UNIVERSITY OF SINDH JOURNAL OF ANIMAL SCIENCES
DOI: https://doi.org/10.57038/usjas.v10i01.8023
Uni. Sindh. J. Anim. Sci., vol. 10(1); 86-99, 2026
Email: editors.usjas@usindh.edu.pk
ISSN (P): 2521-8328
ISSN (E): 2523-6067
Published by University of Sindh, Jamshoro.

PREDICTION OF MORPHOMETRIC CHARACTERISTICS OF POMADASYS MACULATUS FROM THE KARACHI COAST, PAKISTAN, USING PRINCIPAL COMPONENT ANALYSIS (PCA)

SHAHLA SIDDIQUI[1], MUHAMMAD ASIF IQBAL[2]*, TARIQ JAVED[3]

[1]Department of Statistics, University of Karachi, Karachi, Pakistan
[2]Department of Zoology, University of Karachi, Karachi, Pakistan
[3]Government College for Men, Nazimabad, Karachi, Pakistan

ARTICLE INFORMATION

Article History:
Received: 08th April 2026
Accepted: 20th June 2026
Published Online: 30th June 2026

Author Contributions:
SS conducted the study and prepared the manuscript. MAI assisted in data analysis and interpretation. TJ supervised the research and critically reviewed the final manuscript.

Key words:
Principal Component Analysis (PCA), Factor Score (Y_1, Y_2, Y_3), *Pomadysis maculatus*, Karachi Coast, Fisheries Management.

Similarity Index:
14%

SDGs Targeted:
SDG 14: Life Below Water
SDG 13: Climate Action

ABSTRACT

Among the fisheries resources of Pakistan, *Pomadasys maculatus* is an important species due to its economic potential and export value. This study aimed to identify a new set of variables for *P. maculatus* using multivariate statistical methods. Morphometric data of male and female fish were collected from the Korangi Creek area, located at 24°47′N latitude and 67°05′E longitude along the Karachi coast, from January 2013 to December 2014. During the study period, a total of 1,056 specimens, including 503 males and 553 females, were investigated. The dataset included important morphometric variables such as total length (TL, mm), standard length (SL, mm), body weight (Bwt, g), fork length (FL, mm), and other related measurements. Principal Component Analysis (PCA) was applied as a statistical tool to reduce the dimensionality of the data. The first three components, PC1, PC2, and PC3, accounted for 90% of the total variation. PC1 explained 83.96% of the variance and was mainly influenced by TL, Bwt, and SL, representing the overall size of the fish. PC2 explained 3.58% of the variance and was associated with strong swimming muscles, particularly caudal peduncle length (CPL). PC3 contributed 2.5% of the variance and was linked to eye diameter (ED). By applying PCA, 90% of the data variance was preserved while reducing dimensionality. Regression models using factor scores as independent variables were developed to predict future fish characteristics. These models were validated using root mean square error (RMSE) and mean absolute percentage error (MAPE), which showed minimum error values. This study highlights the effectiveness of PCA, and factor score projections in predicting the morphometric characteristics of P. maculatus along the Karachi coast.

1. INTRODUCTION

The importance of fisheries typology has increased significantly in recent years to identify diversity and nutritional value (Aziz *et al.,* 2021). In Pakistan, Grunters, particularly Pomadasys maculatus, are highly valued for their excellent quality and long shelf life (Belton *et al.,* 2018). This species is an essential part of sea trade in the Indo-Pak region (Baset *et al.,* 2020).

*Corresponding Author: asifiqbal@uok.edu.pk

Grunts (Pomadasys maculatus,) are marine fishes recognized for their distinctive sounds. The importance of fisheries classification has increased due to their measurable selection and health benefits (Safi and Khan, 2005; Hargreaves *et al.,* 2017). Fact-based models play a significant role in correlating morphometric factors. However, researchers often rely on distinct measurements. Statistical methodology procedures have gained interest, but few similar studies have been conducted (Aziz *et al.,* 2021). Artificial neural networks have been used to relate fish species distributions to environmental factors (Baset *et* *al.,* 2020). Pomadasyidae family members, including Pomadasys maculatus, are common in Pakistan and have been studied in various countries (Vadziutsina and Riera, 2020).

Globally, fish production reached 174.6 million metric tons in 2020, with a market value of $125.2 billion in 2017, expected to grow to $155.32 billion by 2023 (World Bank, 2022). Pakistan's fish production and valuation have fluctuated, with Pomadasys maculatus being the highest produced and traded species (Marine Fishers Department, 2006). *Pomadysis maculatus* also known as Bloch, is a highly demanded fish species in the Indo-Pacific region, renowned for its excellent quality and long shelf life (Baset *et al.,* 2020). This species is widely distributed in the Southeast Atlantic, Eastern and Western Indian Ocean, and Western Central and Northwest Pacific Oceans, typically inhabiting shallow waters and rocky tide pools (Akhtar and Bilqees, 2011).

The commercial importance of Pomadysis *maculatus* lies in its high-quality flesh, which can be stored for an extended period without spoilage (Safi, 2021). As Asia dominates global fish production, accounting for 68% of total production and 92% of cultivated fish (Maturi *et al.,* 2019), this species is a vital export product for Pakistan. Principal Component Analysis (PCA) is a dimensionality-reduction method used to analyze large datasets by retaining significant information while reducing variables (Abdi and Williams, 2010). A study on "Trichomycterus areolatus" employed PCA to identify morphological variations, revealing that PC1 explained 34.72% of the total variance, with 12 characters related to linear measurements and depth (Colihueque *et al.,* 2017).

2. MATERIALS AND METHODS

Study Area

The investigation area chosen for this study remained Korangi Creek, Karachi coast is situated at 24° 47' N scope 67°5'E longitude (Safi *et al.,* 2014). Random samples of the fish ` *Pomadasys maculatus*` were drawn fortnightly from the commercial area from the commercial area of Korangi Creek during the period of January 2013 to December 2014, on a monthly basis. A total of 1056 specimens were used, where 553 females and 503 males collected primary data.

Dependent and Independent Variables

Morphometric measurements were taken by the same person on the left side of the fish, except for widths and perimeters, following the conventional method described (Diodatti *et al.,* 2008). Animals were measured using a graduated digital calliper. As this study focuses on sex identification therefore, we use fifteen important variables which paly significant role to determine (see table 1 and figure 1).

Statistical Analysis Approaches

Principal Component Analysis (PCA) was performed as a statistical tool to reduce the dimensionality of the data and identify key factors that explain the variation in morphometric measurements.

Regression Models

Factor scores were used as independent variables to predict future morphometric trends. Other than that model validation using Root Mean Square (RMSE) and Mean Absolute Percentage Error (MAPE) used for data validation.

3. RESULTS

The means and standard deviations for the 15 morphometric variables were calculated, and the data showed a clear distinction between male and female morphometric characteristics, with males being generally smaller than females in terms of total length and body weight. First, we calculated the principal component and found the three factor scores labelled as Y_1, Y_2, and Y_3. Secondly, we calculated descriptive statistics with these three scores (Y_1, Y_2, Y_3). Thirdly, checking linearity is checked using graph and calculation. Lastly, found the projection of year (predication next year data on the bases of factor scores).

Furthermore, the predication models are built through factor scores obtained from PCA then we checked validity of predication models validation. The predication model in this study through factor score is as under:

$$\hat{Y}_{T.L\ to\ CFRL} = \beta_o + \beta_1(Y_1\ to\ Y_3)$$

The total variance captured by the first PCA is 0.83, first two PCA is 0.87, first three PCA is 0.90, and for the individual variance captured the variance by first PCA is 12.594, second PCA is 0.538, third PCA is 0.375, and the last PCA is 0. 020.Since 90.047% of the total variance is captured by the first three PCA itself, we take only 3 components of PCA (see table 2).

- Principal Component Analysis (PCA):
  PC1 accounted for 83.96% of the variance, primarily representing fish size and strongly correlated with Total Length (TL), Body Weight (Bwt), and Standard Length (SL). PC2, explaining 3.58% of the variance, was associated with swimming muscles, primarily Caudal Peduncle Length (CPL).
  PC3, explaining 2.5%, was linked to Eye Diameter (ED) (see table 2).

- Interpretation of PCA Results:
  The first principal component (PC1), which accounts for the majority of the variation, is largely driven by Total Length (TL), Body Weight (Bwt), and Standard Length (SL). This component can be interpreted as representing the overall size of *Pomadysis maculatus*. The second principal component (PC2) highlights the importance of swimming muscles, as Caudal Peduncle Length (CPL) is strongly correlated with this component This could suggest that fish with a larger caudal peduncle are better adapted for swimming, which could be an important factor for survival and movement.

The principal component analysis demonstrated three-component. In the first component coefficients (TL, SL, and B.wt) are greater than 0.95 indicating that the important variables. The second component is CPL and the third component is ED (see Figure 2). The principal component shows that the TL, SL, B.wt, CPL, and ED of fish are important variables to identify the fish gender. The coefficients (TL, B.wt and SL) in PC1 are greater than 0.95 indicating that PC1 would be body size of the fish, PC1 score is large for big fish and it is small for small fish.PC2 explains major variation in CPL (0.631) measurement represent the strong swimming muscles of the fish.PC3 shows variation in ED (0.546) is main cause (see table 3).

The result shows the following information for each regression equation fits the data well. This indicates that factor scores (Y_1, Y_2, and Y_3) as independent variables have explanatory power beyond what would be expected by chance. All the results show that the models are important and use to estimate fish variables. Based on R-square, the best predication model for male fish is the one having factor score Y1 as independent variable for the following dependent variables TL, Bwt, HL, SL, BD, LL, POL, AFL, DFL and CFRL. Furthermore, the best predication model for male fish factor score Y_2 as independent variable and CPL as dependent. In female fish factor score Y3 as independent variable and ED as dependent variable is better. Overall, the factor score Y_1 is preferred as compared to Y_2 and Y_3 (see table 4).

The residuals histogram is subsequently plotted gender wise for the morphometry variables TL (mm) to CFRL (mm). A symmetric bell-shaped histogram which is evenly distributed around zero indicates that the normality assumption is likely to be true. The points in a residual plot are randomly dispersed around the horizontal axis, therefore a linear regression model is appropriate for the data. All graphs show that TL to CFRL through factor scores (Y_1, Y_2, and Y_3) are interrelated to each other and mean ±1SD as flows (Figure 3 and 4). Show here TL and B.wt graphs.

The predicted factor scores for female fish are Y_1=-1.01139, Y_2=0.84464, Y_3=0.04221 and for male fish are $Y_{1=}$ -0.06106, Y_2= -1.09666, Y_3= -0.26088. By putting all the values in the above equations, we got the predicted total length as shown in the above table. Similarly, the remaining variables predicted values are calculated.

- Regression Model:
  The regression model using factor scores as independent variables showed high predictive accuracy. RMSE and MAPE values were low, indicating that the model can effectively predict future morphometric trends based on the factor scores.

- Comparison with Other Studies:
  The results are consistent with previous studies that have used PCA for fish morphometry, where size-related variables typically dominate the first component. However, the influence of swimming muscles and eye diameter in the second and third components is a novel finding.

- Implications for Fisheries Management:
  By understanding the key factors that influence fish morphology, fisheries managers can predict population trends and implement more effective management strategies. The regression models based on factor scores can also be used to predict future morphometric changes, which is crucial for maintaining sustainable fish populations Furthermore, predication models are built through factor scores obtained from PCA then we checked validity of predication models validation.

- Validation of the predication performance
  Next we validated all the models in-sample and out-sample by RMSE and MAPE. If, the out-of-sample root mean squared error, also known as the root mean squared prediction error, is substantially lower than the in-sample mean square error, this is a sign of perfection in the model. i.e. RMSE (out sample) less than RMSE (in sample). The model is said to be highly accurate for forecasting if MAPE is less than 10 whereas if MAPE is greater than 50, model should be considered as inaccurate for forecasting (Lewis, 1982).

Based on the results, all variables (TL- CFRL) root mean square error out sample is lower than in sample mean square error so predicated model is better. More over Mean absolute percent error in sample and out sample for the variables TL-CFRL are less than 10 % (see table 5). These morphometric advancements techniques analysis in the fish population include: Principal Component Analysis (PCA). The forecast through factor score has emerged as a new tool.

4. CONCLUSION

This study demonstrates the effectiveness of Principal Component Analysis (PCA) in analysing and predicting the morphometric characteristics of *Pomadysis maculatus* along the Karachi coast. The first three principal components explained 90% of the total variance, enabling accurate predictions of future population trends through regression models.

The study's findings have significant implications for fisheries management and sustainable fishery practices. The prediction models were validated using mean square error and mean absolute percentage error, confirming their statistical significance. The factor score Y1 model was identified as the preferred model. This research contributes to the development of fish variables and provides a valuable framework for future studies on *Pomadysis maculatus* and other fish populations. Moreover, it offers a cost-effective method for predicting *Pomadysis maculatus* populations in Pakistan

5. CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(4), 433–459.

Akhtar, Y., & Bilqees, F. M. (2011). Food and feeding habits of marine edible fish *Pomadasys maculatum* (Bloch, 1797) and description of a new nematode *Dujardinascaris maculatum* n. sp. *International Journal of Biology and Biotechnology*, 8(1), 11–19.

Aziz, M. S., Hasan, N. A., Mondol, M. M., Alam, M. M., & Haque, M. M. (2021). Decline in fish species diversity due to climatic and anthropogenic factors in Hakaluki Haor, an ecologically critical wetland in northeast Bangladesh. *Heliyon*, 7(1), e05861.

Baset, A., Liu, Q., Liao, B., Waris, A., Ahmad, I., Yanan, H., & Qingqing, Z. (2020). Population dynamics of saddle grunt fish, *Pomadasys maculatus* (Bloch, 1793), from Pakistani waters. *Bioprocess Engineering*, 4(1), 1–8.

Belton, B., Bush, S. R., & Little, D. C. (2018). Not just for the wealthy: Rethinking farmed fish consumption in the Global South. *Global Food Security*, 16, 85–92.

Colihueque, N., Corrales, O., & Yáñez, M. (2017). Morphological analysis of *Trichomycterus areolatus* Valenciennes, 1846, from southern Chilean rivers using a truss-based system (Siluriformes, Trichomycteridae). *ZooKeys*, 695, 135–154.

Diodatti, F. C., Tadeu, R., & Freato, T. A. (2008). Parámetros morfométricos en el rendimiento de los componentes corporales de tilapia del Nilo (*Oreochromis niloticus*). *Anales de Veterinaria de Murcia*, 24.

Hargreaves, D., Buckland, A., & Sheaves, M. (2017). Trophic guild concept: Factors affecting within-guild consistency for tropical estuarine fish. *Marine Ecology Progress Series*, 564, 175–186.

Jawad, L. A., Faddagh Ziyadi, M. S., Näslund, J., Pohl, T., & Al-Mukhtar, M. A. (2018). Checklist of the fishes of the newly discovered coral reef in Iraq, north-west Arabian Gulf, with 10 new records to the Arabian Gulf. *Aqua, Journal of Ichthyology and Aquatic Biology*, 24(3), 89–138.

Johnson, P., & Williams, K. (2019). Fisheries management and modeling techniques. *Fishery Management Review*, 47(3), 202–215.

Lewis, C. D. (1982). *Industrial and Business Forecasting Methods*. Butterworths, London.

Marine Fisheries Department. (2006). *Handbook of Fisheries Statistics of Pakistan*. Government of Pakistan, Fish Harbour, West Wharf, Karachi.

Maturi, V., Mabudi, H., & Valinassab, T. (2019). The stock surveys and distribution of Haemulidae in the northern Oman Sea, Sistan and Baluchistan Province. *Journal of Marine Biology*, 11(1), 13–24.

Rutberg, S., & Bouikidis, C. D. (2018). Focusing on the fundamentals: A simplistic differentiation between qualitative and quantitative research. *Nephrology Nursing Journal*, 45(2), 209–213.

Safi, A. (2021). Gonadal morphology, histology and spermatogenesis of striped grunt fish, *Pomadasys stridens* (Forsskål, 1775) (Family: Pomadasyidae). *Journal of Zoological Research*, 3(1).

Safi, A. K., & Khan, M. A. (2005). Food and feeding habits of spotted grunt fish, *Pomadasys maculatum* (Bloch, 1747) (Family: Pomadasyidae), from Karachi coast, Pakistan. *International Journal of Biology and Biotechnology*.

Safi, A., Khan, M. A., & Khan, M. Z. (2014). Study of some morphometric and meristic characters of saddle grunt fish, *Pomadasys maculatus* (Bloch, 1793), off Karachi coast, Pakistan. *American Journal of Marine Science*, 2(2), 38–42.

Siddiqui, S., & Fatima, S. (2019). Assessment of morphometric and meristic characteristics of *Pomadasys maculatum* of Karachi coast. *International Journal of Biosciences*, 14(6), 383–388. https://doi.org/10.12692/ijb/14.6.383-388

Siddiqui, S., & Fatima, S. (2020). Multivariate techniques to *Pomadasys maculatum* (Bloch) of Karachi coast, Pakistan. *Journal of Biodiversity and Environmental Sciences*, 17(1), 85–92.

Smith, J. A., & Lee, R. (2020). Morphometric analysis of fish populations using PCA. *Journal of Marine Science*, 58(2), 123–134.

Talwar, P. K., & Jhingran, A. G. (1991). *Inland Fishes of India and Adjacent Countries* Vols. I–II. Oxford & IBH Publishing Co., New Delhi, India.

Vadziutsina, M., & Riera, R. (2020). Review of fish trap fisheries from tropical and subtropical reefs: Main features, threats and management solutions. *Fisheries Research*, 223, 105432.

World Bank. (2022). *World Bank Open Data*. Retrieved from https://data.worldbank.org

Table 1. Morphometric measurements Pomadasys maculatus used in this study

Character	Description	Acronyms
Independent / Dependent variables		
Total length	Tip of the upper jaw to the caudal end of the caudal fin	TL (mm)
Head length	Distance between the snout or premaxilla and the posterior most edge of the opercular bone.	HL (mm)
Snout Length	Horizontal distance from the tip of the snout to the anterior margin of the pigmented region of the eye.	Sn.L (mm)
Eye diameter	The greatest bony diameter of the orbit	ED (mm)
Standard Length	The greatest bony diameter of the orbit	SL (mm)
Fork Length	The snout to the end of the middle caudal fin rays	FL (mm)
Body Breath	Fish breathe with their gills.	BB (mm)
Body Depth	The vertical distance between body margins (exclusive of fins) through the anterior margin of the pectoral fin base.	BD (mm)
Post-orbit Length	Distance from the posterior margin of the orbit to the end of the two orbitals.	POL (mm)
Caudal peduncle length	Distance from the end anal fin the base of the median caudal rays.	CPL (mm)
Anal fin length	Length of the longest ray from the anterior insertion of anal fin.	AFL (mm)
Lateral Length	The lateral line is a sense organ used to detect movement and vibration in the surrounding water.	LL (mm)
Dorsal fin length	Length of the longest ray from the anterior insertion of anal fin.	DFL (mm)
Caudal Fin Rays length	The caudal fin is the tail fin located at the end of the caudal peduncle.	CFRL (mm)
Factor Scores	Through Principal component	Y_1, Y_2, Y_3
Dependent Variables		
Body weight	Total weight including gut and gonads	B. wt (g)
Categorical variables		
Gender	Male, Female	M, F
Months	January – December	2013-2014

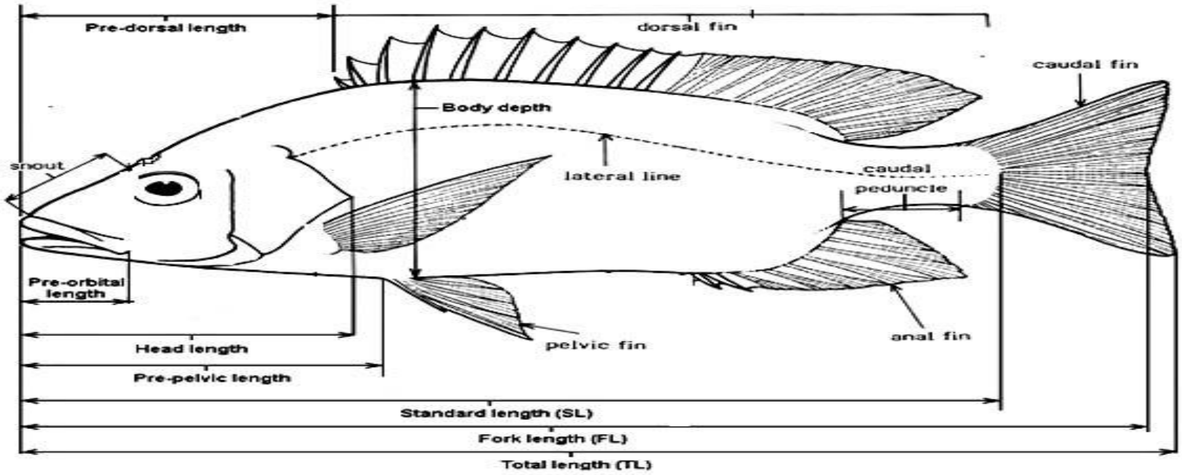

Figure 1. Picto-diagram for morphometry variables

Table 2. Morphometric Total Variance Explained in *Pomadysis maculatus* (Fixed factor-3) Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	12.594	83.960	83.960	12.594	83.960	83.960
2	.538	3.587	87.547	0.538	3.587	87.547
3	.375	2.500	90.047	0.375	2.500	**90.047**
4	.334	2.226	92.273			
5	.227	1.511	93.784			
6	.194	1.291	95.075			
7	.148	.986	96.061			
8	.135	.901	96.962			
9	.116	.776	97.738			
10	.102	.682	98.420			
11	.081	.539	98.959			
12	.063	.418	99.377			
13	.040	.270	99.647			
14	.033	.222	99.869			
15	.020	.131	100.000			

* Extraction Method: Principal Component Analysis'; *3 components extracted; TL: total length, B.wt: Body weight ,HL: Head length,BD:Body depth, Sn.L: Snout length , ED :Eye diameter , SL: Standard length , FL: Fork length, BB: Body breath , POL:Post-orbit length, CPL: Caudal peduncle length,AFL:Anal fin length, LL :Lateral line scales ,DFL: Dorsal fin rays length , CFRL: Caudal fin rays length, sample size :Male 503 and Female 553.

Table 3. Morphometric Component Matrix in *Pomadysis maculatus*

Component Matrix			
	Component*		
	1	2	3
TL (mm)	**0.984**	0.001	-0.031
SL (mm)	**0.974**	0.028	-0.003
Bwt (gm)	**0.949**	-0.12	-0.115
BD (mm)	0.935	-0.065	-0.065
FL (mm)	0.951	-0.01	-0.027
HL (mm)	0.963	-0.027	-0.01
SnL (mm)	0.943	-0.046	0.02
ED (mm)	0.802	-0.224	**0.546**
BB (mm)	0.826	-0.226	-0.089
CPL (mm)	0.73	**0.631**	0.156
POL (mm)	0.939	0.014	-0.114
LL (mm)	0.906	0.069	-0.04
DFL (mm)	0.942	-0.011	-0.01
AFL (mm)	0.972	-0.025	-0.019
CFRL (mm)	0.888	0.102	-0.098

* Extraction Method: Principal Component Analysis'; *3 components extracted; TL: total length, B.wt: Body weight, HL: Head length,BD:Body depth, Sn.L: Snout length , ED :Eye diameter , SL: Standard length , FL: Fork length, BB: Body breath , POL:Post-orbit length, CPL: Caudal peduncle length,AFL:Anal fin length, LL :Lateral line scales ,DFL: Dorsal fin rays length , CFRL: Caudal fin rays length, sample size :Male 503 and Female 553.

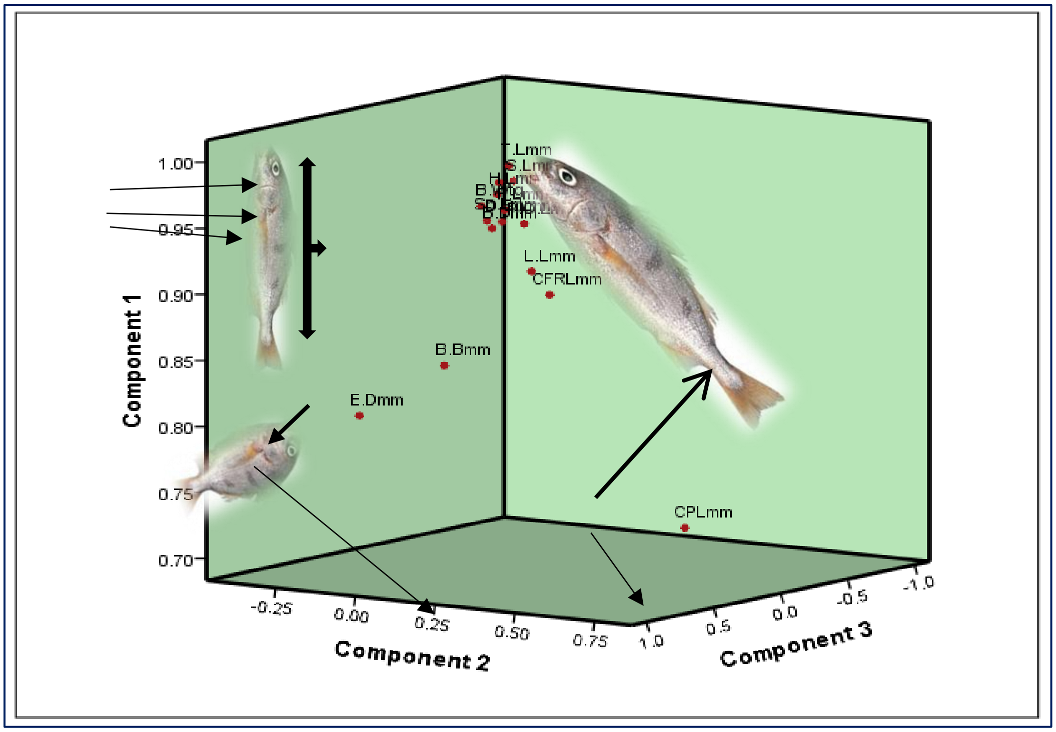

Figure 2. Component plot in Morphometric variables

Table 4. Univariate regression analysis results of various outcome variables and each of factor scores by gender as bellow:

Variable	Factor score	Gender		R Square	
		Male	**Female**	**Female**	**Male**
TL(mm)	Y_1	$168.131+21.462Y_1$	$171.734+17.729Y_1$	0.614	0.697
	Y_2	$164.751+12.910Y_2$	$174.586+7.122\ Y_2$	0.105	0.253
	Y_3	$165.728+13.259\ Y_3$	$174.305+5.948Y_3$	0.09	0.178
BWt(gm)	Y_1	$75.173+24.958\ Y_1$	$82.376+25.174Y_1$	0.694	0.806
	Y_2	$70.785+10.296Y_2$	$86.835+5.477Y_2$	0.035	0.137
	Y_3	$71.579+10.667\ Y_3$	$86.176+7.472Y_3$	0.079	0.098
SL(mm)	Y_1	$132.462+16.697\ Y_1$	$136.247+13.169Y_1$	0.556	0.656
	Y_2	$129.947+11.204Y_2$	$138.327+5.733Y_2$	0.112	0.296
	Y_3	$130.598+10.350\ Y_3$	$138.065+5.019Y_3$	0.105	0.168
HL(mm)	Y_1	$43.027+5.660\ Y_1$	$44.508+4.827Y_1$	0.561	0.687
	Y_2	$42.135+3.390Y_2$	$45.303+1.727Y_2$	0.076	0.247
	Y_3	$42.415+3.607\ Y_3$	$45.178+1.816Y_3$	0.103	0.186
BD(mm)	Y_1	$49.978+7.892\ Y_1$	$51.006+6.040Y_1$	0.565	0.743
	Y_2	$48.602+3.373Y_2$	$52.002+2.145Y_1$	0.076	0.136
	Y_3	$48.847+3.402Y_3$	$51.830+2.362Y_3$	0.112	0.092

Variable	Factor score	Gender		R Square	
		Male	**Female**	**Female**	**Male**
FL(mm)	Y_1	$157.363+19.766\ Y_1$	$161.563+16.755Y_1$	0.633	0.593
	Y_2	$154.310+12.503Y_2$	$164.335+5.863Y_2$	0.082	0.238
	Y_3	$155.203+12.526\ Y_3$	$163.993+5.623Y_3$	0.092	0.159
Sn.L(mm)	Y_1	$13.507+2.006\ Y_1$	$14.277+1.682Y_1$	0.528	0.635
	Y_2	$13.192+1.224Y_2$	$14.559+0.549Y_2$	0.06	0.237
	Y_3	$13.295+1.319\ Y_3$	$14.495+0.737Y_3$	0.132	0.183
ED(mm)	Y_1	$12.687+1.083\ Y_1$	$13.360+0.688Y_1$	0.097	0.315
	Y_2	$12.527+0.760Y_2$	$13.475+0.229Y_2$	0.011	0.156
	Y_3	$12.780+1.942\ Y_3$	$13.234+1.711Y_3$	0.775	0.678
BB(mm)	Y_1	$23.165+3.305\ Y_1$	$23.739+2.823Y_1$	0.53	0.679
	Y_2	$22.528+0.788\ Y_2$	$24.247+0.212Y_2$	0.003	0.039
	Y_3	$22.758+1.822\ Y_3$	$24.149+0.944Y_3$	0.077	0.138
LL(mm)	Y_1	$20.120+2.958\ Y_1$	$20.665+2.681Y_1$	0.445	0.634
	Y_2	$19.651+1.744Y_2$	$21.105+0.984Y_2$	0.133	0.268
	Y_3	$19.716+1.395\ Y_3$	$21.100+0.602Y_3$	0.047	0.156
POL(mm)	Y_1	$106.580+14.132\ Y_1$	$108.097+11.512Y_1$	0.642	0.666
	Y_2	$104.412+9.179Y_2$	$109.818+6.110Y_2$	0.092	0.232
	Y_3	$104.971+8.569\ Y_3$	$109.857+3.277Y_3$	0.042	0.099
CPL(mm)	Y_1	$17.043+1.230\ Y_1$	$17.819+0.639Y_1$	0.046	0.168
	Y_2	$17.044+2.739\ Y_2$	$17.716+2.573Y_2$	**0.799**	**0.832**
	Y_3	$16.\ 956+1.060Y_3$	$17.901+0.278\ Y_3$	0.011	0.083
DFL(mm)	Y_1	$54.605+7.257\ Y_1$	$56.478+5.900Y_1$	0.517	0.664
	Y_2	$53.454+4.283Y_2$	$57.427+2.368Y_2$	0.088	0.232
	Y_3	$53.802+4.542\ Y_3$	$57.297+2.218Y_3$	0.095	0.174
AFL(mm)	Y_1	$88.040+11890\ Y_1$	$90.197+9.965Y_1$	0.611	0.668
	Y_2	$86.187+7.344Y_2$	$91.858+3.345Y_2$	0.073	0.255
	Y_3	$86.735+7.502\ Y_3$	$81.598+3.633Y_3$	0.105	0.178
CFRL(mm)	Y_1	$35.972+4.890\ Y_1$	$37.026+4.001Y_1$	0.493	0.577
	Y_2	$35.243+3.364Y_2$	$37.627+2.091Y_2$	0.143	0.274
	Y_3	$35.245+1.952Y_3$	$37.672+0.915Y_3$	0.033	0.061
Dependent variables: Total length (TL), Bodyweight (Bwt), Head length (HL), Snout length (Sn.L), Eye diameter (ED), Standard length (SL) and, Fork length (FL), Body breath (BB), Body depth (BD), Post-orbit length (POL) and Caudal peduncle length (CPL), Anal fin rays (AFL), Lateral line scales (LL), Caudal fin rays (CFRL), Dorsal fin rays length (DFL), Independent variable: Factor scores (Y_1, Y_2, and Y_3).					

Residual plot for morphometry variables

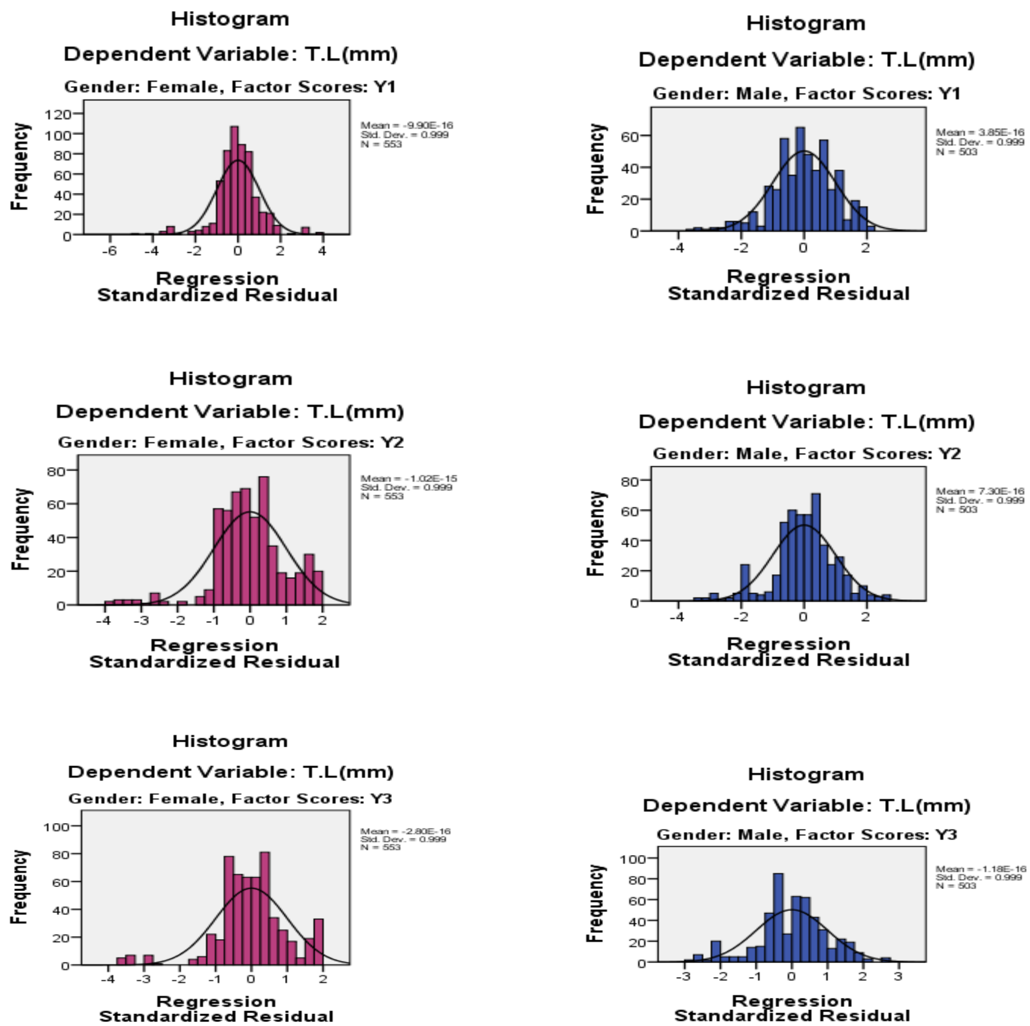

Figure 3. Gender wise representation Factor Scores (Y_1, Y_2, Y_3) between T.L (mm)

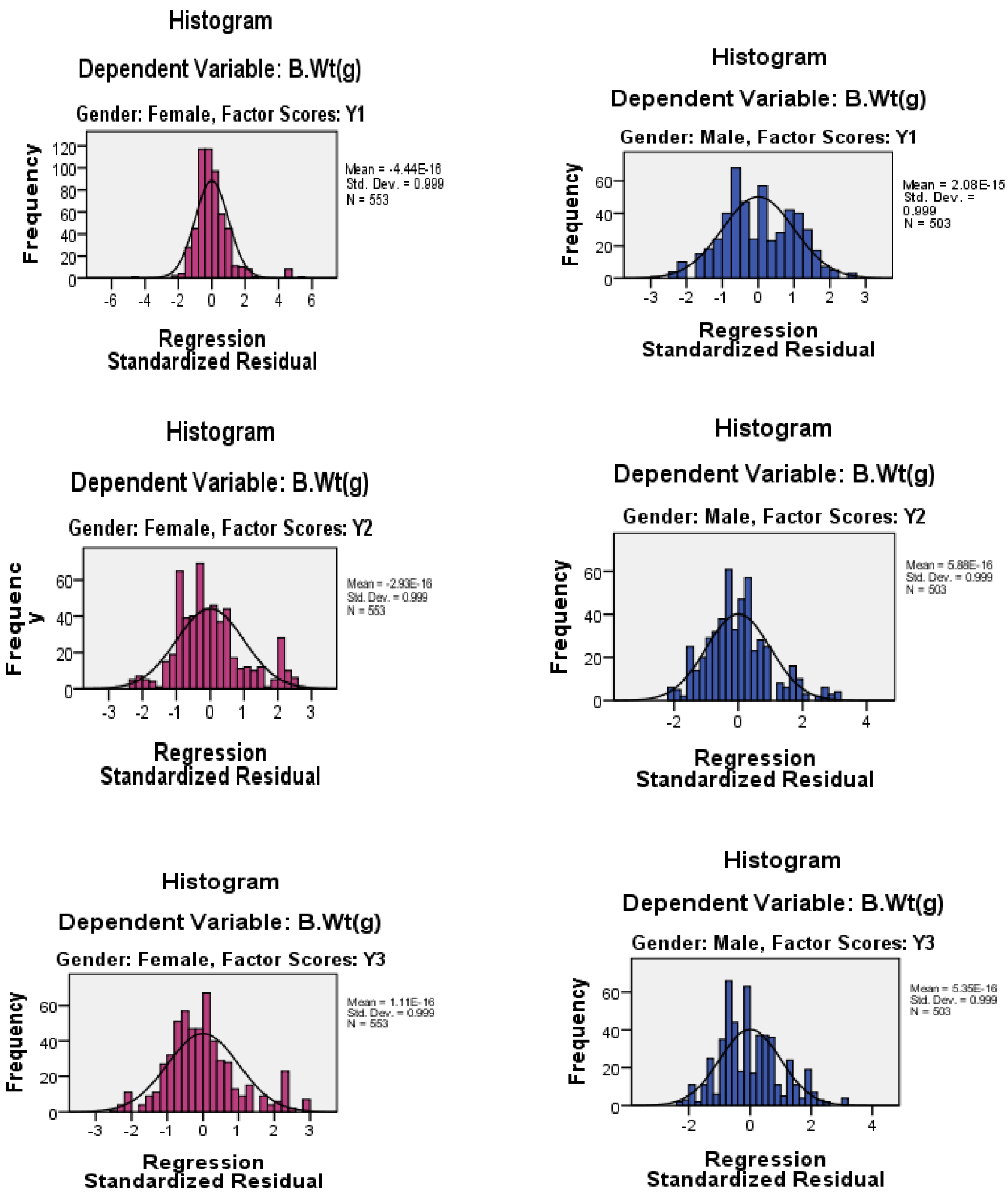

Figure 4. Gender wise representation Factor Scores (Y_1, Y_2, Y_3) between B.wt (g)

- **Prediction model for TL solve an example**

In this section, the total length (TL) of next female and male fish are predicted from the below equation.

Male		Female	
Estimated	**Actual**	**Estimated**	**Actual**
$TL=168.131+21.462Y_1=166.8205$	TL=150 for Y_1	$TL=171.734+17.729Y_1=153.8031$	TL=181 for Y_1
$TL=164.751+12.910Y_2=150.5931$	TL=150 for Y_2	$TL=174.586+7.122Y_2=180.6015$	TL=181 for Y_2
$TL=165.728+13.259Y_3=162.269$	TL=150 for Y_3	$TL=174.305+5.948Y_3=174.5561$	TL=181 for Y_3

Table 5. Validation models in-sample and out-sample by RMSE and MAPE results of various outcome variables and each of factor scores by gender as bellow:

Variable	Factor Score	RMSE				MAPE			
		IN SAMPLE		OUTSAMPLE		IN SAMPLE		OUTSAMPLE	
		Female	Male	Female	Male	Female	Male	Female	Male
TL	Y1	14.1446	13.3814	6.40E-05	1.00E-04	6.12%	7.13%	0.00%	0.00%
	Y2	21.5394	21.0172	7.34984	12.1738	9.96%	10.72%	2.96%	6.26%
	Y3	21.7259	22.0446	6.78774	10.221	9.89%	11.48%	2.18%	4.97%
SL	Y1	11.8442	11.8442	4.80E-05	7.80E-05	6.50%	7.69%	0.00%	0.00%
	Y2	16.7521	16.3628	5.91711	10.565	9.75%	10.57%	3.01%	6.91%
	Y3	16.8193	17.7815	5.72756	7.97796	9.89%	11.63%	2.32%	4.92%
B.wt	Y1	16.8305	11.5799	9.10E-05	0.00012	16.28%	17.79%	0.01%	0.00%
	Y2	29.8846	24.4183	5.65233	9.7095	39.79%	42.88%	4.59%	12.04%
	Y3	29.1884	24.9646	8.52653	8.22273	37.89%	43.76%	5.45%	9.73%
BD	Y1	5.33833	4.38515	2.20E-05	3.70E-05	7.66%	8.60%	0.00%	0.00%
	Y2	7.62382	8.04675	2.21365	3.18052	12.26%	15.94%	2.99%	5.53%
	Y3	7.62382	8.24768	2.69496	2.62257	11.97%	16.79%	2.89%	4.30%
FL	Y1	12.8525	15.4798	6.10E-05	9.20E-05	6.03%	10.62%	0.00%	0.01%
	Y2	20.2001	21.1899	6.05085	11.7903	9.99%	13.74%	2.59%	6.48%
	Y3	20.2001	22.2553	6.41663	9.65558	9.80%	14.66%	2.19%	5.01%
HL	Y1	4.29628	3.61717	1.80E-05	2.60E-05	7.28%	7.48%	0.00%	0.00%
	Y2	6.23358	5.60761	1.7821	3.19634	11.37%	11.21%	2.76%	6.43%
	Y3	6.14254	5.8279	2.07222	2.78018	10.95%	11.62%	2.56%	5.29%
Sn.L	Y1	1.59938	1.43788	6.10E-06	9.30E-06	8.38%	9.77%	0.01%	0.01%
	Y2	2.25836	2.07971	0.56665	1.1538	12.46%	13.50%	2.74%	7.45%
	Y3	2.17026	2.15105	0.84124	1.01659	12.00%	13.99%	3.22%	6.23%
ED	Y1	2.11817	1.50898	2.50E-06	5.00E-06	9.61%	10.39%	0.00%	0.00%
	Y2	2.2158	1.67567	0.23601	0.71664	10.58%	11.00%	1.23%	4.82%
	Y3	1.05665	1.03502	1.95244	1.49696	6.26%	7.30%	8.14%	9.95%
BB	Y1	2.67406	2.14714	1.00E-05	1.50E-05	8.53%	8.15%	0.00%	0.01%
	Y2	3.89597	3.71831	0.21894	0.74352	13.01%	13.81%	0.63%	2.76%
	Y3	3.74912	3.52124	1.0769	1.4043	12.42%	13.52%	2.49%	4.97%

Variable	Factor Score	RMSE				MAPE			
		IN SAMPLE		OUTSAMPLE		IN SAMPLE		OUTSAMPLE	
		Female	Male	Female	Male	Female	Male	Female	Male
CPL	Y1	2.91619	2.59319	2.30E-06	5.70E-06	12.37%	13.97%	0.00%	0.00%
	Y2	1.33756	1.16339	2.6558	2.58332	6.14%	5.81%	11.05%	13.45%
	Y3	2.96912	2.72141	0.31725	0.81711	13.06%	14.65%	1.00%	3.84%
POL	Y1	2.01575	1.98158	9.70E-06	1.40E-05	7.78%	8.86%	0.40%	0.00%
	Y2	3.20977	3.00437	1.0155	1.64495	12.54%	12.60%	3.38%	7.12%
	Y3	3.29671	3.25421	0.68691	1.07513	12.59%	13.83%	1.83%	4.37%
LL	Y1	12.9335	10.159	4.20E-05	6.60E-05	12.83%	9.67%	0.00%	0.00%
	Y2	16.1671	14.3652	6.30557	8.65571	15.71%	12.67%	4.04%	7.05%
	Y3	16.954	15.4258	3.73926	6.60572	16.54%	14.11%	1.91%	5.07%
DFL	Y1	5.74462	4.88747	2.10E-05	3.40E-05	7.55%	8.01%	0.00%	0.00%
	Y2	7.88904	7.38621	2.44423	4.03915	11.33%	11.61%	2.99%	6.41%
	Y3	7.86166	7.6593	2.53101	3.50106	11.08%	12.23%	2.46%	5.25%
AFL	Y1	8.00582	7.92732	3.60E-05	5.50E-05	6.65%	8.34%	0.00%	0.00%
	Y2	12.3556	11.873	3.45256	6.9256	11.02%	11.91%	2.64%	6.83%
	Y3	12.1389	12.4762	4.14565	5.78284	10.85%	12.51%	2.52%	5.39%
CFRL	Y1	4.08652	3.956	1.50E-05	2.30E-05	8.69%	10.38%	0.00%	0.00%
	Y2	5.3121	5.18634	2.15759	3.17243	11.70%	12.96%	4.04%	7.68%
	Y3	5.64121	5.89584	1.0436	1.50482	12.07%	15.16%	1.56%	3.39%

TL: total length, Bwt: Body weight ,HL: Head length,BD:Body depth , Sn.L: Snout length , ED :Eye diameter , SL: Standard length , FL: Fork length ,BB: Body breath , POL: Post-orbit length, CPL: Caudal peduncle length, AFL: Anal fin length, LL: Lateral line scales ,DFL: Dorsal fin rays length , CFRL: Caudal fin rays length, Factor score(Y1,Y2,Y3),RMSE: Root mean square error(*RMSE-Out sample<RMSE-In sample is better model) MAPE: Mean absolute percent error(MAPE less than 5% is considered as an indication that the forecast is acceptably accurate),sample size :Male 503 and Female 553.