ZIBELINE INTERNATIONAL™ PUBLISHING
ISSN: 2521-0831 (Print)
ISSN: 2521-084X(Online)
CODEN: MSMAD

Matrix Science Mathematic (MSMK)

DOI: http://doi.org/10.26480/msmk.01.2025.19.25

RESEARCH ARTICLE

OPTIMAL MEMBERSHIP FUNCTION SELECTION FOR A CO-ACTIVE ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM MODELLING OF RESERVOIR SEDIMENTATION IN NIGERIA

Stephen Olushola Oladosu[a,*] Alfred Sunday Alademomi[b], Samuel Elisha Odonye[c]

[a] *Department of Geomatics, Faculty of Environmental Sciences, University of Benin, P.M.B. 1154, Edo State.*
[b] *Department of Surveying and Geoinformatics, Faculty of Engineering, University of Lagos, P.M.B. 12003, Akoka, Lagos State.*
[c] *Department of Surveying and Geoinformatics, School of Environmental Studies, Ramat Polytechnic, P. M. B 1070, Maiduguri, Borno State.*
**Corresponding Author Email: olushola.oladosu@uniben.edu*

ARTICLE DETAILS

Article History:

Received 10 January 2025
Revised 25 February 2025
Accepted 27 March 2025
Available online 24 April 2025

ABSTRACT

This study evaluates the performance of various fuzzy membership functions (MFs) in predicting volume and bedload rate using sediment data from a bathymetric survey at Ikpoba Dam. Twelve cases with different membership functions: Gaussian, triangular, trapezoidal, and bell-shape were tested across different epochs. The models were assessed based on Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-squared (R^2) values for both training and testing datasets. The Gaussian membership function (Gaussmf), with 7 membership functions and 200 training epochs, outperformed the others, achieving the lowest RMSE of 0.568 (training) and 0.579 (testing), MAE of 0.437 (training) and 0.445 (testing), and highest R^2 values of 0.914 (training) and 0.932 (testing) for volume prediction. For bedload rate, it also achieved the lowest RMSE of 0.509 (training) and 0.517 (testing), MAE of 0.391 (training) and 0.397 (testing), and highest R^2 values of 0.9354 (training) and 0.9496 (testing). In contrast, the Trapezoidal membership function (Trapmf) showed the worst performance with RMSE values of 0.874 (training) and 0.905 (testing), MAE values of 0.652 (training) and 0.677 (testing), and R^2 values of 0.812 (training) and 0.804 (testing). These results emphasize the significance of membership function selection and training epochs in optimizing fuzzy models for environmental and geospatial applications.

KEYWORDS

Ikpoba Dam, Deposition, Sedimentation, Soft-Computing, Nigeria

1. INTRODUCTION

In the field of fuzzy logic and neural networks, the Co-active Neuro-Fuzzy Inference System (CANFIS) model is a powerful tool that combines the strengths of both techniques to handle uncertainty and imprecision in data (Singh et al., 2028). The CANFIS relies on the concept of membership functions, which define how each point in the input space is mapped to a membership value between 0 and 1. The choice of membership functions plays a crucial role in the performance of a CANFIS model, as it directly influences the system's ability to approximate nonlinear functions. In most cases, researchers are faced with challenge of finding the best choice of membership function that will succinctly describe the real-life scenario intended to be modeled. In that case, the comparison of various membership functions for the CANFIS model could reveals significant insight into their effectiveness and adaptability. Research indicates that different membership functions can substantially influence the performance of CANFIS in various applications.

For instance, some researchers highlighted that triangular and trapezoidal membership functions yield distinct results in terms of accuracy and computational efficiency, with trapezoidal functions often providing better performance in complex datasets (Kabir and Kabir, 2021). Some researcher further emphasise the importance of selecting appropriate membership functions, noting that Gaussian functions can enhance the model's ability to generalise from training data, thus improving predictive accuracy (Talpur et al., 2017). In other findings suggest that the choice of membership function also affects the convergence speed of the learning algorithm, with some functions leading to faster adaptation in dynamic environments (Pancardo et al., 2021). However, a group researcher points out that while it is true that certain membership functions may excel in specific scenarios, their performance can vary significantly based on the nature of the input data and the problem domain (Narayan et al., 2021).

According to a study, a hybrid approach integrating a combination of multiple membership functions, may offer a more robust solution, allowing for greater flexibility and improved outcomes across diverse applications (Joseph et al., 2023). The selection of the appropriate membership function is crucial in designing a CANFIS model that accurately captures the underlying patterns in the data (Scherer, 2012). Each of these functions offers unique advantages depending on the specific requirements of the application, such as the nature of the data, the desired smoothness of transitions, and computational constraints. The selection of the appropriate membership function is a critical design decision in CANFIS models, as it impacts both the accuracy and interpretability of the system (Tomasiello et al., 2023).

While each membership function has its strengths, characteristics, and suitability for different applications, the choice between them depends on the specific application requirements. For instance, the bell-shaped and Gaussian functions are preferred in applications requiring smooth transitions and high precision, while the triangular and trapezoidal functions are more suitable for systems where simplicity and computational efficiency are critical (Gupta et al., 2023). Research suggests that in applications like pattern recognition and control systems,

Quick Response Code	Access this article online	
	Website: www.matrixsmathematic.com	**DOI:** 10.26480/msmk.01.2025.19.25

Cite The Article: Stephen Olushola Oladosu, Alfred Sunday Alademomi, Samuel Eliza Odonye (2025). Optimal Membership Function Selection for A Co-Active Adaptive Neuro-Fuzzy Inference System Modelling of Reservoir Sedimentation in Nigeria. *Matrix Science Mathematic, 9(1): 19-25.*

where input data can vary widely, Gaussian and bell-shaped functions tend to perform better due to their smooth and natural representation of data (Janková, et al., 2022). On the other hand, in environments where quick decisions are needed, such as in embedded systems or real-time fault detection, the triangular and trapezoidal functions may offer the necessary speed and simplicity (Sadollah, 2018).

The CANFIS model is a hybrid computational framework that synergises the learning capabilities of neural networks with the reasoning power of fuzzy logic. Central to its functioning are membership functions, which map input values to degrees of membership, thus enabling the system to handle uncertainty and imprecision effectively (Liu and Vuillemot, 2023). This paper aims to examines four widely used membership functions viz: the bell-shape, the Gaussian, the triangular, and the trapezoidal, with a view to selecting the best performing type for implementation in CANFIS model for sediment deposited in a reservoir in Nigeria.

2. MATERIALS AND METHODS

2.1 Study Area, Data Availability and Data Transformation

The geographical location of the study area in UTM Zone 31 North are: (787709.409 mE; 709586.339 mN, and 793326.993mE; 705212.341 mN). By Koppen classification, Benin City has a tropical savanna climate with an average annual rainfall of 2, 681 mm. From 2017 to 2019, data was collected from Ikpoba dam by repetitive bathymetric surveys. Records were taken each time data was available for other predetermined geo-environmental factors for an extended period of three years. Irrelevant and incomplete data points were filtered, and the refined dataset was re-sampled to 480 data points. The sum of computed sediment at the dam was 840127.346 m^3. Other dominant geo-environmental variables contributing to sedimentation like rainfall, topography inlet velocity etc. were made to undergo transformation and conversion into usable format for modelling. Figure 1 show the boundary of the study location and the catchment area.

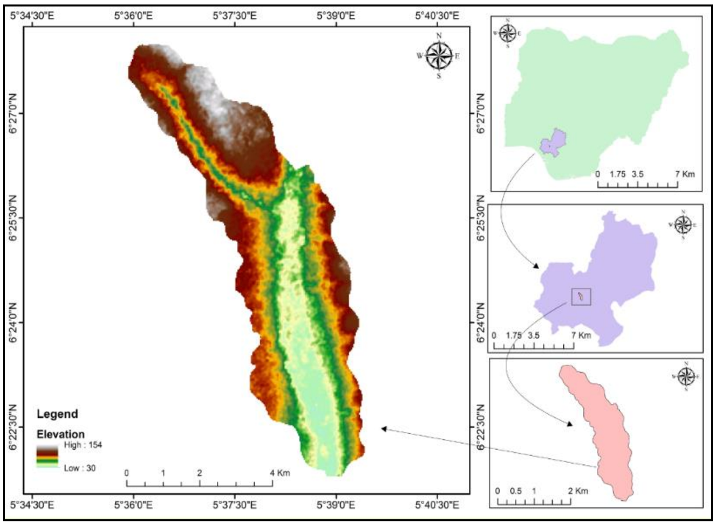

Figure 1: Map of the study area catchment

2.2 Conceptual Framework

The conceptual framework in the flowchart in Figure 2, outlines a stepwise decision-making and process management structure. It begins with input, followed by sequential activities that include decision points to evaluate conditions. Based on the decisions, it branches into alternative paths or loops back for iterative actions. The framework concludes with a defined endpoint that ensures the process is goal-oriented and dynamic.

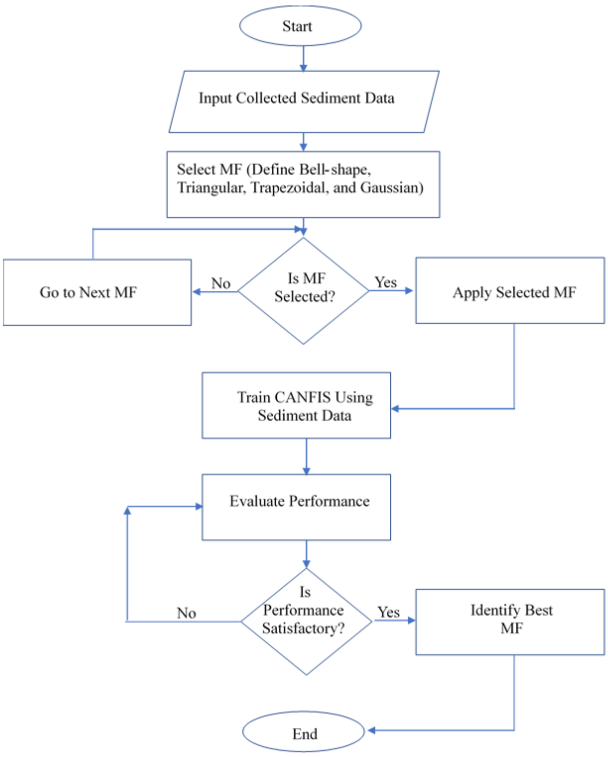

Figure 2: Conceptual diagram of the study.

2.3 Data Cleaning and Testing

A Multicollinearity test using Variance Inflation Factor (VIF) confirmed that all variables were suitable for modeling, with no signs of collinearity (VIF > 5) (Kutner, 2004). Missing values were imputed, and incomplete records were removed. The data was standardized to ensure equal contribution from all variables, while outliers were managed using appropriate transformations to prevent skewed results. Noise was filtered to improve accuracy.

The dataset was split into 75% for training and 25% for testing, ensuring unbiased evaluation. Necessary transformations, including Log, Square, Cubic, Box-Cox, and Yeo-Johnson, were applied to address non-linearity, skewness, and variance issues to improve the model performance. Equations 1–5 show the transformations applied. The transformations are explained a bit further.

The general logarithmic transformation model is typically represented by equation 1. It was used to stabilizes variance and makes the data closer to normal distribution for the variables.

$$Y' = \log(Y) \tag{1}$$

Where: *Y***:** represents the original data or variable to be transformed

The square transformation was adopted amplifies check the differences between larger values in the variables especially when the variance increases with the variable's magnitude. This model is represented by equation 2.

$$Y' = Y^2 \tag{2}$$

Where: *Y***:** is the original data or variable to be transformed

For a dependent variable *Y*, the cubic transformation is defined by equation 3. This model was particularly useful in this study as it caters for situations where relationships in variables are inherently cubic in nature.

$$Y' = Y^3 \tag{3}$$

Where: *Y***:** is the original data or variable to be transformed

The Box-Cox transformation is defined as in equation 4. This was adopted in cases where we have data that are strictly positive and the study required a flexible approach to normalize and stabilize the variance.

$$Y' = \begin{cases} \dfrac{Y^{\lambda} - 1}{\lambda}, if\ \lambda \neq 0 \\ \log(Y), if\ \lambda = 0 \end{cases} \tag{4}$$

Where: λ, represents the transformation parameter

Yeo-Johnson transformation helps make data more symmetric and reduces skewness, which can improve model performance and interpretability. When the data contains zero or negative values and normalization is required, this model is capable of handling such scenario. Equation 5 represent the model.

$$Y' = \begin{cases} \dfrac{((Y+1)^{\lambda} - 1)}{\lambda}, & if\ \lambda \neq 0\ and\ Y \geq 0 \\ \log(Y+1), & if\ \lambda = 0\ and\ Y \geq 0 \\ \dfrac{-((-Y+1)^{2-\lambda} - 1)}{2-\lambda}, & if\ \lambda \neq 2\ and\ Y < 0 \\ -\log(-Y+1), & if\ \lambda = 2\ and\ Y < 0 \end{cases} \tag{5}$$

The groupings like (5,5,5,5,5,5,5), or (7,7,7,7,7,7,7) represent the number of MFs used for each input variable. These configurations are typically achieved by adjusting the fuzzy system's parameters, and the system uses epochs (iterations) during training to refine the positions, shapes, and spans of the MFs. More epochs allow the system to experiment with different configurations and improve its predictive accuracy, selecting the optimal number of MFs based on error metrics such as RMSE, MAE, or R^2 as will be presented later in the results section. Table 1 shows the summary of the input parameters after some modification processes.

Table 1: Brief description of input parameters and MF

Variable Name	Min	Max	No. of MF	Description
rainfall	32.4905	1017.6	7	Refers to the height of the water layer covering the ground in a period of time.
L-S factor	0.0339	43.277	7	Length-slope factor describes the effect of topography on soil erosion.
depth	0.3571	8.2138	7	Mean distances from the bottom of the eroded surface profile.
inletvel	0.0128	1.2429	7	Inlet velocity is the gain of mass of suspended sediment by unit of time.
psize	0.0748	39.188	7	Weighing capacity of eroded soil particles.
twi	0.0022	3.4357	7	Topographic wetness index identifies the potential of runoff generation locations.

Note: node_id, means grid node; inlet. Vel., means inlet velocity; p. size, means particle size (d_{50}); twi, means topographic wetness index

2.4 CANFIS Model Development

To develop a CANFIS model in this research for the prediction of sedimentation in a reservoir, the following data were sourced: hydrological data, including rainfall, streamflow, and water levels. Sediment data such as particle size distribution and bathymetric data, geographical and geological information like watershed characteristics and soil types are carefully gathered and laboratory analysis conducted to determine the needed variables. Seven input variables and two output variables are represented in the CANFIS network architecture in Figure 3.

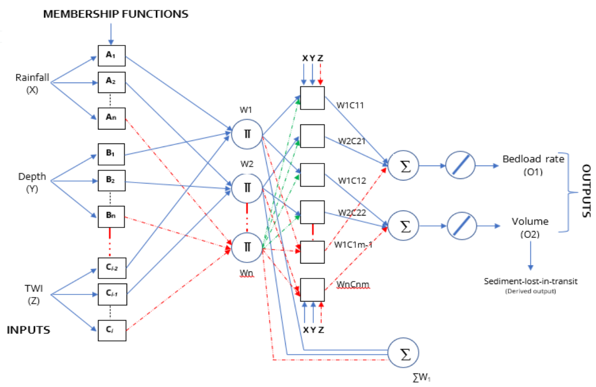

Figure 3: CANFIS architecture with multiple input multiple output (MIMO)

Tables 2 and 3 show the excerpts from raw data and normalised data for 480 rows by 8 columns. The random selection at grid nodes was intentional as all data cannot be presented.

Table 2: Excerpt from raw data before transformation was applied

node_id	rainfall	slope	depth	inlet. vel.	p. size	twi	bedload rate	volume
1	976.810	22.064	1.428	0.338	9.273	1.651	1.796	4633.911
2	1017.685	22.666	1.428	0.338	9.273	2.673	1.796	2411.311
3	869.906	25.000	1.428	0.338	9.273	1.457	1.796	2914.542
4	957.944	23.877	1.428	0.338	9.273	1.775	1.796	2096.793
5	838.463	23.328	1.428	0.338	9.273	1.094	1.796	2914.542
6	378.357	19.777	1.428	0.338	9.273	1.325	1.796	3501.644
7	617.319	16.375	1.071	0.350	8.974	0.619	1.977	3229.060
8	855.233	15.251	1.071	0.350	8.974	0.809	1.977	2767.766
9	561.771	15.419	1.071	0.350	8.974	0.639	1.977	3124.221
10	276.693	15.419	1.071	0.350	8.974	0.460	1.977	3795.194

Note: node_id, means grid node; inlet. Vel., means inlet velocity; p. size, means particle size (d_{50}); twi, means topographic wetness index

Table 3: Excerpt from normalised data after applying transformations

node_id	rainfall	slope	depth	inlet. vel.	p. size	twi	bedload rate	volume
346	0.871	0.050	1.000	0.643	0.914	0.149	0.679	0.500
151	0.872	0.050	0.136	0.698	0.411	0.287	1.000	0.229
38	0.257	0.287	0.136	0.193	0.174	0.232	0.220	0.438
255	0.887	0.106	0.136	0.120	0.145	0.153	0.530	0.847
361	0.990	0.040	1.000	0.643	0.914	0.534	0.581	0.131
287	0.553	0.549	0.091	0.488	0.193	0.389	0.090	0.340
170	0.971	0.132	0.136	0.698	0.411	0.284	0.442	0.313
303	0.817	0.225	0.091	0.339	0.222	0.170	0.314	0.369
437	0.790	0.174	1.000	0.643	0.914	0.420	0.669	0.189
291	0.366	0.461	0.227	0.198	0.270	0.214	0.082	0.378

2.5 CANFIS Membership Functions Algorithms

For the proposed model, the following membership functions were taken into account. Equations 6 – 9 represent each membership function model:

Gaussian MF: The Gaussian membership function is specified by two parameters $\{c,\sigma\}$

$$guassmf\left(x,c,\sigma\right)=\exp\left[-\frac{\left(x-c\right)^{2}}{2\sigma^{2}}\right] \tag{6}$$

c is the center of the MF, and σ determines the width of the MF axons.

Triangular MF: A Triangular membership function is defined by three parameters $\{a,b,c\}$, where a,b and c represents the χ coordinates of the three vertices of $\mu_A\left(x\right)$ in a fuzzy set A. Points a and c represents lower and upper boundary respectively where membership degree is 0 while b is the centre where membership degree is 1. $\mu_A\left(x\right)=trimf\left(x,a,b,c\right)$:

$$=\begin{cases} 0 & if\ \ xa \\ \dfrac{x-a}{b-a} & if\ \ a\le x\le b \\ \dfrac{c-x}{c-b} & if\ \ b\le x\le c \\ 0 & if\ \ c\le x \end{cases} \tag{7}$$

Trapezoidal MF: A Trapezoidal membership function is specified by four parameters as $\{a,b,c,d\}$:

$$trapmf\left(x,a,b,c,d\right)=\begin{Bmatrix} 0 & if\ \ x\le a \\ \dfrac{x-a}{b-a} & if\ \ a\le x\le b \\ 1 & if\ \ b\le x\le c \\ \dfrac{d-x}{d-c} & if\ \ c\le x\le d \\ 0 & if\ \ d\le x \end{Bmatrix}$$

$$=\max\left(\min\left(\frac{x-a}{b-a},1,\frac{d-x}{d-c},0\right)\right) \tag{8}$$

Bell-Shape MF: For a bell-shaped membership function, μ_{A_i} is given as:

$$\mu_{A_i}=\frac{1}{1+\left[\left(x-c_i\right)/a_i\right]^{2b_i}} \qquad i=1,2 \tag{9}$$

Where x is value of input to *i* node, and $\{a_i,b_i,c_i\}$ are the adaptable parameters of membership function of this set. These conditioner parameters always change the bell-shape function on linguistic label A_i

2.6 CANFIS Training Process

CANFIS learn with hybrid algorithm combining gradient descent and least squares methods. The Forward training and the backward pass uses equation 10.

$$\left.\begin{aligned} f &= \frac{w_1}{w_1+w_2} f_1 + \frac{w_2}{w_1+w_2} f_2 \\ f &= \overline{w}\left(p_1x+q_1y+r_1\right)+\overline{w}\left(p_2x+q_2y+r_2\right), \\ f &= \left(\overline{w}_1x\right)p_1+\left(\overline{w}_1y\right)q_1+\left(\overline{w}_1\right)r_1+\left(\overline{w}_2x\right)p_2+\left(\overline{w}_2y\right)q_2+\left(\overline{w}_2\right)r_2 \end{aligned}\right\} \tag{10}$$

where p_1, q_1, r_1, p_2, q_2, and r_2 are the linear consequent parameters.

2.7 CANFIS Production Rules

To fine-tune the parameters of a (FIS), CANFIS uses ANN learning techniques. The fuzzy production rules are in these orders: fuzzification, aggregation, activation, accumulation, and difzzification. Figure 4 shows a typical fuzzy production rule interface.

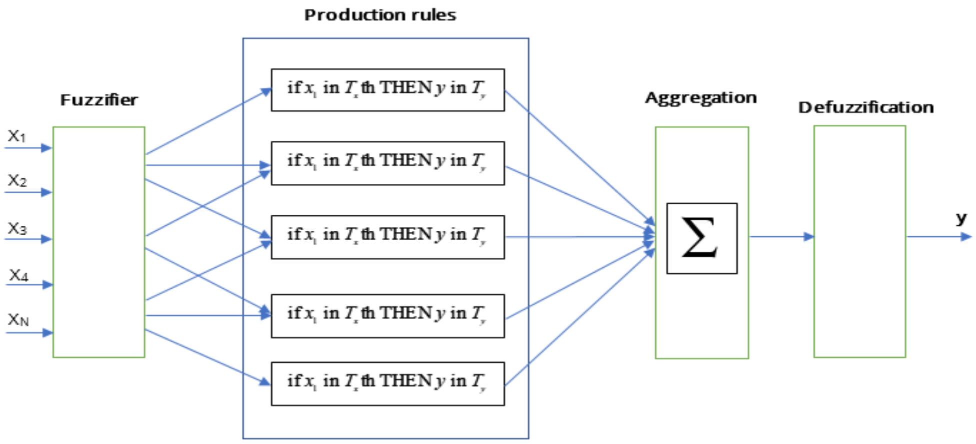

Figure 4: Fuzzy production rule system

An excerpt of the fuzzy production rules from this research is shown in Figure 5.

```
     [,1] [,2]      [,3] [,4]      [,5]  [,6]    [,7] [,8]      [,9]  [,10]   [,11] [,12]     [,13] [,14]      [,15] [,16
 [1,] "IF" "rainfall" "is" "large"    "and" "slope" "is" "v.small"  "and" "depth" "is"  "v.small"  "and" "inletvel" "is"  "sma
 [2,] "IF" "rainfall" "is" "vv.large" "and" "slope" "is" "v.small"  "and" "depth" "is"  "v.small"  "and" "inletvel" "is"  "sma
 [3,] "IF" "rainfall" "is" "large"    "and" "slope" "is" "v.small"  "and" "depth" "is"  "v.small"  "and" "inletvel" "is"  "sma
 [4,] "IF" "rainfall" "is" "medium"   "and" "slope" "is" "v.small"  "and" "depth" "is"  "v.small"  "and" "inletvel" "is"  "sma
 [5,] "IF" "rainfall" "is" "large"    "and" "slope" "is" "v.small"  "and" "depth" "is"  "v.small"  "and" "inletvel" "is"  "v.s
 [6,] "IF" "rainfall" "is" "v.large"  "and" "slope" "is" "v.small"  "and" "depth" "is"  "v.small"  "and" "inletvel" "is"  "sma
 [7,] "IF" "rainfall" "is" "v.large"  "and" "slope" "is" "v.small"  "and" "depth" "is"  "v.small"  "and" "inletvel" "is"  "sma
 [8,] "IF" "rainfall" "is" "large"    "and" "slope" "is" "v.small"  "and" "depth" "is"  "v.small"  "and" "inletvel" "is"  "sma
 [9,] "IF" "rainfall" "is" "large"    "and" "slope" "is" "v.small"  "and" "depth" "is"  "v.small"  "and" "inletvel" "is"  "sma
[10,] "IF" "rainfall" "is" "v.large"  "and" "slope" "is" "v.small"  "and" "depth" "is"  "v.small"  "and" "inletvel" "is"  "lar
[11,] "IF" "rainfall" "is" "medium"   "and" "slope" "is" "small"    "and" "depth" "is"  "v.small"  "and" "inletvel" "is"  "v.s
[12,] "IF" "rainfall" "is" "v.large"  "and" "slope" "is" "small"    "and" "depth" "is"  "v.small"  "and" "inletvel" "is"  "sma
[13,] "IF" "rainfall" "is" "vv.large" "and" "slope" "is" "v.small"  "and" "depth" "is"  "v.small"  "and" "inletvel" "is"  "lar
[14,] "IF" "rainfall" "is" "vv.large" "and" "slope" "is" "vv.small" "and" "depth" "is"  "vv.large" "and" "inletvel" "is"  "lar
```

Figure 5: Excerpt of CANFIS fuzzy production rules.

2.8 Model Performance Metrics

In any model development process, it is important to define the criteria by which the performance of the model and its prediction accuracy is evaluated, (Legates and McCabe, 1999). The following criteria were used in examining the performance metrics of the models which are commonly used for regression modelling.

2.8.1 Root mean squared error (RMSE)

This node estimates the residual between the actual value and predicted value. A model has better performance if it has a smaller RMSE. Equation 11 is used for this purpose.

$$RMSE = \sqrt{\frac{1}{m}\sum_{k=1}^{m}\left(t_k - y_k\right)^2} \tag{11}$$

Where: t_k is the actual value, y_k is the predicted value produced by the model, and m is the total number of observations.

2.8.2 Means absolute error (MAE)

It's the absolute difference between the estimated value and true value.

With equation 12 this can be achieved.

$$MAE = \frac{1}{m}\sum_{k=1}^{m}\left|t_k - y_k\right| \tag{12}$$

Where: t_k is the actual value, y_k is the predicted value produced by the model, and m is the total number of observations.

2.8.3 Correlation coefficient (R)

This criterion reveals the strength of relationships between actual values and predicted values. The correlation coefficient has a range from 0 to 1, and a model with a higher R means it has a better performance. This was achieved using equation 13.

$$R = \frac{\sum_{k=1}^{m}\left(t_k - \bar{t}\right)\left(y_k - \bar{y}\right)}{\sqrt{\sum_{k=1}^{m}\left(t_k - \bar{t}\right)^2 . \sum_{k=1}^{m}\left(y_k - \bar{y}\right)^2}} \tag{13}$$

Where: $\bar{t} = \frac{1}{m}\sum_{k=1}^{m} t_k$ and $\bar{y} = \frac{1}{m}\sum_{k=1}^{m} y_k$ are the average values of t_k and y_k respectively.

3. Results and Discussions

The result obtained from the different membership functions with their respective training and testing errors for the output parameters are displayed in Table 4.

Table 4: The CANFIS results from combinations of shape and number of MF

Output	No	MF Pair	MF type	No of epochs	RMSE		MAE		R^2	
					Training Error	Testing Error	Training Error	Testing Error	Training Error	Testing Error
For volume	1	(5,5,5,5,5,5,5)	Gaussmf	100	0.7967	0.8172	0.56207	0.5633	0.8720	0.8854
	2	(5,5,5,5,5,5,5)	Trimf	100	0.8460	0.8362	0.63920	0.6648	0.8509	0.8487
	3	(5,5,5,5,5,5,5)	Trapmf	100	0.8738	0.9054	0.65229	0.6767	0.8123	0.8035
	4	**(7,7,7,7,7,7,7)**	**Gaussmf**	**200**	**0.5683**	**0.5799**	**0.43649**	**0.4454**	**0.9139**	**0.9326**
	5	(7,7,7,7,7,7,7)	Trimf	200	0.6330	0.6832	0.49637	0.4987	0.8925	0.8975
	6	(7,7,7,7,7,7,7)	Bellmf	200	0.7632	0.7821	0.53390	0.5493	0.8633	0.8591
For bedload rate	7	(5,5,5,5,5,5,5)	Gaussmf	100	0.6238	0.6300	0.57383	0.5612	0.9113	0.9047
	8	(5,5,5,5,5,5,5)	Trimf	100	0.7653	0.7972	0.59223	0.6722	0.8356	0.8399
	9	(5,5,5,5,5,5,5)	Trapmf	100	0.8098	0.8152	0.62292	0.6430	0.8473	0.8570
	10	**(7,7,7,7,7,7,7)**	**Gaussmf**	**200**	**0.5088**	**0.5165**	**0.39074**	**0.3967**	**0.9354**	**0.9496**
	11	(7,7,7,7,7,7,7)	Trimf	200	0.5451	0.5573	0.46559	0.4830	0.9011	0.9274
	12	(7,7,7,7,7,7,7)	Bellmf	200	0.5970	0.6118	0.53710	0.5863	0.8972	0.9053

3.1 Discussions

Table 4 presents a comparison of the performance of different membership functions (MFs) for predicting volume and bedload rate, using different configurations. The performance is evaluated using three key metrics: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-squared (R^2) for both training and testing datasets.

3.1.1 Volume Prediction Performance

For predicting volume, Gaussian Membership Function (Gaussmf), using 7 membership functions (MFs) and trained over 200 epochs, consistently outperforms all other membership functions. This configuration resulted in the lowest RMSE values for both training (0.5683) and testing (0.5799). The MAE values were also the lowest for both training (0.43649) and testing (0.4454), and it achieved the highest R^2 values of 0.9139 for training and 0.9326 for testing. This indicates that the model explains a high percentage of the variance in the data, and the predictions are quite accurate.

On the other hand, Trapezoidal Membership Function (Trapmf), with 5 membership functions and trained for 100 epochs, performed the worst. It had the highest RMSE (0.8738 for training and 0.9054 for testing) and MAE (0.65229 for training and 0.6767 for testing). The R^2 values were also the lowest, with 0.8123 for training and 0.8035 for testing, suggesting poor explanatory power and a less accurate model.

The Triangular Membership Function (Trimf), with 5 membership functions and 100 epochs, performed somewhat better than Trapmf but still lagged behind Gaussmf. Its RMSE (0.846 for training and 0.836 for testing) and MAE (0.639 for training and 0.665 for testing) were higher than those of Gaussmf, and its R^2 values (0.851 for training and 0.849 for testing) were lower. Besides, increasing the number of membership functions from 5 to 7 and training the model for 200 epochs led to noticeable improvements in performance for both Gaussmf and Trimf, with Gaussmf achieving the best results overall.

3.1.2 Bedload Rate Prediction Performance

The performance for predicting bedload rate follows a similar pattern. Gaussmf, using 7 membership functions and trained over 200 epochs, again emerges as the best-performing model. It recorded the lowest RMSE values (0.509 for training and 0.517 for testing) and the lowest MAE values (0.391 for training and 0.397 for testing). This configuration also achieved the highest R^2 values (0.935 for training and 0.949 for testing), suggesting that the model is highly accurate and explains the variance in the data well.

Trapmf, with 5 membership functions and trained for 100 epochs, had the highest RMSE (0.809 for training and 0.815 for testing) and the highest MAE (0.623 for training and 0.643 for testing), leading to lower R^2 values (0.847 for training and 0.857 for testing), indicating poorer predictive performance. Trimf, with 5 membership functions and trained for 100 epochs, also performed better than Trapmf but worse than Gaussmf. Its RMSE values were 0.765 for training and 0.797 for testing, with MAE values of 0.592 for training and 0.672 for testing. Its R^2 values were 0.836 for training and 0.839 for testing, which are lower than those of Gaussmf.

4. Conclusions

Based on the findings from this study, it can be concluded that the optimal selection of membership functions plays a crucial role in improving the performance of Co-active Adaptive Neuro-Fuzzy Inference System (CANFIS) models for reservoir sedimentation prediction in Nigeria. Among the different membership functions tested, the Gaussian membership function (Gaussmf) consistently provided the most accurate predictions for both volume and bedload rate, especially when used with 7 membership functions (MFs) and trained over 200 epochs. This configuration led to higher R^2 values, indicating stronger explanatory power and more reliable results. On the other hand, fewer MFs and fewer training epochs resulted in less accurate predictions, with higher errors and lower R^2 values. Therefore, for optimal model accuracy in predicting reservoir sedimentation, using the Gaussian function with 7 MFs and training for 200 epochs is recommended. This approach can be effectively applied to improve the management of reservoir sedimentation in Ikpoba dam, Nigeria and similar environments.

ACKNOWLEDGEMENTS

The authors are grateful to Iterlen Industrial Services Ltd, Km 1 Effurun-DSC Express Road, Ugbolokposo, Delta State for the release of the Hi-Target echo-sounder & accessories. The Department of Geomatics, University of Benin, for the release of the Total Station for data collection, the Department of Civil Engineering for the laboratory aids for soil analysis, and

DECLARATION OF COMPETING INTEREST

The authors have no conflict of interest to declare.

REFERENCES

Gupta, S., Biswas, P.K., Aljafari, B., Thanikanti, S.B., and Das, S.K., 2023. Modelling, simulation and performance comparison of different membership functions-based fuzzy logic control for an active

magnetic bearing system. https://doi.org/10.1049/tje2.12229

Janková, Z., and Rakovská, E., 2022. Comparison uncertainty of different types of membership functions in T2FLS: Case of international financial market. Applied Sciences, 12 (2), Pp. 918. https://doi.org/10.3390/app12020918

Joseph, L., O'Leary, M.L., and Zagré, B., 2023. Integrating hybrid grid partition and rough set method for fuzzy rule generation: A novel approach for accurate dataset classification. International Journal of Enterprise Modelling, 17 (2), PP. 88–98. https://doi.org/10.35335/emod.v17i2.21

Kabir, M., and Kabir, M.M., 2021. Fuzzy membership function design: An adaptive neuro-fuzzy inference system (ANFIS)-based approach. In 2021 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1–5). IEEE.

Kutner, M., Nachtsheim, C., and Neter, J., 2004. Applied linear statistical models (4th ed.). McGraw-Hill.

Liu, L., and Vuillemot, R., 2023. A generic interactive membership function for categorization of quantities. IEEE Computer Graphics and Applications, 43(In press), Pp. 39–48. https://doi.org/10.1109/mcg.2023.3301449

Martinsen, A.B., and Lekkas, A.M., 2018. Curved path following with deep reinforcement learning: Results from three vessel models. OCEANS 2018 MTS/IEEE Charleston, Pp. 1–8. https://doi.org/10.1109/OCEANS.2018.8604829

Narayan, J., Banerjee, S., Kamireddy, D., and Dwivedy, S.K., 2021. Fuzzy membership functions in ANFIS for kinematic modeling of 3R manipulator. In Hussain, C.M., and Di Sia, P. (Eds.), Handbook of smart materials, technologies, and devices (pp. 1–25). Springer. https://doi.org/10.1007/978-3-030-58675-1_26-1

Pancardo, P., Hernández-Nolasco, J.A., Wister, M., and Garcia-Constantino, M., 2021. Dynamic membership functions for context-based fuzzy systems. IEEE Access, 9, Pp. 29665–29676. https://doi.org/10.1109/ACCESS.2021.3058943

Ramírez, M., Melin, P., and Castillo, O., 2024. A new hybrid approach for clustering, classification, and prediction of world development indicators combining general type-2 fuzzy systems and neural networks. Axioms, 13 (6), Pp. 368. https://doi.org/10.3390/axioms13060368

Sadollah, A., 2018. Introductory chapter: Which membership function is appropriate in fuzzy systems? In Fuzzy logic and applications. InTech. https://doi.org/10.5772/intechopen.79552

Scherer, R., 2012. Introduction to fuzzy systems. In Multiple fuzzy classification systems (pp. 1–18). Springer. https://doi.org/10.1007/978-3-642-30604-4_2

Singh, V.K., Kumar, D., Kashyap, P.S., and Kisi, O., 2018. Simulation of suspended sediment based on gamma test, heuristic, and regression-based techniques. Environmental Earth Sciences, 77, Pp. 708. https://doi.org/10.1007/s12665-018-7892-6

Talpur, N., Abdulkadir, S.J., Alhussian, H., Hasan, M.H., Aziz, N., and Bamhdi, A., 2023. Deep neuro-fuzzy system application trends, challenges, and future perspectives: A systematic survey. Artificial Intelligence Review, 56, Pp. 865–913. https://doi.org/10.1007/s10462-022-10188-3

Tomasiello, S., Uzair, M., Liu, Y., and Loit, E., 2023. Data-driven approaches for sustainable agri-food: Coping with sustainability and interpretability. Journal of Ambient Intelligence and Humanized Computing, 14, Pp. 16867–16878. https://doi.org/10.1007/s12652-023-04702-w

Zadeh, L.A., 1965. Fuzzy sets. Information and Control, 8 (3), Pp. 338–353.