Available online at www.garmian.edu.krd

Passer Journal

Passer 5 (Issue 2) (2023) 304-310

http://passer.garmian.edu.krd/

A Forecasting Time Series Model Based on Entropy and Fuzzy Logic

Farhad Ali Ahmed*[1], Monem Aziz Mohammed[1]

[1]*Department of Statistics and Informatics, College of Administration and Economic, University of Sulaimani, Sulaimani, Kurdistan Region, Iraq.*

Received 12 January 2023; revised 20 June 2023;
accepted 27 June 2023; available online 17 July 2023

DOI: 10.24271/PSR.2023.381058.1230

ABSTRACT

Electricity Power Consumption Forecasting (EPCF) plays an essential role in global electricity distribution systems that has a significant impact on the operation, control, and planning for the production and distribution of electricity. Due to the complexity, and uncertainty of electricity consumption, especially when the amount of load consumed during different hours is not the same, performing forecasting by using the classical method is inaccurate. To strengthen the efficiency, the time series method that uses a fuzzy approach based on refined entropy is presented in the upcoming article. First, given the specified features, the minimization principle approach of entropy (MPAE) is pursued to define the longitude of each interval in the world of discourse. Secondly, a fuzzy relation matrix of time-invariant is constructed according to the first-order model of fuzzy time series, and the minimum fixed amount of time that the data approach the steady state is obtained using the entropy of the fuzzy set, respectively. Eventually, the forecast results are calculated based on the operation of the maximum combination and the principle of full membership. To show the whole forecasting process, hourly data from July 2022 to September 2022 in Sulaymaniyah / Iraq province is used. Results are compared to the traditional statistical (ARIMA) model, and it indicates that the mean squared error and other criteria of the forecasting error in the entropy based on the fuzzy method are significantly better than the traditional statistical model.

Keywords: EPCF, MPAE, Fuzzy relation matrix, ARIMA model, Uncertainty.

1. Introduction

Electricity power Consumption Forecasting (EPCF) plays a significant role in the formulation of energy policies. It contains a crucial impact on the control of resources within the production of electrical energy. Due to the impossibility of storing it, it is considered one of the sensitive issues within the development requirements of any country.

In the past years, many forecasting methods have been utilized in this field, like classical ARIMA models. For instance, Yuan (2016) expressed that the ARMA model's fitted values compared to Gray Model (1,1), combined Gray Model (1,1) ARMA methods, which reacts less to vacillations[1]. and Mahia (2019) suggested that that is superior to expand the ARMA time series model for complex time series data[2]. However, because of the complexity and ambiguity within the electricity consumption, the accuracy and efficiency of those methods are very low. They cannot be used as a stable model in this field (Nichiforov,2017)[3]. For this reason, to extend the efficiency and accuracy of forecasting in this field, a technique of forecasting fuzzy time series of entropy productivity has been used[4]. It is noteworthy that Liang (2010) proposed in the forecast of Stoke Market to produce the minimization principle approach of entropy linguistic values that can simplify and accelerate the extraction of the rough set algorithm. Furthermore, in the view of Hsue Cheng (2005), The MPAE of the discourse is based on the character of data. It is considered an excellent way for those decision-makers who have minimal data[5,6]. MPAE aims to effectively convert accurate data into fuzzy in an objective and solving a problem manner. Therefore, the purpose of a constant time index is to form an optimal fuzzy relationship. These issues and the efficiency and accuracy of the proposed methodology are discussed in part 2, subsequently analyzed and considered carefully, and the results are compared with those of the classical (ARIMA) method in section 3. The final portion, section 4, presents the conclusions.

* Corresponding author
E-mail address: farhad.ahmed@univsul.edu.iq (Instructor).
Peer-reviewed under the responsibility of the University of Garmian.

2. Material and Methods

2.1 Entropy and its Dependence on probability distribution in Determining the information content

The fundamental idea in classical information theory is Shannon's entropy. Claude Shannon (1948) founded information theory by publishing an article, and he is known as the pioneer of information science. Shannon's entropy Y quantitatively determines the average amount of information obtained from the random variable (y), or the measure of ambiguity about Y before knowing its value. These two points of view are complementary to each other. A random variable's informational value should not be modified by the variety of labels it might have. if a random variable has been used as a symbol to represent different possible distributions, the information obtained should not depend on the type of that variable. For this reason, according to the definition, the entropy of a random variable (Y) is a function of the probability of various possible values that the random variable (Y) can take. Shannon showed that the entropy, p_1,p_2,...,p_n, as a function of the probability distribution corresponding to this probability distribution, can be defined as follows[8-10].

$$H(Y) \equiv H(p_1,\ldots,p_n) \equiv -\sum p_y \log p_y \tag{1}$$

2.2 Minimum Entropy Principle Approach

The entropy of a probability distribution is a measurement of its level of ambiguity. Whenever a system is in thermodynamic equilibrium, its entropy will reach its maximum value. Otherwise, the entropy will not be a significant value. Assume that a threshold value between α and b is desired for a sample.

For the area $[a, a+\tau]$ and $[a+\tau, b]$, an entropy equation is given for each value for y, and we designate the beginning area p and the following area q. The formula for entropy at any value of (y) in the area among (a, b) is:

$$S(y) = p(y).\, S_p(y) + q(y).\, S_q(y) \tag{2}$$

∴

$$S_p(y) = -[p_1(y)\, ln\, p_1(y) + p_2(y) \ln p_2(y)] \tag{3}$$
$$S_q(y) = -[q_1(y) \ln q_1(y) + q_2(y) \ln q_2(y)] \tag{4}$$

Where $p_j(y)$, and $q_j(y)$ are conditional probabilities that the class j sample is within the area $[a, a+\tau]$ and $[a+\tau, b]$, successively, p(y) and q(y) are probabilities that every sample is within the area $[a, a+\tau]$ and $[a+\tau, b]$, respectively, then according to the general law of probability, the sum of p(y) and q(y) is equal to one.

A value of y that offers the minimum entropy, is the optimum threshold value. An entropy estimates of $p_j(y)$, $q_j(y)$, p(y) and q(y) are computed as the following:

$$p_j(y) = \frac{(n_j(y)+1)}{(n(y)+1})\tag{5}$$
$$q_j(y) = \frac{(N_j(y)+1)}{(N(y)+1)} \tag{6}$$
$$p(y) = \frac{n(y)}{n} \tag{7}$$
$$q(y) = 1 - p(y) \tag{8}$$

Where;

$n_j(y)$ is the number of class j samples in the area $[a, a+\tau]$.

$n(y)$ is the total number of samples in the area $[a, a+\tau]$.

$N_j(y)$ the number of class j samples in $[a+\tau, b]$.

$N(y)$ is the total number of samples in the area $[a+\tau, b]$.

Moreover, (n) is the total number of samples in$[a, b]$.

Entropy values are calculated for every position of (y), when the (y) moves in$[a, b]$. The value of (y) with the minimum entropy is named the Central Threshold Point (CTP). The secondary threshold values identified as sections one and two will be determined through the iterative partition procedure. To extend the last partition, the tertiary threshold values identified as THR1, THR2, THR3, and THR4 will be determined[11,12].

2.3 The Time Index of Minimum Invariant

In the theory of information, if a system has a feature or characteristic that does not change with a particular transformation, it means that the system has symmetry corresponding to that feature and it is in a stable state. It is also necessary to mention that in a fuzzy system, when the system's fuzzy degrees remain constant as it transits from one state to the next, it is said to be in a steady state. Thus, to determine a system's degree of fuzziness and measure the time (T), the entropy notion can be used where the data are approaching a steady state[13,14].

Definition.1: In a fuzzy set, the entropy is defined as follows

$$H(\tilde{B}) = K \sum_{i=1}^{m} \mu_{\tilde{B}}(Y_i)\,.\,\ln\mu_{\tilde{B}}(Y_i) \tag{9}$$

where the fuzzy set is

$$\tilde{B} = \{(y, \mu_B(Y))\} \tag{10}$$

$$d(\tilde{B}) = H(\tilde{B}) + H(\tilde{B}^c) \tag{11}$$

Where m is the number of units supporting $\tilde{B}$, $\tilde{B}^C$ is the complement of $\tilde{B}$, and K is a positive constant.[15]

Definition 2: Suppose g (t) occurs only by g (t -1) and is defined by g (t-1) = g(t); then there is a relationship of fuzzy between g (t) and g (t-1) and can be expressed as the Equation of fuzzy relational:

$$g(t) = g(t-1)\ o\ R(t, t-1) \tag{12}$$

Where ‘‘o ’’ is the Max – Min composition operator. The relation R is called a first-order model of g (t). Further, if fuzzy relation R (t, t -1) of g(t) is independent of time t, that is to say, for different times t_1 and t_2, R (t_1, t_1- 1) = R(t_2, t_2 - 1), then g(t) is called a time-invariant fuzzy time series.[16,17]

Definition 3: Let R be an (n×n) fuzzy relation matrix; the mth order matrix of fuzzy relation R^m is as described as follows:

$$R^m = R^{m-1} \circ R \tag{13}$$

Where ‘○’ is the ‘max-min’ factor.

According to Hsue Cheng (2005), at steady state, it has $d(R^{T+1})$=d(R^T), TÎ[1,m], subsequently supposes K = 1, we can conclude model (15) to calculate the most negligible value of T as below.

$$\begin{aligned} &\min \quad T \\ &\text{s.t} \quad r_{ji}^{T+1} = r_{ji}^{T} \qquad \forall_{ji} \ 1,2,3,\ldots,n \qquad (14) \\ &\qquad 0 \le r \le 1 \end{aligned}$$

By solving model (14), we can get the objective value of T to get the value of $[\frac{n}{t}]+1$ relations matrices of time-invariant per T, where $[\frac{n}{t}]$ is the supposed value. After that, by applying the relation matrix of fuzzy $R^{(K)}$, where $K = 1, 2, \ldots, [\frac{n}{t}]+1$ the forecasting result G(t) may be achieved by input fuzzy data $\tilde{B}_i$ and i =1,2,…,n. as below model.

$$G(x) = \tilde{B}_i \circ R^{(K)} \qquad \forall\ (t-1) \le K * T \qquad (15)$$

Where '○' is the 'max-min' factor. $K = 1,2,\ldots,[\frac{n}{t}]+1$ [18,19].

2.4 The Proposed Entropy-based Fuzzy Times Series Method

In this part, according to the approach of the minimum entropy principle and the minimum value of the constant time index T, the details and proposed manner processes of the proposed entropy-based fuzzy time series method are as the following; [19,20]

1st step: Describing the universe under study. In such a way that U_{min} is the lowest and U_{Max} is the highest amount of time series data, as well as u_1 and u_2 are two non-zero positive numbers, and we can determine it as follows

$$U= [U_{min}-u_1, U_{max}+u_2] \qquad (16)$$

2nd step: The existing data must be divided into different classes, but due to the unique characteristics of entropy, there is no particular rule for determining the number of classes and limiting each class.

3rd step: Sorting the studied data in an ascending manner and selecting an intermediate value of (y) between two adjacent values that represents a Central Threshold Point (CTP).

4th step: To determine the threshold value, calculating each potential threshold and the desired value will be applied to the equations (1-8). To determine the seventh value, each of the potential threshold points must be selected so that the entropy value is minimized.

This process will continue until all seven thresholds are determined. Therefore, the universe under study is divided into seven unequal overlaps, where

$d_1 = [U_{MIN} - u_1, SEC_1]$, $d_2 = [THR_1, THR_2]$, $d_3 = [SEC_1, CTP]$, $d_4 = [THR_2, THR_3]$, $d_5 = [CTP, SEC_1]$, $d_6 = [THR_3, THR_4]$, $d_7 = [SEC_2, U_{MAX} + u_2]$

5th step: Identify the fuzzy sets in the universe under study. The linguistic values are defined as follows

F1: Not many consumption F5: Very many consumption

F2: Not too many consumption F6: Too many consumption

F3: Many consumption F7: Too many many consumption

F4: Many many consumption

And each set of fuzzy $\tilde{B}_i$ where i= 1,2,3,...,7 is defined as

$$\begin{aligned}
B_1 &= \{ 1/d_1, 0.5/d_2, 0/d_3, 0/d_4, 0/d_5, 0/d_6, 0/d_7 \} \\
B_2 &= \{ 0.5/d_1, 1/d_2, 0.5/d_3, 0/d_4, 0/d_5, 0/d_6, 0/d_7 \} \\
B_3 &= \{ 0/d_1, 0.5/d_2, 1/d_3, 0.5/d_4, 0/d_5, 0/d_6, 0/d_7 \} \\
B_4 &= \{ 0/d_1, 0/d_2, 0.5/d_3, 1/d_4, 0.5/d_5, 0/d_6, 0/d_7 \} \\
B_5 &= \{ 0/d_1, 0/d_2, 0.5/d_3, 1/d_4, 0.5/d_5, 0/d_6, 0/d_7 \} \\
B_6 &= \{ 0/d_1, 0/d_2, 0/d_3, 0.5/d_4, 1/d_5, 0.5/d_6, 0/d_7 \} \\
B_7 &= \{ 0/d_1, 0/d_2, 0/d_3, 0/d_4, 0/d_5, 0.5/d_6, 1/d_7 \}
\end{aligned} \qquad (17)$$

6th step: For building the membership function, the length of intervals must be identified through the fourth step's thresholds as the triangular fuzzy number midpoint.

7th step: The phase of fuzzy in the time series data, using the adapted membership function in the previous stages, is calculated by the grade of the membership of each data studied and then specified to a desired linguistic value.

8th step: The fuzzy set concept will define the most minor extent of the time-invariant value (T). First, the fuzzy relation matrix must be created. After that, the index of time (T) will be marked for invariant time relationships.

9th step: By considering each data's linguistic value, it is possible to obtain the logical fuzzy relations of the data, which are shown as $i \to j$. Where i and j are two consecutive states. These relations are obtained through the matrices of fuzzy relations R^K. Where $K = 1,2,\ldots,[\frac{n}{t}]+1$. Any fixed (invariant) time can be used to obtain the relation matrices in a fuzzy manner.

10th step: Computing the outputs, if the data under study y(t) where $t = [1,2,...,m]$, is set to the fuzzy set B_i, to get the forecastable value $Y(K+1)$, based on the knowledge obtained from the previous steps, and the following relationship will be used.

$$Y(K+1) = B_i \circ R^{[\frac{t}{T}+1]} \qquad t \in [1,2,...,m] \qquad (18)$$

11th step: To convert the obtained results, which are in the form of fuzzy sets, they must be converted into fundamental values. Method of Q, Song (1993) has been used for this conversion.

2.5 Autoregressive Integrated Moving Average Processes (ARIMA)

One of the most well-known time series modeling techniques is the ARIMA model, commonly referred to as the Box-Jenkins model. It is primarily used to forecast time series using the tenet that a variable's future value is a linear function of previous observations and random errors. The process of creating a time series takes the following form:

$$y_t = \theta_0 + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \quad (19)$$

Where

y_t is the time series value

$\varphi_i (i = 1,2,3, \dots, p)$ is the Moving average parameters

$\theta_j (j = 1,2,3, \dots, q)$ is the Autoregressive parameters

ε_t is a random process with zero mean and σ^2 variance.

2.6 Comparison with the Classical Method (ARIMA)

To check the forecast power of the methods used, four standard criteria in this field were used. [16,21]

The first criterion is the average of errors (ME), which is the average of all prediction errors for a data group and is defined as follows.

$$ME = {\sum e_i}/{N} \quad (20)$$

Where e_i is the difference between the forecasted value and its actual value, and N is the number of forecasts.

The second criterion is the mean percentage value of errors (x). In this criterion, the mean percentage value of the errors is used for each of the forecasts. That is mean

$$MPE = \frac{100\%}{N} \sum ({e_i}/{y_i}) \quad (21)$$

Where N is the number of forecasts, and e_i is the difference between the forecasted and actual value.

The third criterion is the mean absolute value of the error (MAPE). This criterion is similar to (MAE), with the difference that the error is expressed as a percentage.

$$MPAE = \frac{100}{N} \sum |e_i| \quad (22)$$

The fourth criterion is the mean squared of error (MSE), which is expressed as follows.

$$MSE = {\sum e_i^2}/{N} \quad (23)$$

3. Result and Discussion

3.1 Implementation of the Suggested Method

This section will utilize the hourly power consumption data for the Sulaymaniyah province that was prepared from July to September 2022, and it will execute the analysis based on the previous section's procedures. The steps in the preceding section can be applied in the following manner.

1. According to the studied data, the maximum and minimum values are 441.55 and 911.73, respectively. Where u_1=1.55, and u_2= 0.27 for simplifying in the calculation, thus, the discourse universe is U [440, 912].
2. By consulting with energy consumption experts, the studied data has been divided into three different groups in terms of consumption, low = 1, medium = 2, and high = 3, as described in Table 1. Due to the large amount of data, the link of all data can be seen in the annex A.1.

Table 1: Part of the hourly data based on the desired classes (1-7-2022 to 31-9-2022).

Hourly time	Data	Classes	Hourly time	Data	Classes	Hourly time	Data	Classes
1 AM	717.93	1	11 AM	805.88	2	9 PM	742.85	3
2 AM	728.18	1	12 AM	732.80	2	10 PM	713.02	1
3 AM	749.05	1	1 PM	732.80	2	11 PM	705.42	1
4 AM	771.93	1	2 PM	722.88	2	5 PM	883.71	2
5 AM	781.35	1	3 PM	745.43	2	6 PM	885.75	3
6 AM	827.02	3	4 PM	766.38	2	7 PM	889.10	3
7 1M	799.45	3	5 PM	752.03	2	8 PM	889.28	3
8 AM	777.72	3	6 PM	727.22	3	9 PM	891.40	3
9 AM	794.88	2	7 PM	726.70	3	10 PM	898.21	1
10 AM	790.92	2	8 PM	732.23	3	11 PM	907.65	1

Sorting the data and calculating the (CTP).

Dividing the studies data into the seventh part by equations (1-8) to find the seven thresholds value as shown in the following table.

Table 2: The Thresholds Points.

Border	THR1	1st section	THR2	CTP	THR3	2nd section	THR4
Value	464.48	502.85	540.16	561.81	714.41	821.6	851.58

According to equations (11-19), define the seven linguistic values.

To identify the triangular fuzzy number midpoint and determine the membership function, the threshold value in step (4) has been used and shown in Table 3.

Table 3: The Membership Function of MPAE.

Linguistic Value	Lower. B	Midpoint	Upper. B	Longitude
B1: Not many consumptions	440	464.4833	502.8500	62.8500
B2: Not too many consumption	464.4833	502.8500	540.1667	75.6834
B3: Many consumption	502.8500	540.1667	561.8167	58.9667
B4: Many many consumption	540.1667	561.8167	714.4167	174.2500
B5: Very many consumptions	561.8167	714.4167	821.6000	259.7833
B6: Too many consumption	714.4167	821.6000	851.5832	137.1665
B7: Too many many consumptions	821.6000	851.5832	912	90.4000

Fuzzing the data under study, based on the membership function values identified in the previous step, and trying to find the dignity of membership for each data is then allotted to a proportional linguistic value, and a part of it is shown in Table 4. (The complete results have been shown in the annex A.2).

Table 4: Part of Fuzzified Data Based on Minimum Entropy.

Time	B_1	B_2	B_3	B_4	B_5	B_6	B_7	State	Relation
1 AM	0	0	0	0	0.289425	0.710575	0	B_6	-----
2 AM	0	0	0	0	0.484915	0.515085	0	B_6	$B_6 \to B_6$
3 AM	0	0	0	0	0.351011	0.648989	0	B_6	$B_6 \to B_6$
4 AM	0	0	0	0	0.88056	0.11944	0	B_5	$B_6 \to B_5$
5 AM	0	0	0	0	0.265319	0.734681	0	B_6	$B_5 \to B_6$
6 AM	0	0	0	0.990825	0.009175	0	0	B_4	$B_6 \to B_4$
7 AM	0	0	0	0.941022	0.058978	0	0	B_4	$B_4 \to B_4$
8 AM	0	0	0	0.064439	0.9355610	0	0	B_5	$B_4 \to B_5$
9 AM	0	0	0	0.808759	0.191241	0	0	B_4	$B_5 \to B_4$
10 AM	0	0	0	0	0.093157	0.906843	0	B_6	$B_5 \to B_6$

According to the sets of fuzzy that were defined in the previous step, and based on the matrix of the below fuzzy relation

$$B=\begin{bmatrix} 1/d_1 & 0.5/d_2 & 0/d_3 & 0/d_4 & 0/d_5 & 0/d_6 & 0/d_7 \\ 0.5/d_1 & 1/d_2 & 0.5/d_3 & 0/d_4 & 0/d_5 & 0/d_6 & 0/d_7 \\ 0/d_1 & 0.5/d_2 & 1/d_3 & 0.5/d_4 & 0/d_5 & 0/d_6 & 0/d_7 \\ 0/d_1 & 0/d_2 & 0.5/d_3 & 1/d_4 & 0.5/d_5 & 0/d_6 & 0/d_7 \\ 0/d_1 & 0/d_2 & 0/d_3 & 0.5/d_4 & 1/d_5 & 0.5/d_6 & 0/d_7 \\ 0/d_1 & 0/d_2 & 0/d_3 & 0/d_4 & 0.5/d_5 & 1/d_6 & 0.5/d_7 \\ 0/d_1 & 0/d_2 & 0/d_3 & 0/d_4 & 0/d_5 & 0.5/d_6 & 1/d_7 \end{bmatrix} \tag{23}$$

Then, because

$$B \circ B = B^2 \neq B \quad , B^2 \circ B = B^3 \neq B^2 \quad \dots \quad , B^{n-1} \circ B = B^n \neq B^{n-1}$$

where n is integer

According to model (15), for the time-invariant relationship, the time index of T has been taken where T=5.

By Equation (23), the forecasted outcome in the fuzzy form is computed and converted to the real numbers. Where part of them is shown in Table (5). (The complete results have been shown in the annex A.3).

Table 5: Part of the Forecasted Result.

Time	B_1	B_2	B_3	B_4	B_5	B_6	B_7	Fuzzy out put	Real out put
1 AM	---	---	---	---	---	---	---	---	---
2 AM	0	0	0	0	0.12846	0.87154	0	B_6	717.9333
3 AM	0	0	0	0	0.323173	0.676827	0	B_6	728.1833
4 AM	0	0	0	0	0.536703	0.463297	0	B_5	749.05
5 AM	0	0	0	0	0.793469	0.206531	0	B_5	732.80
6 AM	0	0	0	0	0.590669	0.409331	0	B_5	722.8833
7 AM	0	0	0	0	0.750856	0.249144	0	B_5	745.4333
8 AM	0	0	0	0	0.713842	0.286158	0	B_5	766.3833
9 AM	0	0	0	0	0.8535	0.1465	0	B_5	752.0333

3.2 Modeling Electricity Consumption Using the ARIMA Model

Since the condition of applying these models and obtaining reliable predictions based on them is the significance of the variable under investigation, the significance of the trend of electricity consumption in the period, time (1-7-2022 to 31-9-2022), as shown in Figure 1, has been obtained with a one-time difference. After checking the stability, it is important to determine the rank of p and q in the ARIMA model; for this purpose, the autocorrelation function (ACF) and partial autocorrelation function (PACF) have been used. Therefore, the ARIMA models (2,1,0) have been chosen as the appropriate models, with AR1= 0.0880 and AR2 = -0.0662 as shown in below

$$y_t = 788.25 + 0.0880\, y_{t-1} - 0.0662 y_{t-2} + \varepsilon_t$$

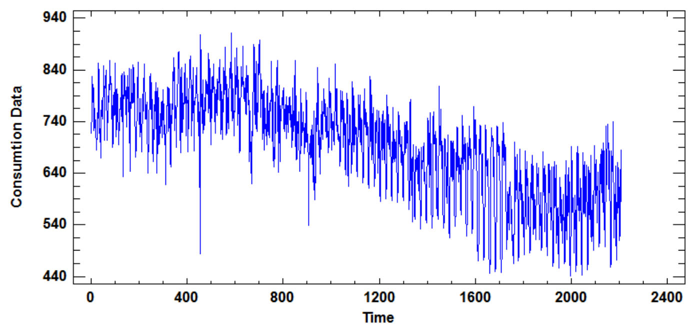

Figure 1: Time sequence of electricity consumption plot.

3.3. Performance Evaluation of Entropy-based Method and ARIMA Method for Forecasting

To show the effectiveness of the entropy-based forecast model, the forecasting performance based on this method and the ARIMA model using the performance evaluation criteria presented in relations (19 –23) is examined, and the results are presented in Table (6).

Table 6: Comparison of the Different Forecasting Methods.

	ME	MPE	MAPE	MSE
ARIMA	0.060	0.17	4.4303	0.1
Entropy	0.010	0.083	0.102	0.002

Conclusions

Traditional statistical methods may not be sufficient to forecast complex and ambiguous data logically. Therefore, in this study, the ARIMA time series model and entropy based on fuzzy time series were used to compare and evaluate time series data forecasting methods. To evaluate the methods during the studied period, the expected value was calculated, and the predicted values were compared with the actual values using four different criteria. The results showed that the entropy method based on fuzzy time series is more efficient for predicting energy consumption in hourly and long-term data. In contrast, the ARIMA method required several different normalization methods to address the presence of nonlinear trends in the data. It is worth noting that the entropy method showed a high ability to compare different data and provide better results than classical models. Although the analysis was conducted on short-term data, the results suggest that this method can also be effective for long-term data and data without overall coherence.

Conflict of interests

None

Author contribution

F.A and M.A suggested the idea, F.A suggested the outline of the proposal. Then F.A did each part, extracted the results and A.K supervised this article as a consultant professor.

Funding

The authors received no financial support for the research, authorship, and publication of this article.

References

1. Yuan Chaoqing, Liu Sifeng, Fang Zhigeng. Comparison of China's primary energy consumption forecasting using the ARIMA (the autoregressive integrated moving average) model and GM (1, 1) model. *Energy*, **100**, 384-390. (2016).
2. Mahia, F, Dey, A.R, Masud. Md. A, Mahmud M. S. Forecasting electricity consumption using ARIMA model. International Conference on Sustainable Technologies for Industry 4.0 (STI), *IEEE*, 1- 6, (2019).
3. Nichiforov. Cristina, Stamatescu. Iulia, Făgărăşan. Ioana, Stamatescu. Grigore. Energy consumption forecasting using ARIMA and neural network models. 5th International Symposium on Electrical and Electronics Engineering (ISEEE), *IEEE*, 1-4, (2017).
4. Su, Chung. Ho, Chen, Tai. Liang, Cheng, Ching. Hsue. Chen, Ya. Ching. Forecasting the stock market with linguistic rules generated from the minimize entropy principle and the cumulative probability distribution approaches. *Entropy* **12**, 2397-2417, (2010).
5. Cheng, C.H. Chang, J.R. Yeh, C.A. Entropy-based and trapezoid fuzzification-based fuzzy time series approaches for forecasting IT project cost. Elsevier Inc. *Technological forecasting & social change*, **72**, 563-574, (2005).
6. Song, Q., Chissom, B.S. Forecasting enrollments with fuzzy time series—Part I. Elsevier Science Publishers, AL, USA. *Fuzzy Sets Syst*, **54**, 1–10, (1993).
7. Ross, T.J. *Fuzzy Logic with Engineering Applications*. McGraw-Hill, USA, (2000).
8. Miller,G. A, The magical number seven, plus or minus two: some limits on our capacity of processing information, Harvard University. *Psychol. Rev*. **63**, 81–97, (1956).
9. Zadeh, L. A, Fuzzy sets. California, USA, *Inf. Control* , **8**, 338–353, (1965).
10. Christensen, R. *Entropy Minimax Sourcebook*, Entropy Ltd., Lincoln, Mass, (1980).
11. Yager, R. Filev, D.P. Template-based fuzzy systems modeling. John Wiley & Sons, N.Y, USA, *Journal of Intelligent and Fuzzy Systems*, **2**, 39-54 (1994).
12. Zhang, C. Ma, C.B. Xu, J.D. A new fuzzy MCDM method based on trapezoidal fuzzy AHP and hierarchical fuzzy integral. Springer-Verlag, Berlin Heidelberg. *Lecture Notes in Artificial Intelligence (LNAI)* 3614, 466–474, (2005).
13. Wu, B.L. Lin, Y.C. Fuzzy time series analysis and forecasting: with an example to Taiwan weighted stock index. *Acta Mathematics Application*, **25**, 67-76, (2002)
14. Tsaur, R.C. Yang, J.C. Wang, H.F. Fuzzy relation analysis in fuzzy time series model. Elsevier Ltd. *Computers and mathematics with the application*, **49**, 539-548, (2005).
15. Chen, S.M. Forecasting enrollments based on fuzzy time series. Elsevier Science Publishers, AL, USA. *Fuzzy Sets Syst*,**81** ,311–319 (1996).
16. Tseng, F.M. Tzeng, G.H.Yu, H.C. Fuzzy seasonal time series for forecasting the production value of the mechanical industry in Taiwan. Elsevier Science Inc, N.Y, USA. Technol. *Forecast. Soc. Change*, **60,** 263–273, (1999).
17. Wang, H.F, Lee, C.T, A method for fuzzy time series analysis-an example for telecommunication demands. Vancouver, Canada, IFORS,**96**, 8–12, (1996).
18. Cox, E. *The Fuzzy Systems Handbook: A Practitioner’s Guide to Building, Using, and Maintaining Fuzzy Systems.* Academic Press Professional, Inc., San Diego, CA, (1994).
19. Song, Q. Chissom, B.S. Forecasting enrollments with fuzzy time series—Part II, Elsevier Science Publishers, AL, USA. *Fuzzy Sets Syst*,**62**, 1-8, (1994).
20. Sullivan, J. Woodall, W.H. A comparison of fuzzy forecasting and Markov modeling. Elsevier Science Publishers, AL, USA. *Fuzzy Sets Syst*, **64,** 279–293 (1994).
21. Huarng, K. Effective lengths of intervals to improve forecasting in fuzzy time series. Elsevier Science B.V. *Fuzzy Sets Syst*,**123**, 387–394, (2001).