International Journal of Innovative Computing, Information and Control
Volume **22**, Number **2**, April **2026**

ICIC International ©2026 ISSN 1349-4198
pp. **475–486**

CONVEX OPTIMIZATION MODELING AND TIME SERIES PREDICTION OF HIGH-ORDER FUZZY COGNITIVE MAPS

Dan Shan[1], Guotao Cong[1] and Wei Lu[2]

[1]Department of Electronic Engineering
Dalian Neusoft University of Information
No. 8, Software Park Road, Dalian 116023, P. R. China
{ shandan; congguotao }@neusoft.edu.cn

[2]School of Control Science and Engineering
Dalian University of Technology
No. 2, Linggong Road, Ganjingzi District, Dalian 116024, P. R. China
luwei@dlut.edu.cn

Received June 2025; revised October 2025

Abstract. *Fuzzy Cognitive Mapping (FCM), as a soft computing methodology, has emerged as a prominent research area in recent years. However, existing FCM construction methods exhibit significant limitations in weight learning, particularly when applied to long-term or complex time series prediction tasks, demonstrating insufficient prediction accuracy and limited capability for logical knowledge representation and reasoning. This paper proposes a novel approach for constructing high-order FCMs that extracts concepts from data using triangular membership functions and transforms the FCM learning problem into a constrained convex optimization problem, thereby effectively determining the weights of high-order FCMs. The proposed method is applied to large-scale time series prediction tasks. Ten diverse real-world time series datasets are employed to validate the predictive performance of the proposed approach. Experimental results demonstrate that the proposed time series numerical prediction method is effective, achieving superior prediction accuracy: it yields an average accuracy improvement of 55% compared to baseline methods and 25% compared to the latest methods. Additionally, it exhibits enhanced FCM representation capabilities and improved inference performance relative to existing approaches.*

Keywords: Fuzzy cognitive map, Convex optimization, Time series prediction

1. **Introduction.** FCM [1, 2] is a fuzzy weighted directed graph, where nodes represent concepts and edges reflect causal relationships between nodes. As an effective soft computing tool, it has become a prominent research focus for many researchers in common time series modeling [3, 4] or some improved version [5, 6], prediction [7, 8] in the domain of education [9], pharmaceuticals [10], life sciences [11] and decision-making applications in the field of epidemics [12], school enrolment [13], agriculture [14] and so on. Most research related to time series modeling and prediction focuses on the development of FCM weight learning methods.

Among these approaches, several population-based algorithms, such as those proposed by Zou and Liu [15], have demonstrated good performance in modeling and prediction effectiveness. By virtue of simulating the mechanisms of biological population evolution and group collaboration, it enables parallel search in complex solution spaces and exhibits strong adaptability to nonlinear problems. Through information interaction and

DOI: 10.24507/ijicic.22.02.475

iterative updating among individuals within the population, it effectively explores the potential patterns of data, and demonstrates excellent adaptability when dealing with high-dimensional, dynamically changing modeling and prediction tasks. In recent years, several new methods have emerged, including Quantum Fuzzy Cognitive Mapping (QFCM) [16], Interval Type-2 Fuzzy Cognitive Mapping (IT2FCM) [17], and Neuro Fuzzy Cognitive Mapping (NFCM) [18]. These novel methods demonstrate significant advantages when addressing complex nonlinear systems. QFCM, by leveraging the principles of quantum computing, enables efficient parallel search to explore multiple solution spaces simultaneously, effectively capturing subtle patterns in data. IT2FCM incorporates interval type-2 fuzzy sets, which enhances its robustness when handling uncertainty and fuzziness in system inputs, ensuring more stable modeling results in dynamic environments. NFCM, by integrating neural network mechanisms, possesses strong self-learning and adaptive capabilities, allowing it to dynamically adapt to the evolution of system characteristics and improve prediction accuracy during the iterative process. In the tasks of modeling and predicting complex nonlinear systems, compared with traditional models, these novel methods can dig deeper into the inherent nonlinear correlations in data, exhibit better flexibility and adaptability, and continuously elevate the performance of system analysis and prediction.

However, population-based algorithms have inherent limitations. During the iterative process, particle diversity decays rapidly near local optima, which adversely affects the algorithm's overall optimization capability and prediction accuracy. This may cause the algorithm to prematurely converge to local optima, making it difficult to explore better solution spaces and limiting its further application in modeling and prediction scenarios for complex nonlinear problems. While latest algorithms exhibit better prediction accuracy, their computational burden increases significantly. Complex mechanisms such as quantum state operations in QFCM, double-layer fuzzy set calculations in IT2FCM, and neural network training in NFCM require substantial computational resources. This raises hardware requirements, potentially restricting their widespread use in resource-constrained scenarios. In the modeling and prediction process of FCM, most methods face the challenge of balancing performance improvement and computational cost constraints, which constitutes a practical application challenge, particularly in scenarios where computational resources are limited. Therefore, the necessity of further optimizing algorithm design becomes prominent, especially in the most critical weight learning process which is the most computationally intensive. In light of this, a novel FCM weight learning method in this paper is proposed that transforms the FCM weight learning problem into a constrained standard convex optimization problem [19], which can significantly reduce computational complexity while improve accuracy.

Furthermore, current FCM modeling approaches demonstrate notable limitations in practical time series applications. Specifically, the aforementioned modeling methods rely exclusively on node values at time t as their sole input that stands in fundamental conflict with the intrinsic properties of dynamic systems. In terms of the operational mechanism of dynamic systems, their state evolution is characterized by distinct cumulativeness and correlation. Modeling based solely on node values at a single time point essentially disrupts the temporal interconnectedness of dynamic systems, making it challenging to capture their inherent evolutionary patterns. As a matter of fact, the value of each node in an FCM at time $t+1$ results from the combined influence of multiple pieces of historical information: it includes both the immediate impact of time t and the lag effects of earlier time points such as $t-1$, and $t-2$. This multi-scale temporal dependency represents a core manifestation of the complexity inherent in dynamic systems. To effectively integrate such historical information, high-order FCM incorporates multi-order time lag

terms, integrating node states at different moments into a unified modeling framework. This allows the model to quantify the cumulative impact of earlier time instances on the current state, thereby significantly improving prediction accuracy.

Additionally, raw historical time series primarily exist in the form of precise numerical values, while the reasoning process of FCM relies on fuzzy expressions. The accuracy of the conversion between these two forms directly affects the reliability of the model. A typical triangular membership function establishes a zero-error mapping relationship between raw time series and fuzzy time series by constructing rigorous fuzzy partitioning rules: on the one hand, it converts precise numerical values into fuzzy linguistic variables that conform to human cognitive habits, providing appropriate inputs for FCM reasoning; on the other hand, it can restore fuzzy results to precise numerical values through inverse transformation. This ensures the consistency of data semantics and statistical characteristics throughout the entire modeling process, providing reliable support for subsequent analysis.

2. **Expansion of FCM from 1-order to High-order.** The dynamic characteristics of 1-order FCM can be described mathematically as shown in Equation (1).

$$A_i(t+1) = f\left(\sum_{j=1}^{n} (w_{ij} A_j(t))\right) \tag{1}$$

where $A_j(t)$ represents the activation degree of the jth concept at time t, $w_{ij} \in [-1, 1]$ denotes the weight from the jth concept to the ith concept, i.e., the causal relationship from A_j to A_i, n is the number of concepts, and the activation function f is a nonlinear continuous non-decreasing function such as the sigmoid function Equation (2) [20].

$$f(x) = \frac{1}{1 + e^{-\lambda x}} \tag{2}$$

where λ is the steepness parameter; the larger this parameter value, the steeper the activation function shape and the more sensitive it becomes to the value of x.

For dynamic systems, high-order fuzzy cognitive maps can better reflect their dynamic characteristics. Typically, the closer the distance to the current moment, the greater the impact, and vice versa. The dynamic characteristics of high-order FCM can be described mathematically as shown in Equation (3). Here, to minimize computational load as much as possible, it is assumed that the conceptual values at times t and $t-1, t-2, \ldots, 1$ have equal impact on time $t+1$.

$$A_i(t+1) = f\left(\sum_{j=1}^{n} (w_{1ij} A_j(t)) + \sum_{j=1}^{n} (w_{2ij} A_j(t-1)) + \cdots + \sum_{j=1}^{n} (w_{tij} A_j(1))\right) \tag{3}$$

where $w_{1ij}, \ldots, w_{(t-1)ij}, w_{tij}$ are the weights from the jth node to the ith node at times $t, t-1, \ldots, 1$, respectively. After FCM modeling is completed, FCM starts from the initial state and performs continuous iterations according to the formula to implement prediction.

3. **Proposed Method.** This section provides a detailed introduction to the proposed novel method for large-scale and complex time series modeling and prediction. The highlight of this method is the rapid and effective determination of FCM weights. The proposed method is summarized in the following Algorithm 1.

3.1. **Fuzzy time series generation and restoration.** The generation of fuzzy time series [21] consists mainly of two steps: normalization and triangular fuzzification. Normalization eliminates the influence of dimensionality, while triangular fuzzification converts

Algorithm 1 Learning the weight matrix of q-order FCM by least square method.

Require:

$TFtype$: the type of activation function for FCM to be learned;
λ: the value of parameter of activation function corresponding to $TFtypes$, its value is larger than 0;
β: a regularization parameter, its value is larger than 0;
$\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k, \ldots, \mathbf{x}_n$: actual time series data used to learn FCM;

Ensure:

$\mathbf{w}_{kopt}$: well-learned m-by-m weight matrix for FCM at order k, and $k = 1, \ldots, q$;
$\mathbf{w}_{qopt} \leftarrow \mathbf{0}^{m\times m}, \ldots, \mathbf{w}_{1opt} \leftarrow \mathbf{0}^{m\times m}$;
$H_k \leftarrow [d_1[k : n-q+k, :]; \ldots; d_n[k : n-q+k, :]]$;
$A_i \leftarrow [H_k]$;
$Y \leftarrow [d_1[q+1 : n, :]; \ldots; d_n[q+1 : n, :]]$;
$i \leftarrow 0$;
while $i + q \leq n$ **do**
    $Y_{i+q} \leftarrow -\ln\left(\frac{1}{Y[:,i]} - 1\right)$;
    $objection \leftarrow \min\{\|\lambda(A_t\mathbf{w}_{1j} + A_{t-1}\mathbf{w}_{2j} + \cdots + A_{t-q+1}\mathbf{w}_{qj}) - \mathbf{Y}_j\|_2 + \|\beta * \mathbf{w}_{1j}\|_1 + \|\beta * \mathbf{w}_{2j}\|_1 + \cdots + \|\beta * \mathbf{w}_{qj}\|_1\}$;
    $constraint \leftarrow \{\|\mathbf{w}_{1j}\|_\infty \leq 1, \|\mathbf{w}_{2j}\|_\infty \leq 1, \ldots, \|\mathbf{w}_{qj}\|_\infty \leq 1\}$;
    $problem \leftarrow (variable, objection, constraint)$;
    $\mathbf{w}_{1j}, \ldots, \mathbf{w}_{qj} \leftarrow$ SOLVER($problem$); ▷ SOLVER($\cdot$) indicates an interior-point method based on the convex optimization solver.
    $i \leftarrow i + 1$;
end while

the original time series to fuzzy time series. The observed values $X = [x(1), x(2), \ldots, x(n)]$ are normalized to $Z = [z(1), z(2), \ldots, z(n)]$. Subsequently, the triangular membership function is used to fuzzify Z into fuzzy sequence A_i, $(i = 1, 2, \ldots, n)$ shown in Equation (4). For the triangular membership function, we assume there are m equidistant intervals on $[0, 1]$, meaning the corresponding FCM has n concepts.

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} = \begin{bmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1m} \\ \mu_{21} & \mu_{22} & \cdots & \mu_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{n1} & \mu_{n2} & \cdots & \mu_{nm} \end{bmatrix} \tag{4}$$

In contrast to the process of generating fuzzy time series, defuzzification takes the predicted fuzzy value $\hat{A_{(t+1)}} = \left[\hat{\mu_{(t+1)1}}, \hat{\mu_{(t+1)2}}, \ldots, \hat{\mu_{(t+1)n}}\right]$ as input and produces the numerical value $\hat{z_{(t+1)}}$ prior to denormalization. Following denormalization, the original predicted value $\hat{x_{(t+1)}}$ is obtained.

3.2. **Solving high-order FCM weights and predictions.** Due to the utilization of high-order FCM in this algorithm, assuming the FCM order is q, the weights comprise q components: $\mathbf{w}_1, \ldots, \mathbf{w}_q$, which lie in the interval $[-1, 1]$.

For fuzzy time series A, assuming there are c sets of different initial state vectors $\mu^s(r) = [\mu_1^s(r), \mu_2^s(r), \ldots, \mu_n^s(r)]_{r=1,2,\ldots,q;s=1,2,\ldots,c}$ serve as inputs to q-order FCM. FCM generates k iterations on each initial vector to obtain c response sequences according to the specified Equation (5).

$$\mu^{\hat{c}}_{(t+1)j} = \frac{1}{1 + e^{-\lambda(A_t \mathbf{w}_{1j} + A_{t-1}\mathbf{w}_{2j} + \cdots + A_{t-q+1}\mathbf{w}_{qj})}} \tag{5}$$

A one-to-one correspondence exists between the response value $\mu^{\hat{c}}_{(t+1)j}$ and the actual measured value $\mu^{c}_{(t+1)j}$ for all j and c. The accuracy of the FCM is typically assessed by calculating the difference between $\mu^{\hat{c}}_{(t+1)j}$ and $\mu^{c}_{(t+1)j}$. To further clarify the system optimization objective, a commonly employed minimum error function Equation (6) is incorporated into the system.

$$\arg\min J_1 = \frac{1}{cm(n-q)\sum_{s=1}^{c}\sum_{t=q}^{n}\sum_{j=1}^{m}\left(\mu^{\hat{s}}_{(t+1)j} - \mu^{s}_{(t+1)j}\right)^2} \tag{6}$$

For the convenience of subsequent algorithm implementation, Equation (5) is transformed into a more suitable form, as shown in Equation (7).

$$\lambda(A_t \mathbf{w}_{1j} + A_{t-1}\mathbf{w}_{2j} + \cdots + A_{t-q+1}\mathbf{w}_{qj}) = -\ln\left(\frac{1}{\mu^{\hat{c}}_{(t+1)j}} - 1\right) \tag{7}$$

For each initial vector, the FCM performs r iterations starting from the initial vector, and according to Equation (5), r equations can be obtained. Since there are c sets of initial vectors, a total of $c * r$ equations can be derived.

Let $Y_i = -\ln\left(\frac{1}{\mu^{\hat{c}}_{(t+1)j}} - 1\right)$ from Equation (14), Equation (7) can be expressed in Equation (8).

$$\lambda(A_t \mathbf{w}_{1j} + A_{t-1}\mathbf{w}_{2j} + \cdots + A_{t-q+1}\mathbf{w}_{qj}) = \mathbf{Y}_j \tag{8}$$

Considering the existence of errors, the constrained linear equation system in Equation (8) typically does not have exact solutions. Therefore, seeking an approximate solution that satisfies the constraints while minimizing the error is a feasible approach to address this problem. Consequently, the equation system in Equation (8) can be converted into a constrained least squares problem, as shown in Equation (9).

$$\begin{aligned} &\min \|\lambda(A_t \mathbf{w}_{1j} + A_{t-1}\mathbf{w}_{2j} + \cdots + A_{t-q+1}\mathbf{w}_{qj}) - \mathbf{Y}_j\|_2 \\ &\text{s.t. } \|\mathbf{w}_{1j}\|_\infty \le 1, \|\mathbf{w}_{2j}\|_\infty \le 1, \ldots, \|\mathbf{w}_{qj}\|_\infty \le 1, \end{aligned} \tag{9}$$

where $\|\cdot\|_2$ denotes the 2-norm, $\|\cdot\|_\infty$ denotes the infinity norm, and $\lambda > 0$ is the steepness parameter of the sigmoid function. The constraints ensure that the solved weight vectors lie within the interval $[-1, 1]$. Considering that sparse matrices in large-scale FCMs also offer significant computational efficiency advantages, penalty terms $\|\mathbf{w}_{1j}\|_1, \|\mathbf{w}_{2j}\|_1, \ldots, \|\mathbf{w}_{qj}\|_1$ shown in Equation (10) are introduced to promote sparsity in the weight matrices. Furthermore, the degree of sparsity increases as β increases.

$$\begin{aligned} &\min \|\lambda(A_t \mathbf{w}_{1j} + A_{t-1}\mathbf{w}_{2j} + \cdots + A_{t-q+1}\mathbf{w}_{qj}) - \mathbf{Y}_j\|_2 + \beta\|\mathbf{w}_{1j}\|_1 + \beta\|\mathbf{w}_{2j}\|_1 + \cdots \\ &\qquad + \beta\|\mathbf{w}_{qj}\|_1 \\ &\text{s.t. } \|\mathbf{w}_{1j}\|_\infty \le 1, \|\mathbf{w}_{2j}\|_\infty \le 1, \ldots, \|\mathbf{w}_{qj}\|_\infty \le 1. \end{aligned} \tag{10}$$

Equation (10) represents a classic convex optimization problem. Therefore, when λ and β are treated as hyperparameters, it is appropriate to employ interior point methods [22], such as the barrier function method, primal-dual method, and their enhanced variants, to solve Equation (10). Since interior point methods guarantee convergence to the global optimal solution, they can obtain the globally optimal solution for each weight matrix. After obtaining the weight matrices of the high-order FCM, the FCM model is constructed and used for fuzzy time series prediction.

TABLE 1. Brief introduction to 10 data sets in study

Number	Data set	Description	Train set	Test set	Source
1	Soybean Price	the prices of soybean in Guangdong from January 4, 2010 to June 6, 2014	1013	338	https://download.csdn.net/download/qq40797107/10477157
2	Yahoo stock	recording the finance Yahoo stock data (high point) from September 28, 2010 to September 25, 2015	944	314	https://download.csdn.net/download/weixin42126677/18402174
3	sunspots	records 13-month smoothed monthly total sunspot number	2468	822	https://www.sidc.be/silso/datafiles
4	Image Recognition Times (Mac1)	image recognition task execution times in mobile edge computing of MacBookPro1	750	250	http://archive.ics.uci.edu/
5	Hydraulic systems Cooling efficiency	condition monitoring of hydraulic systems for cooling efficiency	1654	551	http://archive.ics.uci.edu/
6	Appliances energy (Press)	pressure in a low energy building	14801	4934	http://archive.ics.uci.edu
7	Appliances energy (Power)	power consumption in a low energy building	14801	4934	http://archive.ics.uci.edu/
8	Tetouan city (humidity)	humidity of Tetouan city	39311	13104	http://archive.ics.uci.edu/
9	jena climate (rho)	gas density from Max Planck Institute for Biogeochemistry	53359	17786	https://blog.csdn.net/m06362362/article/details/127721083
10	jena climate (Tpot)	temperature from Max Planck Institute for Biogeochemistry	2468	822	https://blog.csdn.net/m06362362/article/details/127721083

4. **Experiment.** In this section, the effectiveness of the proposed method is evaluated through two approaches: first, by demonstrating performance on real-world time series; second, by comparing performance with other classic and state-of-the-art methods in terms of accuracy. In the experiments, the "ECOS" solver integrated in the Python package was adopted as a feasible approach for solving convex optimization problems. Additionally, the maximum number of iterations was set to 100 to ensure convergence. For comparison purposes, commonly used and recently developed representative methods in FCM weight learning were employed, including AR, MA, ARIMA, NARnet, LSTM, TLSP-DE, QFCM, NFCM and IT2FCM algorithm. Ten datasets representing four typical types of time series from different fields, as shown in Table 1, were introduced for the experiments, including trend time series, seasonal time series, random time series, and composite time series. The cross-validation method was adopted, where for each dataset, the first 75% of the data was used for training and the last 25% was used for prediction to evaluate forecasting performance. All experiments were implemented on the same computer: HP ProBook 440 G7 with Intel Core i7-10510U CPU @ 1.80GHz-2.30GHz and 8.00GB RAM. The RMSE shown in Equation (11) was chosen to evaluate the accuracy of FCM predictions. The smaller the RMSE value, the lower the FCM prediction error, and vice versa.

$$RMSE = \sqrt{\frac{1}{r}\sum_{t=1}^{r}(x_t - \hat{x}_t)} \tag{11}$$

Among them, r is the prediction step, that is, the prediction level, x_t and $\hat{x}_t$ represent the actual observed value and predicted value at time t, respectively.

4.1. **Finance Yahoo stock time series.** The Finance Yahoo stock time series shown in Table 1 are used to verify the feasibility and effectiveness of the proposed method in detail. To demonstrate the process clearly, a 5-order FCM is chosen for modeling and prediction implementation. Furthermore, the influence of various parameters on prediction accuracy is displayed and analyzed. Table 2 shows the RMSE prediction performance under different values of the concept number h, steepness parameter λ, and regularization parameter β. The value of h ranges within $[20, 200]$, the value of λ ranges within $[2.5, 5.0)$ with a step size of 0.5, and the value of β ranges within $[0.0001, 0.0005)$ with a step size of 0.0001. Figure 2 and Figure 3 illustrate the variation of RMSE performance under different parameter values of h, λ, and β.

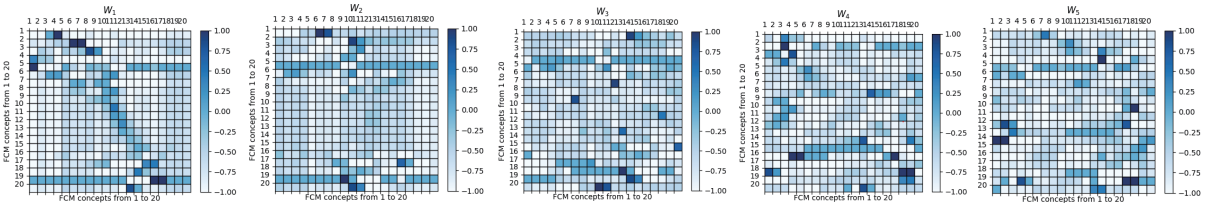
Table 2. The experimental result of RMSE under various parameters

h	λ	RMSE with different β				h	λ	RMSE with different β			
		0.0001	0.0002	0.0003	0.0004			0.0001	0.0002	0.0003	0.0004
20	2.5	0.764	0.764	0.764	0.764	100	2.5	0.699	0.688	0.718	0.722
	3.0	0.747	0.747	0.747	0.747		3.0	0.703	0.704	0.706	0.705
	3.5	0.758	0.758	0.758	0.757		3.5	0.63	0.63	0.548	0.549
	4.0	0.757	0.757	0.757	0.757		4.0	0.545	0.545	0.545	0.545
	4.5	0.756	0.756	0.756	0.756		4.5	0.541	0.541	0.541	0.541
40	2.5	0.681	0.681	0.689	0.681	150	2.5	0.607	0.609	0.611	0.634
	3.0	0.818	0.818	0.818	0.818		3.0	0.573	0.567	0.567	0.574
	3.5	0.765	0.765	0.765	0.765		3.5	0.558	0.558	0.558	0.558
	4.0	0.768	0.768	0.768	0.768		4.0	0.558	0.565	0.565	0.565
	4.5	0.765	0.766	0.766	0.766		4.5	0.565	0.565	0.565	0.565
60	2.5	0.902	0.898	0.898	0.898	200	2.5	0.594	0.595	0.595	0.606
	3.0	0.749	0.745	0.741	0.739		3.0	0.588	0.59	0.593	0.593
	3.5	0.75	0.75	0.75	0.75		3.5	0.552	0.552	0.552	0.554
	4.0	0.749	0.749	0.749	0.749		4.0	0.566	0.566	0.565	0.565
	4.5	0.743	0.596	0.596	0.596		4.5	0.576	0.576	0.576	0.576

According to Table 2, the maximum RMSE is 0.902, achieved when $h = 60$, $\lambda = 2.5$, and $\beta = 0.0001$. Conversely, the minimum RMSE is 0.541, obtained when $h = 100$, $\lambda = 4.5$, and β ranges between 0.0001 and 0.0004. Therefore, it is evident that improved prediction accuracy can be achieved through appropriate adjustment of the parameters h, λ, and β.

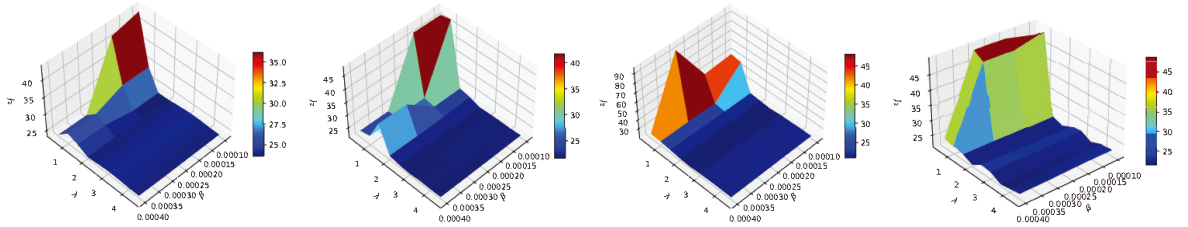
Specifically, taking $h = 20$, $\lambda = 4.5$, $\beta = 0.0002$ as an example, after normalization and fuzzification, the weight matrices $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4, \mathbf{w}_5$ of the 5-order FCM are obtained as shown in Figure 1 during the modeling process. Subsequently, prediction is performed, followed by defuzzification and denormalization in sequence. Finally, the RMSE between the predicted data and the actual test data is calculated to be 0.756.

Figure 2 demonstrates that for a fixed number of FCM concepts h, the RMSE value exhibits a decreasing trend as λ increases from low to high values. As λ increases, RMSE shows a significant decline from high to low values. When λ reaches a certain threshold,



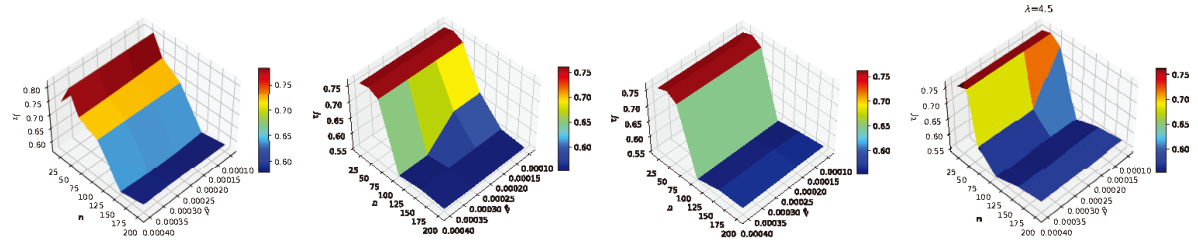
(a) $\mathbf{w}_1$ (b) $\mathbf{w}_2$ (c) $\mathbf{w}_3$ (d) $\mathbf{w}_4$ (e) $\mathbf{w}_5$

FIGURE 1. Weight matrices $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4, \mathbf{w}_5$ of the 5-order FCM



(a) $h = 20$ (b) $h = 60$ (c) $h = 150$ (d) $h = 200$

FIGURE 2. Plots of RMSE under parameters λ and β with fixed concept number h for Finance Yahoo stock data (high)



(a) $\lambda = 3.0$ (b) $\lambda = 3.5$ (c) $\lambda = 4.0$ (d) $\lambda = 4.5$

FIGURE 3. Plots of RMSE under parameters h and β with fixed steepness parameter λ for Finance Yahoo stock data (high)

such as 2.0, it maintains a relatively stable low value. The explanation for this phenomenon is that the steeper the S-shaped activation function becomes, the more sensitive the FCM is to data changes, meaning that the FCM's perceptual ability becomes more flexible. However, when λ reaches a certain value, the FCM's perceptual ability stabilizes and no longer exhibits significant changes. Similarly, Figure 3 shows that under fixed λ values, that is, $\lambda = 3.0, 3.5, 4.0, 4.5$, RMSE demonstrates a decreasing trend as h increases from low to high values, although RMSE may fluctuate slightly with increasing h. Both Figure 2 and Figure 3 exhibit a common characteristic: the influence of β on RMSE is negligible. This occurs because the function of β is not to affect prediction accuracy, but rather to regulate the sparsity of the weight matrices.

4.2. **Comparison with other methods.** To compare the FCM modeling and prediction method proposed in this paper (i.e., TLSP) with other classic and state-of-the-art methods, including AR, MA, ARIMA, NARnet, LSTM, TLSP-DE, QFCM, NFCM and IT2FCM algorithms, comprehensive comparisons were conducted on the real-world time series shown in Table 1 to further evaluate the robustness and broad applicability of TLSP.

TABLE 3. Experimental results obtained by TLSP, AR, MA, ARIMA, NARnet, LSTM, TLSP-DE, QFCM, NFCM and IT2FCM

Data set	TLSP	AR	MA	ARIMA	NARnet	LSTM	TLSP-DE	QFCM	NFCM	IT2FCM
Soybean Price	17.14	78.99	75.776	82.923	68.388	79.72	53.37	49.16	37.94	45.61
Yahoo stock	0.51	1.859	1.867	2.154	1.774	1.61	1.74	0.81	0.98	1.89
sunspots	9.24	17.29	16.875	19.146	12.343	11.31	10.13	10.77	8.46	9.84
Image Recognition Times (Mac1)	0.0034	0.006	0.0058	0.0076	0.0043	0.0041	0.0038	0.0036	0.0037	0.0035
Hydraulic systems Cooling efficiency	0.14	0.301	0.292	0.324	0.227	0.21	0.17	0.16	0.25	0.16
Appliances energy (Press)	0.014	0.049	0.047	0.051	0.032	0.02	0.017	0.021	0.018	0.016
Appliances energy (Power)	108.4	301.24	263.87	324.52	209.73	117.54	114.2	108.3	134.4	119.7
Tetouan city (humidity)	0.63	1.812	1.794	1.845	1.206	0.94	0.81	0.74	0.79	0.68
jena climate (rho)	0.34	0.985	0.941	1.049	0.827	0.38	0.42	0.45	0.38	0.41
jena climate (Tpot)	0.13	0.368	0.375	0.369	0.335	0.21	0.22	0.14	0.19	0.16

The comparison results between the proposed TLSP prediction method and other methods shown in Table 3 for ten real-world datasets demonstrate satisfactory predictive performance. TLSP achieves higher prediction accuracy without manual intervention. To further validate the superior predictive accuracy of the TLSP method, the average proportion of accuracy improvement relative to both the baseline methods and the latest methods was calculated. In comparison with typical baseline methods (AR, MA, ARIMA, NARnet, and LSTM), the average proportion of accuracy improvement is approximately 55%. Similarly, when compared with the latest methods (TLSP-DE, QFCM, NFCM and IT2FCM), it also demonstrates favorable performance, with the average proportion of accuracy improvement being approximately 25%. This superior performance can be attributed to the inherent advantages and limitations of various methods. AR, MA, and ARIMA models have simple structures and low computational complexity. However, while they can effectively capture linear relationships, they are unable to handle the complex nonlinear relationships inherent in large-scale time series. Consequently, their prediction accuracy is relatively low for large-scale time series. NARnet, LSTM and NFCM are essentially neural network-based algorithms that can effectively capture nonlinear features in large-scale time series, thereby improving prediction accuracy. For TLSP-DE, QFCM, NFCM, and IT2FCM, utilizing FCM as the model structure can effectively simulate the intrinsic relationships of complex systems, explain and predict system behavior, and possess strong knowledge reasoning capabilities. However, these methods employ population-based algorithms for weight optimization. Although population-based algorithms offer the advantage of fast convergence, they cannot guarantee convergence to the global optimal solution. While TLSP-DE, QFCM, NFCM, and IT2FCM have achieved considerable predictive performance, TLSP still outperforms them in prediction accuracy. This superiority is attributed to adopting a convex optimization approach, which ensures that the weights converge to the globally optimal solution.

Furthermore, to demonstrate the influence of FCM order on prediction performance, Table 4 presents RMSE values of the TLSP method for FCM predictions with orders ranging from 1 to 200, while Figure 4 illustrates RMSE variation trends across datasets at different orders. Theoretically, higher FCM orders should enhance prediction accuracy by incorporating more comprehensive historical information. However, Table 4 and Figure 4 show that while accuracy improves initially with increasing order, it stops improving and even decreases once a certain order is reached. This is because, for complex large-scale time series, older historical data has diminishing impact on current predictions; excessive inclusion weakens the critical influence of recent data, which is detrimental to time series prediction.

TABLE 4. Experimental results obtained by TLSP with various order FCM

Data set	1-order	2-order	10-order	20-order	40-order	60-order	80-order	100-order	110-order	130-order	150-order	200-order
Soybean Price	20.87	19.43	17.14	16.92	16.86	15.94	15.42	15.54	17.83	24.57	27.62	30.21
Yahoo stock	0.58	0.54	0.51	0.49	0.48	0.49	0.49	0.54	0.59	0.62	0.65	0.65
sunspots	9.41	9.35	9.24	9.24	9.23	9.26	9.27	9.22	9.45	9.51	9.66	9.73
Image Recognition Times (Mac1)	0.0037	0.0036	0.0034	0.0035	0.0035	0.0036	0.0035	0.0034	0.0032	0.004	0.0045	0.0046
Hydraulic systems Cooling efficiency	0.16	0.15	0.14	0.14	0.14	0.15	0.15	0.16	0.17	0.17	0.19	0.19
Appliances energy (Press)	0.018	0.016	0.014	0.015	0.016	0.017	0.017	0.018	0.018	0.019	0.021	0.022
Appliances energy (Power)	116.46	110.6	108.4	110.1	114.2	116.8	118.1	116.4	119.2	118.7	119.4	121.6
Tetouan city (humidity)	0.76	0.69	0.63	0.64	0.63	0.61	0.61	0.62	0.59	0.68	0.72	0.73
jena climate (rho)	0.37	0.35	0.34	0.33	0.35	0.31	0.32	0.31	0.37	0.37	0.46	0.45
jena climate (Tpot)	0.14	0.14	0.13	0.12	0.12	0.11	0.11	0.13	0.15	0.18	0.16	0.19

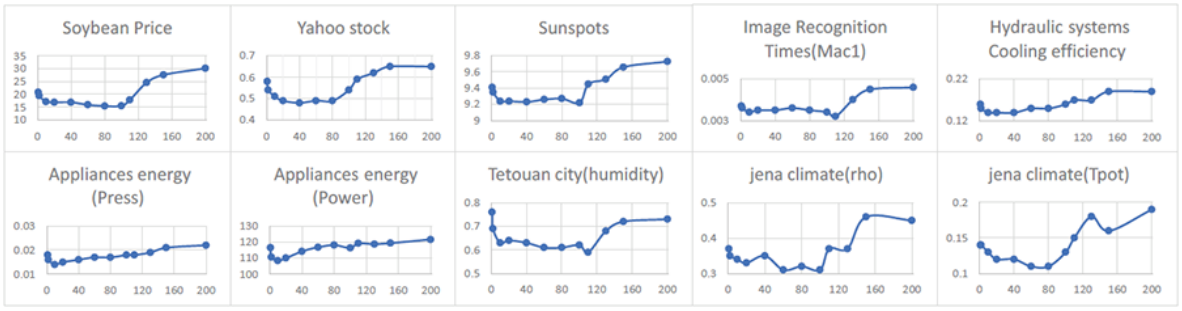


FIGURE 4. Plots of RMSE trend for 10 datasets as FCM order increasing

5. **Conclusions.** This paper proposes a high-order FCM modeling and prediction method for large-scale complex time series. To address the challenge of obtaining high-order FCM weights, a constrained convex optimization approach is utilized to build high-quality models. To validate TLSP's effectiveness, four typical large-scale time series are tested: trend, seasonal, random, and composite. Its experimental methodology involves four steps: 1)

transforming ordinary time series into fuzzy ones via normalization and triangular fuzzification (as FCM training data); 2) computing FCM weights by solving constrained convex optimization problems; 3) performing prediction with the trained FCM model; and 4) validating accuracy against other methods. Experimental results show TLSP outperforms representative alternatives. The study also analyzes the impact of key parameters (FCM concept number h, steepness λ, regularization β, and order) on prediction performance. However, this method has a limitation: although multiple parameter combinations are provided, each parameter is treated as a hyperparameter and has not been fully optimized. Therefore, future work will consider comprehensive parameter optimization to further improve prediction accuracy.

REFERENCES

[1] B. Kosko, Fuzzy cognitive maps, *International Journal of Man-Machine Studies*, vol.24, no.8, pp.65-75, 1986.

[2] M. Hagiwara, Extended fuzzy cognitive maps, *Proc. of IEEE International Conference on Fuzzy Systems*, pp.795-801, 1992.

[3] W. Pedrycz, A. Jastrzebska and W. Homenda, Design of fuzzy cognitive maps for modeling time series, *IEEE Trans. Fuzzy Systems*, vol.24, no.1, pp.120-130, 2016.

[4] M. Bose and K. Mali, Designing fuzzy time series forecasting models, *Int. J. Approx. Reason.*, vol.111, pp.78-99, 2019.

[5] E. Bozdag and C. Kadaifci, A distance-based approach to fuzzy cognitive maps using pythagorean fuzzy sets, *International Journal of Fuzzy Systems*, vol.27, no.1, pp.93-109, 2025.

[6] J. Zhang, Z. Tao, J. Liu, X. Liu and H. Chen, A hybrid interval-valued time series prediction model incorporating intuitionistic fuzzy cognitive map and fuzzy neural network, *Journal of Forecasting*, vol.44, no.1, pp.93-111, 2025.

[7] D. Shan, W. Lu and J. Yang, The data-driven fuzzy cognitive map model and its application to prediction of time series, *International Journal of Innovative Computing, Information and Control*, vol.14, no.5, pp.1583-1602, 2018.

[8] Y. Liu and L. Wang, Long-term prediction of time series based on fuzzy time series and information granulation, *Granular Computing*, vol.9, no.2, pp.1-13, 2024.

[9] M. Lepore, A holistic framework to model student's cognitive process in mathematics education through fuzzy cognitive maps, *Heliyon*, vol.10, no.16, pp.3652-3667, 2024.

[10] G. Napoles, I. Grau, R. Bello and R. Grau, Two-steps learning of fuzzy cognitive maps for prediction and knowledge discovery on the HIV-1 drug resistance, *Expert Systems with Applications*, vol.41, no.3, pp.821-830, 2014.

[11] A. Amirkhani, M. Kolahdoozi, C. Wang and L. Kurgan, Prediction of DNA-binding residues in local segments of protein sequences with fuzzy cognitive maps, *IEEE/ACM Trans. Computational Biology and Bioinformatics*, vol.17, no.4, pp.1372-1382, 2020.

[12] S. Mei, Y. Zhu, X. Qiu, X. Zhou, Z. Zu, A. V. Boukhanovsky and P. M. A. Sloot, Individual decision making can drive epidemics: A fuzzy cognitive map study, *IEEE Trans. Fuzzy Systems*, vol.22, no.2, pp.264-273, 2014.

[13] N. H. Shafii, R. Alias, S. R. Shamsuddin and D. S. M. Nasir, Fuzzy time series for projecting school enrolment in Malaysia, *Journal of Computing Research and Innovation*, vol.6, no.1, pp.11-21, 2021.

[14] F. Consentino, I. Peri, M. Litrico, D. Spina and G. Vindigni, Mapping young farmers' choice to pursue geographical indication in a rural context: Application of fuzzy cognitive map, *Agricultural and Food Economics*, vol.12, no.1, pp.1-18, 2024.

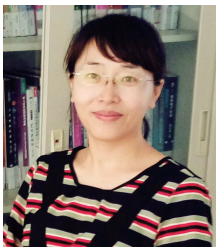
[15] X. Zou and J. Liu, A mutual information-based two-phase memetic algorithm for large-scale fuzzy cognitive map learning, *IEEE Trans. Fuzzy Systems*, vol.26, no.4, pp.2120-2134, 2018.

[16] M. Kolahdoozi, A. Amirkhani, M. H. Shojaeefard and A. Abraham, A novel quantum inspired algorithm for sparse fuzzy cognitive maps learning, *Applied Intelligence*, vol.49, no.10, pp.3652-3667, 2019.

[17] A. Amirkhani, M. Shirzadeh, T. Kumbasar and B. Mashadi, A framework for designing cognitive trajectory controllers using genetically evolved interval type-2 fuzzy cognitive maps, *International Journal of Intelligent Systems*, vol.37, no.1, pp.305-335, 2022.

[18] A. Amirkhani, M. Shirzadeh, M. H. Shojaeefard and A. Abraham, Controlling wheeled mobile robot considering the effects of uncertainty with neuro-fuzzy cognitive map, *ISA Transactions*, vol.100, pp.454-468, 2020.
[19] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge Univ. Press, 2004.
[20] A. Rehman, Neural computing for online arabic handwriting recognition using hard stroke features mining, *International Journal of Innovative Computing, Information and Control*, vol.17, no.1, pp.177-191, 2021.
[21] M. N. Alemu, A fuzzy model for chaotic time series prediction, *International Journal of Innovative Computing, Information and Control*, vol.14, no.5, pp.1767-1786, 2018.
[22] G. Feng, W. Lu and J. Yang, The modeling of time series based on least square fuzzy cognitive map, *MDPI AG Algorithms*, vol.14, no.3, 69, 2021.

Author Biography



Dan Shan received M.S. degree in Control Theory and Control Engineering from Dalian University of Technology, China, in 2004.

She joined Dalian Neusoft University of Information, China, in 2009. She is currently a professor in Department of Electronic Engineering. Her current research interests include computational intelligence, fuzzy modeling and granular computing, knowledge discovery and data mining, fuzzy intelligent systems and low-cost integrated automation systems.

Guotao Cong received M.S. degree in Signal and Information Processing from Northeastern University, China, in 2008.

He joined Dalian Neusoft University of Information, China, in 2016 and is currently an associate professor in the Department of Electronic Engineering. His current research interests include integrated circuit design and validation, motor control, image processing technology, and neural network systems.

Wei Lu received the M.S. degree and the Ph.D. degree in Control Theory and Control Engineering from Dalian University of Technology, China, in 2004 and 2015, respectively.

He joined Dalian University of Technology in 2004. He is currently a professor in the School of Control Science and Engineering, Dalian University of Technology. His current research interests include computational intelligence, fuzzy modeling and granular computing, knowledge discovery and data mining, fuzzy intelligent systems and low-cost integrated automation systems. He also serves as a Frequent Reviewer for many international journals including IEEE Transactions on Fuzzy Systems, IEEE Transactions on Cybernetics, Knowledge-Based Systems, Applied Soft Computing, Experts Systems and Applications, Int. J. of Approximate Reasoning, Int. J. of Granular Computing (Springer) as well as international conferences.