Journal of Fisheries
Volume 8 Issue 2 Pages 837–842 August 2020
Peer Reviewed | Open Access | Online First

eISSN 2311-3111
pISSN 2311-729X

Original Article

Development of a model for predicting mussel weight: a comparison of traditional and artificial intelligent methods

Kaent Immanuel N. Uba

Department of Fisheries Science and Technology, School of Marine Fisheries and Technology, Mindanao State University at Naawan, Pedro Pagalan St., Poblacion, Naawan, Misamis Oriental, Philippines

Correspondence
Kaent Immanuel N. Uba; Department of Fisheries Science and Technology, School of Marine Fisheries and Technology, Mindanao State University at Naawan, Pedro Pagalan St., Poblacion, Naawan, Misamis Oriental, Philippines.

kaentimmanuel.uba@msunaawan.edu.ph

Manuscript history
Received 30 March 2020 | Revised 20 July 2020 | Accepted 22 July 2020 | Published online 27 July 2020

Citation
Uba KIN (2020) Development of a model for predicting mussel weight: a comparison of traditional and artificial intelligent methods. Journal of Fisheries 8(2): 837–842.

Abstract

The relationship between length and weight is non-linear. Predictive modelling using linear regression methods subjects these variables to transformation which results in models of poor predictive value. Hence, a comparative study on developing a predictive model using traditional (length-weight relationship, LWR; multiple linear regression, MLR) and artificial intelligent (artificial neural networks, ANN) methods was conducted. Specimens (n = 320) of the horse mussel *Modiolus modulaides* were randomly collected from October 2018 to March 2019 at the coastal area of Dumangas, Iloilo, Philippines. Shell length, shell width and shell height were used as predictor variables for total weight. A multi-layer perceptron architecture model was used and the values were determined by the ANNs model using the actual data. In addition, LWR and MLR models were generated from the same data after log-transformation. The results indicated superiority of the ANN model to predict mussel weight to traditional LWR and MLR models. The ANNs model had the highest correlation coefficient and lowest errors among the predictive models. The ANNs model generated from this study can be a good alternative to existing models and may be useful in sustainable fisheries management.

Keywords: Artificial neural networks; fisheries biology; length-weight relationship; multilayer perceptron; predictive modelling

1 | INTRODUCTION

The horse mussel *Modiolus modulaides* (Röding 1798) is a semi-infaunal species of mytilid mussel distributed throughout the Indo-Pacific region (Poutiers 1998). Throughout this range, the horse mussel is not only an important component of the coastal ecosystem but also an important commodity in small-scale fisheries (Morton 1977; Tumanda Jr *et al.* 1997; Poutiers 1998; Ozawa 2001; Napata and Andalecio 2011; Uba *et al.* 2019).

In bivalves, the shell size (length, height and width), total weight and combined calculation of the size-weight ratio are generally used for analyses of the morphological development of the shell and its condition (Anderson and Gutreuter 1983; Rainer and Mann 1992; Gosling 2015). In addition, studies on bivalve growth are essential for generating useful information for managing resources, understanding changing environmental conditions and pollution as well as phenotypic variability in populations (Palmer 1990; Boulding and Hay 1993; Caceres-Martinez *et al.* 2003).

The most important variables in modelling the growth of an aquatic organism are the length and weight. Basically, a length-weight relationship model is achieved through linear regression which is used in the calculation of an equation of growth in length into an equation of growth in weight (Pauly 1983). Moreover, they are of special interest for underwater visual censuses, in order to transform length data into weight data and, as a result, obtain biomass estimates (Samoilys 1997). However, the relationship of length and weight is non-linear (Froese 2006) and transformations in linear regression methods result in poor predictive value. Thus, the traditional methods of statistical analysis (*i.e.* linear regression models, both single and multiple) may be inadequate for quantification (Suryanarayana *et al.* 2008; Tureli Bilen *et al.* 2011; Benzer *et al.* 2017; Benzer and Benzer 2019).

At present, the application of artificial neural networks (ANNs) in predictive modelling offers a promising alternative to traditional statistical approaches when non-linear patterns exist. ANNs are computer algorithms that simulate the activity of neurons and information processing in the human brain. Unlike the more commonly used regression models, neural networks do not require a particular functional relationship or distribution assumptions about the data and do not need data transformation. This makes neural network modelling a powerful tool for exploring complex, nonlinear biological problems such as in fisheries research (Suryanarayana *et al.* 2008).

Several studies have been conducted on predictive modelling using neural networks on the growth of fishes (Benzer and Benzer 2016, 2017, 2020; Özcan and Serdar 2018) and crustaceans (Türeli Bilen *et al.* 2011; Benzer *et al.* 2015) but none on bivalves. Thus, recognising these problems and the ability of neural network models to approximate non-linear relationships, an ANN model was developed in this study to predict the weights of individual horse mussel based on three predictor variables (*i.e.* shell length, shell width and shell height).

2 | METHODOLOGY

2.1 Data collection

Horse mussel specimens (n = 320) were randomly collected from the subtidal area (< 1.5 m depth during low tide) of Dumangas, Iloilo, Philippines from October 2018 to March 2019 (Figure 1). Specimens were immediately preserved in 10% seawater-buffered formalin and brought to the laboratory for morphometric analysis. In the laboratory, the sample specimens were cleaned free of epibionts before measurement. Morphometric measurements such as shell length (widest part across the shell at 90° to the height), shell width (thickest part of the two shell valves), and shell height (distance from the hinge line to the shell margin) were measured to the nearest 0.01 cm using a digital Vernier Caliper while the total weight was measured to the nearest 0.01 g using a digital top-loading balance after blot drying in paper towels.

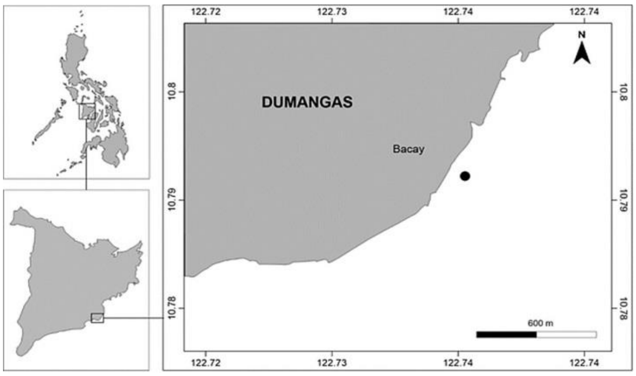

FIGURE 1 Location of the study site. Black dot indicates the horse mussel sampling area.

2.2 Mathematical models

2.2.1 Length-weight relationship (LWR) and multiple linear regression (MLR) equations

The LWR was calculated following the equation of Keys (1928),

$$W = aL^b$$

where W represents the total weight (g), L is the shell length (cm), and a and b are the coefficients denoting functional regression between W and L. The logarithmic transformation of the formula above is Log W = log a + b log L.

In addition, MLR was calculated following the equation,

$$\hat{y} = a + b_1X_1 + b_2X_2 + b_3X_3$$

where $\hat{y}$ is the total weight (g), a is the regression intercept, b is the regression slope, and X is the predictor variable (*i.e.* shell length, shell width and shell height). The data were log-transformed before calculations and analyses were performed in the SigmaPlot version 12 software.

2.2.2 Artificial Neural Networks (ANNs)

ANNs are simulations of biological nervous systems using mathematical models. They are networks with simple processor units, interconnections, adaptive weights and scalar measurement functions (Rumelhart *et al.* 1986). In this study, a multilayer feed-forward neural network was used during the ANN operation and consisted of three interconnected layers of nodes (Figure 2) including an input layer containing one node per independent variable (*i.e.* shell length, shell width and shell height), two hidden layers and an output layer with one node (the total weight). Each layer was connected to another layer with interconnections and adaptive weight values. The nodes were connected to next layer nodes with adjustable weights.

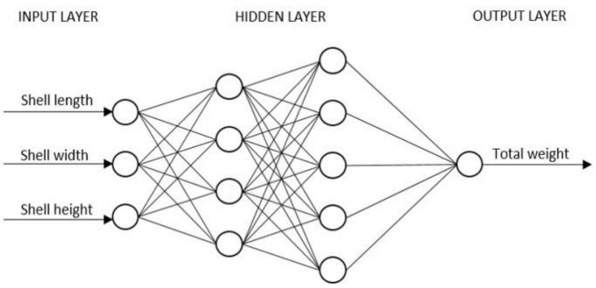

FIGURE 2 Topology of the four-layer feed-forward neural network used in this study consisting of an input layer with three nodes (predictor variables), two hidden layers with four and five nodes respectively and an output layer with one node (total weight) to be predicted.

The ANN calculations were expressed in the following formula (Krenker *et al.* 2011),

$$y = f(n)\left(\sum_{i=1}^{p} W_i x_i + b\right)$$

where x_i is the input, $f(n)$ is the activation function, y is the output value, and W_{i} is the weight. To adjust the connection weights and minimise the error between observed and predicted values, network training (split: 70.0% training, 30.0% testing) was conducted using a back propagation training procedure through the multilayer perceptron algorithm in the Weka version 3.8.4 software (Witten *et al.* 2017). The data used in the ANNs were subjected to normalisation process for range of [0,1] using the following formula,

$$V_N = 0.8 \times \left(\frac{V_R - V_{min}}{V_{max} - V_{min}}\right) + 0.1$$

where V_N is the normalised data, V_R is the data to be normalised, V_{min} is the minimum value of the data, and V_{max} is the maximum value of the data.

2.3 Performance of predictive models

To assess the performance of the predictive models generated in this study, four criteria were used for evaluation including (1) the correlation coefficient (r) between the observed weights and predicted weights of the horse mussel, (2) the mean absolute percentage error (MAPE), (3) mean absolute error (MAE) and (4) mean square error (MSE). These were defined as,

$$r = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y} - \hat{\bar{y}})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}\sqrt{\sum_{i=1}^{n}(y_i - \hat{\bar{y}})^2}}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}|}{y_i}(100\%)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}|$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y})^2$$

where y_i is the observed value, $\hat{y}$ is the predicted value, $\bar{y}$ is the average of observed values, $\hat{\bar{y}}$ is the average of predicted values, and n is the total number of observations. Using these four criteria, it is possible to rank the performance of each predictive model. The optimal model should have larger r and smaller MAPE, MAE, and MSE.

3 | RESULTS AND DISCUSSION

A simple linear regression was calculated to predict the total weight of the horse mussel based on the shell length. A significant regression equation was found ($F_{1,318}$ = 6273.57, P < 0.001) that explains 97.6% of the variance. The final LWR model is shown in Figure 3.

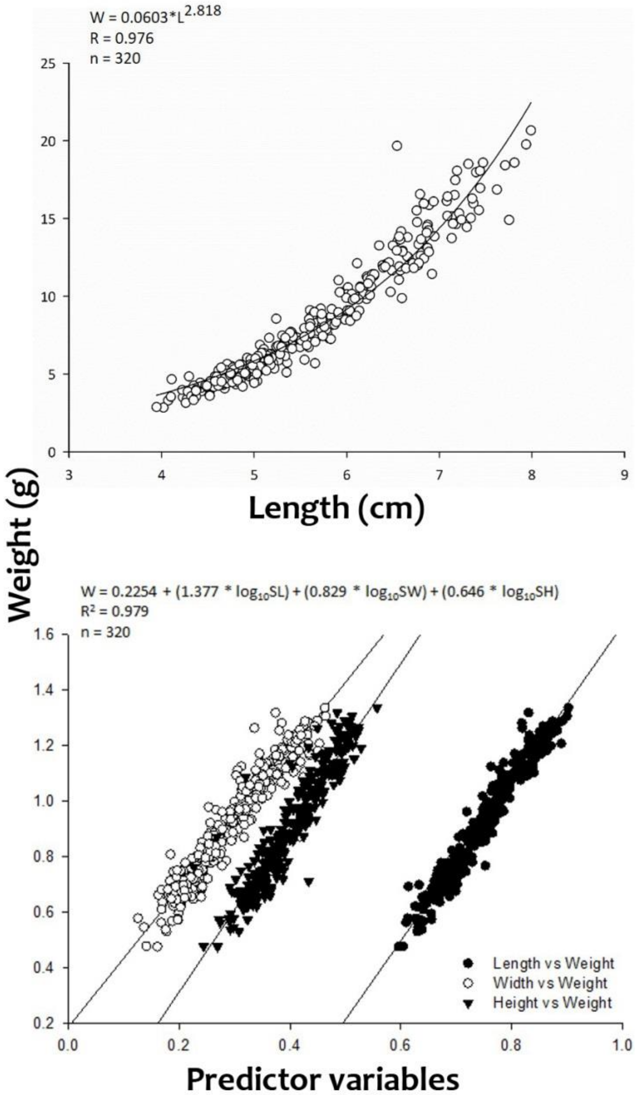

FIGURE 3 The length-weight relationship model (above) and multiple regression model (below) between the predictor variables and weight of the horse mussel (*Modiolus modulaides*).

Moreover, a MLR was carried out to investigate whether shell length (SL), shell width (SW) and shell height (SH) could significantly predict total weight of mussel (W). The results of the regression indicated that the model explained 97.9% of the variance and that the model was a significant predictor of mussel total weight ($F_{3,316}$ = 3094.28, $P < 0.001$; Figure 3). Moreover, all predictor variables contributed significantly to the model ($P < 0.001$). The final predictive model was:

$$W=0.2254+(1.377\times\log_{10}\text{SL})+(0.829\times\log_{10}\text{SW})+(0.646\times\log_{10}\text{SH})$$

On the other hand, the ANNs used shell length, shell width and shell height as input data to predict the total weight of the horse mussel. Among the studied individuals, 224 were used in the training process and 96 were used in the testing process of the ANNs. As shown in Figure 4, the prediction of the ANNs was consistent with the actual weight values of the horse mussel.

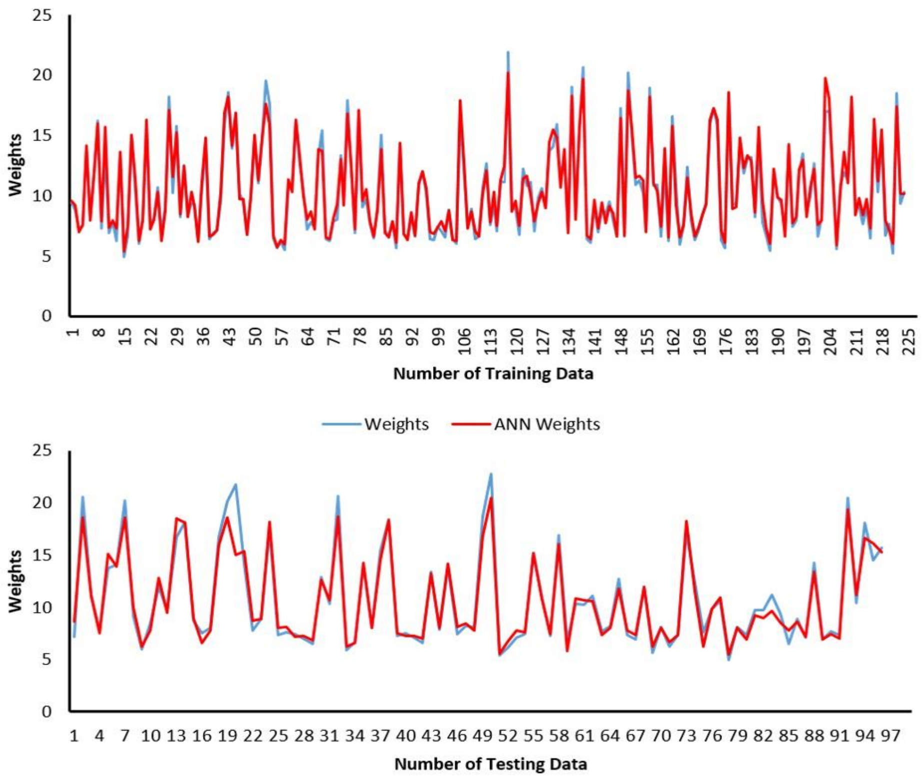

FIGURE 4 The actual and predicted weights of the horse mussel (*Modiolus modulaides*) on the training (n = 224) and testing (n = 94) sets.

Table 1 shows the correlation coefficient, mean absolute percentage error, mean absolute error, and mean square error values of the relationships between observed and estimated values obtained for the length-weight relationship, multiple linear regression and ANNs. On the basis of these criteria, the ANNs model demonstrated a higher predictability and best performance compared to the other regression models used in this study. The MAPE estimates errors as percentage, therefore it can be easily used for comparison of models developed in studies with different units. In the present study, the ANNs predictive model can be classified as highly accurate while the linear regression models are classified as acceptable (Witt and Witt 1991).

Comparative studies on different regression models for length-weight relationship in marine organisms have reported that ANNs give better results with higher correlation coefficient and lower error values (Benzer and Benzer 2016, 2017, 2020; Türeli Bilen *et al.* 2011; Benzer *et al.* 2015). Several authors have also reported greater performances of ANNs compared to other predictive models in fisheries, marine science and ecological researches (Suryanarayana *et al.* 2008; Cabreira *et al.* 2009; Wang 2010).

TABLE 1 Performance of the three predictive models in this study based on the correlation coefficient (r), mean absolute percentage error (MAPE), mean absolute error (MAE), and mean square error (MSE) between the observed weights and predicted weights of the horse mussel.

Predictive models	r	MAPE	MAE	MSE
Length-weight relationship	0.976	37.08	0.660	0.934
Multiple linear regression	0.979	37.66	0.530	0.627
Neural network regression	0.984	5.786	0.0235	0.001

4 | CONCLUSIONS

The length-weight relationship model has been traditionally used to estimate length and weight. The results from this study pointed out that the use of this model is of poor predictive value when used for bivalves. In addition, it was demonstrated that the ANNs model is a good alternative to existing models due to its higher accuracy in predicting the weights of horse mussel. Furthermore, based on the results of the present study, ANNs as predictive models may be a better option to use in predictive modelling in other species of aquatic organisms.

ACKNOWLEDGEMENT

The author extends his gratitude to the Commission on Higher Education and the University of the Philippines Visayas for the scholarship grant and thesis grant respectively. Moreover, the support of the UPV-DOST Invasive Mussel Project and the local government of Dumangas, Iloilo is highly appreciated. This paper was written while at home quarantine during the COVID-19 pandemic.

CONFLICT OF INTEREST

The author declares no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author.

REFERENCES

Anderson RO, Gutreuter SJ (1983) Length, weight, and associated structural indices, In: Nielsen LA, Johnson DL (Eds) Fisheries techniques. American Fisheries Society, Bethesda, MD., USA. pp. 283–300.

Benzer S, Benzer R (2016) Evaluation of growth in pike (*Esox lucius* L., 1758) using traditional methods and artificial neural networks. Applied Ecology and Environmental Research 14: 543–554.

Benzer S, Benzer R (2017) Comparative growth models of big-scale sand smelt (*Atherina boyeri* Risso, 1810) sampled from Hirfanlı Dam Lake, Kırsehir, Ankara, Turkey. Computational Ecology and Software 7(2): 82–90.

Benzer S, Benzer R (2019) Alternative growth models in fisheries: artificial neural networks. Journal of Fisheries 7(3): 719–725.

Benzer S, Benzer R (2020) Growth properties of *Pseudorasbora parva* in Süreyyabey Reservoir: traditional and artificial intelligent methods. Thalassas: An International Journal of Marine Sciences 36: 149–156.

Benzer S, Benzer R, Gunal AC (2017) Artificial neural networks approach in morphometric analysis of crayfish (*Astacus leptodactylus*) in Hirfanli Dam Lake. Biologia 72(5): 527–535.

Benzer S, Karasu Benli Ç, Benzer R (2015) The comparison of growth with length-weight relation and artificial neural networks of crayfish, *Astacus leptodactylus*, in Mogan Lake. Journal of Black Sea 21(2): 208–223.

Boulding EG, Hay TK (1993) Quantitative genetics of shell form of an intertidal snail: constraints on short-term response to selection. Evolution 47(2): 576–592.

Cabreira AG, Tripode M, Madirolas A (2009) Artificial neural networks for fish-species identification. ICES Journal of Marine Sciences 66(6): 1119–1129.

Caceres-Martinez J, Rio-Portilla MA, Gutierrez SCR, Humaran IMG (2003) Phenotypes of the California mussel, *Mytilus californianus*, Conrad (1837). Journal of Shellfish Research 22: 135–140.

Froese R (2006) Cube law, condition factor and weight–length relationships: history, meta-analysis and recommendations. Journal of Applied Ichthyology 22(4): 241–253.

Gosling E (2015) Marine bivalve mollusks (2nd Ed). John Wiley & Sons Ltd., West Sussex, UK. 524 pp.

Keys AB (1928) The weight-length relationship in fishes. Proceedings of the National Academy of Science 14(12): 922–925.

Krenker A, BešTer J, Kos A (2011) Introduction to the Artificial Neural Networks. In: Suzuki K (Ed) Artificial Neural Networks - methodological advances and biomedical applications. InTech Open, London, UK.

Morton B (1977) The biology and functional morphology of *Modiolus metcalfei* (Bivalvia, Mytilacea) from the Singapore mangrove. Malacologia 16: 501–517.

Napata RP, Andalecio MN (2011) Exploitation and management of brown mussel (*Modiolus philippinarum*) resources in Iloilo, Philippines. Philippine Journal of Social Sciences Human 16(2): 22–34.

Ozawa H (2001) Reproductive cycle and spawning of *Modiolus philippinarum* (Bivalvia: Mytilidae) in a seagrass bed in Kin Bay, Okinawa Island, Southern Japan. Venus 60(3): 173–181.

Özcan EI, Serdar O (2018) Artificial neural networks as new alternative method to estimating some population paramters of Tigris loach (*Oxynoemacheilus tigris* (Heckel, 1843)) in the Karasu River, Turkey. Fresenius Environmental Bulletin 27(12B): 9840–9850.

Palmer AR (1990) Effect of crab effluent and scent of damaged conspecifics on feeding, growth, and shell morphology of the Atlantic dogwhelk *Nucella lapillus* (L.). Hydrobiologia 193: 155–182.

Pauly D (1983) Some simple methods for the assessment of tropical fish stock. FAO Fisheries Technical Paper 234: 1–52.

Poutiers JM (1998) Bivalves. Acephala, Lamellibranchia, Pelecypoda. pp. 123–362. In: Carpenter KE, Niem VH. FAO species identification guide for fishery purposes. The living marine resources of the Western Central Pacific. Volume 1. Seaweeds, corals, bivalves, and gastropods. Rome, FAO.

Rainer JS, Mann R (1992) A comparison of methods for calculating condition index in eastern oysters, *Crassostrea virginica* (Gmelin, 1791). Journal of Shellfish Research 11(1): 55–58.

Rumelhart DE, Hinton GE, Williams RJ (1986) Learning internal representations by error propagation in parallel distributed processing. Explorations in the microstructure of cognition, MIT Press 1: 318–362.

Samoilys MA (1997) Manual for assessing fish stocks on Pacific coral reefs. Training series QE97009. Queensland Department of Primary Industries, Townsville. 75 pp.

Suryanarayana I, Braibanti A, Rao RS, Ramam VA, Sudarsan D, Rao GN (2008) Neural networks in fisheries research. Fisheries Research 92: 115–139.

Tumanda Jr MI, Yap HT, McManus LT, Ingles JA, López MG (1997) Growth, mortality and recruitment pattern of the brown mussel, *Modiolus metcalfei* (Bivalvia: Mytilacea), in Panguil Bay, Southern Philippines. Aquaculture 154: 233–145.

Tureli Bilen C, Kokcu P, Ibrikci T (2011) Application of artificial neural networks (ANNs) for weight predictions of blue crabs (*Callinectes sapidus* Rathbun, 1896) using predictor variables. Mediterranean Marine Science 12(2): 439–446.

Uba KIN, HM Monteclaro, Noblezada MM, Quinitio GF (2019) Sexual dimorphism, asymmetry, and allometry in the shell shape of *Modiolus metcalfei* (Hanley, 1843) collected from Dumangas, Iloilo, Philippines: a geometric morphometric approach. Computational Ecology and Software 9(3): 107–120.

Wang W (2010) Computational ecology: artificial neural networks and their applications. World Scientific and Publishing Co. Pvt. Ltd. 296 pp

Witt SF, Witt CA (1991) Modelling and forecasting demand in tourism. Academic Press, London. 195 pp.

Witten IH, Frank E, Hall MA, Pal CJ (2017) Data mining: practical machine learning tools and techniques (4th Edition). Elsevier, Cambridge, USA.

KIN Uba https://orcid.org/0000-0001-7197-1814