup, K., Alerstam, T., Hake, M., Kjellén, N., 2006. Traveling or stopping of migrating birds in relation to wind: An illustration for the osprey. *Behav. Ecol.* 17, 497–502. <http://dx.doi.org/10.1093/beheco/arj054>.
- Türker, D., Zengin, K., Tünay, Ö.K., 2018. Length-weight relationships for nine chondrichthyes fish species from Edremit Bay (North Aegean Sea). *Turk. J. Fish Aquat. Sci.* 19 (1), 71–79. [http://dx.doi.org/10.4194/1303-2712-v19\\_1\\_09](http://dx.doi.org/10.4194/1303-2712-v19_1_09).
- Wagenmakers, E.J., Farrell, S., 2004. AIC model selection using Akaike weights. *Psychon. Bull. Rev.* 11 (1), 192–196. <http://dx.doi.org/10.3758/BF03206482>.
- Wang, Y., Liu, Q., 2006. Comparison of Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) in selection of stock-recruitment relationships. *Fish. Res.* 77 (2), 220–225. <http://dx.doi.org/10.1016/j.fishres.2005.08.011>.
- Xiao, X., White, E.P., Hooten, M.B., Durham, S.L., 2011. On the use of log-transformation vs. nonlinear regression for analyzing biological power laws. *Ecol.* 92 (10), 1887–1894. <http://dx.doi.org/10.1890/11-0538.1>.
- Zhu, L., Li, L., Liang, Z., 2009. Comparison of six statistical approaches in the selection of appropriate fish growth models. *Chin. J. Oceanol. Limnol.* 27 (3), 457–467. <http://dx.doi.org/10.1007/s00343-009-9236-6>.