



# Sustainable management of water demand using fuzzy inference system: a case study of Kenyir Lake, Malaysia

Nor Najwa Irina Mohd Azlan<sup>1</sup> · Marlinda Abdul Malek<sup>2</sup> · Maslina Zolkepli<sup>3</sup> · Jamilah Mohd Salim<sup>4,5</sup> · Ali Najah Ahmed<sup>6</sup>

Received: 14 September 2020 / Accepted: 30 November 2020 / Published online: 6 January 2021  
© The Author(s), under exclusive licence to Springer-Verlag GmbH, DE part of Springer Nature 2021

## Abstract

Sustainable water demand management has become a necessity to the world since the immensely growing population and development have caused water deficit and groundwater depletion. This study aims to overcome water deficit by analyzing water demand at Kenyir Lake, Terengganu, using a fuzzy inference system (FIS). The analysis is widened by comparing FIS with the multiple linear regression (MLR) method. FIS applied as an analysis tool provides good generalization capability for optimum solutions and utilizes human behavior influenced by expert knowledge in water resources management for fuzzy rules specified in the system, whereas MLR can simultaneously adjust and compare several variables as per the needs of the study. The water demand dataset of Kenyir Lake was analyzed using FIS and MLR, resulting in total forecasted water consumptions at Kenyir Lake of 2314.38 m<sup>3</sup> and 1358.22 m<sup>3</sup>, respectively. It is confirmed that both techniques converge close to the actual water consumption of 1249.98 m<sup>3</sup>. MLR showed the accuracy of the water demand values with smaller forecasted errors to be higher than FIS did. To attain sustainable water demand management, the techniques used can be examined extensively by researchers, educators, and learners by adding more variables, which will provide more anticipated outcomes.

**Keywords** Water demand · Fuzzy inference system · Multiple linear regression · Lake · Sustainable water demand management

## Introduction

Approximately 54% of the world population lived in cities in 2014, which is expected to rise to two-thirds by 2050, generating 55% of additional water demand worldwide (Marchal et al. 2011). Future water conditions will become harder to

manage unless current water challenges, namely, demand management, water security, conservation, sustainable consumption, and water efficiency, can be addressed globally (Arfanuzzaman and Rahman 2017). In addition, in the coming years, there will be limited ways to increase water supply and a detailed approach to water use management will be required

---

Responsible Editor: Marcus Schulz

---

✉ Nor Najwa Irina Mohd Azlan  
SC22714@utn.edu.my

Marlinda Abdul Malek  
Marlinda@uniten.edu.my

Maslina Zolkepli  
masz@upm.edu.my

Jamilah Mohd Salim  
jamilah@umt.edu.my

Ali Najah Ahmed  
Mahfoodh@uniten.edu.my

<sup>2</sup> Institute of Sustainable Energy, Universiti Tenaga Nasional, 43000 Kajang, Selangor, Malaysia

<sup>3</sup> Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

<sup>4</sup> Institute of Tropical Biodiversity & Sustainable Development (Bio-D Tropika), Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu, Malaysia

<sup>5</sup> School of Science and Marine Environment (FSME), Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu, Malaysia

<sup>1</sup> Department of Civil Engineering, College of Engineering, Universiti Tenaga Nasional, 43000 Kajang, Selangor, Malaysia

<sup>6</sup> Institute of Energy Infrastructure (IEI), Universiti Tenaga Nasional, 43000 Kajang, Selangor, Malaysia

(Haddad and Lindner 2001). Sustainable water demand management guarantees efficient water usage to maintain economic growth, household consumption, food production, industry, and energy of a country (WWAP 2015). This study takes up the challenges by analyzing water demand at Kenyir Lake meticulously using two techniques: fuzzy inference system (FIS) and multiple linear regression (MLR).

The analysis of water demand is becoming a challenging step because of the changes caused by climatic conditions, land use patterns, human activities, and changes in technology. To know thoroughly the water demand of an area, its data collection and techniques are reviewed, defined, identified, and executed carefully. A number of techniques have been proposed, but they revolved in academics only. This is because of their complexity and nature, the quality of the available data, and the number of variables (Alamanos et al. 2019). FIS is a method that conveys a significant relationship between variables, and MLR gives the direction of each variable. Both methods have their ways and contributions in analyzing current water demand to be compared in this study. There are other techniques used by other researchers, for example, the logistic model (Ren et al. 2019), system dynamics model (Huang and Yin 2017), and water evaluation and planning (Ospina-Noreña et al. 2009), in which their analyses revolve on the prediction of future water demand.

In the problem of demand and supply of water at Kenyir Lake, FIS is performed by defining the fuzzy rules of system input and output variables of the study. Three steps are involved in the fuzzy process: regulate the rule conclusion into a suitable degree of sample and rule premise; all rule conclusions are clustered with results spread from potential output values; and lastly, the defuzzification process involves choosing a value among a set of potential output values (Dubois and Prade 1996). Through the fuzzy process, water demand values are defined, and the difference between the values is portrayed. MLR, conversely, classifies the data according to dependent and independent variables. Instead, of using simple linear regression, MLR is chosen because the study adopted several independent variables to discuss the variations in the water demand (Uriel 2013). In this study, the strategies involved in MLR analysis are a selection of input data, regression analysis, and assessment of model performance. MLR also produced numerical values of water demand that can be compared with those generated via FIS analysis and actual water consumption.

The proposed FIS method creates a flowchart (Fig. 2) on how the input variables go through a fuzzy process (fuzzification, rule evaluation, and defuzzification) to provide results associated with the membership functions of each variable. The model of the fuzzy system essentially includes defining the assumptions and implications of the system, separating variables into subsequent sets, and determining linguistic terms for identified output for each set (Moorthi et al. 2018). FIS analysis can be done easily and is user-friendly

especially to policy makers, as it provides both linguistic and numerical results. As for the comparison method, that is, MLR, the results give the direction and size of the effect of each variable relative to a dependent variable (Neuman 2011). The effect is precisely measured in value. For instance, when the value gets higher, the effect is larger on a variable that predicts the dependent variable.

Input variables of the number of people, income generated, and water consumption at Kenyir Lake are applied in FIS to produce an output of the level of water demand numerically relative to the 27 rules confirmed by experts. For MLR analysis, the independent variables are the number of people and income generated, and the dependent variable is water consumption at Kenyir Lake. The datasets of the population of people and income generation for all premises at Kenyir Lake are collected thoroughly from on-site investigation and interviews. Water consumption at Kenyir Lake for each premise is estimated and used in these analyses. Both result analyses are evaluated using the statistical parameters of mean absolute deviation (MAD), mean square error (MSE), root mean square error (RMSE), and mean absolute percentage error (MAPE) with the actual water consumption at Kenyir Lake to determine their accuracy and performance.

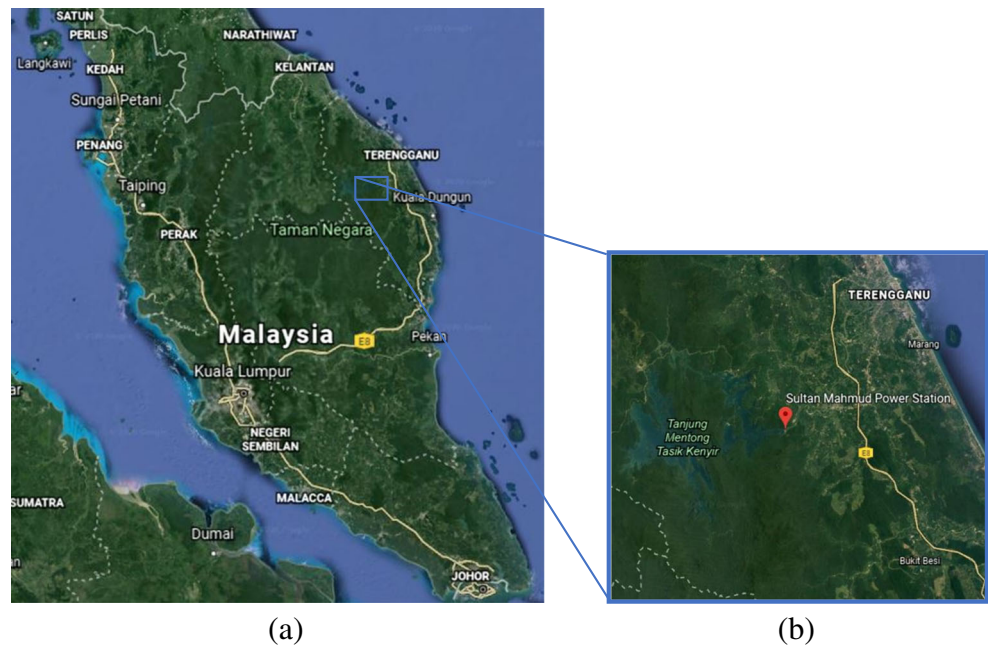
## Materials and methods

### Case study

Kenyir Lake is the largest man-made lake in Southeast Asia surrounded by the greenery of diverse flora and fauna. This man-made lake is located in the district of Hulu Terengganu of the state of Terengganu on the east coast of Malaysia (Fig. 1). Its geographical coordinates are the latitudes of between 4° 40' N and 5° 15' N and the longitudes of between 102° 32' E and 102° 55' E (Rouf et al. 2010). The lake covers an area of 260,000 hectares (Yusof et al. 2009) with a water catchment area of 38,000 hectares (Lola et al. 2017). Its mean depth is 37 m, with a top water level of 145 m above mean sea level (Rouf et al. 2010).

Kenyir Lake was created in the year 1985 by the damming of two main rivers, Terengganu and Terenggan Rivers, to generate hydroelectricity in Malaysia (Suratman et al. 2017). Underwater current from the inflow of river water and the generation of hydroelectric power stations causes a combination of top and bottom water columns (Rosle et al. 2018). Water in the lake replenishes from nine major rivers diverting to the lake, namely, Lasir, Belimbing, Tembat, Ketiar, Leper, Pertang, Kerbat, Terenggan, and Kenyir Rivers. The water level of the lake varies as it works similar to a reservoir. As reported by Rauf et al. (2015), water level is high during monsoon and low during dry season.

**Fig. 1** The location of the study area: **a** the map of Malaysia, **b** Kenyir Lake



The beauty of Kenyir Lake has caught the attention of tourists and has become an ecotourism destination. Several ecotourism activities, namely, fishing, camping, jungle trekking, water sports, staying in houseboats, and exploring the attraction parks, can be done there. It is also known as the haven of sport-fishing and intensive cage aquaculture activities and has a high market value of fish stocks (Sharip and Zakaria 2008). These advantages of Kenyir Lake have contributed to an increase in water demand since there has been an increase in the number of tourists.

**Data description**

To analyze the water demand at Kenyir Lake, important data, including water consumption values, the number of people, and the income generated from all premises at the lake, should be collected. Information on the respective data were gathered during the year 2018 from the on-site investigation and interviews of respective authorities and local communities. Tables 1, 2, and 3 respectively show the data of average daily water consumption, average daily number of people, and average income generated from the premises at Kenyir Lake.

Table 1 establishes the average actual water consumption of every premise at Kenyir Lake. Of all the premises, the highest and lowest water consumptions recorded are 662 m<sup>3</sup> for houseboats and 1.9 m<sup>3</sup> for camping sites, respectively. Houseboats are one of the famous attractions at Kenyir Lake for which locals and tourists are willing to pay for a superior package offered by houseboat operators (Muhamad et al. 2014). The lowest water consumption is from the camping sites because people visited the premise mostly during public and school holidays only.

Table 2 displays the average daily number of people who visited the premises in Kenyir Lake. Houseboats received the

**Table 1** Average daily water consumption at Kenyir Lake

Building/premise	Average daily water consumption (m <sup>3</sup> )
Lake Land Resort	20.3
Musang Kenyir Resort	17.9
Surya Rakit Resort	15.5
Kenyir Eco Resort	22.7
Petang Island Resort	24.5
Tanjung Mentong Resort	22.7
Rumah Rehat Persekutuan	42.8
Lawit Lodge	9.5
Houseboats	662.0
Herb Park	9.3
Orchid Park	3.4
Butterfly Park	7.8
Tropical Park	2.0
Kenyir Water Park	34.5
Kenyir Elephant Conservation Village (KECV)	12.0
Kelah Sanctuary Park	10.6
National Park	4.1
Camping sites	1.9
KETENGAH and Tourist Information Center	5.4
Public toilets at Pengkalan Gawi	291.0
Prayer rooms at Pengkalan Gawi	10.0
Food stalls at Pengkalan Gawi	7.7
Aquaculture at Sungai Como	12.2

**Table 2** Average daily number of people at Kenyir Lake

Building/premise	Average daily number of people
Lake Land Resort	60
Musang Kenyir Resort	52
Surya Rakit Resort	44
Kenyir Eco Resort	68
Petang Island Resort	74
Tanjung Mentong Resort	68
Rumah Rehat Persekutuan	135
Lawit Lodge	24
Houseboats	2700
Herb Park	147
Orchid Park	26
Butterfly Park	120
Tropical Park	14
Kenyir Water Park	167
Kenyir Elephant Conservation Village (KECV)	139
Kelah Sanctuary Park	61
National Park	33
Camping sites	8
KETENGAH and Tourist Information Center	69
Public toilets at Pengkalan Gawi	738
Prayer rooms at Pengkalan Gawi	369
Food stalls at Pengkalan Gawi	11
Aquaculture at Sungai Como	40

**Table 3** Average daily income generated at Kenyir Lake

Building/premise	Average daily income generated (MYR)
Lake Land Resort	6300.00
Musang Kenyir Resort	6000.00
Surya Rakit Resort	3800.00
Kenyir Eco Resort	900.00
Petang Island Resort	8750.00
Tanjung Mentong Resort	6400.00
Rumah Rehat Persekutuan	5760.00
Lawit Lodge	800.00
Houseboats	24,000.00
Herb Park	0.00
Orchid Park	61.10
Butterfly Park	571.00
Tropical Park	0.00
Kenyir Water Park	750.00
Kenyir Elephant Conservation Village (KECV)	3430.00
Kelah Sanctuary Park	478.00
National Park	264.40
Camping sites	13.90
KETENGAH and Tourist Information Center	0.00
Public toilets at Pengkalan Gawi	0.00
Prayer rooms at Pengkalan Gawi	0.00
Food stalls at Pengkalan Gawi	2750.00
Aquaculture at Sungai Como	5000.00

1 United States dollar (1 USD) = 4.16 Malaysian Ringgit (4.16 MYR)

highest number of people daily, whereas camping sites received the lowest. The data collected are the data on actual water consumption. Hence, the average number of people responds directly proportional to the average actual water consumption at Kenyir Lake. As years go by, the population will continue to grow and expand, which will be coupled with increasing per capita consumption, leading to a continuous increase in water consumption (Yuanzheng et al. 2012).

The premises at Kenyir Lake generate income mainly from tourism activities. The higher the number of visitors, the more income the premises generate. Table 3 shows that some of the premises, including Herb Park, Tropical Park, KETENGAH and Tourist Information Center, public toilets at Pengkalan Gawi, food stalls at Pengkalan Gawi, and prayer rooms at Pengkalan Gawi, do not generate any income. Both Herb Park and Tropical Park offer free-of-charge entrance as tourists already pay for boating service to go to these premises. Conversely, other premises are public amenities that can be used and enjoyed for free. The level of income generated has a high correlation with the number of people and water consumption. Makki et al. (2015) found that higher water consumption comes most likely from premises with higher income levels and a higher number of visitors.

## Model description

FIS and MLR are the two techniques used for water demand analyses. These methods involved variables or factors that influence water demand, that is, water consumption, the number of visitors, and income generated; however, the application of the variables is different between the techniques.

## FIS

The fundamental fuzzy logic was initiated by Zadeh (1965). Fuzzy logic is a form of logic that processes and gives reasoning in the rough estimation of linguistic terms (Wu 2015). Some studies of a fuzzy logic approach focusing on the water resource domain include those of Shrestha et al. (1996), who implemented a fuzzy rule-based on reservoir operation, and Faye et al. (2000), who used a fuzzy modeling approach for the long-term management of water resources systems. Fuzzy logic comprises fuzzy sets, fuzzy rules, fuzzy reasoning, and FISs (Jang et al. 1997).

The basis of fuzzy logic is to consider the system states in subsets or fuzzy set form with an example of each labeled with “low,” “medium,” and “high.” The fuzzy set is constituted of





converted into literature means based on their level of water demand, that is, low, medium, or high. The defuzzification strategy used in this study is the centroid method, which is the most common and preferable among all defuzzification methods (Takagi and Sugeno 1985). The algebraic expression used (Eq. 1) to obtain the volume of water demand is shown below:

$$y^* = \frac{\int f(x, y) \cdot y \, dy}{\int f(x, y) \, dy}, \quad (1)$$

where  $y^*$  is the defuzzified output of volume water demand produced in FIS analyses.

### Multiple linear regression

Linear regression is a linear modeling that portrays the connection between a dependent variable and an independent variable. For cases with more than one independent variable, the linear regression is known as MLR. MLR analysis is the most efficient tool in utilizing the relationship between the explanatory variable and the outcome (Sharma et al. 2020). The development of MLR model includes the selection of input data, regression analysis, and evaluation criteria on model performance. The relationship in the MLR model is between two or more independent variables and a dependent variable (Abba et al. 2017). MLR equation with both dependent and independent variables is defined in Eq. 2 (Chen and Liu 2015). The usefulness of MLR is meant for a small number of variables, not significantly collinear, and has a strong relationship of variables in the system (Hanrahan et al. 2004).

$$Y = B_0 + B_1x_1 + B_2x_2 \quad (2)$$

where  $Y$  is the dependent variable (predicted water demand),  $B_0$  is the  $y$ -intercept or constant (value of dependent variable “ $y$ ” when  $x_1$  and  $x_2 = 0$ ),  $B_1$  and  $B_2$  are multiple regression coefficients, and  $x_1$  and  $x_2$  are independent variables (water consumption, number of people, and income generated).

The effectiveness of the developed MLR model was evaluated using the coefficient of determination ( $R^2$ ), adjusted  $R^2$ , and  $p$  value test. A fraction shows the proportion of the total variation of the selected dependent variable and the explanatory variable (Sahoo and Jha 2013). Adjusted  $R^2$  is defined similarly as  $R^2$  except for the fact that the number of degrees of freedom is considered. The probability of error involved in approving the validity of observed results is defined as the  $p$  value. A  $p$  value of 0.05, which is 5% of error and 95% confidence interval, is commonly known as a “border-line acceptable” error level (Haan 2002). The computations of  $R^2$  and adjusted  $R^2$  are presented in Eqs. 3 and 4, respectively.

$$R^2 = \left( \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{N \sigma_x \sigma_y} \right)^2, \quad (3)$$

where  $R^2$  is the coefficient of determination,  $x_i$  and  $y_i$  are  $i$ th variables,  $\bar{x}$  and  $\bar{y}$  are mean;  $\sigma_x$  and  $\sigma_y$  are standard deviation of the  $x$  and  $y$  values of variables, respectively; and  $N$  is the total sample size.

$$\text{Adjusted } R^2 = 1 - \left( \frac{(1 - R^2)(N - 1)}{N - p - 1} \right), \quad (4)$$

where  $R^2$  is the coefficient of determination and  $p$  is the number of variables.

In the MLR model, it is clear that as the number of variables increases, the accuracy of the model increases (Swain and Patel 2017). A good model with several variables has a large percentage of  $R^2$  (> 70%) (Neuman 2011). However, high values of  $R^2$  due to many variables can be misleading. To overcome this problem, adjusted  $R^2$  is computed because the difference between the value of  $R^2$  and adjusted  $R^2$  determines the accuracy of the model too. Yu et al. (2015) mentioned that a good model has a value of adjusted  $R^2$  lesser than that of  $R^2$  and the difference between the values of  $R^2$  and adjusted  $R^2$  is very small. In addition, the  $p$  value test was also examined for every regression factor as part of multiple regression analysis. The  $p$  value test is used to explore the chance of mistakenly rejecting a null hypothesis, in which the value can be between 0 and 1 and commonly expressed as a percentage (Lawens and Mutsvangwa 2018). A  $p$  value of < 5% indicates that the resulting factor is significant.

### Statistical evaluation of analysis methods

Statistical evaluation plays an essential role in choosing the right analysis methods. From the evaluation, the forecasting error, which is the difference between the actual value and forecasted values at a given time, was defined (Hanke and Wichern 2014). The common indicators applied in this study to evaluate the accuracy of the analysis are MAD, MSE, RMSE, and MAPE. Regardless of the measure being applied, the most accurate analysis method resulted in the lowest value of the statistical evaluation (Ryu and Sanchez 2003).

MAD measured the overall forecast error of the analysis. The value is computed by dividing the sum of the absolute value of single forecast errors by the sample size (Heizer and Render 2011). Next, MSE is a typically accepted method for the evaluation of exponential smoothing and other techniques. It measures error as the sum of squares of differences between actual and forecast values divided by the number of samples. RMSE is the square root of MSE. Ryu and Sanchez (2003) stated that RMSE is the computed error in terms of units that

are equal to the original values. Another statistical evaluation chosen in this study is MAPE, which is usually used in quantitative techniques of analysis (Makridakis et al. 1997). MAPE is the average of the sum of all the percentage errors for the sample dataset, disregarding the signs. Equations 5, 6, 7, and 8 present the MAD, MSE, RMSE, and MAPE equations, respectively:

$$MAD = \frac{1}{n} \sum_{i=0}^n |Y_i - F_i|, \tag{5}$$

$$MSE = \frac{1}{n} \sum_{i=0}^n (Y_i - F_i)^2, \tag{6}$$

$$RMSE = \frac{1}{n} \sum_{i=0}^n \sqrt{(Y_i - F_i)}, \tag{7}$$

$$MAPE = \frac{1}{n} \sum_{i=0}^n \frac{|Y_i - F_i|}{Y_i}, \tag{8}$$

where  $Y_i$  is the actual value of water consumption,  $F_i$  is the forecast value of water consumption using FIS and MLR, and  $n$  is the number of premises.

## Results and discussion

### FIS

Water demand at Kenyir Lake was analyzed using FIS with the information on the input, output, and rules selected accordingly. Each process analyzed in FIS was executed using Matlab. Figures 4, 5, 6, and 7 depict the inputs and the outputs used in the water demand analysis.

The ranges of values for input water consumption shown in Fig. 4 are 0–8 m<sup>3</sup> (low), 8–25 m<sup>3</sup> (medium), and 25–662 m<sup>3</sup>

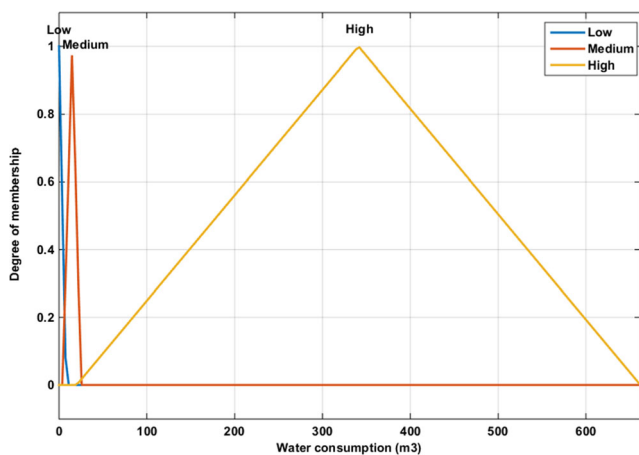


Fig. 4 Input water consumption

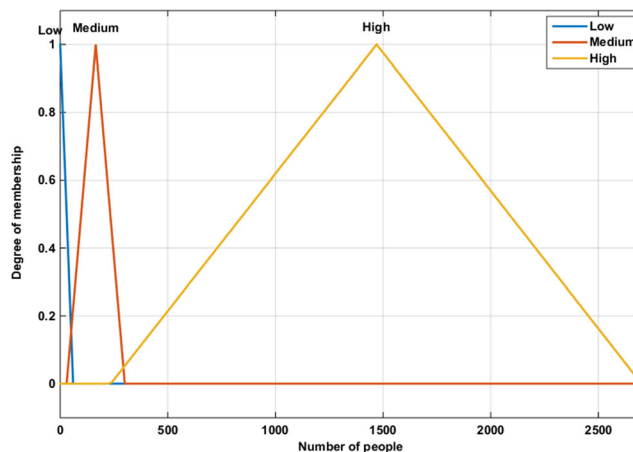


Fig. 5 Input number of people

(high). The water consumption values ranging from 0 to 662 m<sup>3</sup> are used as an input for the FIS process.

Figure 5 establishes the number of people ranging from 0 to 2700 as a second input in the FIS process. There are three levels for the number of people, namely, low, medium, and high, with ranges of 0–60, 60–300, and 300–2700 people, respectively.

The third input used is the income generated, which ranged from RM 0 to RM 24,000 as depicted in Fig. 6. The levels of income generated used are low (RM 0–RM 300), medium (RM 300–RM 6000), and high (RM 6000–RM 24,000).

Figure 7 illustrates the ranges for the output levels of water demand for this FIS analysis. The values of the low, medium, and high levels of water demand are 0–8, 8–25, and 25–670 m<sup>3</sup>, respectively. Figure 8 shows the entire process of the FIS for this water demand analysis. Three left columns and the most-right column in the figure refer to the inputs utilized and the output inferred by FIS, respectively. The process undergoes linguistic variable fuzzification through defuzzification of the aggregate

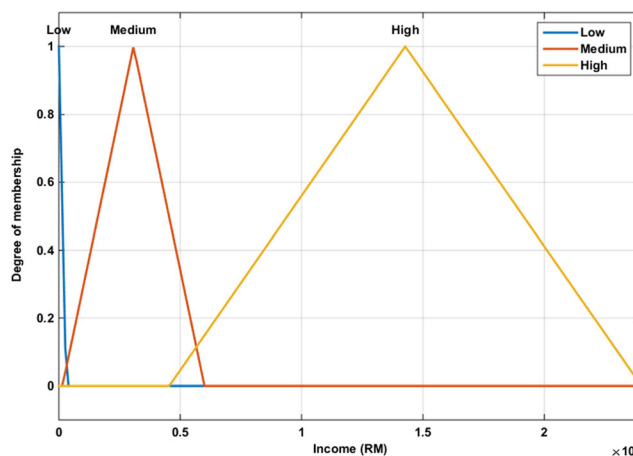


Fig. 6 Input income generated

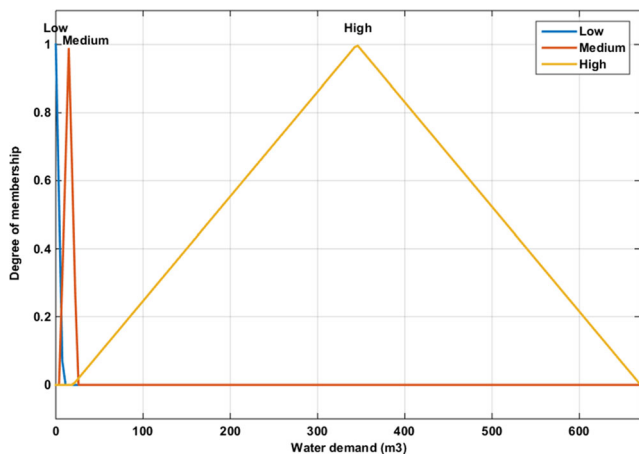


Fig. 7 Output levels of water demand

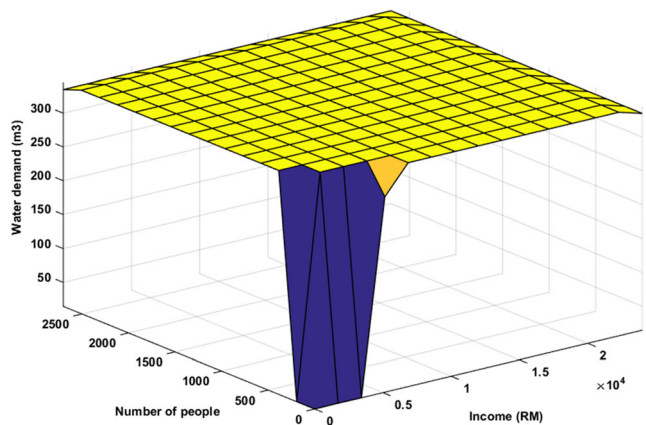


Fig. 9 Three-dimensional curved surface for the input number of people and income generated

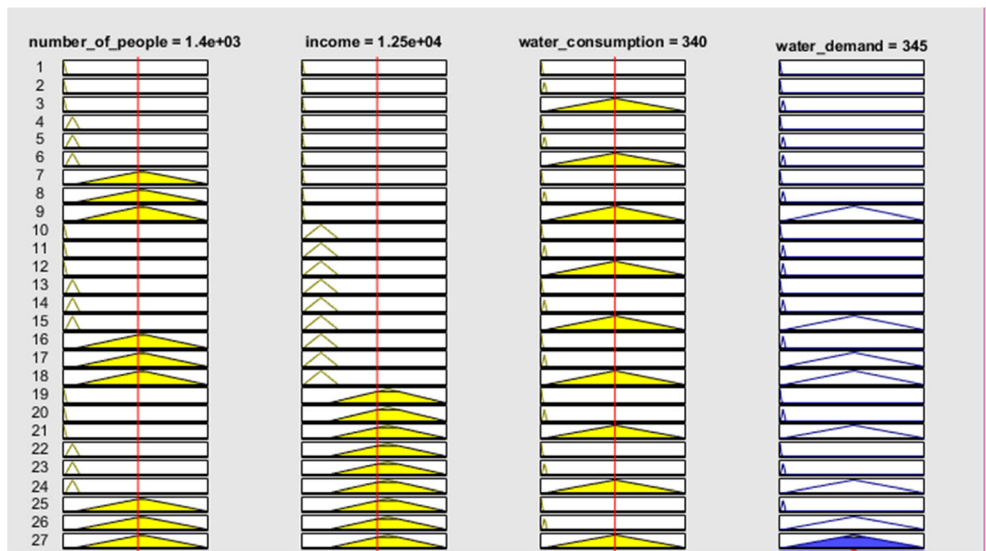
output to evaluate the numerical values of the expected water demand analyzed.

From Fig. 8, the yellow area in each set shows which degree of the input value is determined as a member of the associated fuzzy set. The blue area depicts which degree of a related fuzzy set is chosen based on the input data used, whereas the blue area at the most bottom represents the aggregated output of the fuzzy set. The red line at the middle of the input variables of the number of people, income generated, and water consumption shows their numerical values. Nevertheless, the red line on the output column from Fig. 8 depicts the numerical values of the water demand inferred by FIS through the entire process. One example of the FIS process related to Fig. 8 is as follows: the input number of 1400

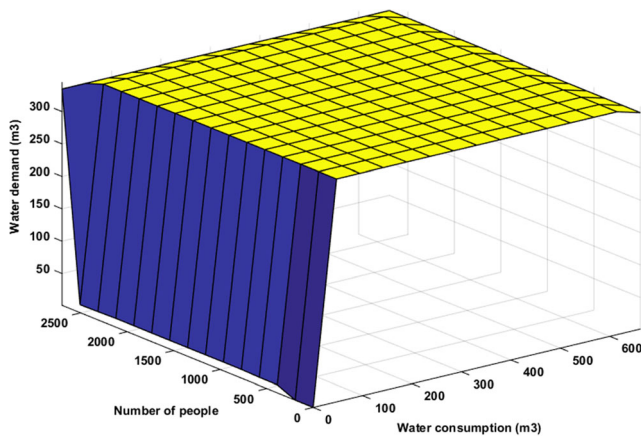
people, RM 12,500 income generated, and 340 m<sup>3</sup> water consumption that resulted in 345 m<sup>3</sup> of estimated water demand.

The three-dimensional curved surface shown in Figs. 9, 10, and 11 illustrates the output variables of water demand analysis at different input variables. Specific values of water demand correspond to each operating condition. The summary that can be drawn from the figures is water demand reacts to the number of people and water consumption rather than to income, since as the number of people and water consumption increase, water demand increases. Admission to some of the premises at Kenyir Lake is free of charge; hence, no income is generated from them. Thus, income is less influential than the number of people and water consumption. Table 4 presents

Fig. 8 FIS process of water demand analysis with 27 IF–THEN rules







**Fig. 10** Three-dimensional curved surface for the input number of people and water consumption

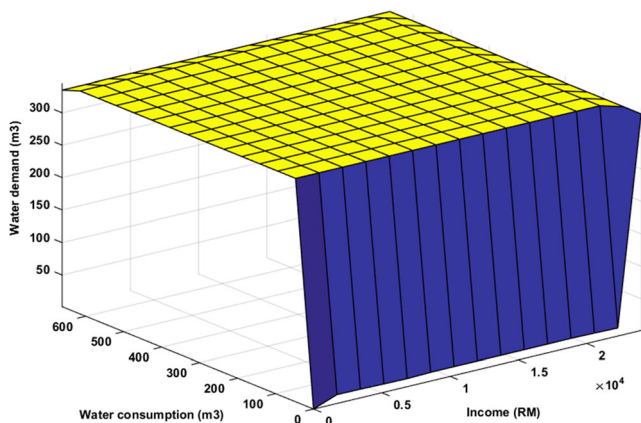
the full results of forecasted water consumption at Kenyir Lake using FIS analysis.

**MLR**

The MLR model used in this study comprises two independent variables, namely, the number of people and the income generated. The regression function constructed from MLR model (Eq. 2) is given as:

$$Y = -22.66854 + 0.15743 x_1 + 0.00916 x_2. \tag{9}$$

The constant  $B_0$  of  $-22.66854$  represents water demand when the number of tourists and income generated are 0. In this study, historical data showed an increase of tourists (the number of people) in the study area yearly, which corresponds perpendicularly to the income growth anticipated yearly.



**Fig. 11** Three-dimensional curved surface for the input water consumption and income generated

**Table 4** Actual and forecasted water consumption at Kenyir Lake

Building/premise	Actual water consumption (m <sup>3</sup> )	Forecasted water consumption (m <sup>3</sup> )	
		FIS	MLR
Lake Land	20.32	350.00	20.12
Musang Kenyir	17.91	350.00	17.74
Surya Rakit	15.51	17.50	11.90
Kenyir Eco	22.72	17.50	12.18
Petang Island	24.52	17.50	27.79
Tanjung Mentong	22.72	17.50	22.15
Rumah Rehat Persekutuan	42.84	362.33	36.43
Lawit Lodge	9.51	17.50	1.85
Houseboats	662.00	350.00	660.76
Herb	9.31	17.50	0.47
Orchid	3.42	1.27	0.98
Butterfly	7.83	3.50	23.57
Tropical	2.02	1.00	20.46
Kenyir Water Park	34.48	362.27	34.73
KECV	12.01	16.88	33.13
Kelah Sanctuary	10.60	17.50	9.80
National Park (jungle trekking, camp sites, caves, and fishing)	4.14	3.50	2.96
Other camp sites	1.91	0.94	3.26
KETENGAH and Kenyir Information Center Office	5.39	3.50	11.81
Public toilets	291.00	362.49	326.03
Prayer rooms	9.97	3.20	64.55
Food stalls	7.66	3.50	2.39
Aquaculture	12.19	17.50	13.16
Total	1249.98	2314.38	1358.22

From the MLR model, the results show that the number of people is more influential to water demand than to income generated. This can be proven by the fact that premises with a low number of people have negative water demand values. Those premises include offices of KETENGAH and Tourist Information Center, camping sites, National Park, Kelah Sanctuary Park, Tropical Park, Orchid Park, Lawit Lodge, and Kenyir Eco Resort. As explained by Neuman (2011), the positive and negative signs of the resulting values only show the direction and effect of the variables to the dependent variable. In this study, Table 4 establishes the forecasted water consumption for every premise at Kenyir Lake using MLR analysis.

The MLR results are also being validated with the coefficient of determination ( $R^2$ ), adjusted  $R^2$ , and  $p$  value tests. The computed values of  $R^2$  and adjusted  $R^2$  are found to be 0.9684 and 0.9653, respectively.  $R^2$  and adjusted  $R^2$  values are also expressed as 97% of the variation in water demand values estimated to be relative to the variation in the independent variables (the number of people and the income generated), which

shows an extremely satisfactory agreement. In addition, the resulting  $R^2$  value in this study has a good linear approximation of the regression function because only 3% error was found in the model. Differences between  $R^2$  and adjusted  $R^2$  are very small and sufficient to justify the predictability of the model.

The  $p$  values for variable  $x_1$  (the number of people) and  $x_2$  (the income generated) are found to be 0.01008 and 0.00132, respectively. Thus, the  $p$  values for both variables are significant ( $< 5\%$ ) in presenting the target variable, and the null hypothesis can be rejected. The validation from  $R^2$  and adjusted  $R^2$  of the model and the  $p$  value of the independent variables show that the final model and variables used can be considered notable.

## Summary

Water demand values for all premises at Kenyir Lake have been analyzed using FIS and MLR. Table 4 presents the results for both analyses and the actual water consumption at Kenyir Lake. The results of water demand analyses via FIS and MLR range from 0.47 to 660.76 m<sup>3</sup>. The values are compared with the actual water consumption at Kenyir Lake to determine the accuracy and performance of each analysis. FIS analysis relies on membership functions, range of membership functions, and fuzzy rules to estimate water demand. Conversely, the MLR technique accesses both independent and dependent variables to produce water demand values.

As shown in Table 4, there is a large difference between the water demand values from FIS analysis and the actual water consumption. However, this is not applicable to all premises being analyzed because of the free-of-charge entrance for some premises. This also led to a low connection between the membership function and the variable. Premises that have a large gap between the actual water consumption are public toilets at Pengkalan Gawi, Kenyir Water Park, houseboats, Rumah Rehat Persekutuan, Musang Kenyir Resort, and Lake Land Resort. One of the premises that have a slight gap between FIS result and actual water consumption is Tropical Park at 1.00 m<sup>3</sup> and 2.00 m<sup>3</sup>, respectively. Although this study has developed a FIS model to analyze water demand, the results obtained are responsive to the type of membership function utilized. Thus, FIS has dimensionality problems that come from the increased number of fuzzy sets (Moorthi et al. 2018).

The outcome of water demand analysis at Kenyir Lake using MLR is more reliable than using FIS, since most of the water demand values from MLR are almost similar to the actual water consumption, as shown in Table 4. For instance, the water demand value using MLR and the actual water consumption for houseboats are very close, that is, 660.76 m<sup>3</sup> and 662.00 m<sup>3</sup>, respectively. Total water demand values for all premises at Kenyir Lake for FIS and MLR analyses are 2314.38 m<sup>3</sup> and 1358.22 m<sup>3</sup>, respectively. Also, the total actual water consumption is 1249.98 m<sup>3</sup>, which is the least difference with

**Table 5** Statistical evaluations for FIS analysis and MLR analysis

Statistics	FIS analysis	MLR analysis
Mean absolute deviation (MAD)	76.83	8.96
Mean square error (MSE)	23103.41	244.97
Root mean square error (RMSE)	152.00	15.65
Mean absolute percentage error (MAPE)	262.83%	108.87%

water demand value using MLR. Lawens and Mutsvangwa (2018) agreed that water demand forecasts using statistical analysis (MLR) have great benefits, since the significant driver can be identified. In addition, the MLR analysis used is valuable for future research on water modeling (Sundari et al. 2013).

Forecasting errors for FIS and MLR analysis were evaluated using statistical indices. Table 5 presents the statistical results for both analyses. The values of MAD, MSE, RMSE, and MAPE for 23 premises using FIS analysis are 76.83, 23103.41, 152.00, and 262.83%, respectively. Conversely, the MAD, MSE, RMSE, and MAPE of MLR analysis are found to be 8.96, 244.96, 15.65, and 108.87%, respectively. Given the statistical parameter, MLR analysis predicts water consumption more accurately than FIS analysis because the forecast error values of MLR analysis are smaller than those of FIS analysis. MLR is the most appropriate analysis method in terms of accuracy (statistically) and simplicity. Similar to other studies (Karmaker et al. 2017; Mehri 2013; Ryu and Sanchez 2003), the best forecasting method is chosen based on the smallest forecasted errors of MAD, MSE, RMSE, and MAPE.

The methods of FIS and MLR are valid analysis tools and great visualization for the relationships between the variables. Researchers, decision makers, and the community can have a better understanding of results via FIS because this method provides verbal terms. Factors that influence water demand can also be analyzed using MLR. Water demand analysis can be extended using other techniques by gathering more variable datasets from the study area. A longer duration of research study provides an opportunity for the data to be widely explored, hence rendering meaningful effects on the study.

The application of the analysis techniques to determine the supply and demand at Kenyir Lake, Terengganu, contributes to sustainable water demand management. Water is closely related to most paths of sustainable development, but whether water is used or abused depends on water managers and communities (Cosgrove and Loucks 2015). Hence, the participation of communities in improving water demand management is very crucial as they can develop an environmentally friendly behavior and reconstruct their intention toward it. A study by Aliabadi et al. (2020) proved that rural people focus more on the advantages and benefits of sustainable water management because they know the importance of water resources, human health, and environmental impacts.

## Conclusions

The values of water demand analysis at Kenyir Lake, as revealed by the experimental results of FIS and MLR, are 2297 m<sup>3</sup> and 1471 m<sup>3</sup>, respectively. It can be shown that MLR analysis converges closer to the actual water consumption of 1249.8 m<sup>3</sup> than does FIS analysis. Water demand analysis by MLR is found to be valid and highly applicable as the method also can be proven by its  $R^2$  of 0.9684 and adjusted  $R^2$  of 0.9653.

The techniques proposed in this study, namely, FIS and MLR analyses, have their distinctive benefits in analyzing the water demand. FIS analysis interpreted the input variables and produced linguistic and numerical terms by considering the fuzzy rules confirmed by the experts. Experts' knowledge contributes to the great reliability of the results. Conversely, MLR results provide verifications from  $R^2$ , adjusted  $R^2$ , and the numerical value of the effect on each variable. From this, the resulting water demand can be verified easily. MLR is a very direct and quick technique that requires only a few variables (Luna et al. 2017). The statistical evaluations also showed that MLR analysis outperformed FIS analysis by smaller values of forecasted errors. The MAD, MSE, RMSE, and MAPE values of MLR analysis are 8.96, 244.96, 15.6, and 108.87%, respectively.

**Acknowledgments** The authors would like to acknowledge technical support from Universiti Putra Malaysia, Universiti Malaysia Terengganu, and access to data from Syarikat Air Terengganu (SATU) and Development Authority of Terengganu Tengah (KETENGAH).

**Authors' contributions** Conceptualization: Nor Najwa Irina Mohd Azlan, Marlinda Abdul Malek, Maslina Zolkepli, and Jamilah Mohd Salim. Methodology: Nor Najwa Irina Mohd Azlan, and Marlinda Abdul Malek. Software: Nor Najwa Irina Mohd Azlan, Maslina Zolkepli, Jamilah Mohd Salim, and Ali Najah Ahmed. Validation: Nor Najwa Irina Mohd Azlan, Maslina Zolkepli, Jamilah Mohd Salim, and Ali Najah Ahmed. Formal analysis: Nor Najwa Irina Mohd Azlan and Ali Najah Ahmed. Investigation: Nor Najwa Irina Mohd Azlan. Resources: Marlinda Abdul Malek. Data curation: Nor Najwa Irina Mohd Azlan, and Marlinda Abdul Malek. Writing—original draft preparation: Nor Najwa Irina Mohd Azlan and Ali Najah Ahmed. Writing—review and editing: Nor Najwa Irina Mohd Azlan, Marlinda Abdul Malek, and Ali Najah Ahmed. Visualization: Nor Najwa Irina Mohd Azlan and Ali Najah Ahmed. Supervision: Marlinda Abdul Malek, Maslina Zolkepli, Jamilah Mohd Salim, and Ali Najah Ahmed. Project administration: Marlinda Abdul Malek. All authors read and approved the final manuscript.

**Data availability** Authors do not have permission to share data.

## Compliance with ethical standards

**Ethics approval and consent to participate** Not applicable

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Abba SI, Jasim S, Abdullahi J (2017) River water modelling prediction using multi-linear regression, artificial neural network, and adaptive neuro-fuzzy inference system techniques. *Procedia Comput Sci* 120: 75–82
- Alamanos A, Sfyris S, Fafoutis C, Mylopoulos N (2019) Urban water demand assessment for sustainable water resources management, under climate change and socioeconomic changes. *Water Supply* 20:1–9. <https://doi.org/10.2166/ws.2019.199>
- Aliabadi V, Gholamrezai S, Ataei P (2020) Rural people's intention to adopt sustainable water management by rainwater harvesting practices: application of TPB & HBM models. *Water Supply* 20:1–15. <https://doi.org/10.2166/ws.2020.094>
- Arfanuzzaman M, Rahman AA (2017) Sustainable water demand management in the face of rapid urbanization and ground water depletion for social–ecological resilience building. *Glob Ecol Conserv* 10: 9–22. <https://doi.org/10.1016/j.gecco.2017.01.005>
- Baldovino RG, Dadios EP (2016) Fuzzy logic control: design of a “Mini” fuzzy associative matrix (FAM) table algorithm in motor speed control. *IEEE Region 10 Annual International Conference, Proceedings/TENCON*. <https://doi.org/10.1109/TENCON.2015.7372923>
- Chen W, Liu W (2015) Water quality modeling in reservoirs using multivariate linear regression and two neural network Models. *Advances in Artificial Neural Systems*
- Cosgrove WJ, Loucks DP (2015) Water management: current and future challenges and research directions. *Water Resour Res* 51:4823–4839. <https://doi.org/10.1002/2014WR016869>. Received
- Dubois D, Prade H (1996) What are fuzzy rules and how to use them. *Fuzzy Sets Syst* 84:169–185
- Faye R, Sawadogo S, Mora-camino F, Achaibou K (2000) A fuzzy modeling approach for the long term management of water resource systems. *FUZZ 2000, 9th IEEE International Conference on Fuzzy Systems*, 499–504. San Antonio, United States
- Haan C (2002) *Statistical methods in hydrology*, 2nd edn. Iowa State Press, Ames Iowa
- Haddad M, Lindner K (2001) Sustainable water demand management versus developing new and additional water in the Middle East : a critical review. *Water Policy* 3:143–163
- Hanke JE, Wichern D (2014) *Business forecasting (Ninth Edition)*. In Pearson education (Vol. 5). Retrieved from <https://ejournal.poltektegal.ac.id/index.php/siklus/article/view/298%0Ahttp://repositorio.unan.edu.ni/2986/1/5624.pdf%0Ahttp://dx.doi.org/10.1016/j.jana.2015.10.005%0Ahttp://www.biomedcentral.com/1471-2458/12/58%0Ahttp://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=refe>
- Hanrahan G, Udeh F, Patil DG (2004) Chemometrics and statistics - multivariate calibration techniques. *Encycl Anal Sci Second Ed* 1: 27–32. <https://doi.org/10.1016/B0-12-369397-7/00077-7>
- Heizer J, Render B (2011) *Operations management (Tenth Edition)*. Pearson Education
- Huang L, Yin L (2017) Supply and demand analysis of water resources based on system dynamics model. *J Eng Technol Sci* 49(6):705–720. <https://doi.org/10.5614/j.eng.technol.sci.2017.49.6.1>
- Jang J-SR, Sun C-T, Mizutani E (1997) *Neuro-fuzzy and soft computing*. Prentice-Hall International
- Karmaker CL, Halder PK, Sarker E (2017) A study of time series model for predicting jute yarn demand: case study. *J Ind Eng* 2017:1–8. <https://doi.org/10.1155/2017/2061260>
- Lawens M, Mutsvangwa C (2018) Application of multiple regression analysis in projecting the water demand for the City of Cape Town. *Water Pract Technol* 13(3):705–711. <https://doi.org/10.2166/wpt.2018.082>



- Lola MS, Hussin MF, Yusoff IM, Ramlee MNA, Isa SH, Kamil AA, ... Abdullah MT (2017) A system dynamic model for sustainable ecotourism in Tasik Kenyir, Terengganu, Malaysia. Preprints (2):1–13. <https://doi.org/10.20944/preprints201702.0005.v1>
- Luna I, Hidalgo IG, Pedro PSM, Barbosa PSF, Francato AL, Correia PB (2017) Fuzzy inference systems for multi-step ahead daily inflow forecasting. *Pesquisa Operacional* 37(1):129–144. <https://doi.org/10.1590/0101-7438.2017.037.01.0129>
- Makki AA, Stewart RA, Beal CD, Panuwatwanicha K (2015) Novel bottom-up urban water demand forecasting model : revealing the determinants, drivers and predictors of residential indoor end-use consumption novel bottom-up urban water demand forecasting model : revealing the determinants, drivers and predictors. *Resour Conserv Recycl* 95:15–37
- Makridakis SG, Wheelwright SC, Hyndman RJ (1997) Forecasting methods and applications (Third). Wiley
- Marchal V, Dellink R, Van Vuuren D, Clapp C, Chateau J, Lanzi E, ... Van Vliet J (2011) OECD environmental outlook to 2050 Chapter 3: climate change. (November), 90. <https://doi.org/10.1787/9789264122246-en>
- Mehri M (2013) A comparison of neural network models, fuzzy logic, and multiple linear regression for prediction of hatchability. *Poult Sci* 92(4):1138–1142. <https://doi.org/10.3382/ps.2012-02827>
- Moorthi PVP, Singh AP, Agnivesh P (2018) Regulation of water resources systems using fuzzy logic: a case study of Amaravathi dam. *Appl Water Sci* 8(5):1–11. <https://doi.org/10.1007/s13201-018-0777-8>
- Muhamad WNH, Radam A, Yacob MR (2014) Using choice experiments to understand visitors preferences for the man-made lake ecotourism services in Terengganu. *J Market Consum Res* 4(September 2016):41–50
- Neuman WL (2011) Social research methods: qualitative and quantitative approaches. *Teach Sociol* 30:380. <https://doi.org/10.2307/3211488>
- Ospina-Noreña JE, Gay-García C, Conde AC, Sánchez-Torres Esqueda G (2009) Analysis of the water supply-demand relationship in the Sinú-Caribe basin, Colombia, under different climate change scenarios. *Atmosfera* 22(4):399–412
- Rauf M, Ullah S, Haseeb A, Shah H, Khan M (2015) Physiochemical investigation of River Kabul at Michini, Khyber Pakhtunkhwa. *Pakistan*. 7(3):280–291
- Ren C, Dong Y, Xue P, Xu W (2019) Analysis of water supply and demand based on logistic model. *IOP Conf Ser Earth Environ Sci* 300(2). <https://doi.org/10.1088/1755-1315/300/2/022013>
- Rosle S, Ibrahim S, Yusof MN (2018) Relationship of water quality parameters and depth with fish density in Kenyir Lake, Malaysia. *Int J Adv Sci Res* 22(November):22–25 Retrieved from <http://www.ramp-alberta.org>
- Rouf AJMA, Phang SM, Ambak MA (2010) Depth distribution and ecological preferences of periphytic algae in Kenyir Lake, the largest tropical reservoir of Malaysia. *Chin J Oceanol Limnol* 28(4):856–867. <https://doi.org/10.1007/s00343-010-9088-0>
- Ryu K, Sanchez A (2003) The evaluation of forecasting methods at an institutional food service dining facility. *J Hosp Financ Manage* 11(1):27–45. <https://doi.org/10.1080/10913211.2003.10653769>
- Sahoo S, Jha MK (2013) Groundwater-level prediction using multiple linear regression and artificial neural network techniques : a comparative assessment. *Hydrogeol J* 21:1865–1887. <https://doi.org/10.1007/s10040-013-1029-5>
- Sharip Z, Zakaria S (2008) Lakes and reservoir in Malaysia : management and research challenges. The 12th World Lake Conference, pp 1349–1355
- Sharma P, Sood S, Mishra SK (2020) Development of multiple linear regression model for biochemical oxygen demand (BOD) removal efficiency of different sewage treatment technologies in Delhi, India. *Sustain Water Resour Manag* 9:1–13. <https://doi.org/10.1007/s40899-020-00377-9>
- Shrestha BP, Duckstein L, Stakhiv EZ (1996) Fuzzy rule-based modeling of reservoir operation. *J Water Resour Plan Manag* 122:262–269
- Sundari R, Hadibarata T, Fakhri Abdul Malik R, Aziz M (2013) Multiple linear regression (MLR) modeling of wastewater in urban region of Southern Malaysia. *J Sustain Sci Manag* 8(1):93–102
- Suratman S, Bedurus E, Seng TH (2017) A preliminary study of the distribution of nitrogen compounds in Tasik Kenyir, Malaysia. *Orient J Chem* 33(3):1325–1330. <https://doi.org/10.13005/ojc/330332>
- Swain S, Patel P (2017) A multiple linear regression model for precipitation forecasting over Cuttack District, Odisha, India. 2017 2nd International Conference for Convergence in Technology (I2CT), pp 355–357
- Takagi T, Sugeno M (1985) Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans Syst Man Cybern* 1: 116–132
- Uriel E (2013) Multiple linear regression: estimation and properties. Retrieved from [https://www.uv.es/uriel/3 Multiple linear regression estimation and properties.pdf](https://www.uv.es/uriel/3%20Multiple%20linear%20regression%20estimation%20and%20properties.pdf)
- Wu W (2015) Oil and gas pipeline risk assessment model by fuzzy inference systems and artificial neural network. University of Regina
- (WWAP), U N W W A P (2015) Water for a sustainable world
- Yu S, Kang W, Ko S, Paik J (2015) Single image super-resolution using locally adaptive multiple linear regression. *J Opt Soc Am A* 32(12): 2264. <https://doi.org/10.1364/josaa.32.002264>
- Yuanzheng Z, Jinsheng W, Yanguo T, Rui ZUO (2012) Water demand forecasting of Beijing using the time series forecasting method. *J Geogr Sci* 22(5):919–932. <https://doi.org/10.1007/s11442-012-0973-7>
- Yusof A, Omar Fauzee M, Mohd Shah P, Soh K (2009) Exploring small-scale sport event tourism in Malaysia. *Res J Int Stud* 9:47–58
- Zadeh LA (1965) Fuzzy sets. *Inf Control* 8:338–353
- Zolkepli M, Dong F, Hirota K (2014) Automatic switching of clustering methods based on fuzzy inference in bibliographic big data retrieval system. *Int J Fuzzy Logic Intell Syst* 14(4):256–267

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.