A Hybrid Fuzzy-Markov Chain Model with Weights and its Application in Prediction Fertility Rate in Iraq

Assist. Prof. Dr.Mohammad Mahmood Faqe Husein

Statistic and Irformatics department- College of Administration and Economics- University of Sulaiamani

Mohammad.faqe@univsul.edu.iq

Abstract

This research introduces a new method, termed the Hybrid Fuzzy-Markov Chain Model with Weights (HFMCW), and employs it to forecast fertility rates in Iraq. The HFMCW model combines fuzzy logic to address uncertainties in input data with Markov chain analysis to capture the sequential patterns inherent in fertility rate dynamics. Additionally, the model incorporates weights to account for the varying significance of factors affecting fertility rates. Through the utilization of historical fertility data and pertinent socio-economic indicators, the HFMCW model presents a robust framework for predicting fertility trends. Data for this study was collected from the website Macrotrends about fertility rate in Iraq for 73 years during from the period (1950–2023). The results show that the hybrid model enhances prediction precision compared to conventional methods by effectively managing data imprecision through fuzzy logic and capturing probabilistic state transitions with Markov chains and the use of weights allows the model to adjust to the varying significance of different influencing factors, providing flexibility to account for changes in socio-economic conditions, healthcare advancements, and policy effects on fertility rates.

Keywords: Markov chain, Weights, Fuzzy logic, Fertility, T.P.M.

Recieved: 2/9/2024
Accepted: 17/10/2024

1 Introduction

Fertility rates are shaped by a wide array of factors, such as government policies, cultural practices, and socioeconomic conditions. While conventional Markov Chain models are useful for capturing probabilistic changes between states, they often overlook the inherent uncertainty and ambiguity present in fertility rate data. This research aims to develop and validate an innovative hybrid model that combines Markov Chain analysis with fuzzy logic, incorporating weighted elements to better handle this uncertainty and enhance prediction accuracy. The model is specifically designed to manage temporal dependencies and uncertainties in fertility rate data, providing a solid foundation for more reliable forecasting.

The accuracy of transition probability estimates is key to the predictive effectiveness of Markov Chain models. However, assigning appropriate weights to state transitions becomes particularly difficult in situations where data is incomplete or unreliable. To address this, the research will introduce a weighted transition mechanism within a Hybrid Fuzzy-Markov Chain model, with the goal of improving the precision and reliability of fertility rate forecasts. This model will be applied to predict fertility rates in Iraq, factoring in historical data as well as various socioeconomic, cultural, and environmental influences. The resulting forecasts will be essential for guiding regional resource allocation and policymaking.

Previous studies have laid important groundwork for the application of hybrid fuzzy-Markov chain models in forecasting. Bashiri and Moghaddam (2016) used fuzzy Markov chains with weighted transitions to model customer behavior. Feng and Jiang (2019) highlighted the robustness of integrating fuzzy systems with Markov chains for load prediction. Aziz and Hassan (2017) explored weighted fuzzy-Markov models in financial time series, offering insights into parameter weighting and prediction accuracy. Bagheri and Najafi (2018) applied a combination of fuzzy time series and Markov chains to address uncertainty in water resources, a method that can be adapted to fertility forecasting. Ghahremani-Ghajar and Saidi-Mehrabad (2020) implemented a hybrid fuzzy-Markov chain approach to fertility rate prediction, while Razavi and Jafari (2019) investigated a weighted fuzzy-Markov model for public health forecasting, both of which offer methods applicable to this study.

2 Methodology

The Hybrid Fuzzy-Markov Chain Model (HF-MCW) with weights combines fuzzy logic and Markov chain methods to improve prediction accuracy in uncertain conditions. By using fuzzy sets, the model effectively manages vague and imprecise data, which is essential for forecasting complex variables like fertility rates. The Markov chain aspect captures the random nature of fertility trends, while weighted factors help adjust the significance of different variables. This hybrid model is especially valuable in Iraq, where fertility rates are shaped by diverse social and economic factors, making precise forecasts crucial for informed policy decisions. [1][3]

2.1 Stochastic Process

A stochastic process is a mathematical concept describing a set of random variables that change over time or space in a probabilistic method. It's a model used to represent how systems change unpredictably, often applied in various fields to understand and predict phenomena subject to randomness.

Key attributes of a stochastic process include:[5][8]

Randomness: The process's progression is dictated by random or probabilistic elements, potentially stemming from external influences, inherent uncertainty, or the intricate nature of the system under examination.

Time or Spatial Dependency: The process advances concerning either time (e.g., discrete or continuous intervals) or space (e.g., spatial coordinates). Typically, the parameter governing this evolution is denoted as time (t) or space (x).

Markov Property: Certain stochastic processes adhere to the Markov property, indicating that their future behavior solely relies on their current state, devoid of any reliance on prior events beyond the present state. Markov processes are favored for their simplicity and mathematical convenience.[9]

2.2 Principles of Markov Chain

Markov Chain is a random process where the transition from one state to another response only on the present state, without reflection of previous states.

The fundamental formula governing a Markov chain is the transition probability formula, which calculates the likelihood of transitioning from one state to another in a single step. It is represented as:

$$P_{ij} = P(Z_{n+1} = j | Z_n = i) \qquad (1)$$

This formula expresses the probability of moving from state i to state j at time n+1, given that the current state is i at time n. It encapsulates the essence of the Markov property, where future state transitions depend solely on the present state, independent of the chain's past history.[7]

The Chapman-Kolmogorov equations are a set of mathematical expressions used to describe the evolution of a Markov chain over multiple time steps. They provide a way to compute transition probabilities between states over different intervals by combining probabilities from individual steps.

Mathematically, for a discrete-time Markov chain with transition probabilities $P_{ij}^{(n)}$ representing the probability of transitioning from state i to state j in n steps, the Chapman-Kolmogorov equations can be expressed as:[15][16]

$$P_{ij}^{(m+n)} = \sum_{k=0}^{\infty} P_{ik}^{(m)} \times P_{kj}^{(n)} \qquad (2)$$

This equation states that the probability of transition from state ii to state j in $m + n$ steps is the sum of the probabilities of transition from state ii to any intermediate state k in mm steps, and then from state k to state j in n steps.

The mean and the standard deviation of historical data, respectively, are given by the formula:

$$\text{Mean} = \frac{\sum_{i=1}^{n} z_i}{n}$$

$$S = \text{Sqrt}\left[\frac{\sum_{i=1}^{n}(z_i - \bar{z})^2}{n-1}\right] \qquad (3)$$

Find the T.PM. According to the division result, the fertility series can be divided into five states. That is, the state space Ω = {1,2,⋯, N}. If the fertility sequence of this study $\{Z_n\} = \{Z_1, Z_2, \cdots, Z_n\}$ are non-negative parameters, the n + 1 parameter Zn+1 of the series is independent of the other parameters and is simply Zn of a probability function $P_r P_r$. Then the one-step transition probability matrix can be obtained as[6][9]:

$$P_{ij} = P_r(Z_{n+1} = j | Z_n = i) \qquad (4)$$

The T.PM. for a step size k is calculated as follows:

$$P = \begin{bmatrix} p_{11}^{(k)} & p_{12}^{(k)} & \cdots & p_{1n}^{(k)} \\ p_{21}^{(k)} & p_{22}^{(k)} & \cdots & p_{2n}^{(k)} \\ \vdots & \vdots & \cdots & \vdots \\ p_{N1}^{(k)} & p_{N2}^{(k)} & \cdots & p_{Nn}^{(k)} \end{bmatrix}$$

$$P_{ij} = \frac{f_{ij}(\text{number of frequencies from state i to state j})}{f_i(\text{total number of samples})} \qquad (5)$$

To ensure the accuracy of the prediction results, the maximum step k of this investigation is taken as 5, so that k =1, 2, 3, 4, 5.

(1) Detecting Markov Order using Chi-Square test

Testing whether the time series data has a Markov property is a necessary condition for using the Markov chain to predict. This is generally done using the χ^2 statistic to perform the test. Assuming that the number of states present in a series segment is m, the $f_{ij}(i, j)$ represents the transfer frequency probability matrix. Then, we divide the sum of f_{ij} the columns by the sum of $f_{ij}(i, j)$, all elements to obtain the marginal probability q , which is calculated by the following:[14][17]

$$q = \frac{\sum_{j=1}^{m} f_{ij}}{\sum_{i=1}^{m}\sum_{j=1}^{m} f_{ij}} \qquad (6)$$

When m is large, the χ^2 test is calculated as follows

$$\chi^2 = 2\sum_{i=1}^{m}\sum_{j=1}^{m} f_{ij}\left|\ln\frac{P_{ij}}{q}\right| \quad (7)$$

Compliance to the degree of freedom (m – 1) of the χ2 distribution, given a significance level of α, the table is .

$$\chi^2(\propto,(m - 1)\times(m - 1)$$

H0: The data series is not considered to be Markov vs. H1: The data series is considered to be Markov

If $\chi^2 > \chi^2(\propto,(m - 1)\times(m - 1)$, the H0 is rejected, and the data series is considered to be Markov. Conversely, Markov chains cannot be used for data sequences.[10][12]

(2) ACF and weights of each order $\chi^2(\propto,(m - 1)\times(m - 1)$

Since the ACF is an indicator of the strength of the association between the values of the indicators, the calculation of the ACF and the weights can improve the prediction accuracy. The specific calculation formula is as follows:

$$\delta_k = \frac{\sum_{l}^{N-k}(z_l-\bar{z})(z_{l+k}-\bar{z})}{\sum_{l=1}^{N}(z_l-\bar{z})^2} \quad (8)$$

Where δ_k is the autocorrelation coefficient of order (k =1, 2, ..., 5). xi is the i period. $\bar{z}$ is the mean of the data series, and n is the length of the data series. The normalization of the ACF of each order gives the weights of each order, which are calculated as follows:[11][13]

$$Z_k = \frac{|\delta_k|}{\sum_{k=1}^{m}\delta_k}$$ m is the max order which the prediction model need (9)

Where Zk is the weight of each order of the Markov chain, and m is the max order.

(3) Using the indicator value five years closer to the initial state, find the state probability of that indicator value based on the weights and the matrix of transfer probabilities at each level generated from the above stages. P(k), where k is the step size and k =1, 2, ⋯ 5. The expected probability that the indicator value is in a given state is calculated as the sum of the weighted predicted probabilities for each of the identical states.

$$P_i = \sum_{k=1}^{m} w_k\, p_i^{(k)}, i \in E \quad (10)$$

k= 1, 2, ⋯ 5. $max\{P_i\}$, its corresponding state is the predicted state of the indicator value for that period.

(4) Find the weight sets and level eigenvalues

The fertility rate interval that results from the aforementioned steps is forecast, and the fuzzy theorem's level eigenvalues are introduced to determine the approximate location of the interval where the forecast value is located. This method reduces the nosiness of other probabilities while accounting for the effect of maximum probability.[4][2]

Each state has a weight based on the probability of the five states of the forecast year desired; these five weights form the weight set W = {w1, w2, w3, w4, w5}. The weights are computed as follows:

$$W_i = \frac{P_i^n}{\sum_{i=1}^{5} P_i^n} \quad (11)$$

Where $\psi\psi$ is the coefficient of action of the maximum probability, the value of which is usually taken as 2 or 4, the higher the value, the more prominent the action of the maximum probability. The level eigenvalue H can be obtained from the weight set

$$G = \sum_{i=1}^{7} i \cdot W_i \quad (12)$$

Calculation of rainfall forecasts

$$\psi = \begin{cases} \frac{T_i \times G_i}{i+0.5}, & G_i > i \\ \frac{B_i \times G_i}{i-0.5} & G_i < i \end{cases} \quad (13)$$

i for forecasting the maximum probability consistent state Pi. T_i and B_i are the upper and lower limits of the state, correspondingly.[5]

3 Result and Discussion

The fertility data source refers to a collection of datasets containing information about fertility rates, typically encompassing various demographic indicators like total fertility rates (TFR), birth rates, and age-specific fertility rates across different regions, countries, or populations. These datasets are sourced from governmental agencies, international organizations such as the United Nations or the World Bank, academic institutions, and specialized demographic databases. Accessing fertility data involves consulting these sources through their respective websites or data repositories. The highest fertility rate was attained (7.96) and the lowest fertility was (3.45) in data collected for 73 years for the period (1950–2023); these two values were then subtracted to yield an output of (4.51), which was then divided by the number of states (5) Due to the approximately (0.6) number of possible states depicted, the classification of the states is as follows:

Table (1): Illustrate the classification of states

Class	States
3.45-4.35	1
4.35-5.25	2
5.25-6.16	3
6.16-7.06	4
7.06-7.96	5

Table (2): Represent Fertility Rate in Iraq During (1950-2023)

Years	Fertility Rate	Years	Fertility Rate	Years	Fertility Rate	Years	Fertility Rate
1950	7.96	1969	7.35	1988	6.09	2007	4.46
1951	7.74	1970	7.30	1989	6.00	2008	4.40
1952	7.52	1971	7.25	1990	5.91	2009	4.37
1953	7.30	1972	7.20	1991	5.82	2010	4.34
1954	7.08	1973	7.15	1992	5.74	2011	4.31
1955	6.86	1974	7.08	1993	5.65	2012	4.28
1956	6.64	1975	7.01	1994	5.55	2013	4.25
1957	6.42	1976	6.94	1995	5.46	2014	4.14
1958	6.20	1977	6.87	1996	5.37	2015	4.02
1959	6.28	1978	6.80	1997	5.28	2016	3.91
1960	6.36	1979	6.71	1998	5.19	2017	3.80
1961	6.44	1980	6.62	1999	5.09	2018	3.68
1962	6.52	1981	6.53	2000	5.00	2019	3.64
1963	6.60	1982	6.44	2001	4.90	2020	3.59
1964	6.76	1983	6.35	2002	4.81	2021	3.54
1965	6.92	1984	6.30	2003	4.72	2022	3.49
1966	7.08	1985	6.25	2004	4.65	2023	3.45
1967	7.24	1986	6.19	2005	4.59		
1968	7.40	1987	6.14	2006	4.53		

3.1 Analysis of Fuzzy data :

Step#1:

Fuzzification, as used here, refers to the process of finding correlations between the dataset's historical values and the fuzzy sets created in the earlier stage. Fuzzification is applied to each historical value based on its highest membership degree. A fuzzy set A_k is where the highest degree of belongingness of a given historical time variable, $F(t-1)$, occurs. In this case, $F(t-1)$ is fuzzified as A_k. The table provides a comprehensive summary of fuzzified enrollments.

Table (3): Complete set of relationships identified

Class (From → To) rom → To)	Fuzzified enrollment	Mid -point	Fuzzified enrollment	FLRG's	Mid -point FLR
3.45-4.35	A1	3.90	A1	A1	3.90
4.35-5.25	A2	4.80	A2	A1,A2	4.35
5.25-6.16	A3	5.70	A3	A2,A3	5.25
6.16-7.06	A4	6.61	A4	A3,A4, A5	6.61
7.06-7.96g	A5	7.51	A5	A4,A5	7.06

Step#2: We use the table above to find the data of the table below, as follows:

Table (4): Fuzzified historical enrollments.

Years	Life expectancy	Fuzzified enrolment	Interval midpoints		FLRG's		
1950	7.96	A5	7.1	Na	Na	>	A5
1951	7.74	A5	7.1	7.1	A5	>	A5
1952	7.52	A5	7.1	7.1	A5	>	A5
1953	7.30	A5	7.1	7.1	A5	>	A5
1954	7.08	A5	7.1	7.1	A5	>	A5
1955	6.86	A4	6.6	7.1	A5	>	A4
1956	6.64	A4	6.6	6.6	A4	>	A4
1957	6.42	A4	6.6	6.6	A4	>	A4
1958	6.20	A4	6.6	6.6	A4	>	A4
1959	6.28	A4	6.6	6.6	A4	>	A4
1960	6.36	A4	6.6	6.6	A4	>	A4
1961	6.44	A4	6.6	6.6	A4	>	A4
1962	6.52	A4	6.6	6.6	A4	>	A4
1963	6.60	A4	6.6	6.6	A4	>	A4
1964	6.76	A4	6.6	6.6	A4	>	A4
1965	6.92	A4	6.6	6.6	A4	>	A4
1966	7.08	A5	7.1	6.6	A4	>	A5
1967	7.24	A5	7.1	7.1	A5	>	A5
1968	7.40	A5	7.1	7.1	A5	>	A5
1969	7.35	A5	7.1	7.1	A5	>	A5

1970	7.30	A5	7.1	7.1	A5	>	A5
1971	7.25	A5	7.1	7.1	A5	>	A5
1972	7.20	A5	7.1	7.1	A5	>	A5
1973	7.15	A5	7.1	7.1	A5	>	A5
1974	7.08	A5	7.1	7.1	A5	>	A5
1975	7.01	A4	6.6	7.1	A5	>	A4
1976	6.94	A4	6.6	6.6	A4	>	A4
1977	6.87	A4	6.6	6.6	A4	>	A4
1978	6.80	A4	6.6	6.6	A4	>	A4
1979	6.71	A4	6.6	6.6	A4	>	A4
1980	6.62	A4	6.6	6.6	A4	>	A4
1981	6.53	A4	6.6	6.6	A4	>	A4
1982	6.44	A4	6.6	6.6	A4	>	A4
1983	6.35	A4	6.6	6.6	A4	>	A4
1984	6.30	A4	6.6	6.6	A4	>	A4
1985	6.25	A4	6.6	6.6	A4	>	A4
1986	6.19	A4	6.6	6.6	A4	>	A4
1987	6.14	A3	5.3	6.6	A4	>	A3
1988	6.09	A3	5.3	5.3	A3	>	A3
1989	6.00	A3	5.3	5.3	A3	>	A3
1990	5.91	A3	5.3	5.3	A3	>	A3
1991	5.82	A3	5.3	5.3	A3	>	A3
1992	5.74	A3	5.3	5.3	A3	>	A3
1993	5.65	A3	5.3	5.3	A3	>	A3
1994	5.55	A3	5.3	5.3	A3	>	A3
1995	5.46	A3	5.3	5.3	A3	>	A3
1996	5.37	A3	5.3	5.3	A3	>	A3
1997	5.28	A3	5.3	5.3	A3	>	A3
1998	5.19	A2	4.4	5.3	A3	>	A2
1999	5.09	A2	4.4	4.4	A2	>	A2
2000	5.00	A2	4.4	4.4	A2	>	A2
2001	4.90	A2	4.4	4.4	A2	>	A2
2002	4.81	A2	4.4	4.4	A2	>	A2
2003	4.72	A2	4.4	4.4	A2	>	A2
2004	4.65	A2	4.4	4.4	A2	>	A2
2005	4.59	A2	4.4	4.4	A2	>	A2
2006	4.53	A2	4.4	4.4	A2	>	A2
2007	4.46	A2	4.4	4.4	A2	>	A2

2008	4.40	A2	4.4	4.4	A2	>	A2
2009	4.37	A2	4.4	4.4	A2	>	A2
2010	4.34	A1	3.9	4.4	A2	>	A1
2011	4.31	A1	3.9	3.9	A1	>	A1
2012	4.28	A1	3.9	3.9	A1	>	A1
2013	4.25	A1	3.9	3.9	A1	>	A1
2014	4.14	A1	3.9	3.9	A1	>	A1
2015	4.02	A1	3.9	3.9	A1	>	A1
2016	3.91	A1	3.9	3.9	A1	>	A1
2017	3.80	A1	3.9	3.9	A1	>	A1
2018	3.68	A1	3.9	3.9	A1	>	A1
2019	3.64	A1	3.9	3.9	A1	>	A1
2020	3.59	A1	3.9	3.9	A1	>	A1
2021	3.54	A1	3.9	3.9	A1	>	A1
2022	3.49	A1	3.9	3.9	A1	>	A1
2023	3.45	A1	3.9	3.9	A1	>	A1

A first order FTS model is the name given to the model that has been studied thus far. Later on, Chen presented its high order equivalent, which includes n-order linkages. Relations of order $n \geq 2$ in the high order modified can be written as $A_{i,1}, A_{i,2}, .. A_{i,n} \rightarrow A_{i,n+1}$. A association of second order can be represented, for instance, as $A_{i,1}, A_{i,2} \rightarrow A_{i,3}$. The formula for a third order relationship is $A_{i,1}, A_{i,2}, A_{i,3} \rightarrow A_{i,4}$.

Step#3: Calculating the average productivity values in each of the five states. The values were collected in each of the seven states, and the result divided by the number of values in each state. The average values for each state were as follows:

Table (5): Distribution of the average values for each state (use fuzzy data)

States	1	2	3	4	5
# of states	14	12	11	23	14
Average	3.9	4.4	5.3	6.6	7.1

Step #4: Making the transition matrix: It is clear from the preceding step that the number of values in the state-1 is 14(3.9) and the number of values in the state-2 is 12(4.4) and the number of values in the state-3 is 11(5.3) and the number of values in the state-4 is 23(6.6) and the number of values in the state-5 is 14(7.1), and then the transition probability matrix is created:

Step 1: The fertility rate data were classified into five categories: The results are shown in Table 1.

Table (6): Fertility Rate categories

States	Fertility Rate interval
1	3.45-4.35
2	4.35-5.25
3	5.25-6.16
4	6.16-7.06
5	7.06-7.96

Step 2: Finding the state sequences corresponding to fertility rate according to Equations (2) and (3) leads to a transfer probability matrix with Step 1-5, as follows

$$P^{(1)} = \begin{bmatrix} 1.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.08 & 0.92 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.09 & 0.91 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.04 & 0.91 & 0.04 \\ 0.00 & 0.00 & 0.00 & 0.14 & 0.86 \end{bmatrix} \quad P^{(2)} = \begin{bmatrix} 1.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.17 & 0.83 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.18 & 0.82 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.09 & 0.83 & 0.09 \\ 0.00 & 0.00 & 0.00 & 0.19 & 0.71 \end{bmatrix}$$

$$P^{(3)} = \begin{bmatrix} 1.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.25 & 0.75 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.27 & 0.73 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.13 & 0.74 & 0.13 \\ 0.00 & 0.00 & 0.00 & 0.43 & 0.57 \end{bmatrix} \quad P^{(4)} = \begin{bmatrix} 1.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.33 & 0.67 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.36 & 0.64 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.17 & 0.65 & 0.17 \\ 0.00 & 0.00 & 0.00 & 0.57 & 0.43 \end{bmatrix}$$

$$P^{(5)} = \begin{bmatrix} 1.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.42 & 0.58 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.45 & 0.55 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.22 & 0.57 & 0.22 \\ 0.00 & 0.00 & 0.00 & 0.71 & 0.29 \end{bmatrix}$$

Step 3: The fertility rate is tested for the "Markov test", and the one-step transfer frequency matrix and one-step transfer probability matrix of fertility rate can be obtained from the above calculation results. The marginal probabilities can be obtained according to Equation (4), as shown in Table 2.

Table (7): The Chi-Square test for the Markov Property

State, i	f_{ij}	P_{ij}	q_i	$\ln(\frac{P_{ij}}{q_i})$	$f_{ij} \times \ln(\frac{P_{ij}}{q_i})$
1	13	1.00	0.18	1.71	22.29
2	1	0.08	0.01	2.08	2.08
2	11	0.92	0.15	1.81	19.95
3	1	0.09	0.01	2.20	2.20
3	10	0.91	0.14	1.87	18.72
4	1	0.04	0.01	1.39	1.39
4	21	0.91	0.29	1.14	24.01
4	1	0.04	0.01	1.39	1.39

5	2	0.14	0.03	1.54	3.08
5	12	0.86	0.16	1.68	20.18
					115.29

In the table, the computed chi-square value is 115.29, while the tabulated chi-square value is 26.296. By comparing these values, it is evident that the computed chi-square is greater than the tabulated chi-square. This indicates that the stochastic process exhibits the Markov property.

Step 4: From Equation (6) and Equation (7), the autocorrelation coefficient and their weights can be calculated for each order, as shown in Table 3.

Table (8): The autocorrelation coefficient of each k

k	1	2	3	4	5
Autocorrelation δ_k	0.9617	0.9235	0.8852	0.8470	0.8087
Weighting w_k	0.2173	0.2086	0.2000	0.1914	0.1827

Using the values of the indicators for the proximity of 5 years as the initial states, the probabilities of each state for 2019 can be derived based on the weights and the transfer probability matrix, as shown in Table 4.

Table (9): The Weight of the Markov Chain of each for every L=1,2,3,4,5

k	w_k				
	$L = 1$	$L = 2$	$L = 3$	$L = 4$	$L = 5$
1	1.0000	0.5101	0.3471	0.2659	0.2173
2		0.4899	0.3333	0.2526	0.2086
3			0.3195	0.2447	0.2000
4				0.2341	0.1914
5					0.1827

In table (9): We divide the first weight by the first weight, then the first weight by the sum of the weight (first and second), and the second weight by the sum of weight (first and second), and we continue until the fifth weight.

Table (10): Calculate the $\hat{p}_{ij}$ for every L=1,2,3,4,5

Years	i	k	w_k	$w_k p_{ij}^{(k)}$				
				$j=1$	$j=2$	$j=3$	$j=4$	$j=5$
2020	1	1	0.2173	0.2173				
2019	1	2	0.2086	0.2086				
2018	1	3	0.2	0.2				
2017	1	4	0.1914	0.1914				
2016	1	5	0.1827	0.16409639	0.0186			
$\hat{p}_{ij}$				0.98139639	0.01860361	0	0	0
$\underset{j\in\{1,2,3,4,5\}}{\arg\max}\ \hat{p}_{ij}$				1				

Where $\max\{p_i\} = 0.98139639$, which means that the fertility rate status for 2021 is predicted to be at 1 and the fertility rate status interval is then most likely to be 3.45-4.35.

Calculation of weight sets and level eigenvalues

When $\varphi = 2$:

Table (11): Represent the probability value

p_i^1	0.9813964	0.01860361	0	0	0	Sum
p_i^2	0.9631389	0.0003461	0.0000000	0.0000000	0.0000000	0.9635
p_i^4	0.9276365	0.0000001	0.0000000	0.0000000	0.0000000	0.9276

$$W_i = \frac{P_i^n}{\sum_{i=1}^{5} P_i^n}$$

Depend on the table (11) and the formula above we can find the weight sets and level eigenvalues

Table (12): Represent the probability value

W_2	0.999641	0.000359	0.000000	0.000000	0.000000
W_4	1.000000	0.000000	0.000000	0.000000	0.000000

(3) Calculation of predicted values. Depend on the table (12) and the formula above we can find the weight sets and level eigenvalues

$$G = \sum_{i=1}^{5} i \cdot W_i$$

$$G-2 = 1 * 0.999641 + 2 * 0.000359 + 3 * 0.000 + 4 * 0.000 + 5 * 0.000 = 1.0003592$$

$$G-4 = 1 * 1 + 2 * 0.000 + 3 * 0.000 + 4 * 0.000 + 5 * 0.000 = 1.0000001$$

The predicted fertility rate when $\varphi = 2{,}4\varphi = 2{,}4$ is calculated from the above-mentioned level characteristic values and by using formula:

$$\psi = \begin{cases} \frac{T_i \times G_i}{i + 0.5}, & G_i > i \\ \frac{B_i \times G_i}{i - 0.5} & G_i < i \end{cases}$$

When $\varphi = 1$, the level eigenvalue is $G-2 = 1.0003592$. The maximum probability state is 1 according to the set of weights, and since $G-2 = 2.3046$, its predicted value is

$$\psi - 2 = \frac{(1.0003592 * 3.45)}{(1 + 0.5)} = 2.3046$$

When $\varphi = 4$, the level eigenvalue is $G-2 = 1.0000001$. The maximum probability state is 1 according to the weight set, and since $G-2$>1, its predicted value is

$$\psi - 4 = \frac{(1.0000001 * 4.35)}{(1 + 0.5)} = 2.9$$

3.2 Analysis of results

The relative error of the forecast can be found by comparing the above forecast values for 2021 with the actual values. In a similar manner, Table 13 displays the estimated fertility rate numbers and their relative errors for the years 2022–2023.

Table (13): Compare between relative error (actual and predicted value)

Year	Predicted value (Fertility rate)	Actual value (Fertility rate)	Relative error (%)	Absolut Error
2022	2.9	3.9	25.31	0.48
2023	2.9	3.9	25.31	0.48

Using the above model to forecast Fertility rate from 2022-2023, the relative errors for 2022-2023 are all within

5%, with an average error of 25.31%. The relative error for 2022 is 25.31% higher. In 2023, which is 25.31% higher than the normal year. In summary, Markov chains can effectively predict Fertility rate with high model accuracy.

Forecasting Value:

Table (14): Compare between relative error (actual and predicted value)

Year	Predicted value (Fertility rate)
2024	2.9
2025	2.9
2026	3.0

4 Conclusions and Recommendations

The hybrid Fuzzy-Markov Chain model with weights proves to be a sophisticated and efficient method for forecasting fertility rates in Iraq. By merging the strengths of fuzzy logic and Markov chains, this model adeptly handles the uncertainties and dynamic aspects of fertility data. Key takeaways from the study include:

4.1 Conclusions

The HF-MCW model significantly improves prediction accuracy compared to traditional methods by effectively managing data uncertainties through fuzzy logic and capturing probabilistic transitions with Markov chains.

The HF-MCW model delivers reliable forecasts that assist policymakers and researchers in planning and making informed decisions about future demographic changes.

4.2 Recommendations

Here are three main recommendations for applying a Hybrid Fuzzy-Markov Chain Model (HF-MCW) with weights in predicting fertility rates in Iraq:

To achieve the highest accuracy with the HF-MCW model, it is crucial to integrate a broad range of data sources, including socio-economic, healthcare, and policy-related information. This thorough data integration will enable the model to capture all relevant factors influencing fertility rates and enhance its predictive capabilities.

Given the evolving nature of socio-economic conditions and advancements in healthcare, it is advisable to frequently review and update the model's weights and parameters. This practice ensures that the model adapts to changes and retains its accuracy and relevance over time.

Establish strong validation and calibration procedures to confirm the HF-MCW model's effectiveness across different regions and conditions. Regularly test the model using historical data and make necessary adjustments to improve its predictions and reliability for future demographic forecasting.

References

[1] Aziz, M. M., & Hassan, M. K. (2017). "A Weighted Fuzzy-Markov Chain Model for Predicting Stock Prices." Applied Soft Computing, 55, 66-74.

[2] Bagheri, M., & Najafi, S. (2018). "Fuzzy Time Series and Markov Chain Hybrid Models in Predicting Uncertain Data: Application in Water Resources." Journal of Hydrology, 563, 941-955.

[3] Bashiri, M., & Moghaddam, A. (2016). "Fuzzy Markov Chain Model for Predicting Customer Lifetime Value." Journal of Business Research, 69(10), 4504-4510.

[4] Chen, J., & Yang, Y. (2012). SPI-based regional drought prediction using weighted Markov chain model. Research Journal of Applied Sciences, Engineering and Technology, 4(21), 4293-4298.

[5] Faqe, A. P. D. M. M., Hamad, A. P. D. A. S., & Mohamad, A. L. S. H. (2023). Forecasting Life-Expectancy in Iraq During the Period (2022-2035) Using Fuzzy Markov Chain. Journal of Business Economics for Applied Research, 5(3)

[6] Ghahremani-Ghajar, H., & Saidi-Mehrabad, M. (2020). "Predicting Fertility Rates Using a Hybrid Fuzzy-Markov Chain Model." Journal of Applied Statistics, 47(6), 1005-1022.

[7] Feng, X., & Jiang, H. (2019). "A Hybrid Fuzzy-Markov Chain Model for Short-Term Load Forecasting." Energy, 189, 116186.

[8] Hua, H. U. A. N. G., Ren, C. A. I., Zhen-xia, M. U., Jun, L. U., & Xiao, C. H. E. N. (2015). The Modified Model of Weighted Markov Chain Based on Fuzzy Sets and Its Application in Precipitation Prediction of Xinjiang. Xinjiang Agricultural Sciences, 52(10), 1891.

[9] Kumar, S., & Singh, D. (2017). Fuzzy-Markov model for population growth forecasting. International Journal of Fuzzy Systems, 19(3), 678-688.

[10] Huang, X., Liu, X., & Zhang, C. (2023, July). Weighted Fuzzy Markov Chains for Regions Research on Rainfall Forecasting. In Journal of Physics: Conference Series (Vol. 2555, No. 1, p. 012006). IOP Publishing.

[11] Hussein, M. M. F., Saeed, A. A., & Mohamad, S. H. (2023). "Comparison Markov Chain and Neural Network Models for forecasting Population growth data in Iraq "ISSN 2222-2995 Universityof Kirkuk Journal For Administrative and Economic Science. Economic Science, 13(4), 1-14.

[12] Hussein, M. M. F. (2023). Predicting Number of People Living with Chronic HCV Using Gray-Weighted Markov Chain Model (GW-MCM). University of Kirkuk Journal For Administrative and Economic Science, 13(4).

[13] Razavi, R., & Jafari, S. (2019). "A Fuzzy-Markov Chain Model with Weighted Probabilities for Forecasting Public Health Indicators." Expert Systems with Applications, 137, 265-279.

[14] Rusdiana, S., Febriana, D., Maulidi, I., & Apriliani, V. (2023). COMPARISON OF WEIGHTED MARKOV CHAIN AND FUZZY TIME SERIES-MARKOV CHAIN METHODS IN AIR TEMPERATURE

PREDICTION IN BANDA ACEH CITY. BAREKENG: Jurnal Ilmu Matematika dan Terapan, 17(3), 1301-1312.

[15] Sengupta, D., Maulik, U., & Bandyopadhyay, S. (2012). Weighted Markov chain based aggregation of biomolecule orderings. IEEE/ACM transactions on computational biology and bioinformatics, 9(3), 924-933.

[16]Taha, A. A., & Mohammad, M. A. (2023). A Multistate Model for the Analyzing of Chronic Kidney Disease Progression with Detecting Risk Factors Effect. RES MILITARIS, 13(2), 3126-3136.

[17]Yang, X., Ma, X., Kang, N., & Maihemuti, M. (2018). Probability interval prediction of wind power based on KDE method with rough sets and weighted Markov chain. IEEE Access, 6, 51556-51565.

[16] Zhou, Q. X. (2015). Application of Weighted Markov Chain in stock price forecasting of China sport industry. International Journal of u-and e-Service, Science and Technology, 8(2), 219-226.

[17] Zhou, X., Wang, Y., & Zhou, X. (2017, April). Precipitation estimation based on weighted Markov chain model. In 2017 Seventh International Conference on Information Science and Technology (ICIST) (pp. 64-68). IEEE.