https://doi.org/10.48047/AFJBS.6.12.2024.4323-4329

African Journal of Biological Sciences

Journal homepage: http://www.afjbs.com

Research Paper Open Access

A Standard Deviation Method of Fuzzy Time Series Forecasting using Coal Production Data based on Trapezoidal Fuzzification

[1]S.Imalin, [2]V. Anithakumari, [3]V.M. Arul Flower Mary

[1]Research Scholar, Register No.21113042092011 Department of Mathematics, Holy Cross College (Autonomous), Nagercoil - 629004

[2]Assistant Professor, Department of Statistics, Government Arts and Science College, Nagercoil-629004

[3]Associate Professor, Department of Mathematics, Holy Cross College (Autonomous), Nagercoil – 629004.

[1]*imalinjas25@gmail.com*, [2]*anithakumari_v@yahoo.co.in*, [3]*arulflowermary@gmail.com*

Affiliated to ManonmaniamSundranar University, Abishekapatti, Tirunelveli-627012, Tamilnadu, India.

Article History
Volume 6, Issue 12, 2024
Received: 20 June 2024
Accepted: 10 July 2024
doi:
10.48047/AFJBS.6.12.2024.4323-4329

Abstract

In current decades, a number of forecasting models based on the ideas of fuzzy time series have been put forth. These models have been widely used to address a general range of problem domains, particularly those involving forecasting when the historical data are linguistic values. The development of fuzzy time seriesand its application to coal production forecasts are presented in this study. It includes an analysis of the Standard Deviation Method along with estimates conducted to predict India's coal production. The 43 years of historical data serve as the foundation for the Fuzzy time series forecasting. Error estimates are used to analyse this model. The suggested approach is used to forecast data on coal production.

Keywords: Standard Deviation Method, Linguistic Values, Coal production, Fuzzy time series, Forecasting.

1.Introduction

A time series is a sequence of discrete temporal data that is uniformly spaced. It may include some or all of the following components: trend, cyclical, seasonal, and irregular. A trend is a long-term pattern, whereas a cyclical is a sequence of up and down movements.

Seasonal, on the other hand, is a consistent change within the same month or quarter, whereas irregular is an unexplained random component.Forecasting is the prediction of future values based on previous and current time series data trends. Short-term forecasting (STF), medium-term forecasting (MTF), and long-term forecasting (LTF) are the three types of forecasting.

Zadeh firstly proposed the concept of fuzzy set theory in 1965 [7]. Song and Chissom, based on Zadeh's work, were the first to apply the concept of the fuzzy set to time series and build a first-order time-invariant Fuzzy time series model in 1993. Song and Chissom[4] later applied time-invariant and time-variant FTS on Alabama University enrollment from 1971 to 1992.

2.Basic Definitions

Definition 1:

Let $U = \{u_1, u_2, \ldots \ldots, u_n\}$ be a universe of discourse (universal set); a fuzzy set A of U is defined $A = f_A(u_1)/u_1 f_A(u_2)/u_2 \ldots\ldots\ldots\ldots f_A(u_n)/u_n$, where f_A is a membership function of a given set A, $f_A : U \rightarrow [0,1]$.

Definition 2 [4]:

Let $\mathrm{Y(t)}, (\mathrm{t} = \cdots, 0,1,2, \ldots)$, a subset of U, be the universe of discourse defined by the fuzzy set $\mu_i(t)$. If $\mathrm{F(t)}$ consists of $\mu_i(t)$ (i=1,2,3,…), $\mathrm{F(t)}$ is called a fuzzy time series on $\mathrm{Y(t)}$.

Definition 3:

If there exist a fuzzy relationship $\mathrm{R\,(t-1,t)}$, such that $\mathrm{F(t) = F(t-1)^\circ\, R\,(t-1,t)}$ where $^\circ$ is an arithmetic operator, then $\mathrm{F(t)}$ is said to be caused by $\mathrm{F(t-1)}$. The relationship between $\mathrm{F(t)}$ and $\mathrm{F(t-1)}$ can be denoted by $\mathrm{F(t-1) \rightarrow F(t)}$.

Definition 4:

Suppose F (t) is calculated by F (t-1) only, and $\mathrm{F(t) = F(t-1)^\circ\, R\,(t-1,t)}$. For any t, if $\mathrm{R\,(t-1,t)}$ is independent of t, then F (t) is considered a time invariant fuzzy time series. Otherwise, F (t) is time – variant.

Definition 5 [6]:

Suppose $\mathrm{F(t-1)} = A_i$ and F (t) $= A_j$, a fuzzy logical relationship can be defined as $A_i \rightarrow A_j$ Where A_i and A_j are called the left-hand side and the right-hand side of the fuzzy

Definition 6: A membership function for a fuzzy set A on the universe of discourse X is defined as μ_A:X→ [0,1],where each element of X is mapped to a value between 0 and 1.This value, called membership value or degree of membership, quantifies the grade of membership of the element in X to the fuzzy set A.

Trapezoidal Fuzzy Number:

A Trapezoidal fuzzy number (TrFN) [3] denoted by $\check{A}$ is defined as (a_1, a_2, a_3, a_4) where the membership function

$$\mu_{\check{A}}(y) = \begin{cases} 0 & y_i < a \\ \dfrac{y_i - a}{b - a}, & a \le y_i \le b \\ 1 & b \le y_i \le c \\ \dfrac{d - y_i}{d - c}, & c \le y_i \le d \\ 0 & y_i > d \end{cases}$$

Proposed Method: Algorithm

- ❖ Separate the actual values.
- ❖ Evaluate the interval using standard deviation.
- ❖ Define the universe of discourse.
- ❖ To Find the membership degree.
- ❖ Define fuzzy set.
- ❖ To calculate the FLRG.
- ❖ Evaluate the Forecasted values.

Table1: Estimation of Actual values and Range values of the Coal Production Data

Year	Actual values	Sort	R	$\|R-AR\|^2$	Year	Actual values	Sort	R	$\|R-AR\|^2$
1980	113.9	113.9	-	-	2002	341.3	341.3	13.5	0.3594
1981	124.2	124.2	10.3	16.8059	2003	361.3	361.3	20	34.8159
1982	130.5	130.5	6.3	60.8322	2004	382.6	382.6	21.3	51.8472
1983	138.2	138.2	7.7	40.9536	2005	407	407	24.4	106.1003
1984	147.4	147.4	9.2	24.0051	2006	430.8	430.8	23.8	94.0997
1985	154.2	154.2	6.8	53.2827	2007	457.1	457.1	26.3	148.8522
1986	165.8	165.8	11.6	6.2475	2008	492.8	492.8	35.7	466.5816
1987	179.7	179.7	13.9	0.0393	2009	532	532	39.2	630.0351
1988	194.6	194.6	14.9	0.6408	2010	532.7	532.7	0.7	179.5466
1989	200.9	200.9	6.3	60.8322	2011	540	540	7.3	46.2332
1990	211.7	211.7	10.9	10.2368	2012	556.4	556.4	16.4	5.2923
1991	229.3	229.3	17.6	12.2535	2013	565.8	565.8	8.4	32.4843
1992	238.3	238.3	9	5.0995	2014	609.2	609.2	43.5	864.3894
1993	246	246	7.7	40.9536	2015	639.2	639.2	30	252.8259
1994	253.8	253.8	7.8	39.6837	2016	657.8	657.8	18.6	20.2545
1995	270.1	270.1	16.3	4.8422	2017	675.4	675.4	17.6	12.2535
1996	285.7	285.7	15.6	2.2515	2018	728.7	728.7	53.3	1536.6792
1997	295.9	292.3	10.2	15.2061	2019	730.8	730.8	2.1	143.9880
1998	292.3	295.9	3.6	110.2395	2020	756.494	756.494	25.694	134.4324
1999	300	300	7.7	40.9536	2021	778.21	778.21	21.716	58.0110
2000	309.6	309.6	9.6	20.2455	2022	893.08	893.08	114.87	10154.6936
2001	327.8	327.8	18.2	16.8141	-	-	-	-	-

Table 2. Trapezoidal Fuzzy Numbers (a, b, c, d)

Fuzzy set	**A**	**b**	**C**	**d**
A_1	99.8005	113.9	127.9995	142.099
A_2	127.9995	142.099	156.1985	170.298
A_3	156.1985	170.298	184.3975	198.497
A_4	184.3975	198.497	212.5965	226.696
A_5	212.5965	226.696	240.7955	254.895
A_6	240.7955	254.895	268.9955	283.094
A_7	268.9955	283.094	297.1935	311.293
A_8	297.1935	311.293	325.3925	339.492
A_9	325.3925	339.492	353.5915	367.691
A_{10}	353.5915	367.691	381.7905	395.81
A_{11}	381.7905	395.81	409.9895	424.089
A_{12}	409.9895	424.089	438.1855	452.288
A_{13}	438.1855	452.288	466.3875	480.487
A_{14}	466.3875	480.487	494.5865	508.686
A_{15}	494.5865	508.686	522.7855	536.8805
A_{16}	522.7855	536.8805	550.98	565.0795
A_{17}	550.98	565.0795	579.179	593.2785
A_{18}	579.179	593.2785	607.378	621.4775
A_{19}	607.378	621.4775	635.5725	649.672
A_{20}	635.5725	649.672	663.7715	677.871
A_{21}	663.7715	677.871	691.9705	706.07
A_{22}	691.9705	706.07	720.1695	734.269
A_{23}	720.1695	734.269	748.3685	762.468
A_{24}	748.3685	762.468	776.5675	790.667
A_{25}	776.5675	790.667	804.7665	818.866
A_{26}	804.7665	818.866	832.9655	847.065
A_{27}	832.9655	847.065	861.1645	875.264
A_{28}	861.1645	875.264	889.3635	903.463
A_{29}	889.3635	903.463	917.5625	931.662

Table 2 Shows that the Fuzzy sets of this proposed method.

Table 3: Fuzzy Logical Relationship Groups

Year	Actual Data	Fuzzified	FLRG	Year	Actual Data	Fuzzified	FLRG
1980	113.9	A_1	-	2002	341.3	A_9	$A_8 \to A_8, A_9$
1981	124.2	A_1	$A_1 \to A_1, A_2$	2003	361.3	A_{10}	$A_9 \to A_{10}$
1982	130.5	A_1	$A_1 \to A_1, A_2$	2004	382.6	A_{10}	$A_{10} \to A_{10}, A_{11}$
1983	138.2	A_2	$A_1 \to A_1, A_2$	2005	407	A_{11}	$A_{10} \to A_{10}, A_{11}$
1984	147.4	A_2	$A_2 \to A_2, A_3$	2006	430.8	A_{12}	$A_{11} \to A_{12}$
1985	154.2	A_2	$A_2 \to A_2, A_3$	2007	457.1	A_{13}	$A_{12} \to A_{13}$
1986	165.8	A_3	$A_2 \to A_2, A_3$	2008	492.8	A_{14}	$A_{13} \to A_{14}$
1987	179.7	A_3	$A_3 \to A_3, A_4$	2009	532	A_{16}	$A_{14} \to A_{16}$
1988	194.6	A_4	$A_3 \to A_3, A_4$	2010	532.7	A_{16}	$A_{16} \to A_{16}, A_{17}$
1989	200.9	A_4	$A_4 \to A_4, A_5$	2011	540	A_{16}	$A_{16} \to A_{16}, A_{17}$
1990	211.7	A_4	$A_4 \to A_4, A_5$	2012	556.4	A_{17}	$A_{16} \to A_{16}, A_{17}$
1991	229.3	A_5	$A_4 \to A_4, A_5$	2013	565.8	A_{17}	$A_{17} \to A_{17}, A_{18}$
1992	238.3	A_5	$A_5 \to A_5, A_6$	2014	609.2	A_{18}	$A_{17} \to A_{17}, A_{18}$
1993	246	A_5	$A_5 \to A_5, A_6$	2015	639.2	A_{19}	$A_{18} \to A_{19}$
1994	253.8	A_6	$A_5 \to A_5, A_6$	2016	657.8	A_{20}	$A_{19} \to A_{20}$
1995	270.1	A_6	$A_6 \to A_6, A_7$	2017	675.4	A_{21}	$A_{20} \to A_{21}$
1996	285.7	A_7	$A_6 \to A_6, A_7$	2018	728.7	A_{23}	$A_{21} \to A_{23}$
1997	295.9	A_7	$A_7 \to A_7, A_8$	2019	730.8	A_{23}	$A_{23} \to A_{23}, A_{24}$
1998	292.3	A_7	$A_7 \to A_7, A_8$	2020	756.494	A_{24}	$A_{23} \to A_{23}, A_{24}$
1999	300	A_7	$A_7 \to A_7, A_8$	2021	778.21	A_{24}	$A_{24} \to A_{24}, A_{27}$
2000	309.6	A_8	$A_7 \to A_7, A_8$	2022	893.08	A_{27}	$A_{24} \to A_{24}, A_{27}$
2001	327.8	A_8	$A_8 \to A_8, A_9$	-	-	-	-

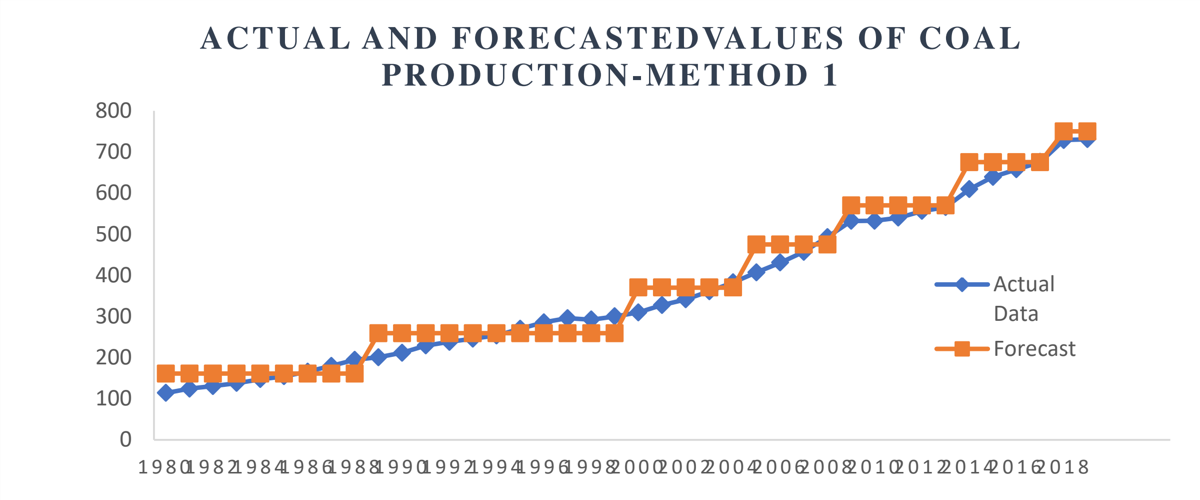

Figure 1: Comparison ofActual and Forecasted values of Coal Production Data

ERROR ANALYSIS:

Mean Squared Error [5]:

Mean squared error of an estimator measures the average of the squares of the errors that is, the average squared difference between the estimated values and the actual values. MSE is a risk function, corresponding to the expected value of the squared error loss. The fact that MSE is almost always strictly positive is because of randomness or because the estimator does not account for information that could produce a more accurate estimate.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

n - number of data points

Y_i -observed values

$\hat{Y}_i$ - predicted values

The Mean Squared Error of the proposed method is 54.9030

Conclusion:

This paper presents the study of Standard deviation method of fuzzy time series forecasting using Trapezoidal and Triangular membership function and average range to replace arbitrary numbers of R_1 and R_2 was applied on Coal production data from the year 1980 to 2022. The Mean Square Error of the Forecasted method-1 of fuzzy time series is 54.9030.Hence it is proven that better prediction can be achieved by applying this Trapezoidal membership function fuzzy time series method.

References:

1. AladagCagdas H., et.al. Forecasting in high order fuzzy times series by using neural networks to define fuzzy relations, Expert Systems with Applications,2009; Vol.36, pp. 4228– 4231.
2. Aznarte José Luis, et.al. Financial time series forecasting with a bio-inspired fuzzy model, Expert Systems with Applications, (2012); Vol.39, pp.12302–12309.
3. Cheng C H, Chang J R and Yeh CA ,2006, Entropy-based and trapezoid fuzzification-based fuzzy time series approaches for forecasting IT project cost Technological Forecasting & Social Change 73:524–542.
4. Song Q., Brad S Chissom., Fuzzy time series and its models. Fuzzy Sets and Systems. 1993; 54:269 – 277.

5. Tay K G., Choy Y Y& Chew CC, 2018, "Forecasting Electricity Consumption Using Fuzzy Time Series", International Journal of Engineering and Technology 7(4.30) 342-346.
6. Wang J. R. H., Chen S.M., Lee C. H., Handing forecasting problems using fuzzy time series. Fuzzy Sets and Systems. 1998; 100: 217-228.
7. Zadeh L A1965 Fuzzy sets Information and Control 8 338- 353.