The benefits of social insurance system prediction using a hybrid fuzzy time series method

Ahmed Abdelreheem Khalil[1,2], Mohamed Abdelaziz Mandour[3] and Ahmed Ali[4,5]

[1] School of Mathematics and Statistics, Central South University, Changsha, Hunan, China
[2] Faculty of Commerce, Assiut University, Assiut, Egypt
[3] Department of Business Administration, Faculty of Commerce, Mansoura University, Mansoura, Egypt
[4] Department of Computer Science, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia
[5] Higher Future Institute for Specialized Technological Studies, Cairo, Egypt

ABSTRACT

Decision-making in many industries relies heavily on accurate forecasts, including the insurance sector. The Social Insurance System (SIS) in Egypt, operating under a fully funded paradigm, depends on reliable predictions to ensure effective financial planning. This research introduces a hybrid predictive model that combines fuzzy time series (FTS) Markov chains with the tree partition method (TPM) and difference transformation to forecast total pension benefits within Egypt's SIS. A key feature of the proposed model is its ability to optimize the partitioning process, resulting in the creation of nine intervals that reduce computational complexity while maintaining forecasting accuracy. These intervals were consistently applied across all fuzzy time series models for comparison. The model's performance is evaluated using established metrics such as MAPE, Thiels' U statistic, and RMSE. Additionally, prediction interval coverage probability (PICP) and mean prediction interval length (MPIL) are used to assess the quality of prediction intervals, with a 95% prediction interval serving as the baseline. The proposed model achieved a PICP of approximately 95%, indicating well-calibrated prediction intervals, although the MPIL of 424.5 reflects a wider uncertainty range. Despite this, the model balances coverage accuracy and interval precision effectively. The results demonstrate that the proposed model significantly outperforms traditional models like linear regression, ARIMA, and exponential smoothing and conventional FTS models like Song, Chen, Yu, and Cheng by achieving the lowest MAPE with the value of 11.8% for training and 10.65% for testing. This superior performance highlights the model's reliability and potential applicability to further forecasting tasks in the field of insurance and beyond.

Subjects Artificial Intelligence, Data Mining and Machine Learning, Social Computing
Keywords Fuzzy time series, Markov chain, Predictive method, Insurance, Forecasting

INTRODUCTION

Forecasting is a critical aspect for solving problems, particularly in various businesses and the insurance industry, where predictions of future events must inform decision making. Therefore, forecasting techniques enable the insurance authorities to anticipate

future results and formulate strategies accordingly. Thus, accurate and efficient forecasting systems are highly sought after in this industry (*Khalil, Liu & Ali, 2022a*).

The social insurance system (SIS) plays a pivotal role in countries around the world as a cornerstone of social welfare and economic stability. Its significance lies in its ability to provide a safety net for individuals and families during times of need, ensuring that unforeseen events such as illness, unemployment, or disability do not lead to financial ruin (*Liu et al., 2023*; *Raouf & Elsaieed, 2022*). Consequently, the SISs serve as a means for individuals to actively participate in safeguarding their future. Thus, these systems rely on contributions from both workers and employers, effectively distributing risks among a wide population. By doing so, social insurance programs promote social cohesion and mitigate income disparities, thereby cultivating a sense of security and well-being among the populace (*Chen et al., 2022*).

SIS holds profound significance in countries as a critical component of their economic and social fabric. Additionally, these programs often stimulate economic growth by maintaining consumer spending even in times of economic downturns, as individuals have a safety net to fall back on. Moreover, they contribute to the overall health and productivity of the workforce, as access to healthcare and unemployment benefits can prevent individuals from falling into poverty traps. In essence, social insurance embodies a nation's commitment to the welfare and resilience of its citizens, promoting economic stability and social justice. Simultaneously, it is widely acknowledged as a fundamental issue that numerous nations grapple with due to the intricate interplay between this matter and its implications on public finances (*Raouf & Elsaieed, 2022*).

In Egypt, SIS operates on a fully funded model where employee contributions are invested for pension payouts, evolving over time to include elements of a defined benefit plan. Governed by laws, the latest being Law 148/2019, this legal framework aims to modernize the system. However, Egypt's SIS faces challenges due to fragmentation and diverse benefit packages for different workforce segments, covering areas like old age, disability, and illness. Unsustainability and inefficiency arise from low real interest rates for pension fund investments, generous pensions, and flexible pension adjustments based on recent contributions. Many workers under-report their income throughout most of their working careers to reduce their payments, threatening the system's financial viability (*Kassem, 2021*; *Loewe & Westemeier, 2018*). In light of Egypt's significant and increasing total benefits (pension), which have exhibited a notable upward trajectory in recent years, there is a need to forecast these benefits for the coming years.

Actuarial models are employed to estimate annual benefits, forming the basis for demographic and financial predictions in pension systems. These models rely on factors such as demographics and economics and are vital for supporting policy decisions. Consequently, the integration of actuarial methodologies in tandem with predictive analytics assumes paramount importance in comprehending and forecasting future behavioral trends or occurrences, thereby furnishing the requisite support for informed policy formulation and decision-making. In essence, the prediction and evaluation of benefits have utmost significance for actuaries in their capacity as advisors for strategic

decision-making within the realm of social insurance (*Hassani, Unger & Beneki, 2020*; *Raouf & Elsaieed, 2022*).

In our study, we utilized a summative modeling approach to predict retirement benefits, primarily based on historical data trends within the Egyptian social insurance system. Unlike traditional actuarial methodologies, which often employ mortality projections, such as cohort-based mortality tables to estimate future pension payments, our method focused on the available data. Actuarial literature, particularly from the Government Actuary's Department and other studies, highlights the importance of mortality assumptions for accurately forecasting long-term pension liabilities. These actuarial models predict benefit changes by incorporating life expectancy data, where pension growth is typically modeled multiplicatively, reflecting the natural increase in benefit payouts as retirees age and longevity improves (*Booth & Tickle, 2008*).

However, due to the unavailability of detailed mortality data for Egypt's population, we were unable to apply these traditional methods. Instead, we relied on a summative model, in which retirement benefits are predicted based on observed incremental changes in benefits, rather than factoring in mortality or life expectancy. Although our approach may seem basic by comparison, it offers a practical solution given the constraints of the available data. Incorporating more complex actuarial models in future work, including mortality rates, could provide a more refined projection of retirement benefits (*Belloni & Maccheroni, 2013*; *Ingale & Paluri, 2023*).

The study's motivation and contribution

Fuzzy time series (FTS) models have exhibited more effectiveness in comparison to traditional forecasting methods in the field of forecasting, particularly when dealing with datasets that are incomplete, ambiguous, or have a limited number of observations (*Khalil, Liu & Ali, 2024a*). Furthermore, these models do not require the imposition of statistical assumptions. They use mathematical reasoning to build models and forecast future events based on historical linguistic data. In contrast to the traditional approaches taken for predicting time series, the inherent unpredictability of temporal observations is acknowledged and accounted for by fuzzy time series forecasting techniques. This eliminates the requirement for strict assumptions and a substantial amount of prior knowledge about the observed data. Moreover, it is important to highlight that fuzzy time series forecasting techniques demonstrate competence in handling a restricted number of data, emphasizing their flexibility and usefulness (*Mao & Xiao, 2019*; *Wu et al., 2021*).

In recent times, there has been a notable proliferation in the adoption of hybrid models as a means to enhance predictive accuracy. It has been demonstrated that the utilization of hybrid approaches frequently results in enhanced performance outcomes. The fundamental concept underlying hybridization models is their ability to synergistically enhance the identification of patterns in data. Numerous studies, both theoretical and empirical, have provided evidence supporting the notion that combining multiple models through hybridization leads to enhanced performance as compared to the utilization of a single model in isolation (*Chen & Chen, 2015*; *Hadwan et al., 2022*; *Khairalla & Al-Jallad, 2017*; *Khozani et al., 2022*).

Numerous fuzzy time series models have been advanced for predictive purposes across diverse domains. Building upon the merits inherent in these models and their demonstrated efficacy in predicting, this study presents a hybrid FTS that builds on the strengths of the fuzzy time series Markov chain (FTSMC), the difference transformation data technique, and the Tree Partition technique (TPM) for making predictions. The primary aim of this model is to forecast the total benefits (pensions) of the social insurance system in Egypt, reflecting a commitment to harness the capabilities of advanced methodologies for enhancing predictive accuracy in this specific application domain. The research presents several significant contributions to literature as outlined below:

1-   The efficacy of FTS in modeling non-linear, with inherent uncertainties and accommodating linguistic variables, and providing adequate performance with limited data availability. Thus, this study investigates a hybrid prediction model that combines fuzzy time series with Markov chain, by leveraging the tree partition method (TPM) and data transformation method to enhance the accuracy of forecasting the total benefits (pensions) of the social insurance system in Egypt. Three metrics are proposed to evaluate the prediction performance of the model. This method can enhance the accuracy and effectively manage challenges in insurance data, such as uncertainty and limited dataset size, and assist actuaries in fulfilling their professional obligation to engage in strategic decision-making inside insurance businesses.

2-   To the best of our knowledge, this work is the first empirical investigation to employ a hybrid FTS that builds on the strengths of the fuzzy time series Markov chain (FTSMC), the difference transformation data technique, and the tree partition technique (TPM) for predicting the total benefits (pensions) of social insurance system in Egypt. This research addresses a large gap in existing insurance literature, making it a valuable contribution to the field.

3-   Finally, the author's primary objective is to provide insurance authorities with a thorough and practical technique that can assist actuaries in fulfilling their professional duty of making strategic decisions within the realm of the social insurance system. Overall, the study offers several important contributions to the field of insurance and gives crucial insights into the Egyptian social insurance system.

The remainder of this work is outlined as follows: 'Background and Literature Review' provides a brief introduction to prior studies. 'Data & Evaluation Methods' provides a detailed account of the research methodology, including the research design, describe the specific techniques and approaches that were employed in the study and the evaluation metrics. The findings are presented and analyzed in 'Results and Discussion'. Finally, we present the conclusion in 'Conclusions'.

BACKGROUND AND LITERATURE REVIEW

Forecasting within the domain of insurance holds significant importance, necessitating the application of various statistical and soft computing methodologies. These methods

play a pivotal role in the prediction of events and trends within the insurance industry. Enhanced forecasting capabilities are instrumental in aiding actuaries in their professional responsibilities, enabling the informed and strategic decision-making processes inherent to the management of social insurance systems. For instance, *Ying et al. (2017)* present a deep learning based RNN framework for forecasting payment behaviour using real data from Taiwan's Ministry of Health and Welfare. This model outperforms SVM and HMM in various settings, although it addresses individual payment predictions, which differs from our research focus. *Khalil, Liu & Ali (2022a)* tried to forecast the loss ratios in petroleum insurance using both ANFIS and ARIMA models. They found that the ANFIS models provided more accurate results than the ARIMA models.

Fuzzy time series methods

In recent years, an increasing number of researchers have employed fuzzy time series methodologies to address prediction challenges as in *Alyousifi et al. (2020)*, *Alyousifi et al. (2021)*, *Arnita, Afnisah & Marpaung (2020)*, *Jilani, Burney & Ardil (2008)*, *Lee et al. (2017)*, *Mao & Xiao (2019)*, *Rahim et al. (2018)*, *Raouf & Elsaieed (2022)*, *Tsaur (2012)* and *Wu et al. (2021)*. These fuzzy time series models have demonstrated notable efficacy when juxtaposed with conventional forecasting approaches, particularly within the realm of forecasting methodology. *Tsaur (2012)* designed a new model that integrates Fuzzy Time Series analysis with Markov chain simulation, referred to as FTSMC. This model was devised to ascertain suitable weights for fuzzy relationships within time series patterns, thereby enhancing model accuracy. However, FTSMC exhibits a limitation as it relies on arbitrary partitioning of intervals and does not possess the capability to determine the optimal interval length. This limitation constitutes a drawback in the context of effectively determining interval length within this hybrid model.

Fuzzy time series in insurance and actuarial science

Fuzzy time series (FTS) has become an invaluable tool in insurance and actuarial science, particularly for dealing with the uncertainty, vagueness, and imprecision inherent in many actuarial datasets. Actuaries, who traditionally rely on precise statistical methods, often face challenges when working with datasets where variability and ambiguity prevail common in areas like mortality rates, claims forecasting, and pension liabilities. The integration of fuzzy logic within time series models offers a flexible and powerful approach to overcoming these challenges.

Several studies have employed FTS methods to address various predictive challenges in insurance and actuarial science, including car accidents, profitability, insurance claims forecasting, and mortality rates. For example, *Jilani, Burney & Ardil (2008)* developed a multivariate fuzzy time series model for predicting car accidents in Belgium. Their approach demonstrated superior performance compared to existing techniques, offering actuaries a robust tool for forecasting in both car and life insurance contexts. Similarly, *De Andrés-Sánchez & Puchades (2019)* proposed a fuzzy-random extension of the Lee–Carter model, which integrates uncertainties in time-dependent mortality trends (using ARIMA) and age-dependent coefficients (represented by triangular fuzzy numbers). This enhanced

model significantly improved the accuracy of mortality predictions, outperforming the original Lee–Carter model as well as other fuzzy-based extensions in predicting central mortality rates, death probabilities, and life expectancies.

Hong et al. (2021) expanded on this work by combining the Lee–Carter model with machine learning and statistical techniques such as random forest (RF), artificial neural networks (ANN) and ARIMA to forecast mortality rates. Their findings indicated that the LC-ANN model achieved the highest accuracy in countries with less efficient healthcare systems, such as Malaysia, whereas the LC-ARIMA model performed better in regions with more advanced healthcare infrastructures. This integration of machine learning methodologies further enhanced the precision of mortality predictions.

Raouf & Elsaieed (2022) contributed to the field by developing several fuzzy time series models, including those proposed by Chen and others, aimed at forecasting social insurance benefits. They recommended the use of the Huarng partitioning method to determine optimal interval lengths. Empirical validation confirmed that the Chen model, combined with Huarng partitioning, provided high forecasting accuracy with low error rates during both training and testing phases. This underscores the effectiveness of FTS methods in enhancing the accuracy and reliability of actuarial forecasts across diverse applications.

Fuzzy time series in other fields

On the other hand, the FTS method proved efficiency for predicting in other fields. For instance, *Lee et al. (2017)* improved FTS models by using the F-transform technique to reduce prediction errors, achieving greater accuracy in forecasting, such as student enrollment and Taiwanese patent data, and underscoring the value of dataset transformation. *Arnita, Afnisah & Marpaung (2020)* predicted rainfall in Medan using fuzzy time series models, comparing *Chen (1996)*, Markov chain, and *Cheng, Chen & Chiang (2006)* variations. The Chen method proved most accurate based on the lowest MAPE values, with intervals determined using average-based principles. *Alyousifi et al. (2020)* used a fuzzy time series Markov chain (FTSMC) model with a grid technique to optimize partitions for predicting daily air pollution in Klang, Malaysia. This model improved prediction accuracy and outperformed traditional statistical models, making it a potentially more reliable method for forecasting air pollution. *Alyousifi, Othman & Almohammedi (2021)* developed a hybrid forecasting model using Markov chain and C-Means clustering for fuzzy time series data. This model showed superior performance over existing models on real datasets, including TAIEX trading data and PM10 concentrations in Melaka, Malaysia.

FTS is a broad field that comprises various models used in different fields. These models might vary based on aspects such as the choice of model, characteristics of the dataset, methodology for partitioning, and the selection of tools for assessing correctness. Therefore, the primary objective of our study is to anticipate the total benefits (pensions) within SIS in Egypt using the proposed hybrid predictive model. The primary objective of this approach is to attain exceptional forecast accuracy, as supported by empirical evidence. The primary benefit provided by the suggested model is in its ability to optimize

the partitioning process, therefore mitigating computational complexity and decreasing the associated computational load in the forecasting task.

FUZZY TIME SERIES AND PROPOSED METHOD

The concept of fuzzy set theory was first proposed by *Zadeh (1965)*. This theory is utilized to tackle the challenges of ambiguity and uncertainty that arise in practical situations, by providing a framework for representing and analyzing linguistic fuzzy information. The utilization of FTS models has been widespread and has contributed to significant advancements in various fields. The original FTS model is developed by *Song & Chissom (1993)*; *Song & Chissom (1994)* utilized the max–min composition technique, but it encountered computational difficulties. *Chen's (1996)* subsequent model introduced simplifications to the calculations; however, it did not incorporate suitable weight mechanisms for fuzzy logical connections (FLRs).

Numerous researchers have since made modifications and enhancements to FTS models in order to boost the accuracy of predictions such as *Chen & Hwang (2000)*, *Huarng (2001a)*, *Huarng (2001b)*, *Huarng & Yu (2006)*, *Hwang, Chen & Lee (1998)* and *Yu (2005)*. These improvements encompass a range of weighing mechanisms, extended linguistic intervals, and other division algorithms. The application of fuzzy theory, as refined by *Yu (2005)* has been observed in several conventional FTS models, such as those used for stock price predictions. The application of various weighted technologies was utilized by Yu to improve the accuracy of forecasting.

On the other hand, *Cheng, Chen & Chiang (2006)* merged FTS with a trend-weighting technique to make forecasts for both real stock prices and university enrolment. Combining fuzzy sets with the time series model generates the idea of fuzzy time series. The essential procedures for constructing fuzzy time series models encompass the establishment of a defined universe of discourse U, the partitioning of U into a uniform number of intervals, the process of fuzzification, the specification of fuzzy logic relations, the determination of forecasted values, and the subsequent defuzzification.

Previous literature has provided several key definitions of fuzzy time series that have been constructed. These definitions are outlined as follows (*Chen, 1996*; *Song & Chissom, 1993*; *Song & Chissom, 1994*):

Definition 1

Let $X(t)(t = 1, 2, 3, .., n,)$ represent a collection of real numbers, serving as the universe of discourse within which fuzzy sets $f_i(t)$ are defined. The collection $F(t)$, denoted as $f_1(t), f_2(t), \ldots, f_n(t)$, is referred to as a fuzzy time series defined on $X(t)$.

Definition 2

A fuzzy set refers to a category of entities that possess a continuous range of membership grades. Let U denote the Universe of discourse, which is defined as $U = u_1, u_2, \ldots, u_n$, where u_i represents the potential linguistic values within U. Then a fuzzy set of linguistic variables A_i of U is defined in Eq. (1) by

$$A_i = \frac{f_{A_i}(u_1)}{u_1} + \frac{f_{A_i}(u_2)}{u_2} + \ldots + \frac{f_{A_i}(u_n)}{u_n}, \tag{1}$$

where f_{A_i} represents the membership function of fuzzy set A_i, and $f_{A_i}(u_1)$ represents the degree of belonging of u_1 to A_i, if u_i is a member of A_i.

Definition 3

The fuzzy logic relationship (FLR) between $F(t-1)$ and $F(t)$, indicated as $R(t, t-1)$, can be represented as $F(t-1) \rightarrow F(t)$. If $R(t, t-1)$ is independent of t for any given t value, then,

$$R(t, t-1) = R(t-1, t-2). \tag{2}$$

Definition 4

Let us assume that the value of F at time $(t-1)$ is denoted as A_i, and the value of F at time (t) isdenoted as A_j. The FLR, or Forward-Looking Relationship, is a term used to describe the connection between two consecutive observations, $F(t-1)$ and $F(t)$. It is denoted as $A_i \rightarrow A_j$, where A_i represents the left-hand side of the FLR and A_j represents the right-hand side.

Tree partition method (TPM)

TPM methodology is a partition approach introduced by *Alyousifi, Othman & Almohammedi (2021)* to linguistic partitioning that has been devised as a way for re-partitioning, utilizing the average length of the initial sub-intervals, particularly the average inter-quartile range. The method outlined in this study is characterized by its simplicity and effectiveness in determining the optimal interval lengths for obtaining an ideal partition of the universe of discourse (U), in contrast to clustering methods. In the majority of research conducted on fuzzy time series, U is partitioned into intervals of equal duration. However, fuzzy time series models may not provide aesthetically pleasant prediction outcomes in situations where the distribution of the universe of speech is not uniform. Therefore, the tree partition method is implemented in this study. Figure 1 shows the flowchart of TPM, and the steps of the suggested partition method are outlined in detail as follows:

First step: The universe of discourse, denoted as U, is defined.

Second step: This universe is partitioned into a minimum of three and a maximum of five equal-length intervals.

Third step: Calculate the frequency of each sub-interval based on the observations in data. Then, the average of sub-intervals' frequencies is calculated.

Fourth step: To optimize the granularity of the intervals, if any of the sub-intervals exceed the computed average length, a further division of the specific sub-interval into two equal parts is undertaken. This procedure is repeated until no individual sub-interval is longer than the average length at the outset.

Fifth step: In the end, the number of partitions where the smallest sub-interval is shorter than the average is taken into account for subsequent calculations and analysis.

The proposed model

In this study, a hybrid predictive model that integrates fuzzy time series Markov chain (FTSMC) with the TPM and the difference transformation data method for predicting
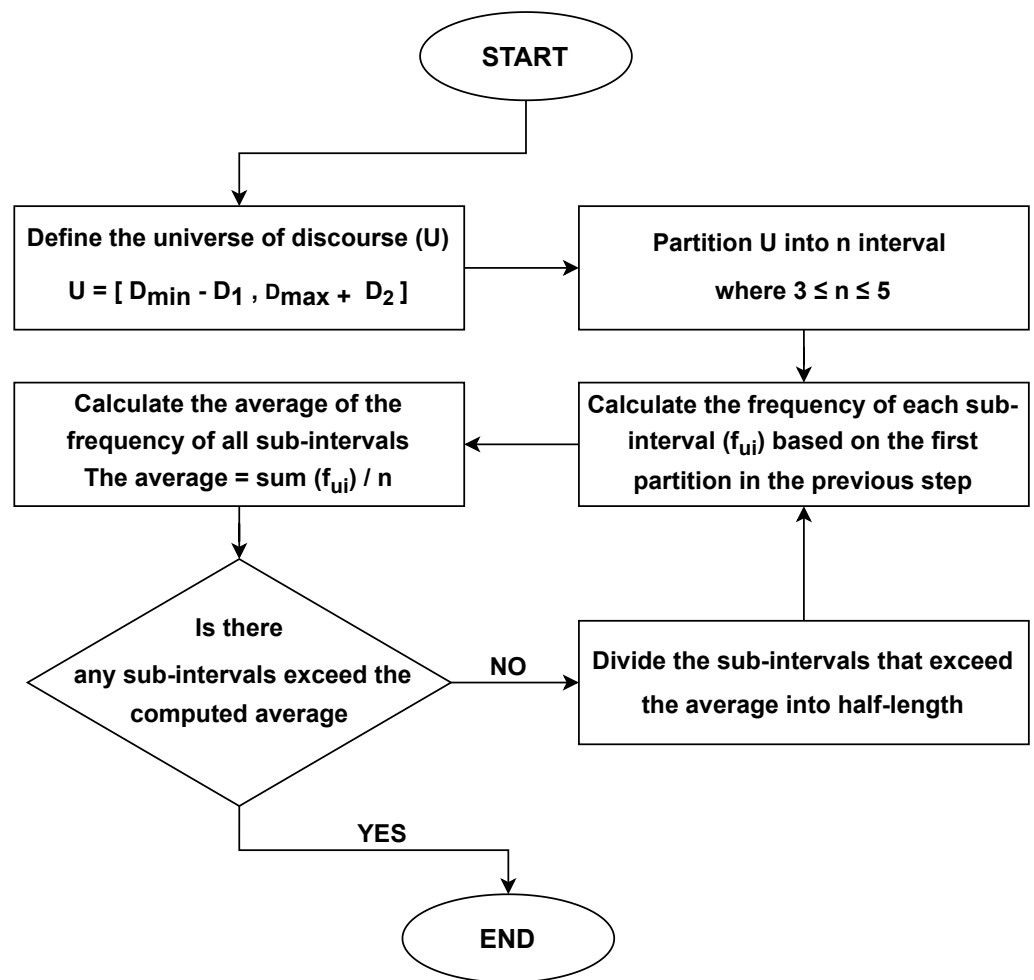
Khalil et al. (2024), *PeerJ Comput. Sci.*, DOI 10.7717/peerj-cs.2500

8/28

Figure 1 **The flowchart of TPM algorithm.**

Full-size ⬚ DOI: 10.7717/peerjcs.2500/fig-1

the total benefits (pensions) of the SIS in Egypt. The proposed framework utilizes the simplified arithmetic operations of the FTSMC model, as developed by *Tsaur (2012)*. The phases of the suggested model are illustrated in Fig. 2 and can be delineated in detail as follows:

Step 1:

This step collects the total benefits (pensions) of SIS in Egypt and preprocessing the data using the difference transformation data method and splitting data into training and testing datasets using the holdout (80:20) splitting approach.

Step 2:

The universe of discourse (U) is defined and established based on the dataset collection using Eq. (3) below:

$$U = [D_{min} - D_1, D_{max} + D_2],$$

(3)

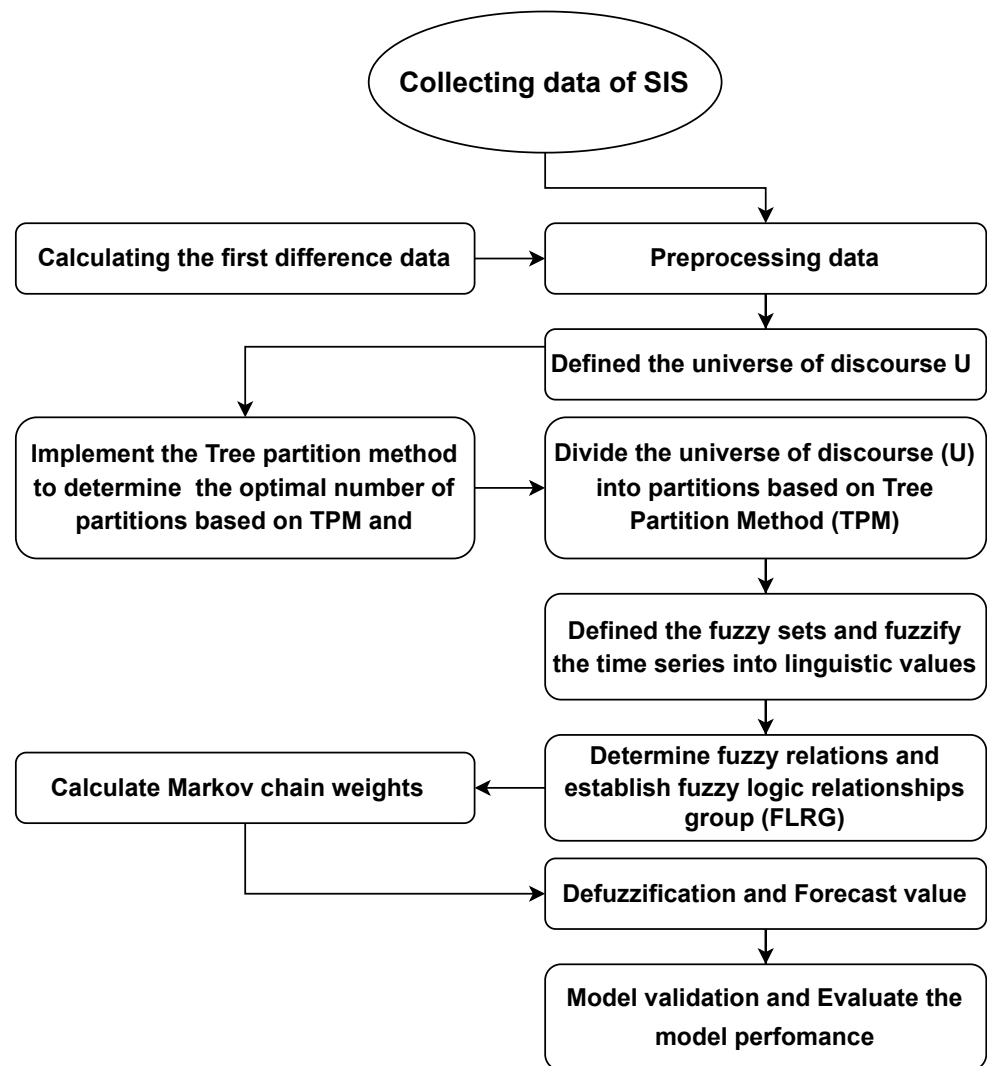Khalil et al. (2024), *PeerJ Comput. Sci.*, DOI 10.7717/peerj-cs.2500

9/28

Figure 2 **The proposed model's framework.**

where the minimum number of observations is D_{min} and the maximum is D_{max}. The numbers D_1 and D_2 are both positive integers.

Step 3:

In this step, the universe of discourse (U) is divided into numerous partitions, and the optimal number of partitions is determined using the TPM algorithm, as discussed in subsection (tree partition method).

Step 4:

The fuzzy set, denoted as A_i, is a collection of things that possess a continuous membership grade. This membership grade is determined using the universe of discourse (U), which is defined as $u_1, u_2, \ldots, u_n$, as specified in Eq. (1). Next, the fuzzy sets are established for each real value of the time series by assigning them fuzzy numbers

determined by the maximum membership value, following the intervals outlined in Step 3.

Step 5:

In this step, the fuzzy logical relationships (FLRs) are established between the linguistic values $F(t-1) = A_i$ and $F(t) = A_j$, denoted as $A_i \rightarrow A_j$, for all FLRs. For instance, FLRs such as $A_1 \rightarrow A_1$, $A_1 \rightarrow A_2$, and $A_1 \rightarrow A_3$ are determined. Subsequently, the fuzzy logical relation groups (FLRGs) are categorized and transformed into groups $A_1 \rightarrow A_1$, A_2 and A_3.

Step 6:

The Markov weights, also known as the transition probability matrix, are derived from the frequencies of the established FLRGs in Step 6. The matrix P is $P_{n \times n}$. The state transition probability, denoted as P_{ij}, is the probability of transitioning from state A_i to state A_j. This probability can be determined using Eq. (4) as follows:

$$P_{ij} = \frac{N_{ij}}{Ni}, \qquad i,j = 1,2,3,\dots,n, \tag{4}$$

where N_{ij} represents the count of transitions from state A_i to state A_j, while Ni represents the overall count of transitions in state A_i.

Step 7:

Determine the forecasted values through the process of forecasting. The calculation of forecasts considers the following rules. Two cases are being considered: one-to-one and one-to-many.

Rule 1, when the fuzzy logical relationship group of A_i exhibits a one-to-one relationship, meaning there is only one transition for A_i, the forecasting of $F(t)$ is determined by m_k, which represents the midpoint of u_k. Here, k ranges from 1 to n. This calculation may be performed using the Eq. (5) as follows:

$$F(t+1) = m_k.P_{ij} = m_k. \tag{5}$$

Rule 2. when A_i fuzzy logical relationship group contains more than one transition, also known as "many-to-one". To calculate the predicted value $F(t+1)$, the following Eq. (6) can be used if the state is A_i for the actual value $Y(t)$ at time t.

$$F(t+1) = m_1.P_{i1} + m_2.P_{i2} + \dots + Y(t).P_{ii} + \dots + m_k.P_{in}, \tag{6}$$

where $m_1, \dots, m_k$ represents the midpoint of $u_1, \dots, u_k$.

Step 8:

Calculate the final value of prediction after adding adjusted values by using Eq. (7) as shown below:

$$F^*(t+1) = F(t+1) \pm D, \tag{7}$$

where $F^*(t)$ represents the final prediction value after the adjustment, $F(t)$ is the prediction value before the adjustment. D is the adjustment value according to the differences to reduce the estimated error.

Table 1 The dataset description.

Year	Private sector	Government sector	Total benefits in millions	Year	Private sector	Government sector	Total benefits in millions
1976	58	121	179	1998	5003	5252	10255
1977	40	145	185	1999	5846	6034	11880
1978	52	162	214	2000	6798	6689	13487
1979	69	195	264	2001	7563	7742	15305
1980	105	243	348	2002	8314	9590	17904
1981	121	207	328	2003	8996	10791	19787
1982	352	401	753	2004	9756	12240	21996
1983	425	436	861	2005	10590	13988	24578
1984	485	557	1042	2006	12865	15441	28306
1985	553	641	1194	2007	13398	16867	30265
1986	631	779	1410	2008	15170	19311	34481
1987	688	841	1529	2009	18139	19488	37627
1988	824	1013	1837	2010	18456	22660	41116
1989	969	1275	2244	2011	22124	28724	50848
1990	1134	1417	2551	2012	28164	35568	63732
1991	1386	1714	3100	2013	35777	33640	69417
1992	1706	2037	3743	2014	43175	40558	83733
1993	2085	2353	4438	2015	51816	48554	100370
1994	2563	3019	5582	2016	59169	55495	114664
1995	3000	3587	6587	2017	67600	62300	129900
1996	3470	4107	7577	2018	77600	72800	150400
1997	4091	4810	8901	2019	92800	85700	178500

DATA & EVALUATION METHODS

Data description and preprocessing of pension payments in Egypt

The dataset presented in this study was acquired from authoritative sources, namely the National Fund for Social Security and the Ministry of Insurance and Social Affairs (MOISA) (https://www.capmas.gov.eg/Pages/Publications.aspx?page_id=5104&Year=23578). It encompasses data from a substantial 44-year period, from 1976 to 2019, with annual statistics specifically drawn from chapter/section 17 titled "Social Care" during the period 1979–2019. The records apply to the total benefits that are required by the social insurance system, which include contributions from the government sector, public and private sector. Table 1 and Fig. 3 show the total benefits of Egyptian social insurance system in million during this study period.

To preprocess the data, we utilize the first difference transformation technique and select the holdout (80:20) splitting approach. The division of data into training and testing sets is a fundamental and essential practice in statistical modelling. This process plays a critical role in evaluating the performance of the model and assuring its ability to effectively generalize to new, unseen data. By training the model on one subset of the data and then evaluating it on a distinct subset that it hasn't seen before (*Joseph, 2022*; *Kahloot & Ekler,*
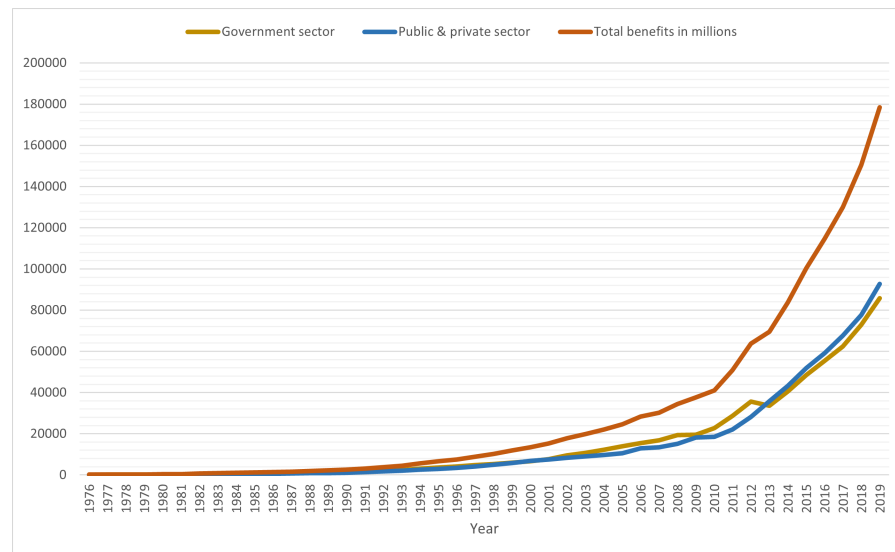
Figure 3 **The total benefit (in millions) graph.**

Full-size ⬜ DOI: 10.7717/peerjcs.2500/fig-3

2021). Based on the outcomes of our extensive review of the academic literature, it has been observed that the holdout (80:20) splitting methodology is a widely employed method for data partitioning, as previously employed in various studies (*Hasanov, Wolter & Glende, 2022*; *Joseph & Vakayil, 2022*), to divide our time series data into two independent sets. The initial dataset comprised 80% of 2009 and covering the time frame from 1976 to 2009 and was employed for the development of the models deployed in the training process. The second subset consisted of 20% of the data collected between 2010 and 2019. This subset was utilized to validate and assess the effectiveness of the proposed model.

In this study, we chose to model pension benefits using an additive approach, rather than the commonly used multiplicative model, for several key reasons. First, the historical data we analyzed demonstrated non-linear and often irregular growth patterns, particularly during periods of economic instability and policy changes. These variations were not adequately captured by multiplicative models, which tend to assume more consistent exponential growth. The additive model, on the other hand, allowed us to represent incremental changes more accurately, particularly during periods of slower or more erratic growth, where the assumption of constant percentage increases would lead to overestimations. Additionally, the additive approach provides greater flexibility in scenarios where pension increments are driven by fixed monetary adjustments rather than proportional increases, which was observed in some segments of our data. By employing an additive model, we aimed to reflect the real-world dynamics of the pension system under study, acknowledging that while multiplicative models may be more suitable for scenarios with consistent growth, the additive model better fits the specific characteristics of our dataset.

The evaluation metrics of forecasting models

The evaluation of a model's performance is dependent on its ability to generate forecast values that are closely aligned with the observed values for testing dataset. To evaluate the efficacy of the proposed model, four distinct forecast consistency measures were utilized. Four statistical criteria, namely mean absolute percent error (MAPE), Thiels' U statistic, and root mean square error (RMSE), were adopted as measures of forecasting accuracy to assess the models (*Khalil, Liu & Ali, 2024b*; *Khalil et al., 2022b*; *Rashidpoor Toochaei & Moeini, 2023*), as defined in Eqs. (8)–(10), respectively. The measures are presented as follows:

$$MAPE = \frac{100}{n} \sum_{t=1}^{n} \left| \frac{y_t - y_t^*}{y_t} \right|, \tag{8}$$

$$\text{Thiels}' \text{U} = \frac{\sqrt{\sum_{t=1}^{n}(y_t - y_t^*)^2}}{\sqrt{\sum_{t=1}^{n}(y_t)^2} + \sqrt{\sum_{t=1}^{n}(y_t^*)^2}}, \tag{9}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n}(y_t - y_t^*)^2}, \tag{10}$$

where y_t and y_t^* denote actual and predicted model data samples, respectively. n is the sample size. $\overline{y}_t$ is the average of actual data.

Evaluation metrics for the prediction interval of the proposed model

Prediction intervals are typically assessed by their ability to accurately encompass the target values (reliability) and their overall width (sharpness). Two widely used, independent metrics for evaluating precise prediction intervals are the prediction interval coverage probability (PICP), which measures reliability, and the mean prediction interval length (MPIL), which measures sharpness. These metrics are briefly described below.

Prediction interval coverage probability (PICP)

PICP measures the proportion of observed data points that fall within the prediction intervals generated by a model. It indicates how well the prediction intervals capture the true values of the target variable (*Adjenughwure & Papadopoulos, 2020*). The PICP is defined by using Eqs. (11) and (12) as follows:

$$PICP = \sum_{t=1}^{n} \frac{C_t}{n}, \tag{11}$$

where,

$$C_t = \begin{cases} 1 & if \, y_t^L \leq y_t \leq y_t^U \\ 0 & \text{otherwise} \end{cases}, \tag{12}$$

where C_t refers to the number of observations within the prediction intervals, and n refers to the total number of observations.

Mean prediction interval length (MPIL)

MPIL quantifies the average width of the prediction intervals for all observations. It reflects the precision of the predictions, with narrower intervals indicating greater prediction accuracy (*Tsao, Leu & Chou, 2021*). The MPIL is defined by using Eq. (13) as follows:

$$\text{MPIL} = \frac{1}{n}\sum_{t=1}^{n}(y_t^U - y_t^L), \tag{13}$$

where y_t^U and y_t^L are the upper and lower prediction limits for y_t.

RESULTS AND DISCUSSION

This section provides a comprehensive calculation of the forecasting methods used in the study, including the FTS and statistical employed. The aim is to provide a clear account of the methodologies employed for data analysis, ensuring transparency and precision.

Experimental setup

The experiments were conducted on a computer equipped with a 2.60 GHz Intel(R) Core (TM) i7-12700F CPU and 32 GB of RAM, running a 64-bit version of Windows 11. The framework was implemented using Python, with the dataset loaded *via* the Pandas data frame. FTS methods were executed using the PYFTS library. The source code and data associated with the proposed work have been made publicly accessible on the author's GitHub page (https://github.com/AhmedKhalil91/The-Benefits-of-Social-Insurance-System-prediction.git).

The proposed model application

The calculations and implementation of the suggested method are fully illustrated in this subsection. The procedure is as follows:

(**1**) Define the universe of discourse (U) from SIS data using Eq. (3).

$$U = [D_{min} - D_1, D_{max} + D_2]$$

$$U = [-20 - 5, 4216 + 4]$$

$$U = [-25, 4220].$$

(**2**) Divide the universe of discourse U based on TMP algorithm. Table 1 shows the implementation of the proposed partition method based on the step that is illustrated in the methodology section. Consequently, it has been determined that the frequencies of u_1, and u_2 are greater than the mean. Subsequently, the intervals are divided into halves. Then, we checked that again until we found there was not any sub-interval larger than the average. According to the result of TPM implementation, the final number of partitions of U is nine sub-intervals as shown in Table 2.

(**3**) Convert each time series' real value into a fuzzy number using the intervals you just put up as a guide, and then define the corresponding fuzzy sets using the maximum

Table 2   TPM training dataset procedures description.

1st partition		2nd partition		3rd partition		4th partition	
Interval	Frequency	Interval	Frequency	Interval	Frequency	Interval	Frequency
[−25, 824]	**18**	[−25, 339.5]	**13**	[−25, 187.25]	**10**	[−25, 81.125]	5
[824, 1673]	**7**	[399.5, 824]	5	[187.25, 399.5]	3	[81.125, 187.25]	5
[1673, 2522]	4	[824, 1248.5]	3	[399.5, 824]	5	[187.25, 399.5]	3
[2522, 3371]	3	[1248.5, 1673]	4	[824, 1248.5]	3	[399.5, 824]	5
[3371, 4220]	2	[1673, 2522]	4	[1248.5, 1673]	4	[824, 1248.5]	3
		[2522, 3371]	3	[1673, 2522]	4	[1248.5, 1673]	4
		[3371, 4220]	2	[2522, 3371]	3	[1673, 2522]	4
				[3371, 4220]	2	[2522, 3371]	3
						[3371, 4220]	2
Repartition	>6.8	Repartition	>6.8	Repartition	>6.8	Repartition	>6.8

The average of frequency for the first partition = $\frac{18+7+4+3+2}{5} = 6.8$

If the frequency of all sub intervals is less than 6.8, then the partitioning process will end. Otherwise, the process of repartition should be continued

Notes.
The bold numbers represent the interval frequency that exceeds the average frequency of all sub-intervals.

Table 3   The fuzzy-numbered intervals.

Number of intervals	Interval u_i	Mid-point of interval (m_i)	Interval code	Fuzzy number
1	[−25, 81.125]	28.0625	u_1	A_1
2	[81.125, 187.25]	134.1875	u_2	A_2
3	[187.25, 399.5]	293.375	u_3	A_3
4	[399.5, 824]	611.75	u_4	A_4
5	[824, 1248.5]	1036.25	u_5	A_5
6	[1248.5, 1673]	1460.75	u_6	A_6
7	[1673, 2522]	2097.5	u_7	A_7
8	[2522, 3371]	2946.5	u_8	A_8
9	[3371, 4220]	3795.5	u_9	A_9

membership value. Table 3 shows the fuzzy sets with the mid-point of each interval and its fuzzy numbers. The linguistic time series of maximum membership is shown in Table 4.

(**4**) The fuzzy logic relationships (FLRs) are established based on the training dataset, for instance, FLRs such as $A_1 \rightarrow A_1$, $A_1 \rightarrow A_2$, as shown in Table 5. Subsequently, the fuzzy logical relation groups (FLRGs) are categorized and transformed into groups $A_1 \rightarrow A_1$, A_2 and A_3, as shown in Table 6.

(**5**) The transition probability matrix is calculated based on FLRGs that were created in the previous step. Table 7 shows the Markov transition probability matrix based on the probability of transitioning calculated by the frequencies of the established FLRGs using Eq. (4).

(**6**) Calculate the forecasted values by using the rules that are illustrated in the proposed methodology, with Eqs. (5) and (6) based on Markov transition probability matrix. For example, the forecast value for year (1978) is calculated as following:

Table 4 Linguistic time series.

No	Linguistic time series values A_i
1	$A_1 = \frac{1}{u_1} + \frac{0.5}{u_2} + \frac{0}{u_3} + \frac{0}{u_4} + \frac{0}{u_5} + \frac{0}{u_6} + \frac{0}{u_7} + \frac{0}{u_8} + \frac{0}{u_9}$
2	$A_2 = \frac{0.5}{u_1} + \frac{1}{u_2} + \frac{0.5}{u_3} + \frac{0}{u_4} + \frac{0}{u_5} + \frac{0}{u_6} + \frac{0}{u_7} + \frac{0}{u_8} + \frac{0}{u_9}$
3	$A_3 = \frac{0}{u_1} + \frac{0.5}{u_2} + \frac{1}{u_3} + \frac{0.5}{u_4} + \frac{0}{u_5} + \frac{0}{u_6} + \frac{0}{u_7} + \frac{0}{u_8} + \frac{0}{u_9}$
4	$A_4 = \frac{0}{u_1} + \frac{0}{u_2} + \frac{0.5}{u_3} + \frac{1}{u_4} + \frac{0.5}{u_5} + \frac{0}{u_6} + \frac{0}{u_7} + \frac{0}{u_8} + \frac{0}{u_9}$
5	$A_5 = \frac{0}{u_1} + \frac{0}{u_2} + \frac{0}{u_3} + \frac{0.5}{u_4} + \frac{1}{u_5} + \frac{0.5}{u_6} + \frac{0}{u_7} + \frac{0}{u_8} + \frac{0}{u_9}$
6	$A_6 = \frac{0}{u_1} + \frac{0}{u_2} + \frac{0}{u_3} + \frac{0}{u_4} + \frac{0.5}{u_5} + \frac{1}{u_6} + \frac{0.5}{u_7} + \frac{0}{u_8} + \frac{0}{u_9}$
7	$A_7 = \frac{0}{u_1} + \frac{0}{u_2} + \frac{0}{u_3} + \frac{0}{u_4} + \frac{0}{u_5} + \frac{0.5}{u_6} + \frac{1}{u_7} + \frac{0.5}{u_8} + \frac{0}{u_9}$
8	$A_8 = \frac{0}{u_1} + \frac{0}{u_2} + \frac{0}{u_3} + \frac{0}{u_4} + \frac{0}{u_5} + \frac{0}{u_6} + \frac{0.5}{u_7} + \frac{1}{u_8} + \frac{0.5}{u_9}$
9	$A_9 = \frac{0}{u_1} + \frac{0}{u_2} + \frac{0}{u_3} + \frac{0}{u_4} + \frac{0}{u_5} + \frac{0}{u_6} + \frac{0}{u_7} + \frac{0.5}{u_8} + \frac{1}{u_9}$

$$F(1978) = Y(t).p_{11} + m_2.p_{12} + m_4.p_{14}$$
$$= 6*0.6 + 134.187*0.2 + 611.75*0.2 = 152.78$$

(7) Utilize Eq. (7) to compute the final predicted values. For instance, the subsequent formula is utilized to compute the final forecast value for total benefits in the year 1978:

$$F^*(1978) = 152.78 + 23 = 175.78$$

Then, the forecasted total benefits $(1978) = 175.78 + 185 = 360.78$

Table 8 illustrates the forecasted values for the total benefits of Egyptian social insurance system in million. The fitted value of the proposed model for training and testing datasets is shown in Fig. 4.

The proposed model's evaluation for the prediction interval

The proposed model's outputs are evaluated for training and testing data using the established metrics of Prediction Interval Coverage Probability (PICP) and mean prediction interval length (MPIL) to assess the quality of the prediction intervals and compared to other traditional fuzzy and statistical time series models. A 95% prediction interval is used as the baseline, with fuzzy numbers representing all intervals from 95% down to 0%. As presented in Table 9, the PICP is approximately 95%, suggesting that the prediction intervals are well-calibrated with near-perfect coverage. However, the MPIL of 424.5 indicates that the model generates relatively wide intervals, reflecting a substantial degree of uncertainty in the data. This result suggests that the proposed model effectively balances coverage accuracy and interval precision.

The results of the proposed model's evaluation

In this subsection, the proposed model's performance is evaluated using three statistical criteria, namely mean absolute percent error (MAPE), Thiels' U statistic, and root mean square error (RMSE) for the training and testing datasets. In addition, to validate the proposed model, we introduce a comparison of the proposed model with existing traditional fuzzy time series models that were proposed by *Song & Chissom (1994)*, *Chen (1996)*, *Yu*

Table 5 Dataset training fuzzy number and fuzzy logic relationships.

Year	Total benefits in millions	Difference	Fuzzy number	Fuzzy logic relationships
1976	179	0	A_1	–
1977	185	6	A_1	$A_1 \rightarrow A_1$
1978	214	29	A_1	$A_1 \rightarrow A_1$
1979	264	50	A_1	$A_1 \rightarrow A_1$
1980	348	84	A_2	$A_1 \rightarrow A_2$
1981	328	-20	A_1	$A_2 \rightarrow A_1$
1982	753	425	A_4	$A_1 \rightarrow A_4$
1983	861	108	A_2	$A_4 \rightarrow A_2$
1984	1042	181	A_2	$A_2 \rightarrow A_2$
1985	1194	152	A_2	$A_2 \rightarrow A_2$
1986	1410	216	A_3	$A_2 \rightarrow A_3$
1987	1529	119	A_2	$A_3 \rightarrow A_2$
1988	1837	308	A_3	$A_2 \rightarrow A_3$
1989	2244	407	A_4	$A_3 \rightarrow A_4$
1990	2551	307	A_3	$A_4 \rightarrow A_3$
1991	3100	549	A_4	$A_3 \rightarrow A_4$
1992	3743	643	A_4	$A_4 \rightarrow A_4$
1993	4438	695	A_4	$A_4 \rightarrow A_4$
1994	5582	1144	A_5	$A_4 \rightarrow A_5$
1995	6587	1005	A_5	$A_5 \rightarrow A_5$
1996	7577	990	A_5	$A_5 \rightarrow A_5$
1997	8901	1324	A_6	$A_5 \rightarrow A_6$
1998	10255	1354	A_6	$A_6 \rightarrow A_6$
1999	11880	1625	A_6	$A_6 \rightarrow A_6$
2000	13487	1607	A_6	$A_6 \rightarrow A_6$
2001	15305	1818	A_7	$A_6 \rightarrow A_7$
2002	17904	2599	A_8	$A_7 \rightarrow A_8$
2003	19787	1818	A_7	$A_8 \rightarrow A_7$
2004	21996	1818	A_7	$A_7 \rightarrow A_7$
2005	24578	2599	A_8	$A_7 \rightarrow A_8$
2006	28306	3728	A_9	$A_8 \rightarrow A_9$
2007	30265	1959	A_7	$A_9 \rightarrow A_7$
2008	34481	4216	A_9	$A_7 \rightarrow A_9$
2009	37627	3146	A_8	$A_9 \rightarrow A_8$

(2005), and *Cheng Chen & Chiang (2006)*. It is important to note that all the fuzzy time series models in the comparison, including the proposed model, were evaluated based on the same number of intervals; specifically, nine intervals. This uniformity ensures that the comparative results are fair and allows for a direct assessment of the relative performance of the models under the same interval partitioning structure.

Table 10 reveals that the proposed model had the best overall performance in predicting the total benefits of social insurance system in Egypt, compared to other traditional fuzzy

Khalil et al. (2024), *PeerJ Comput. Sci.*, DOI 10.7717/peerj-cs.2500

18/28

Table 6 Fuzzy logic relationships groups (FLRG) for training dataset.

No	FLRG
G1	$A_1 \rightarrow (3)A_1, (1)A_2, (1)A_4$
G2	$A_2 \rightarrow (1)A_1, (2)A_2, (2)A_3$
G3	$A_3 \rightarrow (1)A_2, (2)A_4$
G4	$A_4 \rightarrow (1)A_2, (1)A_3, (2)A_4, (1)A_5$
G5	$A_5 \rightarrow (2)A_5, (1)A_6$
G6	$A_6 \rightarrow (3)A_6, (1)A_7$
G7	$A_7 \rightarrow (1)A_7, (2)A_8, (1)A_9$
G8	$A_8 \rightarrow (1)A_7, (1)A_9$
G9	$A_9 \rightarrow (1)A_7, (1)A_8$

Table 7 Markov transition probability matrix and count matrix based on FLRGs.

		1	2	3	4	5	6	7	8	9
1	F	**3**	**1**	0	**1**	0	0	0	0	0
	P	**0.6**	**0.2**	0	**0.2**	0.0	0.0	0.0	0.0	0.0
2	F	**1**	**2**	**2**	0	0	0	0	0	0
	P	**0.2**	**0.4**	**0.4**	0.0	0.0	0.0	0.0	0.0	0.0
3	F	0	**1**	0	**2**	0	0	0	0	0
	P	0.0	**0.33**	0.0	**0.67**	0.0	0.0	0.0	0.0	0.0
4	F	0	**1**	**1**	**2**	**1**	0	0	0	0
	P	0.0	**0.2**	**0.2**	**0.4**	**0.2**	0.0	0.0	0.0	0.0
5	F	0	0	0	0	**2**	**1**	0	0	0
	P	0.0	0.0	0.0	0.0	**0.67**	**0.33**	0.0	0.0	0.0
6	F	0	0	0	0	0	**3**	**1**	0	0
	P	0.0	0.0	0.0	0.0	0.0	**0.75**	**0.25**	0.0	0.0
7	F	0	0	0	0	0	0	**1**	**2**	**1**
	P	0.0	0.0	0.0	0.0	0.0	0.0	**0.25**	**0.5**	**0.25**
8	F	0	0	0	0	0	0	**1**	0	**1**
	P	0.0	0.0	0.0	0.0	0.0	0.0	**0.5**	0.0	**0.5**
9	F	0	0	0	0	0	0	**1**	**1**	0
	P	0.0	0.0	0.0	0.0	0.0	0.0	**0.5**	**0.5**	0.0

Notes.

The bold numbers indicate the Markov transition probability for each interval based on FLRGs.

time series models, as it had lower values of MAPE, RMSE, and Thiels' U for training dataset with values of 11.8, 472.67, and 0.016 respectively. As well as, for testing dataset with values of 10.65, 12999.6, and 0.064 respectively, as shown in Fig. 5.

The evaluation results indicate that the proposed model significantly outperforms conventional models like linear regression, ARIMA, and exponential smoothing in both training and testing datasets as shown in Table 11. With the lowest RMSE (472.67 for training and 12,999.64 for testing), MAPE (10.653% for testing), and Thiels' U (0.064 for testing), the proposed model demonstrates superior accuracy and generalization ability. In contrast, linear regression exhibits the poorest performance, particularly in the testing set,

Table 8 Dataset training fuzzy number and fuzzy logic relationships.

Year	Total benefits	Difference	Forecast value	Adjusted value	Adjusted Forecast value	Forecasted Total benefits
1976	179	0	–	–	–	–
1977	185	6	149.18	6	149.18 + 6 = 155.18	155.18 + 179 = 334.18
1978	214	29	152.78	23	175.78	360.78
1979	264	50	166.58	21	187.58	401.58
1980	348	84	142.96	34	176.96	440.96
1981	328	−20	199.58	−104	95.58	443.58
1982	753	425	104.76	445	549.76	877.76
1983	861	108	292.96	−317	−24.03	728.96
1984	1042	181	166.16	73	239.16	1100.16
1985	1194	152	195.36	−29	166.36	1208.36
1986	1410	216	454.15	64	518.15	1712.15
1987	1529	119	209.36	−97	112.36	1522.36
1988	1837	308	454.15	189	643.15	2172.15
1989	2244	407	415.96	99	514.96	2351.96
1990	2551	307	454.15	-100	354.15	2598.15
1991	3100	549	415.56	242	657.56	3208.56
1992	3743	643	512.36	94	606.36	3706.36
1993	4438	695	549.96	52	601.96	4344.96
1994	5582	1144	947.69	449	1396.69	5834.69
1995	6587	1005	1248.50	−139	1106.53	6688.53
1996	7577	990	1155.40	−15	1140.40	7727.40
1997	8901	1324	1266.88	334	1600.88	9177.88
1998	10255	1354	1517.38	30	1547.38	10448.38
1999	11880	1625	1539.88	271	1810.88	12065.88
2000	13487	1607	1743.13	−18	1725.13	13605.13
2001	15305	1818	2823.88	211	3034.88	16521.88
2002	17904	2599	2946.50	781	3727.50	19032.50
2003	19787	1818	3071.88	−781	2290.88	20194.88
2004	21996	1818	2876.63	0	2876.63	22663.63
2005	24578	2599	2946.50	781	3727.50	25723.50
2006	28306	3728	2522	1129	3651	28229
2007	30265	1959	3354.13	−1769	1585.13	29891.13
2008	34481	4216	2522	2257	4779	35044
2009	37627	3146	2946.50	−1070	1876.50	36357.50

with an RMSE of 78,080.63. While ARIMA and exponential smoothing show comparable results, their errors are notably higher than the proposed model's, highlighting the proposed model's reliability and predictive efficiency.

Furthermore, Table 11 emphasizes the broader implications of this study for the insurance sector, particularly in enhancing operational efficiency through an evaluation of Egypt's social insurance system and the development of precise analytical and predictive
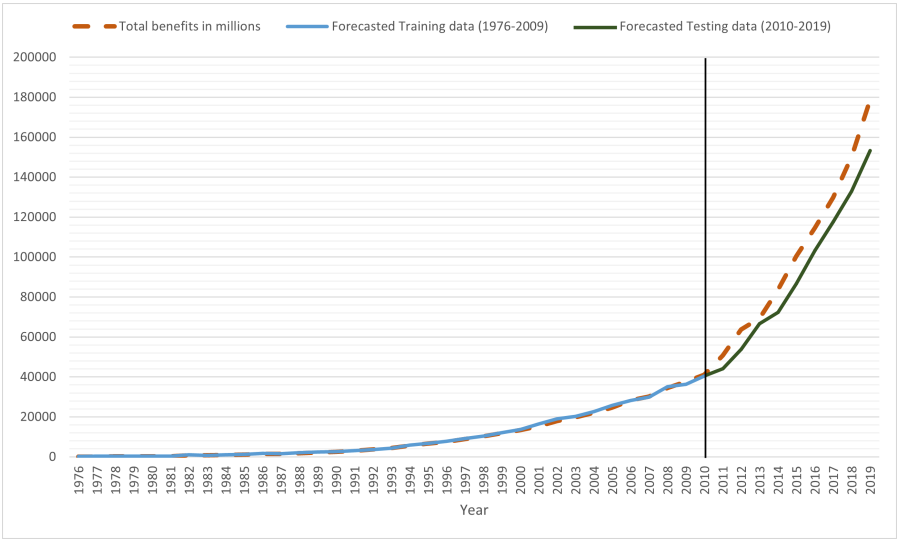
Figure 4 Plots of fit proposed model for training and testing datasets with the actual values of total benefits.

Full-size ⬜ DOI: 10.7717/peerjcs.2500/fig-4

Table 9 Evaluating model prediction intervals.

	PICP	MPIL
Linear regression	42.86%	1119.01
ARIMA	60%	2250.63
Exponential Smooth	60%	2250.63
Song & Chissom (1994)	100%	471.67
Chen (1996)	100%	471.67
Yu (2005)	100%	471.67
Cheng, Chen & Chiang (2006)	100%	471.67
The proposed model	100%	424.5

Table 10 The evaluation of the proposed model with traditional fuzzy time series models.

Model	Evaluation metric						Rank
	Training dataset			Testing dataset			
	RMSE	MAPE%	Thiels' U	RMSE	MAPE%	Thiels' U	
Song & Chissom (1994)	879.41	17.31	0.033	13314.2	11.25	0.069	5th
Chen (1996)	553.26	16.76	0.025	13306.6	11.05	0.068	4th
Yu (2005)	501.50	12.92	0.02	13305.5	10.98	0.067	2nd
Cheng, Chen & Chiang (2006)	504.14	13.24	0.023	13306.6	11.05	0.068	3rd
The proposed model	**472.67**	**11.80**	**0.016**	**12999.6**	**10.65**	**0.064**	1st

Notes.
The bold numbers denote the lowest values of the evaluation metrics achieved by the proposed model.

methods. The study's focus on the FTS model, known for its ability to capture complex, non-linear patterns and uncertainties, further highlights its effectiveness in the Egyptian
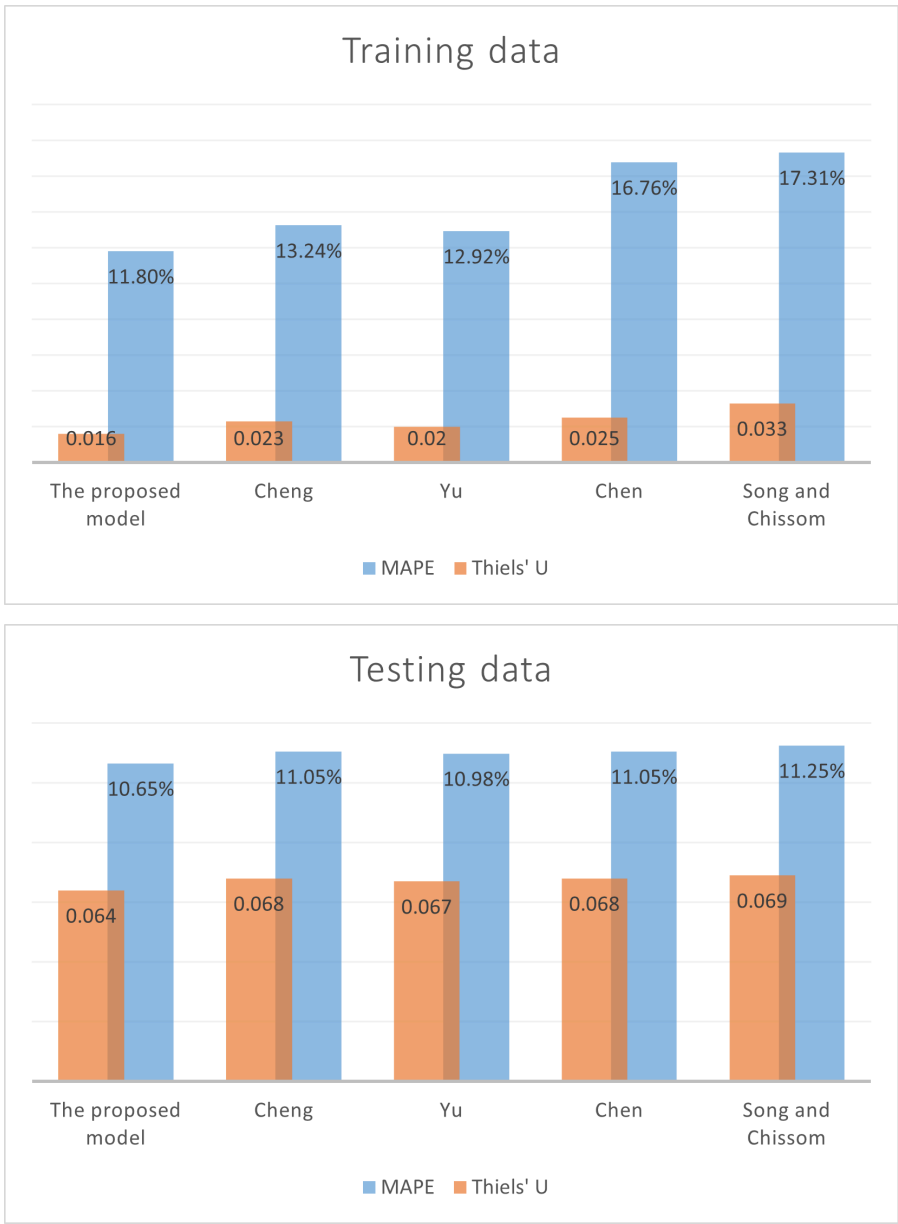
Figure 5 Plots of the evaluation metrics for training and testing datasets for models.
Full-size ⬜ DOI: 10.7717/peerjcs.2500/fig-5

insurance market, where conventional statistical approaches often face challenges due to data constraints. The insights gained from this research provide valuable tools for insurers aiming to improve risk management and profitability, while also holding relevance for stakeholders such as regulators and consumers, potentially leading to more accurate insurance pricing and improved policy frameworks.

Table 11 The evaluation of the proposed model with conventional statistical models.

Model	Evaluation metric					
	Training dataset			Testing dataset		
	RMSE	MAPE%	Thiels' U	RMSE	MAPE%	Thiels' U
Linear Regression	4745.73	452.78	0.169	78080.63	62.82	0.564
ARIMA	805.08	12.03	0.028	52078.54	34.13	0.314
Exponential Smooth	551.10	13.86	0.019	52503.97	34.46	0.317
The proposed model	**472.67**	**11.80**	**0.016**	**12999.64**	**10.653**	**0.064**

Notes.
The bold numbers also represent the lowest values of the evaluation metrics achieved by the proposed model.

CONCLUSIONS

In conclusion, enhancing the accuracy of predicting social insurance benefits in Egypt has far-reaching implications for countries around the world. Social insurance systems are vital components of social welfare and economic stability, serving as a safety net for individuals and providing economic security during times of need. Therefore, improving prediction accuracy in this sector is essential for the efficient allocation of resources and the sustainability of these systems. Accurate forecasting enables policymakers and insurance providers to anticipate future needs, allocate resources more effectively, and adjust policies proactively, ensuring that the system remains robust and responsive to changing social and economic conditions.

Integrating actuarial methods with predictive analytics enhances understanding and anticipation of future trends, supporting better decision-making and strengthening social insurance systems worldwide. Recently, hybrid predictive methods have garnered significant attention. This research introduces a novel hybrid predictive model combining the fuzzy time series Markov chain, TPM, and the difference transformation data method to forecast the total benefits (pensions) within Egypt's social insurance system. The datasets for this study were sourced from the National Fund for Social Security and the Ministry of Insurance and Social Affairs (MOISA). The main goal of this model is to achieve high forecast accuracy, as demonstrated by empirical evidence. The proposed model's key advantage lies in optimizing the partitioning process, thereby reducing computational complexity and load during forecasting.

The model's performance was assessed using standard statistical evaluation metrics such as RMSE, MAPE, and Thiel's U statistic, on both training and testing datasets. The results were compared with traditional fuzzy time series models proposed by *Song & Chissom (1994)*, *Chen (1996)*, *Yu (2005)*, and *Cheng, Chen & Chiang (2006)*. Findings indicated that the proposed hybrid model significantly improved forecasting accuracy by minimizing randomness and variability in the data. The model proved to be more robust and outperformed traditional fuzzy time series approaches. This study contributes to the literature by presenting a mathematical model applicable to estimating crucial values in the insurance sector. It also opens up several avenues for future research on enhancing predictive accuracy in social insurance systems.

The main limitations of our study include the use of a first-order FTS model, which, while simpler and less prone to overfitting, may limit the complexity of the predictions. Additionally, the dataset was constrained to 44 years of data, focusing exclusively on the Egyptian market, which may restrict the generalizability of the findings. The interpretability of the fuzzy rules is also a potential limitation, as understanding the underlying reasons for the model's predictions could be improved.

Future research should explore alternative partition methods and higher-order or hybrid FTS models, potentially integrating techniques such as ARIMA or exponential smoothing, to enhance the interpretability and robustness of the model. Applying these methods to datasets from different countries would also help increase the model's generalizability. Furthermore, the approach presented in this study could be extended to predict other critical ratios within the insurance sector, potentially offering strong performance in forecasting time series data in this field.

ADDITIONAL INFORMATION AND DECLARATIONS

Competing Interests
The authors declare there are no competing interests.

Author Contributions
- Ahmed Abdelreheem Khalil conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Mohamed Abdelaziz Mandour conceived and designed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
- Ahmed Ali performed the experiments, performed the computation work, authored or reviewed drafts of the article, and approved the final draft.

Data Availability
The following information was supplied regarding data availability:
The raw data and code are available in the Supplementary File.
The dataset is available at GitHub and Zenodo:
- https://github.com/AhmedKhalil91/The-Benefits-of-Social-Insurance-System-prediction.git

- Khalil, A. (2024). The Benefits of Social Insurance System prediction using a hybrid Fuzzy Time Series method. Zenodo. https://doi.org/10.5281/zenodo.13987690.

The dataset used in the research was acquired from the National Fund for Social Security and the Ministry of Insurance and Social Affairs (MOISA): https://www.capmas.gov.eg/Pages/Publications.aspx?page_id=5104&Year=23578.

REFERENCES

Adjenughwure K, Papadopoulos B. 2020. Fuzzy-statistical prediction intervals from crisp regression models. *Evolving Systems* **11(2)**:201–213 DOI 10.1007/s12530-019-09285-6.

Alyousifi Y, Othman M, Almohammedi AA. 2021. A novel stochastic fuzzy time series forecasting model based on a new partition method. *IEEE Access* **9**:80236–80252 DOI 10.1109/ACCESS.2021.3084048.

Alyousifi Y, Othman M, Husin A, Rathnayake U. 2021. A new hybrid fuzzy time series model with an application to predict PM10 concentration. *Ecotoxicology and Environmental Safety* **227**:112875 DOI 10.1016/j.ecoenv.2021.112875.

Alyousifi Y, Othman M, Sokkalingam R, Faye I, Silva PCL. 2020. Predicting daily air pollution index based on fuzzy time series markov chain model. *Symmetry* **12(2)**:293 DOI 10.3390/sym12020293.

Arnita OOO, Afnisah N, Marpaung F. 2020. A comparison of the fuzzy time series methods of Chen, Cheng and Markov chain in predicting rainfall in medan. *Journal of Physics: Conference Series* **1462(1)**:012044 DOI 10.1088/1742-6596/1462/1/012044.

Belloni M, Maccheroni C. 2013. Actuarial fairness when longevity increases: an evaluation of the italian pension system. *The Geneva Papers on Risk and Insurance - Issues and Practice* **38(4)**:638–674 DOI 10.1057/gpp.2013.27.

Booth H, Tickle L. 2008. Mortality modelling and forecasting: a review of methods. *Annals of Actuarial Science* **3(1–2)**:3–43 DOI 10.1017/S1748499500000440.

Chen M-Y, Chen B-T. 2015. A hybrid fuzzy time series model based on granular computing for stock price forecasting. *Information Sciences* **294**:227–241 DOI 10.1016/j.ins.2014.09.038.

Chen S-M. 1996. Forecasting enrollments based on fuzzy time series. *Fuzzy Sets and Systems* **81(3)**:311–319 DOI 10.1016/0165-0114(95)00220-0.

Chen S-M, Hwang J-R. 2000. Temperature prediction using fuzzy time series. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **30(2)**:263–275 DOI 10.1109/3477.836375.

Chen Z, Xu Y, Tian Z, Meng X. 2022. The impact of social insurance law on corporate innovation: evidence from a quasi-natural experiment. *Economic Modelling* **111**:105830 DOI 10.1016/j.econmod.2022.105830.

Cheng C-H, Chen T-L, Chiang C-H. 2006. Trend-weighted fuzzy time-series model for TAIEX forecasting. In: King I, Wang J, Chan L-W, Wang D, eds. *Neural information processing.* Berlin, Heidelberg: Springer, 469–477.

Khalil et al. (2024), *PeerJ Comput. Sci.*, DOI 10.7717/peerj-cs.2500

25/28

De Andrés-Sánchez J, Puchades LG-V. 2019. A fuzzy-random extension of the lee—carter mortality prediction model. *International Journal of Computational Intelligence Systems* **12(2)**:775–794 DOI 10.2991/ijcis.d.190626.001.

Hadwan M, Al-Maqaleh BM, Al-Badani FN, Khan RU, Al-Hagery MA. 2022. A hybrid neural network and Box–Jenkins models for time series forecasting. *Computers, Materials and Continua* **70(3)**:4829–4845 DOI 10.32604/cmc.2022.017824.

Hasanov M, Wolter M, Glende E. 2022. Time series data splitting for short-term load forecasting. In: *PESS PELSS 2022; power and energy student summit*. Piscataway: IEEE.

Hassani H, Unger S, Beneki C. 2020. Big data and actuarial science. *Big Data and Cognitive Computing* **4(4)**:40 DOI 10.3390/bdcc4040040.

Hong WH, Yap JH, Selvachandran G, Thong PH, Son LH. 2021. Forecasting mortality rates using hybrid Lee–Carter model, artificial neural network and random forest. *Complex & Intelligent Systems* **7(1)**:163–189 DOI 10.1007/s40747-020-00185-w.

Huarng K. 2001a. Effective lengths of intervals to improve forecasting in fuzzy time series. *Fuzzy Sets and Systems* **123(3)**:387–394 DOI 10.1016/S0165-0114(00)00057-9.

Huarng K. 2001b. Heuristic models of fuzzy time series for forecasting. *Fuzzy Sets and Systems* **123(3)**:369–386 DOI 10.1016/S0165-0114(00)00093-2.

Huarng K, Yu TH-K. 2006. Ratio-based lengths of intervals to improve fuzzy time series forecasting. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **36(2)**:328–340 DOI 10.1109/TSMCB.2005.857093.

Hwang J-R, Chen S-M, Lee C-H. 1998. Handling forecasting problems using fuzzy time series. *Fuzzy Sets and Systems* **100(1)**:217–228 DOI 10.1016/S0165-0114(97)00121-8.

Ingale KK, Paluri RA. 2023. Retirement planning—a systematic review of literature and future research directions. *Management Review Quarterly* DOI 10.1007/s11301-023-00377-x.

Jilani TA, Burney SMA, Ardil C. 2008. Multivariate high order fuzzy time series forecasting for car road accidents. *International Journal of Computer and Information Engineering* **2(6)**:2038–2043.

Joseph VR. 2022. Optimal ratio for data splitting. *Statistical Analysis and Data Mining: the ASA Data Science Journal* **15(4)**:531–538 DOI 10.1002/sam.11583.

Joseph VR, Vakayil A. 2022. SPlit: an optimal method for data splitting. *Technometrics* **64(2)**:166–176 DOI 10.1080/00401706.2021.1921037.

Kahloot KM, Ekler P. 2021. Algorithmic splitting: a method for dataset preparation. *IEEE Access* **9**:125229–125237 DOI 10.1109/ACCESS.2021.3110745.

Kassem NM. 2021. Roles, rules, and controls: an analytical review of the governance of social protection in Egypt. Master's thesis, American University of Cairo, Cairo, Egypt.

Khairalla M, Al-Jallad N. 2017. Hybrid forecasting scheme for financial time-series data using neural network and statistical methods. *International Journal of Advanced Computer Science and Applications* **8(9)**:319–327 DOI 10.14569/IJACSA.2017.080945.

Khalil AA, Liu Z, Ali AA. 2022a. Using an adaptive network-based fuzzy inference system model to predict the loss ratio of petroleum insurance in Egypt. *Risk Management and Insurance Review* **25(1)**:e2348413 DOI 10.1111/rmir.12200.

Khalil AA, Liu Z, Ali A. 2024a. Enhancing operational efficiency of insurance companies: a fuzzy time series approach to loss ratio forecasting in the Egyptian market. *Journal of Business Analytics* **7**:1–19 DOI 10.1080/2573234X.2024.2393609.

Khalil A, Liu Z, Ali A. 2024b. Precision in insurance forecasting: enhancing potential with ensemble and combination models based on the adaptive neuro-fuzzy inference system in the Egyptian Insurance Industry. *Applied Artificial Intelligence* **38(1)**:2348413 DOI 10.1080/08839514.2024.2348413.

Khalil AA, Liu Z, Salah A, Fathalla A, Ali A. 2022b. Predicting insolvency of insurance companies in egyptian market using bagging and boosting ensemble techniques. *IEEE Access* **10**:117304–117314 DOI 10.1109/ACCESS.2022.3210032.

Khozani ZS, Banadkooki FB, Ehteram M, Ahmed AN, El-Shafie A. 2022. Combining autoregressive integrated moving average with Long Short-Term Memory neural network and optimisation algorithms for predicting ground water level. *Journal of Cleaner Production* **348**:131224 DOI 10.1016/j.jclepro.2022.131224.

Lee W-J, Jung H-Y, Yoon JH, Choi SH. 2017. A novel forecasting method based on F-transform and fuzzy time series. *International Journal of Fuzzy Systems* **19(6)**:1793–1802 DOI 10.1007/s40815-017-0354-6.

Liu S, Zhuo Y, Shen X, Cai M, Yang Y. 2023. The impact of declined social insurance contribution rate on enterprise total factor productivity: evidence from China. *International Review of Financial Analysis* **87**:102624 DOI 10.1016/j.irfa.2023.102624.

Loewe M, Westemeier L. 2018. Social insurance reforms in Egypt: needed, belated, flopped. *Available at* https://pomeps.org/social-insurance-reforms-in-egypt-needed-belated-flopped (accessed on 24 October 2018).

Mao S, Xiao F. 2019. Time series forecasting based on complex network analysis. *IEEE Access* **7**:40220–40229 DOI 10.1109/ACCESS.2019.2906268.

Rahim NF, Othman M, Sokkalingam R, Kadir EA. 2018. Forecasting crude palm oil prices using fuzzy rule-based time series method. *IEEE Access* **6**:32216–32224 DOI 10.1109/ACCESS.2018.2846809.

Raouf R, Elsaieed S. 2022. Using fuzzy time series models to estimate the cost of benefits of egyptian social insurance system. *The International Journal of Fuzzy Logic and Intelligent Systems* **22(2)**:169–182 DOI 10.5391/IJFIS.2022.22.2.169.

Rashidpoor Toochaei M, Moeini F. 2023. Evaluating the performance of ensemble classifiers in stock returns prediction using effective features. *Expert Systems with Applications* **213**:119186 DOI 10.1016/j.eswa.2022.119186.

Song Q, Chissom BS. 1993. Forecasting enrollments with fuzzy time series—Part I. *Fuzzy Sets and Systems* **54(1)**:1–9 DOI 10.1016/0165-0114(93)90355-L.

Song Q, Chissom BS. 1994. Forecasting enrollments with fuzzy time series—part II. *Fuzzy Sets and Systems* **62(1)**:1–8 DOI 10.1016/0165-0114(94)90067-1.

Tsao H-H, Leu Y-G, Chou L-F. 2021. A center-of-concentrated-based prediction interval for wind power forecasting. *Energy* **237**:121467 DOI 10.1016/j.energy.2021.121467.

Tsaur R-C. 2012. A fuzzy time series-Markov chain model with an application to forecast the exchange rate between the Taiwan and US dollar. *International Journal of Innovative Computing, Information and Control* **8(7)**:4931–4942.

Wu H, Long H, Wang Y, Wang Y. 2021. Stock index forecasting: a new fuzzy time series forecasting method. *Journal of Forecasting* **40(4)**:653–666 DOI 10.1002/for.2734.

Ying JJ-C, Huang P-Y, Chang C-K, Yang D-L. 2017. A preliminary study on deep learning for predicting social insurance payment behavior. In: *2017 IEEE international conference on big data (big data)*. Piscataway: IEEE, 1866–1875 DOI 10.1109/BigData.2017.8258131.

Yu H-K. 2005. Weighted fuzzy time series models for TAIEX forecasting. *Physica A: Statistical Mechanics and Its Applications* **349(3)**:609–624 DOI 10.1016/j.physa.2004.11.006.

Zadeh LA. 1965. Fuzzy sets. *Information and Control* **8(3)**:338–353 DOI 10.1016/S0019-9958(65)90241-X.