J. Indones. Math. Soc.
Vol. 31, No. 03 (2025), pp. 1–16.

An Imprecise High-Order Fuzzy Time Series Model Forecasting the Stocks Traded Using Central-Log-Ratio Transformation of Compositional Data

Muhammad Shahbaz Khan[1,2*], Mir Ghulam Hyder Talpur[2], and Muhammad Aslam[3]

[1]Mathematics and Statistics, Karachi Institute of Economics and Technology, Pakistan.
[2,3]Mathematics and Statistics, Institute of Business Management, Pakistan.
shahbazk@kiet.edu.pk[1], hyder.talpur@iobm.edu.pk[2], dr.aslam@iobm.edu.pk[3]

Abstract. Malaysia, Indonesia, and Thailand countries have half of the trade in the group of the Association of Southeast Asian Nations (ASEAN) members. An agreement has increased the importance of trade coalition without the US dollar among them. To understand the pattern of the Stock Traded in these countries, this study developed ample fuzzy time series foretelling models. An innovative data transformation approach called compositional data is employed. The Fuzzy Time series models are implemented with different orders of fuzzy logical relationships. The Trapezoidal and Spline S-shaped membership functions are engaged in these models. The performance of forecasting is evaluated through the compositional root mean square error (CRMSE) and compositional mean absolute percentage error (CMAPE). The analysis of forecasted accuracy measurements showed that the Third-Order Fuzzy Time Series model with a Trapezoidal membership function outclassed other orders models. It is also observable that Thailand's stock traded values increased compared to Malaysia and Indonesia.

Key words and Phrases: stock traded total, compositional data, Fuzzy time series, Fuzzy logic relationships, Central-Log-Ratio transformation.

1. INTRODUCTION

The Stocks Traded describes the value of domestic and international shares as the total number of "Shares Traded" multiplied by their corresponding matching prices. The Shares traded are the backbone of any country, which strengthens the

*Corresponding author
2020 Mathematics Subject Classification: 03E72, 62H30.
Received: 13-11-2024, accepted: 25-05-2025.

economy. The improvement of an economy's financial markets is strictly associated with the overall growth of the country Zaho et al. [1].

Malaysia, Indonesia, and Thailand are among the member countries of the Association of Southeast Asian Nations (ASEAN). These three countries have half of the trade among all ASEAN members. An agreement has increased the importance of their trade coalition without the US dollar among them. The forecasting of stock trades that are based on time is a very important part of analyzing the trend of their stock. To forecast these time series data trends, many traditional and non-traditional time series approaches are available and implemented successfully by researchers. In traditional time series models, Autoregressive (AR), Moving Average (MA), Autoregressive Moving Average (ARMA), and Autoregressive Integrated Moving Average (ARIMA) models are implemented. In non-traditional time series models, Machine Learning (ML), Deep Learning (DL), Hybrid Models, and Probabilistic and Bayesian Models are also employed by researchers. The better and more precise predicted values through forecasting models help us to make better and more precise decisions in every sector of the world Liu et al. [2]. These sectors like stock exchange rates, enrollment of students, oil and gas exploration, climate change, finance analysis, currency exchange rates, GDP estimation, tourism demand, and the stock market returns during COVID-19 Chang et al. [3], etc.

In recent studies, another dimension of the novelty of forecasting models is the conversion of data using various transformations. The transformation of data based on the Square root, inverse, logarithmic, and others are used to improve the estimation of parameters. These assumptions help us to fulfill the necessary assumptions for time series models. Amid these types of data transformation, compositional data is comprehensively utilized in recent research for forecasting multivariate time series (MTS) data. Compositional data (CoDa) is the analysis of positive multivariate data. The important property of CoDa analysis is that the data with a constant sum constraint. It means the values for each multivariate sample are whichever witnessed as summing to a constant, usually 1 to 100% Jean Aitchison [4]. In compositional data D-components multivariate sample of positive observations $x_1, x_1, \ldots, x_D$, with property $\sum_{n=1}^{D} x_i = 1$ or 100%, is known as composition. The D-components of a composition are the "D" number of variables of its system. We used compositional data for our multivariate data set related to stocks traded for three members of ASEAN.

The concept of Fuzzy Set Theory was introduced by Lotfi A. Zadeh [5] based on the value between 0 and 1. This value is called the "degree of truth" used as a substitute for Boolean logic 0 or 1 Rusell and Norvig [6]. The fuzziness idea performed efficient performance for prediction output with the high-accuracy rate. The past and recent trends of research work in every discipline of life show the importance of fuzzy logic and its utilization in Fuzzy Time Series. It works like a human mind's intelligence. It has been half a century since fuzzy logic systems were implemented and performed smartly in the scientific and academic community. The algorithms that work on fuzzy logic have tremendous output in modern technologies. The beginning of the FTS idea proposed by Song and Chissom [7] is based on fuzzy logic and fuzzy sets theory. They used Fuzzy logic, Fuzzy sets, and Fuzzy relations theory which was proposed by Lotfi A. Zadeh [5] to forecast the enrollment of Alabama University students.

The S. J. Chen [8] research brought accuracy and effectiveness as compared to work by predicting students' enrollments at Alabama University through a Fuzzy set concept. High-order fuzzy time series improve the accuracy of the forecasted enrollments of students at Alabama University. A study Huarng [9] resolved the issue of the number of intervals based on the universe of discourse.

A type-2 quantum fuzzy neural network (eIT2QFNN) model was developed to predict intelligent hybrid air quality Wang et al. [10]. A study used a grid method with an optimal number of partitions to predict the daily air pollution index using the FTS Markov chain (FTSMC) Alyousufi et al. [11]. A researcher developed a fuzzy multiple linear regression model with Gaussian fuzzy coefficients and Tanaka's minimum fuzziness criterion to predict the air quality index forecast using meteorological factors Gu et al. [12]. A new clustering outlier detection procedure for multivariate data is described by Yusoff and Nur [13]. A researcher predicted that wind speed addresses the issues of low inscrutability and excessive data reprocessing Shi et al. [14].

The fuzzy C-means model is used on a spectral distinction measure with high inequitable power. The divergence quantity matches principal component scores found from estimates of quantile cross-spectral densities López-Oriona [15]. An article proposed an extension to the base model AutoMFIS into e-AutoMFIS with high-dimensionality data Carval [16]. A researcher used a Traffic data set to develop deep attention fuzzy cognitive maps (DAFCM) Qin et al. [17]. A researcher constructed comprehensive time series forecasting models by using renewable energy data as a compositional data set Wei et al. [18]. Log ratios of components provide the natural means of studying compositional data to overcome the issue of spurious effects Pawlowsky et al. [19]. The scale invariance and sub-compositional coherence used in fuzzy clustering methods algorithms are broadly applied in a variety of fields using a nutritional data set Palarea [20]. A researcher proposed various principle component analysis methodologies for zero-inflated CoDa offered on a novel principal compositional subspace. They utilized multi-site microbiome datasets for this research Kim et al. [21]. In this research multi-rule combination prediction of compositional data based on a multivariate fuzzy time series model was applied over the three production sectors in the Chinese economy data Huang et al. [22].

The rest of the study is structured as Section 2, we presented the proposed HO-FTS-CoDa model Methodology. In Section 3, the application of HO-FTS-CoDa models is implemented using the Multivariate Time Series data related to the stocks traded in Indonesia, Malaysia, and Thailand from 2000 to 2022. Results and discussion in Section 4, and conclusion in Section 5.

2. MATERIALS AND METHODOLOGY

In this section, the proposed High-Order FTS model is explained in stages used to analyze and forecast MTS data using the Compositional Data (CoDa) approach. It extends the traditional FTS models by incorporating higher-order relationships and dependencies between variables, leading to more accurate and robust forecasted output.

2.1 Dataset: In this research, the MTS dataset is used, related to data on the Stocks Traded in Malaysia, Indonesia, and Thailand from 2000 to 2022 as reported on the World Bank Group data website [23].

2.2. Stages: The proposed High-Order FTS model stages are as follows:

Stage 1: Need to collect the multivariate time series (MTS) historical data $Y = [y_1, y_2, ..., y_D]^T$.

Stage 2: Convert the MTS data into percentage form $x = [x_1, x_2, ..., x_D]^T$.

Stage 3: Transform percentage data $x = [x_1, x_2, ..., x_D]^T$ into compositional data $c = [c_1, c_2, ..., c_D]^T$ by using the Centered log-ratio transformation (CLR) transformation Aitchison [4] method using Equation (1).

$$clr(x) = [\ln(x_1/g(x)), \ln(x_2/g(x)), \ldots, \ln(x_D/g(x))]^T, \tag{1}$$

where $g(x) = \left(\prod_{i=1}^{D} x_i\right)^{\frac{1}{d}}$ is the geometric mean of x.

Stage 4: To get the desirable number of intervals break the Universe of Discourse (UoD), which is presented as $U = [U_1, U_2, ..., U_D]^T$ using the Huarng [9] method. Note: minimum intervals will be 5 for making smooth utilization of UoD. UoD is defined in Equation (2).

$$U_i = [C_{i(min)} - d_{i(1)}, C_{i(max)} + d_{i(2)}], i = 1, ..., D, \tag{2}$$

where $C_{i(min)}$: transformed data minimum value in i-th variable, $C_{i(max)}$: transformed data maximum value in i-th variable, $d_{(i(1))}$ and $d_{(i(2))}$ are proper values used to smoothly present the UoD.

Stage 5: The number of intervals "K" with a suitable length of interval *"l"* Huarng [9] using Formula (3).

$$K = [(C_{i(max)} + d_{i(2)}) - (C_{i(min)} - d_{i(1)})]/l. \tag{3}$$

The 'K' intervals are presented by $V_{i(1)}$,$V_{i(2)}$, $\ldots$, $V_{i(k)}$ such as $V_{i(1)} = [v_1, v_2]$, $V_{i(2)} = [v_2, v_3]$, $\ldots$, $V_{i(k-2)} = [v_{k-2},\ vk-1]$,$V_{i(k-1)} = [v_{k-1}, v_k]$, and $V_{i(k)} = [v_k, v_{k+1}]$.

Stage 6: The discrete fuzzy sets $\tilde{C_1}$, $\tilde{C_2}$, ... , $\tilde{C_k}$ are defined in Equation (4).

$$\begin{cases} \tilde{C_1} = 1/v_{i(1)} + 0.5/v_{i(2)} + 0/v_{i(3)} + \ldots + 0/v_{i(k-1)} + 0/v_{i(k)} \\ \tilde{C_2} = 0.5/v_{i(1)} + 1/v_{i(2)} + 0.5/v_{i(3)} + \ldots + 0/v_{i(k-1)} + 0/v_{i(k)} \\ \vdots \\ \tilde{C_k} = 0/v_{i(1)} + 0/v_{i(2)} + 0/v_{i(3)} + \ldots + 0.5/v_{i(k-1)} + 1/v_{i(k)}. \end{cases} \tag{4}$$

Note: v_i shows the intervals midpoints $\tilde{C_1}$.

Stage 7: The linguistic term $\tilde{C_k}$ represented the fuzzy sets. The membership function is used to determine the fuzzy numbers (FN). In this algorithm 'S-shape' and 'Trapezoidal' membership functions are utilized using Matlab software (R12a). The S-shape membership function consists of two parameters "a" and "b" [24] as discussed in Equation (5).

$$\mu C(x) = \begin{cases} 0, & x < a; \\ 2((x-a)/(b-a))^2 & a \leq x \leq (a+b)/2); \\ 1-((x-a)/(b-a))^2 & (a+b)/2) \leq x \leq b); \\ 1, & x \geq b. \end{cases} \tag{5}$$

The fuzzy numbers $C_1, C_2, \ldots, C_{k-1}$, and C_k can be defined as $C_1 = (v_1, v_2)$, $C_2 = (v_2, v_3)$, ..., $C_k = (v_{k-1}, v_k)$.

The fuzzy number using the "S-shape" and "Trapezoidal" membership functions [25] are presented in Figure 1 in graphs (a) and (b).

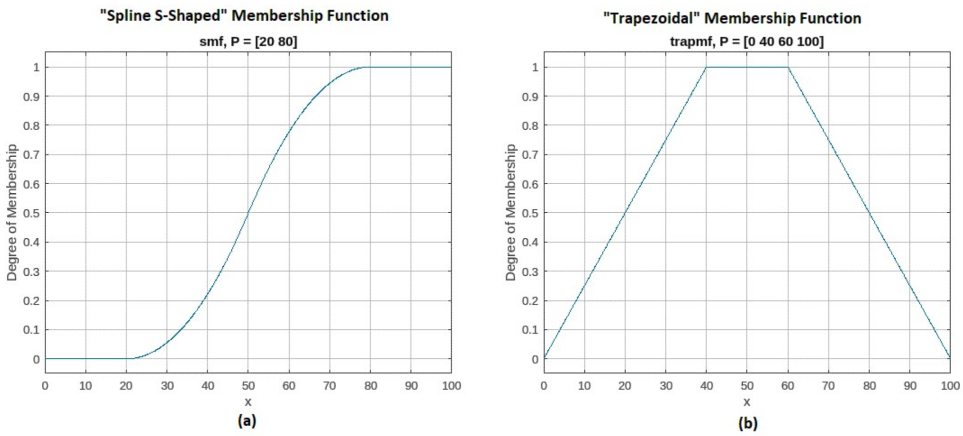

FIGURE 1. Fuzzy Numbers Presented using S-shaped and Trapezoidal Membership Function

Stage 8: The Fuzzy logical relationships (FLRs), there are three types of fuzzy logical relationships considered in this study which are as follows:

(1) First Order FLRs (FO-FLRs)
(2) Second Order FLRs (SO-FLRs)
(3) Third Order FLRs (TO-FLRs)

In FO-LRs, the relationship can be formulated as $\tilde{C_j} \rightarrow \tilde{C_k}$, where $\tilde{C_j}$ identifies the left-hand side (LHS) and $\tilde{C_k}$ denotes the right-hand side (RHS) of the logical relationship. Let us consider, C_{t-1} is a value at a time (t-1) denoted as $\tilde{C_j}$, and C_t is a value at a time (t) denoted as $\tilde{C_k}$. This relationship derives all the FLRs for all the fuzzified data.

In SO-FLRs, the relationship between fuzzified data can be formulated as $\tilde{C_i}$, $\tilde{C_j} \rightarrow \tilde{C_k}$ where $\tilde{C_i}$, and $\tilde{C_j}$ denote the LHS, and $\tilde{C_k}$ denotes the RHS of the logical relationship. Suppose C_{t-2} and C_{t-1} are the values at time (t-2) and (t-1) denoted as $\tilde{C_i}$, and $\tilde{C_j}$ and C_t is the value at time (t) then it is represented by $\tilde{C_k}$.

In TO-FLRs, the relationship between fuzzified data can be formulated as $\tilde{C_i}$, $\tilde{C_j}$, $\tilde{C_k} \rightarrow \tilde{C_l}$ where $\tilde{C_i}$, $\tilde{C_j}$, and $\tilde{C_k}$ are denoted the LHS and $\tilde{C_l}$ denotes the RHS of the logical relationship. Suppose C_{t-3}, C_{t-2}, and C_{t-1} are the values at a time $(t-3), (t-2)$ and $(t-1)$ denoted as $\tilde{C_i}$, $\tilde{C_j}$, $\tilde{C_k}$, and C_t is the value at a time (t) then it is represented by $\tilde{C_l}$.

Stage 9: Defuzzify the linguistic term by FLRs of various orders (FO, SO, and TO) and get the predicted values. $\tilde{C_i} \rightarrow \tilde{C_j}$, $\tilde{C_j}$ will be the forecasted value. In $\tilde{C_i}$, $\tilde{C_j} \rightarrow \tilde{C_k}$, $\tilde{C_k}$ will be the forecasted value. In $\tilde{C_i}$, $\tilde{C_j}$, $\tilde{C_k} \rightarrow \tilde{C_l}$, $\tilde{C_l}$ will be the predicted value. The RHS fuzzy numbers (FN) are the forecasted output. If LHS FN are similar then the average of the RHS Fuzzy numbers is the forecasted output.

Stage 10: Restore the composition data by inverse transformation method.

Stage 11: return $\hat{x} = [\hat{x}_1, \hat{x}_2, ..., \hat{x}_D]^T$.

Stage 12: Convert the percentage data from stage 11 into real data form.

Stage 13: For validation of the model, various accuracy measurements are applied.

The defuzzify value is denoted by "Z" using MATLAB software to defuzzify the predicted fuzzy number. The Centroid formula to defuzzify Fuzzy Numbers is presented in Equation (6).

$$Z = \frac{\Sigma_{i=1}^{n} \mu C(c_i) c_i}{\Sigma_{i=1}^{n} \mu C(c_i)} \tag{6}$$

To measure the accuracy between the actual data "x_i" and forecasting output "$\hat{x}$" in time series uses the mean average percentage error "MAPE" Equation (7), and root mean square error "RMSE" Equation (8).

$$MAPE = \frac{\Sigma_{i=1}^{n} \frac{|x_i - \hat{x}|}{|x_i|}}{n} \times 100 \tag{7}$$

$$RMSE = \sqrt{\frac{\Sigma_{i=1}^{n} (x_i - \hat{x})^2}{n}} \tag{8}$$

In compositional data, time series accuracy measurement is based on the simplex space. The compositional mean absolute percentage error "CMAPE" Equation (9) and the compositional root mean square error "CRMSE" Equation (10) accuracy measurements will be utilized following the Aitchison distance Wei et al. [18].

$$CMAPE = \frac{\Sigma_{i=1}^{n} d_A \frac{(x_i, \hat{x})}{||x_i||}}{N} \times 100 \tag{9}$$

$$CRMSE = \frac{\Sigma_{i=1}^{n} d_A(x_i, \hat{x})}{N} \tag{10}$$

2.3 Algorithm Structure: The proposed method algorithm of the High-Order Fuzzy Time Series model of the Compositional Data (CoDa) "HO-FTS-CoDa" is presented in Figure 2.

3. APPLICATION OF FUZZY TIME SERIES DATA ANALYSIS USING COMPOSITIONAL DATA

This research utilized the compositional data of the Stocks Traded in Malaysia, Indonesia, and Thailand from 2000 to 2022 as reported on the World Bank Group data website [23]. The share's value traded is the shares traded as a whole domestically, and internationally, multiplied by their respective matching prices. Data observations are single-counted (only a single side of the operation is measured) in billions of US dollars. Companies acknowledged to enter, and self-proclaimed to trading are involved in the data. Data observations are converted to U.S. dollars at the corresponding year-end foreign exchange rates. The High-Order Fuzzy Time Series Model of Compositional Data is implemented using a multivariate time series (MTS) data set related to the Stock Traded in Malaysia, Indonesia, and Thailand.

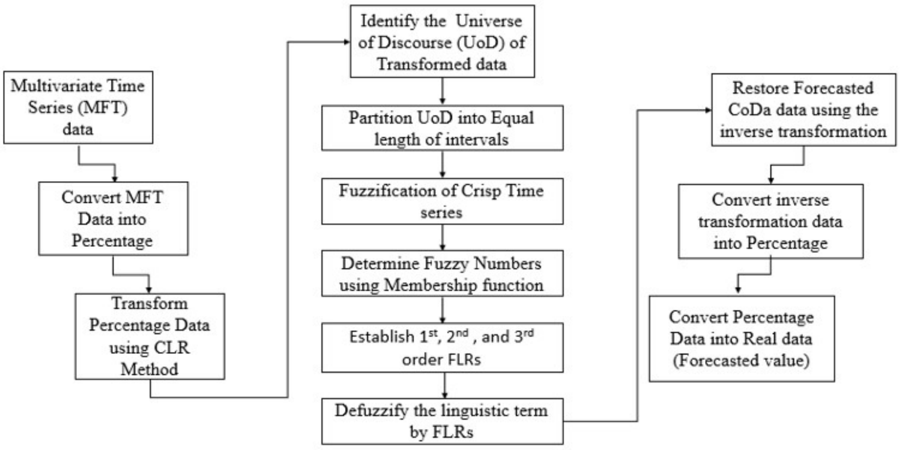

Figure 2. Algorithm of HO-FTS-CoDa model (Constructed by own)

Now apply the stages as described in the methodology. Stage 1, Multivariate time series (MTS) data of The Stock Traded. $S = [s_1, s_2, s_3]^T$ is collected. Stage 2, convert the MTS data into percentage form $x = [p_1, p_2, p_3]^T$. Stage 3, transform $x = [p_1, p_2, p_3]^T$ to $c = [c_1, c_2, c_3]^T$ by using the Centered-log-ratio (CLR) transformation method as discussed in Equation (1) (Aitchison 1982). Table 1 shows the actual and the Compositional data transformation using the central-log-ratio rule.

Table 1. Actual MTS Data, and the CoDa Data

Year	Actual MTS Data			CoDa Data Using CLR		
	Indonesia	Malaysia	Thailand	Indonesia	Malaysia	Thailand
2000	12.902	52.512	19.331	-0.603	0.801	-0.198
2001	9.638	21.033	31.051	-0.650	0.130	0.520
2002	12.388	25.728	41.280	-0.645	0.086	0.559
2003	12.760	45.712	104.550	-1.126	0.150	0.977
2004	21.879	54.191	116.937	-0.861	0.046	0.815
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
2020	131.145	248.608	481.299	-0.647	-0.007	0.654
2021	202.772	204.379	597.944	-0.363	-0.355	0.718
2022	194.860	110.455	478.205	-0.110	-0.678	0.788

Stage 4: Determine the Universe of Discourse (UoD) for each variable in the Compositional data form. For Indonesia UoD, is U = [-1.13, -0.1], for Malaysia UoD, is U = [-0.68, 0.801], and for Thailand UoD, U = [-0.2, 0.98] using Equation (7). Stage 5, determines the suitable length *"l"* Huarng [9] and the number of intervals (K) for Compositional Data for each stock of Indonesia, Malaysia, and Thailand are presented in Table 2 using Equation (8).

TABLE 2. Length of Interval "l" and "K" Intervals for Each Variable

Average-Based Length Approach	For Indonesia	For Malaysia	For Thailand
The average of the FAD between $Y_{(t-i)}$ and $Y_{(t)}$.	0.1486	0.17609	0.169
The half of the average of the FAD	0.0743	0.0880	0.0845
Table base values using the Huarng [9] approach	Less than 1	Less than 1	Less than 1
Length of interval "l" using table base value	0.1	0.1	0.1
The number of intervals "K"	K1 = 11	K2 = 15	K3=12
Average of K1, K2, and K3	K = 13	K = 13	K = 13
	-1.13 – -1.05	-0.68 – -0.56	-0.200 – -0.109
	-1.05 – -0.97	-0.56 – -0.44	-0.109 – -0.018
	-0.97 – -0.89	-0.44 – -0.32	-0.018 – 0.073
	-0.89 – -0.81	-0.32 – -0.20	0.073 – 0.164
	-0.81 – -0.73	-0.20 – -0.08	0.164 – 0.255
	.	.	.
	.	.	.
	.	.	.
	-0.25 – -0.17	0.64 – 0.76	0.801 – 0.892
	-0.17 – -0.09	0.76 – 0.88	0.892 – 0.983

Stage 6: The discrete fuzzy sets $\tilde{C}_{(i,1)}$, $\tilde{C}_{(i,2)}$, . . . , $\tilde{C}_{(i,13)}$ are defined for each variable. For Indonesia Stocks fuzzy sets are $\tilde{C}_{(1,1)}$, $\tilde{C}_{(1,2)}$ to $\tilde{C}_{(1,13)}$. For Malaysia Stocks fuzzy sets are $\tilde{C}_{(2,1)}$, $\tilde{C}_{(2,2)}$ to $\tilde{C}_{(2,13)}$. For Thailand Stocks fuzzy sets are $\tilde{C}_{(3,1)}$, $\tilde{C}_{(3,2)}$ to $\tilde{C}_{(3,13)}$. Stage 7, the fuzzy numbers are determined based on the linguistic terms fuzzy sets using the "S-shaped membership function" and the "Trapezoidal" membership function presented in Table 3. For the "S-shaped" membership function the Fuzzy numbers for Indonesia from $C_{(1,1)}$ = (-1.13, -1.05) to $C_{(1,13)}$ = (-0.17, -0.09), for Malaysia from $C_{(2,1)}$ = (-0.68, -0.56) to $C_{(2,13)}$ = (0.76, 0.88), and for Thailand from $C_{(3,1)}$ = (-0.20, -0.109) to $C_{(3,13)}$ = (0.892, 0.983). For the "Trapezoidal" membership function the Fuzzy numbers for Indonesia are $C_{(1,1)}$ = (-1.21, -1.13, -1.05, -0.97) to $C_{(1,13)}$ = (-0.25, -0.17, -0.09, -0.01), for Malaysia from $C_{(2,1)}$ = (-0.76, -0.68, -0.56, -0.44) to $C_{(2,13)}$ = (0.64, 0.76, 0.88, 1.00), and for Thailand the Fuzzy numbers are from $C_{(3,1)}$ = (-0.291,-0.200, -0.109, -0.018) to $C_{(3,13)}$ = (0.801, 0.892, 0.983, 1.074).

TABLE 3. Fuzzy Numbers of Compositional Data

Year	Compositional Data using CLR Rule			Fuzzy Numbers of Compositional Data		
	Indonesia	Malaysia	Thailand	Indonesia	Malaysia	Thailand
2000	-0.603	0.801	-0.198	C7	C13	C1
2001	-0.650	0.130	0.520	C6	C7	C8
2002	-0.645	0.086	0.559	C7	C7	C9
2003	-1.126	0.150	0.977	C1	C7	C13
2004	-0.861	0.046	0.815	C4	C7	C12
2005	-0.547	-0.076	0.623	C8	C6	C10
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
2021	-0.363	-0.355	0.718	C10	C3	C11
2022	-0.110	-0.678	0.788	C13	C1	C11

In the year 2002, the CoDa value of Indonesia is -0.645 as presented in Table 1. The Fuzzy Number for Indonesia has an interval from -0.73 to -0.65 which is denoted by C_7. Therefore Fuzzy Number for the year 2002 is C_7. Similarly, all

Fuzzy Numbers for each variable were determined using this procedure. Stage 8, the Fuzzy logical relationships (FLRs) and groups at various orders using fuzzified CoDa data Type-1 FTS models are presented in Table 4, Table 5, and Table 6 with respective to First-Order, Second-Order, and Third-Order. Stage 9, defuzzify the linguistic term by FLRs of various orders using MATLAB, using Equation (10).

TABLE 4. First-Order Fuzzy Logical Relationships and Groups

First Order Fuzzy Logical Relationships		
For Indonesia	For Malaysia	For Thailand
$C_7 \to C_6$	$C_{13} \to C_7$	$C_1 \to C_8$
$C_6 \to C_7$	$C_7 \to C_7$	$C_8 \to C_9$
$C_7 \to C_1$	$C_7 \to C_7$	$C_9 \to C_1$
.	.	.
.	.	.
.	.	.
$C_7 \to C_{10}$	$C_6 \to C_3$	$C_{10} \to C_{11}$
$C_{10} \to C_{13}$	$C_3 \to C_1$	$C_{11} \to C_{13}$
First Order FLR Groups		
For Indonesia	For Malaysia	For Thailand
$C_7 \to C_6, C_1, C_{10}$	$C_{13} \to C_7$	$C_1 \to C_8$
$C_6 \to C_7$	$C_7 \to 3C_7,\ C_6,\ C_8$	$C_8 \to C_9,\ C_7$
.	.	.
.	.	.
.	.	.
$C_{10} \to C_7,\ C_{13}$	$C_4 \to C_3$	$C_{11} \to 2C_{10},\ 3C_{11},\ C_{12}$

TABLE 5. Second-Order Fuzzy Logical Relationships and Groups

Second Order Fuzzy Logical Relationships		
For Indonesia	For Malaysia	For Thailand
$C_7, C_6 \to C_7$	$C_{13}, C_7 \to C_7$	$C_1, C_8 \to C_9$
$C_6, C_7 \to C_1$	$C_7, C_7 \to C_7$	$C_8, C_9 \to C_{13}$
$C_7, C_1 \to C_4$	$C_7, C_7 \to C_7$	$C_9, C_{13} \to C_{12}$
.	.	.
.	.	.
.	.	.
$C_{10}, C_7 \to C_{10}$	$C_3, C_6 \to C_3$	$C_{11}, C_{10} \to C_{11}$
$C_7, C_{10} \to C_{13}$	$C_6, C_3 \to C_1$	$C_{10}, C_{11} \to C_{11}$
Second Order FLR Groups		
For Indonesia	For Malaysia	For Thailand
$C_7, C_6 \to C_7$	$C_{13}, C_7 \to C_7$	$C_1, C_8 \to C_9$
$C_6, C_7 \to C_1$	$C_7, C_7 \to 2C_7, C_6$	$C_8, C_9 \to C_{13}$
.	.	.
.	.	.
.	.	.
$C_{13}, C_{13} \to C_{11}$	$C_6, C_3 \to C_1$	$C_{11}, C_{11} \to C_{11}$

In stage 10, the composition data is restored using the inverse transformation method. Stage 11, convert inverse transformation data into percentage form. Stage 12, the forecasted output for each variable of the Stocks Traded using the S-shaped membership function presented in Table 7, and the Trapezoidal membership function is presented in Table 8 using various orders of Fuzzy Time Series models.

Table 6. Third-Order Fuzzy Logical Relationships and Groups

Third-Order Fuzzy Logical Relationships		
For Indonesia	For Malaysia	For Thailand
$C_7, C_6, C_7 \to C_1$	$C_{13}, C_7, C_7 \to C_7$	$C_1, C_8, C_9 \to C_{13}$
$C_6, C_7, C_1 \to C_4$	$C_7, C_7, C_7 \to C_7$	$C_8, C_9, C_{13} \to C_{12}$
$C_7, C_1 \to C_4$	$C_7, C_7, C_7 \to C_6$	$C_9, C_{13}, C_{12} \to C_{10}$
.	.	.
.	.	.
.	.	.
$C_8, C_{10}, C_7 \to C_{10}$	$C_4, C_3, C_6 \to C_3$	$C_{11}, C_{11}, C_{10} \to C_{11}$
$C_{10}, C_7, C_{10} \to C_{13}$	$C_3, C_6, C_3 \to C_1$	$C_{11}, C_{10}, C_{11} \to C_{11}$
Third-Order FLR Groups		
For Indonesia	For Malaysia	For Thailand
$C_7, C_6, C_7 \to C_1$	$C_{13}, C_7, C_7 \to C_7$	$C_1, C_8, C_9 \to C_{13}$
$C_6, C_7, C_1 \to C_4$	$C_7, C_7, C_7 \to C_7, C_6$	$C_8, C_9, C_{13} \to C_{12}$
.	.	.
.	.	.
.	.	.
$C_{10}, C_7, C_{10} \to C_{13}$	$C_3, C_6, C_3 \to C_1$	$C_{11}, C_{10}, C_{11} \to C_{11}$

4. RESULTS AND DISCUSSION

The proposed High-Order Fuzzy time series using the Compositional data (CoDa) approach with central-log-ratio (CLR) transformation is employed on the entire multivariate data of Stocks Traded in Indonesia, Malaysia, and Thailand from 2000 to 2022. The forecasted Multivariate Time Series data set of each country Stocks Traded are analyzed by "S-shaped", and "Trapezoidal" membership functions with various orders of fuzzy logic relationships. The forecasted results in Table 7, and Table 8 indicate that the Stock Traded has continuously grown since 2000. Besides growing trends in these countries, Thailand Stock Traded is leading ahead of Malaysia, and Indonesia. The comparison and assessment of the HO-FTS-CoDa models between actual and forecasted output using CoDa data and real data sets are presented graphically in Figure 3, Figure 4, and Figure 5.

Table 7. Forecasted Output of Real Data Using "S-Shaped" Membership Function

Year	First-Order FTS			Second-Order FTS			Third-Order FTS		
	Indonesia	Malaysia	Thailand	Indonesia	Malaysia	Thailand	Indonesia	Malaysia	Thailand
2000									
2001	9.289	21.446	30.986						
2002	13.073	27.128	39.195	12.486	25.909	41.001			
2003	20.104	46.413	96.505	13.747	44.292	104.983	13.596	45.594	103.832
2004	21.156	55.809	116.041	21.399	54.236	117.372	21.519	53.460	118.029
2005	25.402	48.658	88.263	26.349	48.493	87.481	26.506	47.816	88.002
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
2021	151.380	274.914	578.802	196.738	198.715	609.643	196.738	198.715	609.643
2022	144.680	197.260	441.579	187.623	117.264	478.632	192.863	120.540	470.116

These graphs unveil the stock traded forecasted pattern analysis. These patterns display less deviation in the Trapezoidal membership function in comparison of the S-shaped membership function utilizing various orders of Fuzzy Time Series

Table 8. Forecasted Output of Real Data Using "Trapezoidal" Membership Function

Year	First-Order FTS			Second-Order FTS			Third-Order FTS		
	Indonesia	Malaysia	Thailand	Indonesia	Malaysia	Thailand	Indonesia	Malaysia	Thailand
2000									
2001	9.380	21.226	31.116						
2002	13.198	26.845	39.353	12.603	25.635	41.159			
2003	20.284	45.902	96.836	13.870	43.805	105.346	13.720	45.098	104.204
2004	21.349	55.203	116.455	21.591	53.640	117.776	21.711	52.868	118.428
2005	25.630	48.124	88.569	26.585	47.958	87.780	26.741	47.285	88.297
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
2021	152.689	271.801	580.605	198.173	196.201	610.722	198.173	196.201	610.722
2022	152.091	169.776	461.652	188.815	115.673	479.031	194.093	118.906	470.519

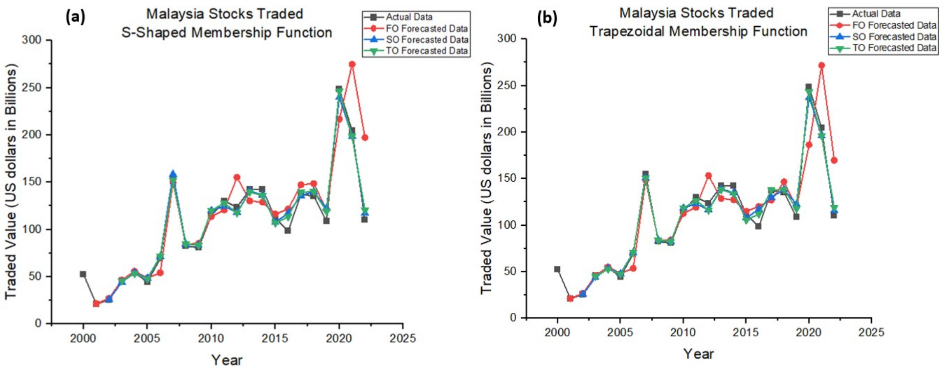

Figure 3. Actual and Forecasted Output of Malaysia Stocks Traded using FO, SO, and TO FTS

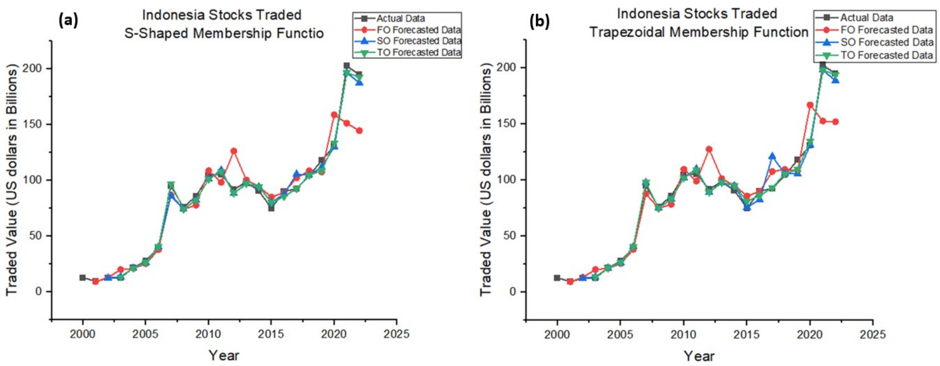

Figure 4. Actual and Forecasted Output of Indonesia Stocks Traded using FO, SO, and TO FTS

models. These figures also show that the predicted output of proposed Third-Order

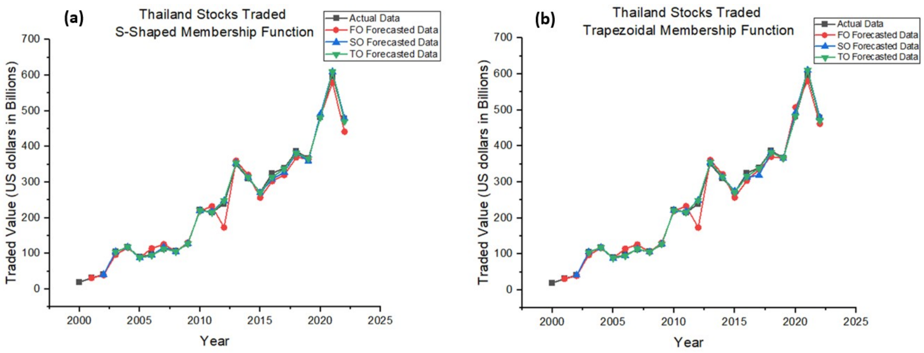

FIGURE 5. Actual and Forecasted Output of Thailand Stocks Traded using FO, SO, and TO FTS

FTS models using the Trapezoidal membership function is superimposed over the actual data.

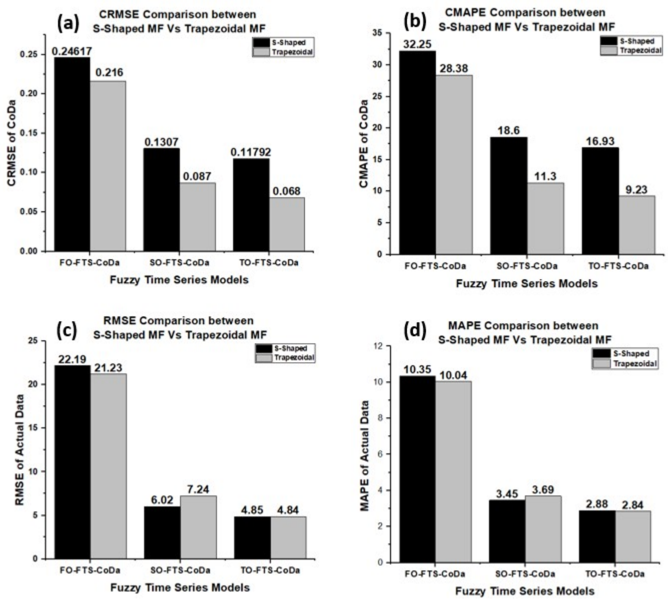

FIGURE 6. CRMSE, CMAPE, RMSE. and MAPE Comparison using FO, SO, and TO FTS Models of S-Shaped, and Trapezoidal Membership Functions

The compositional RMSE, and MAPE accuracy measurements helped us to conclude forecasted performance and identify the membership function's suitability for the forecasted output of each Stocks Traded value. We observed that CRMSE and CMAPE accuracy measurements using Actual data as well as compositional data with central-log-ratio transformation decreases as the order of the Fuzzy Time Series increases, as presented in Table 9 and presented in Figure 6.

TABLE 9. Performance Analysis of Proposed Fuzzy Time Series Models using Membership Functions

	"S-Shaped" Membership Function				"Trapezoidal" Membership Function			
	Compositional Data		Actual Data		Compositional Data		Actual Data	
	CRMSE	CMAPE	RMSE	MAPE	CRMSE	CMAPE	RMSE	MAPE
FO-FTS-CoDa	0.246	32.25%	22.19	10.35%	0.216	28.38%	21.23	10.04%
SO-FTS-CoDa	0.130	18.60%	6.02	3.45%	0.087	11.31%	7.24	3.69%
TO-FTS-CoDa	0.117	16.93%	4.85	2.88%	**0.068**	**9.23%**	**4.84**	**2.84%**

The overall results provide concrete evidence of a robust third-order FTS model with suitable membership, the Trapezoidal membership, in contrast to the S-shaped membership function. Therefore, the Third-Order Fuzzy Time Series model using compositional data with a "Trapezoidal" membership function outclasses other models used in this study. In Figure 6, (a), (b), (c), and (d) graphs show a comparison of CRMSE, CMAPE, RMSE, and MAPE accuracy measurements of FO, SO, and TO FTS CoDa models. These accuracy measurements informed that Second-Order FTS models predicted better results than the First-Order, and the Third-Order FTS models predicted better results than the First-Order, and the Second-Order FTS models using Coda Data.

The interesting facts that are important for the Fuzzy Time Series are achieved in this study using a compositional data approach are classified as:

(1) Determine the number of intervals for each time series data separately. Take the average of these number of intervals for the proposed model.
(2) The order of Fuzzy Logical Relationship is an important factor for Fuzzy Time Series which is identified as the Third-Order Fuzzy Logical Relationship is suitable for the proposed model.
(3) The analysis results suggested that the Trapezoidal membership function performed better as compared to the S-shaped membership function.

5. CONCLUSION

The forecasted output of the Stocks Traded in Malaysia, Indonesia, and Thailand using the compositional data demonstrates outstanding developments from the earlier Fuzzy Time Series approaches. The constraint limit to 0 to 1 property of compositional data keeps the number of intervals up to an optimal level. On the other hand, the number of intervals would be in the hundreds and very hard to present in Fuzzy sets and Fuzzy numbers. The comparison of "S-shaped", and "Trapezoidal" membership functions with various orders of FTS models using compositional data with central-log-ratio transformation finalized that Third-Order FTS Compositional Data with Trapezoidal membership function model performed outclasses as compared to other combinations based on accuracy measurements. The limitation of this research paper is that it applies Composition data over a small number of observations in multivariate time series data. In future research, the researcher may use different numbers of intervals of the universe of discourse such as 5, 10, 15, and 20. Various other membership functions would be used such as Gaussian membership function, Z-Shaped, or others to achieve high accuracy in the Fuzzy Time Series model. This study concluded that the Third-Order Fuzzy Time Series Compositional Data (TO-FTS-CoDa) model using the Trapezoidal membership function achieved the highest accuracy among the models evaluated.

Acknowledgement. We appreciate the editor's valuable feedback, which greatly enhanced the clarity, and coherence of the paper.

REFERENCES

[1] C. Zhao, X. Yuan, J. Long, L. Jin, and B. Guan, "Financial indicators analysis using machine learning: Evidence from chinese stock market," *Finance Research Letters*, vol. 58, p. 104590, 2023. https://doi.org/10.1016/j.frl.2023.104590.

[2] G. Liu, K. Zhong, H. Li, T. Chen, and Y. Wang, "A state of art review on time series forecasting with machine learning for environmental parameters in agricultural greenhouses," *Information Processing in Agriculture*, vol. 11, no. 2, pp. 143–162, 2024. https://doi.org/10.1016/j.inpa.2022.10.005.

[3] H.-W. Chang, T. Chang, and M.-C. Wang, "Revisit the impact of exchange rate on stock market returns during the pandemic period," *The North American Journal of Economics and Finance*, vol. 70, p. 102068, 2024. https://doi.org/10.1016/j.najef.2023.102068.

[4] J. Aitchison, "The statistical analysis of compositional data," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 44, no. 2, pp. 139–160, 1982. https://doi.org/10.1111/j.2517-6161.1982.tb01195.x.

[5] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning—i," *Information sciences*, vol. 8, no. 3, pp. 199–249, 1975. https://doi.org/10.1016/0020-0255(75)90046-8.

[6] S. J. Russell and P. Norvig, *Artificial intelligence: a modern approach.* pearson, 2016.

[7] Q. Song and B. S. Chissom, "Forecasting enrollments with fuzzy time series—part i," *Fuzzy sets and systems*, vol. 54, no. 1, pp. 1–9, 1993. https://doi.org/10.1016/0165-0114(93)90355-L.

[8] S.-M. Chen, "Forecasting enrollments based on high-order fuzzy time series," *Cybernetics and Systems*, vol. 33, no. 1, pp. 1–16, 2002. https://doi.org/10.1080/019697202753306479.

[9] K. Huarng, "Effective lengths of intervals to improve forecasting in fuzzy time series," *Fuzzy sets and systems*, vol. 123, no. 3, pp. 387–394, 2001. https://doi.org/10.1016/S0165-0114(00)00057-9.

[10] J. Wang, H. Li, H. Yang, and Y. Wang, "Intelligent multivariable air-quality forecasting system based on feature selection and modified evolving interval type-2 quantum fuzzy neural network," *Environmental Pollution*, vol. 274, p. 116429, 2021. https://doi.org/10.1016/j.envpol.2021.116429.

[11] Y. Alyousifi, M. Othman, R. Sokkalingam, I. Faye, and P. C. Silva, "Predicting daily air pollution index based on fuzzy time series markov chain model," *Symmetry*, vol. 12, no. 2, p. 293, 2020. https://doi.org/10.3390/sym12020293.

[12] Y. Gu, Y. Zhao, J. Zhou, H. Li, and Y. Wang, "A fuzzy multiple linear regression model based on meteorological factors for air quality index forecast," *Journal of Intelligent & Fuzzy Systems*, vol. 40, no. 6, pp. 10523–10547, 2021. `https://doi.org/10.3233/JIFS-201222`.

[13] N. S. W. YUSOFF, "A new single linkage robust clustering outlier detection procedures for multivariate data," *Sains Malaysiana*, vol. 52, no. 8, pp. 2431–2451, 2023. `https://doi.org/10.17576/jsm-2023-5208-19`.

[14] X. Shi, J. Wang, and B. Zhang, "A fuzzy time series forecasting model with both accuracy and interpretability is used to forecast wind power," *Applied Energy*, vol. 353, p. 122015, 2024. `https://doi.org/10.1016/j.apenergy.2023.122015`.

[15] Á. López-Oriona, P. D'Urso, J. A. Vilar, and B. Lafuente-Rego, "Quantile-based fuzzy c-means clustering of multivariate time series: Robust techniques," *International Journal of Approximate Reasoning*, vol. 150, pp. 55–82, 2022. `https://doi.org/10.1016/S0165-0114(00)00057-9`.

[16] T. Carvalho, M. Vellasco, and J. F. Amaral, "Automatic generation of fuzzy inference systems for multivariate time series forecasting," *Fuzzy Sets and Systems*, vol. 470, p. 108657, 2023. `https://doi.org/10.1016/j.fss.2023.108657`.

[17] D. Qin, Z. Peng, and L. Wu, "Deep attention fuzzy cognitive maps for interpretable multivariate time series prediction," *Knowledge-Based Systems*, vol. 275, p. 110700, 2023. `https://doi.org/10.1016/j.knosys.2023.110700`.

[18] Y. Wei, Z. Wang, H. Wang, and Y. Li, "Compositional data techniques for forecasting dynamic change in china's energy consumption structure by 2020 and 2030," *Journal of Cleaner Production*, vol. 284, p. 124702, 2021. `https://doi.org/10.1016/j.jclepro.2020.124702`.

[19] V. Pawlowsky-Glahn and J. J. Egozcue, "Compositional data and their analysis: an introduction," *Geological Society, London, Special Publications*, vol. 264, no. 1, pp. 1–10, 2006. `https://doi.org/10.1144/gsl.sp.2006.264.01.01`.

[20] J. Palarea-Albaladejo, J. Martín-Fernández, and J. Soto, "Dealing with distances and transformations for fuzzy c-means clustering of compositional data," *Journal of Classification*, vol. 29, pp. 144–169, 2012. `https://doi.org/10.1007/s00357-012-9105-4`.

[21] K. Kim, J. Park, and S. Jung, "Principal component analysis for zero-inflated compositional data," *Computational Statistics & Data Analysis*, vol. 198, p. 107989, 2024. `https://doi.org/10.1016/j.csda.2024.107989`.

[22] H. Huang, Y. Tian, and Z. Tao, "Multi-rule combination prediction of compositional data time series based on multivariate fuzzy time series model and its application," *Expert Systems with Applications*, vol. 238, p. 121966, 2024. `https://doi.org/10.1016/j.eswa.2023.121966`.

[23] T. W. Bank, "Stocks traded, total value (current us$) - indonesia, malaysia, thailand," 2022. Accessed: 2024-08-15.

[24] MathWorks, "S-shaped membership function (smf)," 2023. Accessed: 2024-08-15.

[25] MathWorks, "Trapezoidal membership function (trapezoidalmf)," 2023. Accessed: 2024-08-15.