Networks and Heterogeneous Media

https://www.aimspress.com/journal/nhm

NHM, 21(3): 773–800.
DOI: 10.3934/nhm.2026033
Received: 31 December 2025
Revised: 07 April 2026
Accepted: 13 April 2026
Published: 28 April 2026

Research article

Prediction of systemic risk of China's financial industry based on phase space reconstruction and machine learning

Yulian An[1,*], Zhaojun Liu[1], Zhenkuan Sun[1] and Boan Li[2]

[1] School of Economics and Finance, Shanghai International Studies University, Shanghai 201620, China

[2] NewSchool Of Architecture And Design, San Diego, CA 92101, USA

* **Correspondence:** Email: anyl@shisu.edu.cn; Tel: 86-13636651394.

Abstract: Forecasting systemic financial risks is crucial for effectively preventing and mitigating such risks. This paper introduces an integrated framework for systemic financial risk forecasting, with its core innovation being the novel combination of phase space reconstruction with delay parameterization (RDPM) and a radial basis function neural network (RBFNN). Based on stock trading data from 41 listed companies in the China A-share stock market, a complex financial network is constructed.

Keywords: systemic financial risk; forecast; phase space reconstruction; delay parameterization method; machine learning

1. Introduction

Since the outbreak of the US subprime crisis in 2008, research on financial systemic risk has intensified. The bankruptcy of Lehman Brothers and the European debt crisis exposed structural vulnerabilities within the financial system. Large-scale financial crises inflict disproportionate damage on real economic activity and reveal critical flaws in existing policy frameworks [1]. Although eradicating systemic events remains beyond the reach of prudential authorities, the timely extraction of early-warning signals can significantly improve the calibration of macro-prudential instruments. As global markets become more complex and interconnected, ex-ante quantification of systemic risk has emerged as a primary, yet inherently challenging research frontier [2]. In recent years, a substantial body of research has developed on systemic financial risk, focusing mainly on measurement, contagion, and spillover effects, factors influencing and interactions with macroeconomic conditions [3–6].

Benoit argues that systemic risk simultaneously affects multiple market participants and spreads further throughout the financial system [7]. It is widely recognized that systemic risk spreads rapidly,

has broad impact, and can lead to currency depreciation, market downturns, and economic recession at both national and global levels [8]. Defining systemic risk remains a challenge. Summer [9] notes that the difficulty stems from the lack of a general theoretical framework capable of rigorously distinguishing between crisis contagion and joint crises caused by common shocks. Accurate measurement of systemic financial risk, or design of suitable proxy indicators, forms the foundation for empirical research. Given the interconnectedness of financial markets and the contagious nature of risks, traditional value at risk (VaR) has proven inadequate [10]. Instead, measures such as marginal expected shortfall (MES) and conditional value at risk (CoVaR), which focus on system-wide risk assessment, have gained widespread application [11, 12].

Given the high degree of interconnections within financial markets, complex network theory has been widely applied to the analysis of financial risk contagion. This approach goes beyond the limitations of examining isolated relationships and instead studies the propagation of risk from a holistic, interconnected perspective [13]. Billio et al. systematically applied PCA and Granger causality networks to quantify financial interconnectedness, revealing strong banking-insurance links during crises in [14].

However, due to the inherent complexity of network modeling, most studies construct low-dimensional models that involve a limited number of financial entities. This simplification can introduce certain biases in the research findings [15]. In recent years, many scholars have shifted their focus toward modeling high-dimensional networks [16–18] or graph neural networks (GNNs) [19]. Some researchers employ methods such as cluster analysis to reduce network dimensionality, which helps to effectively identify homogeneous groups within complex systems, as seen in [20–23].

Given the absence of an effective monitoring mechanism, the forward-looking prediction of systemic risk serves as a crucial supervisory tool. Forecasting future fluctuations in systemic risk and detecting points of structural mutation will help formulate macroeconomic policies and implement preventive measures in advance [24].

Systemic risk is measured using high-frequency data from financial markets, thereby inheriting the nonlinear, highly noisy, and chaotic characteristics typical of financial time series [25]. Traditional forecasting models, such as ARIMA, are based on linear assumptions and thus often do not adequately predict systemic risk volatility [26]. In recent years, machine learning techniques—including support vector machines, random forests, neural networks, and deep learning—have demonstrated strong predictive performance in finance [27–30]. Balmaseda et al. examined deep graph learning for systemic risk prediction in financial systems [31]. The review [32] evaluated the methodological diversity and effectiveness of machine learning and hybrid approaches in financial risk management and finds that advanced machine learning techniques consistently demonstrate strong predictive accuracy. Random Forest, XGBoost, LSTM, BiLSTM, and CNN improve prediction by capturing nonlinear dependencies.

Although some explainable AI methods, such as SHAP and feature importance analysis, are used, the limited interpretability arising from the "black box" nature of most machine learning and deep learning methods poses a significant challenge to their practical applicability. On the other hand, machine learning models face challenges with large data requirements and high computational costs [33].

In recent decades, many nonlinear tools, such as chaotic analysis [34–36], differential equations [36–39], and phase space reconstruction [40, 41] have been used to study financial problems. To some extent, these methods based on system dynamics address the shortcomings of purely machine learning approaches. Utilizing techniques such as differential equations and phase space reconstruction, they can intricately

depict a system's structural characteristics and nonlinear behavior by studying its underlying dynamical principles. In addition, they are not reliant on vast amounts of data and demonstrate a certain robustness against non-stationarity.

Phase space reconstruction is a nonlinear dynamic analysis tool derived from chaos theory. By mapping a one-dimensional financial time series to multidimensional space, it is expected to capture the nonlinear relationship and the dynamic evolution process of the financial system in a more comprehensive way [42]. The stochastic distribution embedding based on phase space reconstruction has good performance in predicting high-dimensional short-term data [43, 44]. In [45], the detection of time-varying points was carried out, and abnormal fluctuations caused by the bankruptcy of Lehman Brothers bank before the subprime crisis were successfully verified. The delayed parameterization method shows good characteristics in predicting stock market prices and indexes [40, 44].

Many scholars adopt combined methods to extract effective information from financial time series in order to improve forecasting precision. For example, reference [38] integrated the advantages of neural networks and stochastic differential equations, achieving high-quality predictions for certain stock indices. Phase space reconstruction combined with echo state networks or LSTM has been shown in recent studies [46, 47] to effectively capture hidden nonlinear dependencies between different assets, significantly improving the accuracy of prediction . In the future, risk management and financial analysis will increasingly be driven by artificial intelligence and big data [48–52]. The integration of cross-disciplinary research methodologies is expected to tackle the intricate challenges of the financial system more effectively.

In this study, we measure and forecast systemic financial risk in China's financial sector using stock trading data from 41 listed companies in the A-share market. First, we construct a complex financial network and identify two subsystems through spectral clustering, then employ a weighted average to quantify the systemic risk of each cluster. Second, we predict systemic financial risk by integrating phase-space reconstruction, the delay parameterization method, and a radial basis function neural network. The relative accuracy of the prediction exceeds 94%. The model effectively captures the nonlinear dependencies of financial systemic risks and demonstrates a high R^2 coupled with a low error rate, indicating strong predictive performance.

The remainder of this paper is structured as follows. Section 2 introduces the phase-space reconstruction method. Section 3 presents the measurement framework for systemic risk. The forecasting procedure and the corresponding results are detailed in Section 4. In Section 5, we validate the constructed systemic risk series through analysis of the event study and comparison with the geopolitical risk index (GPR), confirming its credibility and practical relevance. Finally, Section 6 summarizes the main conclusions of this study.

2. Phase space reconstruction and delay parameterization method (RDPM)

The theoretical foundation of phase-space reconstruction for prediction originates from Takens' embedding theorem and the generalized embedding theorem [53, 54]. According to these theorems, the trajectory in the reconstructed phase space retains the topological invariance of the original system's dynamical evolution. This property enables its application in the building of forecasting models. Here is a simplified explanation: suppose that the observed data constitute a vector series $(y_1, y_2, \cdots, y_m)$, and the phase space formed by the original multivariate observations is

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}^T = \begin{pmatrix} y_1(1) & y_2(1) & \cdots & y_m(1) \\ y_1(2) & y_2(2) & \cdots & y_m(2) \\ \cdots & \cdots & \ddots & \cdots \\ y_1(N) & y_2(N) & \cdots & y_m(N) \end{pmatrix} \tag{2.1}$$

where m is the number of vectors and N is the dimension of every vector. The phase space reconstructed by the kth observation vector y_k, is

$$Y_k = \begin{pmatrix} y_k(1) & y_k(1+\tau_k) & \cdots & y_k(1+(m-1)\tau_k) \\ y_k(2) & y_k(2+\tau_k) & \cdots & y_k(2+(m-1)\tau_k) \\ \cdots & \cdots & \ddots & \cdots \\ y_k(N) & y_k(N+\tau_k) & \cdots & y_k(N+(m-1)\tau_k) \end{pmatrix} \tag{2.2}$$

Here, we chose the embedding dimension of the phase space $d = m$ for simplicity, and τ_k is the delay parameter. Suppose that $\Psi_k : Y \to Y_k$ is a linear function $Y_k = \Psi_k(Y) = YP$:

$$Y_k = YP = \begin{pmatrix} y_1(1) & y_2(1) & \cdots & y_m(1) \\ y_1(2) & y_2(2) & \cdots & y_m(2) \\ \cdots & \cdots & \ddots & \cdots \\ y_1(N) & y_2(N) & \cdots & y_m(N) \end{pmatrix} \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \cdots & \cdots & \ddots & \cdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{pmatrix}. \tag{2.3}$$

Here, P is the parameter matrix constituted by the coefficients of coupling relation between Y and Y_k. P and $y_k(N + h\tau_k),\ h = 1, 2, \cdots, (m-1)$ are unknown. Note that the first $N - (m-1)\tau_k$ lines of Y_k are known. Defining $\overline{Y}_k$ and $\overline{Y}$ as

$$\overline{Y}_k = \begin{pmatrix} y_k(1) & y_k(1+\tau_k) & \cdots & y_k(1+(m-1)\tau_k) \\ y_k(2) & y_k(2+\tau_k) & \cdots & y_k(2+(m-1)\tau_k) \\ \cdots & \cdots & \ddots & \cdots \\ y_k(N-(m-1)\tau_k) & y_k(N-(m-2)\tau_k) & \cdots & y_k(N) \end{pmatrix},$$

$$\overline{Y} = \begin{pmatrix} y_1(1) & y_2(1) & \cdots & y_m(1) \\ y_1(2) & y_2(2) & \cdots & y_m(2) \\ \cdots & \cdots & \ddots & \cdots \\ y_1(N-(m-1)\tau_k) & y_2(N-(m-1)\tau_k) & \cdots & y_m(N-(m-1)\tau_k) \end{pmatrix},$$

then we have

$$\overline{Y}_k = \overline{Y}P \tag{2.4}$$

from Eq (2.3). Then, the parameter matrix P can be solved by Eq (2.4). Moreover, we can obtain the prediction of the sequence $y_k(N + h),\ h = 1, 2, \cdots, (m-1)\tau_k$ according to Eq (2.3). Considering that the linear mapping P cannot adequately model the complexity of financial time series, we adopt a radial basis function neural network (RBFNN) to approximate the mapping Ψ_k. The selection of time-delay parameters constitutes a crucial issue.

There are many methods to determine the time delay, such as the autocorrelation function method, mutual information method [48, 49], etc. To fully use the internal dynamics of the system, we treat

all time delay parameters $\tau = (\tau_1, \tau_2, \cdots, \tau_m)$ jointly. Referring to [40, 41], we adopt the delay parameterization method (DPM) to determine the optimal time-delay τ so as to minimize the overall prediction error of the system during training. Defining an object function

$$f = \frac{1}{m}\sum_{k=1}^{m}\left[\text{mean}(\left|\frac{\hat{y}_k(\cdot) - y_k(\cdot)}{y_k(\cdot)}\right|)\Big/\text{std}(\left|\frac{\hat{y}_k(\cdot) - y_k(\cdot)}{y_k(\cdot)}\right|)\right], \tag{2.5}$$

where $\hat{y}_k(\cdot)$ represents the predicted value and $y_k(\cdot)$ represents the true value, the functions mean$(\cdot)$ and std$(\cdot)$ denote the mean value and standard deviation of the time series. The optimal time delay τ^* minimizes the object function f in Eq (2.5) and will be identified by an integer constrained particle swarm optimization (ICPSO) algorithm. The method is briefly described below, and a detailed discussion can be found in [40].

Let $M = N - (m-1)\tau_k$. We divide the observation time series set $(y_1, y_2, \cdots, y_m)$ into a training set and a test set. $\overline{Y}$, $\overline{Y}_k$ belong to the training set and

$$\widetilde{Y} = \begin{pmatrix} y_1(M+1) & y_2(M+1) & \cdots & y_m(M+1) \\ y_1(M+2) & y_2(M+2) & \cdots & y_m(M+2) \\ \cdots & \cdots & \ddots & \cdots \\ y_1(N) & y_2(N) & \cdots & y_m(N) \end{pmatrix},$$

$$\widetilde{Y}_k = \begin{pmatrix} y_k(M+1) & y_k(M+1+\tau_k) & \cdots & y_k(N+1) \\ y_k(M+2) & y_k(M+2+\tau_k) & \cdots & y_k(N+2) \\ \cdots & \cdots & \ddots & \cdots \\ y_k(N) & y_k(N+\tau_k) & \cdots & y_k(N+(m-1)\tau_k)) \end{pmatrix},$$

belong to the test set. Here, the values $y_k(N + h\tau_k)$ for $h = 1, 2, \cdots, (m-1)$ are to be predicted. The prediction is carried out according to the following steps:

- **Step 1:** Observe and normalize the variable

$$y_k(i)/(\max(y_k(i)) - \min(y_k(i)))), \ 1 \le k \le m, 1 \le i \le N,$$

  still denoted by $y_k(i)$ for convenience;
- **Step 2: For** $\tau = 1, 2, \cdots, \tau_{max}$
  1. The normalized matrix $\overline{Y}$, $\widetilde{Y}$, $\overline{Y}_k$ are constructed;
  2. Train the network with $\overline{Y}$ as the input matrix and $\overline{Y}_k$ as the output matrix;
  3. Using the test matrix $\widetilde{Y}$ as a new input matrix and the network structure obtained in the above training process, predict $\hat{y}_k(j)$, $M + 1 \le j \le N$;
  4. Compute the object function f;
     end for
  5. Minimize the objec function f and choose the optimal delay τ^*;
- **Step 3: For** τ^*
  1. The matrix $\overline{Y}$, $\widetilde{Y}$ are reconstructed by using normalized $y_k(i)$, and $\overline{Y}_k$ is constructed by the optimal delay τ^*;

2. Train the network with $\overline{Y}$ as the input matrix and $\overline{Y}_k$ as the output matrix, where $k = 1, 2, \cdots, m$;
3. Using the test matrix $\widetilde{Y}$ as a new input matrix and the network structure obtained in the above training process predicts $\hat{y}_k(N+1)$;
4. Add the predicted value $\hat{y}_k(N+1),\ 1 \le k \le m$ to the observed variable;
   end for

- **Step 4:** Obtain the predicted value $\hat{y}_k(N+j),\ 1 \le k \le m, 1 \le j \le n$. Then, an inverse normalization process is performed;
- **Step 5:** Calculate the values of mean absolute error (MAE), root mean square error (RMSE), and mean relative accuracy (MRA) to evaluate the prediction effect, which are defined as follows:

$$\begin{aligned} MAE &= \frac{1}{n}\sum_{j=1}^{n}\left|\frac{\hat{y}_k(\cdot) - y_k(\cdot)}{y_k(\cdot)}\right|, \\ RMSE &= \sqrt{\frac{1}{n}\sum_{j=1}^{n}(\hat{y}_k(\cdot) - y_k(\cdot))^2}, \\ MRA &= \frac{1}{n}\sum_{j=1}^{n}\left(1 - \left|\frac{\hat{y}_k(\cdot) - y_k(\cdot)}{y_k(\cdot)}\right|\right). \end{aligned} \tag{2.6}$$

Details concerning the implementation of the radial basis function neural nsetwork (RBFNN) used in this procedure are presented in Supplementary B.

3. Systematic risk measurement

3.1. Data

Our study uses financial institutions listed on the China A-share market as of 5 January 2015 as the initial sample, covering 67 companies in the banking, securities, and insurance sectors. After excluding institutions with severely missing annual data (with more than 10% missing data) or labeled "ST", we obtain a final effective sample of 41 financial institutions (see Supplementary for their names, abbreviations, and ID). The sample consists of 16 banks, 21 securities firms, and 4 insurance companies. In 2022, the total market capitalization of the sample institutions accounted for 81.3% of the total market value of all A-share listed financial institutions in China.

Daily closing prices from 5 January 2015 to 31 December 2020 are collected, yielding 1463 trading-day observations. Data are sourced from the CCER database (https://www.ccerdata.cn/). To mitigate the influence of outliers, winsorization is applied at the 1% level on both tails.

The sample period spans several major events, including the 2015 stock-market turbulence, the 2016 circuit-breaker episode, the 2018 China–U.S. trade friction, and the outbreak phases of the COVID-19 pandemic. This provides a representative window for observing the dynamic evolution of systemic financial risk.

3.2. *Measurement of systemic financial risk*

Value-at-risk (VaR) is one of the most widely used measures of an individual institution's market risk. The maximum loss $\mathrm{VaR}^i_{\alpha,t}$ of the institution i at time t at the confidence level $1-\alpha$ (e.g., 95%) is defined as:

$$\Pr(r_i(t) \le \mathrm{VaR}^i_{\alpha,t}) = \alpha, \tag{3.1}$$

where $r_i(t) = \ln P_i(t) - \ln P_i(t-1)$ is the return rate of the institution i at time t and $P_i(t)$ is the closing price of the stock i in period t. While $\mathrm{VaR}^i_{0.5,t}$ defines the median (or baseline) expected loss for the institution i under normal conditions, $\mathrm{VaR}^i_{0.05,t}$ quantifies its tail risk, representing the potential loss at the 95% confidence level ($\alpha = 0.05$) during periods of market stress.

To quantify each institution's contribution to systemic risk, Adrian and Brunnermeier [12] proposed and defined the CoVaR measure as follows:

$$\Pr\left(r_s(t) \le \mathrm{CoVaR}^{s|i}_{\alpha,t} \mid r_i(t) = \mathrm{VaR}^i_{\alpha,t}\right) = \alpha. \tag{3.2}$$

Here, $r_s(t)$ denotes the market return, and $\mathrm{CoVaR}^{s|i}_{\alpha,t}$ is defined as the VaR of the financial system conditional on institution i experiencing an extreme loss equal to its $\mathrm{VaR}^i_{\alpha,t}$ at time t.

To estimate CoVaR, we utilize quantile regression, incorporating two key state variables: volatility and growth rate of the CSI 300 Index*. The spillover of risk from institution i to the financial system s is quantified by the measure ΔCoVaR, defined as:

$$\Delta\mathrm{CoVaR}^{s|i}_{0.05,t} = \mathrm{CoVaR}^{s|i}_{0.05,t} - \mathrm{CoVaR}^{s|i}_{0.5,t}. \tag{3.3}$$

ΔCoVaR has proven to be a valuable tool for assessing systemic risk in financial markets.

In this paper, all estimated values of ΔCoVaR are negative. For ease of interpretation, we use the absolute values |ΔCoVaR| to measure the contributions of systemic risk, where a higher value indicates a greater contribution from the individual institution.

Figure 1 shows the sequence |ΔCoVaR| of China Pacific Insurance (CPIC,insurance), Shenwan Hongyuan (SWGH, securities), and Agricultural Bank of China (ABC, bank). Three |ΔCoVaR| sequences have a strong irregular volatility.

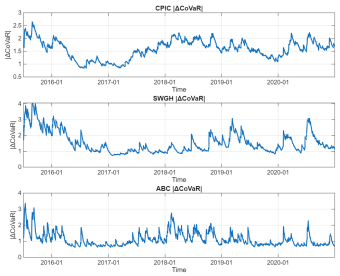

Figure 1. |ΔCoVaR| time sequence diagram: China Pacific Insurance (CPIC), Shenwan Hongyuan (SWGH), and Agricultural Bank of China (ABC).

Table 1 shows the daily average |ΔCoVaR| of each financial institution during the sample period. The calculated average values |ΔCoVaR| are 1.2657 (banking), 1.2763 (securities sector), and 1.2836 (insurance

*The CSI 300 Index, comprising the 300 largest and most liquid stocks listed on the Shanghai and Shenzhen stock exchanges, was launched on 8 April 2005 and serves as a benchmark reflecting the overall performance of China's A-share market.

sector). The insurance industry exhibits the highest systemic risk contribution, substantially exceeding the other two sectors. Among the top 10 highest-risk institutions, securities firms account for 7, insurance companies for 2, and banks for 1. This distribution indicates that systemic risk within the current financial architecture is predominantly concentrated in the securities sector. Although the number of insurance companies is small, accounting for only 9.7% of the total number of samples, Ping An of China (PING) and China Pacific Insurance (CPIC) ranked 1st and 10th respectively, implying that the insurance sector also constitutes an important source of systemic risk.

Table 1. Average value of financial institutions |ΔCoVaR|.

Financial institution	|ΔCoVaR|	Financial institution	|ΔCoVaR|
CPIC	1.5836	BONB	1.2834
SWGH	1.5506	SPDB	1.2817
UTSC	1.5119	CCB	1.2640
NESC	1.4696	CMS	1.2509
GFSC	1.4439	ORNT	1.2373
IBC	1.4167	GUOS	1.2268
SXSC	1.4105	CLIC	1.2125
NCI	1.3985	BOBJ	1.2081
SCSC	1.398	CMBC	1.1764
PING	1.3725	EBSC	1.1501
PAB	1.3714	CEBC	1.1486
GYSC	1.3607	WST	1.1241
SWS	1.3545	HXCB	1.1238
SINO	1.3523	ABC	1.121
CITIC	1.347	FOUND	1.1207
CJSC	1.3458	SEAL	1.0832
BONJ	1.343	ICBC	1.0434
DXSC	1.3404	BOC	1.0383
HTSC	1.3018	BOCOM	0.9968
CMB	1.2874	CNCB	0.9886
IND	1.2868		

Potential explanations are as follows: (1) High leverage and market sensitivity: Securities firms typically operate with high leverage, which inherently amplifies their exposure to market volatility. The procyclical nature of leverage means that during periods of financial stress, risk escalates rapidly, leading to a swift accumulation of losses. (2) Risk transmission through core business activities: As central players in margin-based businesses, most notably equity pledge financing, securities firms are deeply embedded in the market's risk transmission mechanism. A decline in the value of pledged collateral not only directly weakens their balance sheets, but can also trigger forced liquidation, setting off a negative feedback loop that ripples through the market. (3) Regulatory arbitrage and the shift of risk: Driven by regulatory arbitrage, banks often transfer credit activities to shadow banking entities, including securities and insurance firms, effectively shifting risk off their balance sheets. However, these institutions typically maintain thinner capital buffers, allowing risk to accumulate in less tightly

regulated corners of the financial system. (4) Structural vulnerabilities from common asset holdings: Due to substantial overlap in asset holdings, securities firms, and some insurers, face synchronized risk exposure when asset prices fall. (5) Interconnectedness: These institutions are deeply interconnected with commercial banks through various business channels, making them powerful conduits for contagion that can transmit localized shocks across the entire financial system.

This result is consistent with the findings of reference [14, 55]. The findings [55] indicate that non-bank financial institutions (NBFI) negatively affect financial stability and a lagged effect of NBFI expansion on financial stability, which becomes more pronounced in the medium to long term. Billio et al. [14] analyzed the connectivity among hedge funds, banks, and insurers, finding that the contribution of the insurance sector to systemic risk has grown significantly over time.

3.3. Construction of complex network model of financial system

Complex network models provide an effective framework for analyzing financial system interconnection. We construct a network of 41 financial institutions to examine the structural connectivity of China's financial industry.

(1) Nodes: The network nodes are the 41 financial institutions listed in Supplementary.

(2) Edge: We establish edges between nodes based on the transfer entropy of the series of return rate of 41 financial institutions . Transfer entropy is an important tool for analyzing causal relationships in nonlinear systems [56, 57]. It can capture directional and dynamic information without relying on any specific functional form to describe the interrelation between variables.

Consider X as a discrete random variable with probability distribution $p(x)$. The uncertainty or amount of information of X can be measured by its entropy, defined as $H(X) = -\sum_x p(x)\log_2 p(x)$. Suppose X and Y are stochastic processes, and the transfer entropy from Y to X is defined by the formula

$$\mathrm{TE}^{(k,l)}_{Y\to X}(n) = \sum_{x_{n+1},x_n^{(k)},y_n^{(l)}} p(x_{n+1}, x_n^{(k)}, y_n^{(l)}) \log_2 \frac{p(x_{n+1} \mid x_n^{(k)}, y_n^{(l)})}{p(x_{n+1} \mid x_n^{(k)})} \tag{3.4}$$

$\mathrm{TE}_{Y\to X}(k,l)$ indicates the reduction in the uncertainty of X when Y is known, highlighting the effect of Y in the prediction of X. The notation $x_n^{(k)} = (x_n, \ldots, x_{n-k+1})$ and $y_n^{(l)} = (y_n, \ldots, y_{n-l+1})$ represent the vectors of k-dimensional and l-dimensional delay embedding, and $p(x_{n+1} \mid x_n^{(k)})$ represents the conditional probability. This study sets $k = l = 1$ to incorporate the most recently available return data [58].

Recognizing the strong influence of the macroeconomy on the financial network, we use conditional transfer entropy (CTE) to quantify the relationship between two institutions [59]. To represent the macroeconomic environment, we utilize the People's Bank of China's (PBoC's) 7-day interbank offered weighted average rate as a proxy variable. The interest rate assigned to each trading day is the weighted average rate published for that month. This method provides a more comprehensive picture of the dynamics among different asset prices in response to changes in the interest rate environment. We calculate the following conditional transfer entropy with the series of the PBoC's 7-day interbank W:

$$CTE^{(1,1)}_{Y\to X|W}(n) = \sum_{x_{n+1},x_n^{(1)},y_n^{(1)}} p(x_{n+1}, x_n, y_n, w_n)\log_2 \frac{p(x_{n+1}|x_n, y_n, w_n)}{p(x_{n+1}|x_n, w_n)} \tag{3.5}$$

To simplify the network, we take the arithmetic mean of $CTE^{(1,1)}_{Y\to X|W}$ and $CTE^{(1,1)}_{X\to Y|W}$ to define the edge weight of the undirected network. Meanwhile, a threshold θ is required for the weight of the edge.

Since the choice of θ directly influences the topological structure of the network, we study the relationship between θ and the density/average degree of the networks built on the training set. By computing the curve curvature, we identify the maximum curvature at $\theta = 0.6$, suggesting that the topological structure of the network is progressively stabilizing when $\theta > 0.6$. The effect of the threshold on network clustering was further investigated. The results indicate that both clustering quality and consistency are highest when $\theta = 0.6$; see Figure 2.

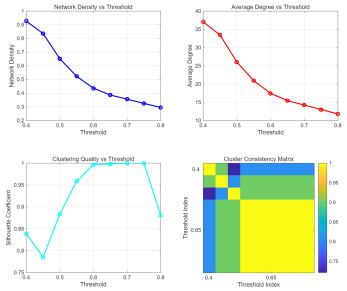

Figure 2. The relationship between threshold θ and the density/average degree of the network.

Then, using a threshold of $\theta = 0.6$, we construct the network from the entire sample dataset; see Figure 3.

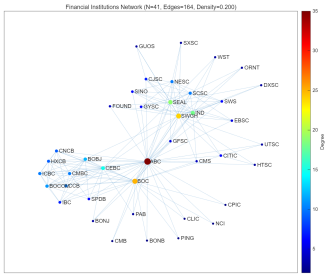

Figure 3. Complex network model of China A-share financial market.

The network topology illustrated in Figure 3 reveals strong interconnectivity among most financial institutions, exhibiting a pronounced "dumbbell" structure. Specifically, banks and insurance companies are predominantly clustered on the left side, and securities firms are concentrated on the right. The key topological characteristics of the constructed network are summarized in Table 2.

Table 2. Key topological characteristics of the complex network.

Node number	Average degree	Average degree centrality	Average path length	Network diameter	Network density
41	8	0.20	1.898	3	0.20

The degree of a node quantifies the number of edges incident to it, while degree centrality is defined as the ratio of its degree to the maximum possible degree in the network, serving as an indicator of the node's relative importance within the network structure. The distance between two nodes is defined as

the number of edges along the shortest path connecting them. The average path length of the network is the mean distance over all pairs of nodes, while the network diameter corresponds to the maximum distance observed between any two nodes. The network density, calculated as the ratio of actual edges to the maximum possible edges, quantifies the overall connectivity of the network. These topological features in Table 2 suggest that the network exhibits relatively strong nodal connectivity. Consequently, in the event of a financial shock, risks are likely to spread more rapidly throughout the system.

Table 3 lists the degree centrality of each node. Banking exhibits the highest mean degree centrality (0.2088), followed by the securities sector (0.1993) and insurance (0.05). Of the top 11 financial institutions in terms of degree centrality, 7 are banks and the remaining 4 are securities firms. This implies that banks, especially the four major state-owned commercial banks, occupy a distinct systemic position, and they fulfill a critical function in safeguarding systemic stability. The significant expansion of securities firms amid sustained national economic and financial development, and has strengthened their economic links with the real economy and other financial entities. The core operations of insurance companies are centered on underwriting and asset management. Consequently, insurers tend to foster deep, stable relationships with a limited number of reinsurance partners rather than establishing broad-based connections.

Table 3. Nodes and degree centrality of the complex network.

Financial institution	Degree centrality	Financial institution	Degree centrality
ABC	0.875	CITIC	0.15
BOC	0.625	EBSC	0.125
SWGH	0.6	GFSC	0.125
SEAL	0.475	FOUND	0.1
IND	0.45	CMS	0.1
CEBC	0.375	GUOS	0.075
BOBJ	0.3	SXSC	0.075
BOCOM	0.275	HTSC	0.075
CMBC	0.25	BONJ	0.075
ICBC	0.25	ORNT	0.075
CCB	0.25	WST	0.075
NESC	0.25	DXSC	0.075
SCSC	0.25	UTSC	0.075
HXCB	0.225	PAB	0.075
CNCB	0.225	CPIC	0.05
CJSC	0.175	NCI	0.05
IBC	0.175	PING	0.05
SWS	0.15	CLIC	0.05
SINO	0.15	BONB	0.05
SPDB	0.15	CMB	0.05
GYSC	0.15		

3.4. Cluster analysis

The value of |ΔCoVaR| captures the marginal risk contribution of individual financial institutions to the system. To proxy systemic risk, one natural approach is to compute a weighted average based on the topological importance of the nodes in the financial network. Financial institutions are closely interconnected due to business operations and other factors. In general, financial risks initially propagate among strongly connected nodes before diffusing and spilling over to other parts of the network. To better capture this dynamic, we employ clustering methods to partition the system into several subgroups, allowing us to examine both the systemic risk within each cluster and the risk spillovers between clusters.

Cluster analysis classifies nodes in a complex network to maximize intra-group homogeneity, thereby enabling the identification of high-risk subgroups and the tracing of risk transmission pathways within the financial system. This approach has been widely applied in areas such as stock index volatility analysis and loan default prevention [12, 13]. Spectral clustering, in particular, serves as an effective tool for uncovering higher-order structures in complex networks and has been extensively adopted in physics, neuroscience, social sciences, and financial risk analysis [14]. For a detailed discussion of spectral clustering algorithms, see [21].

We compared five clustering methods in terms of their silhouette coefficients/modularity and find that spectral clustering achieves relatively strong internal consistency and a well-defined community structure, indicating that financial institutions can be robustly grouped, as shown in Figure 4. The spectral clustering results are visualized in Figure 5 (left). Cluster 1 consists of all 21 securities firms, and Cluster 2 is composed of 16 banks and 4 insurance companies. Cluster stability is shown in Figure 5 (right).

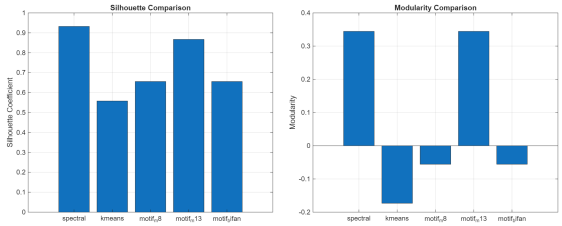

Figure 4. A comparison of five clustering methods.

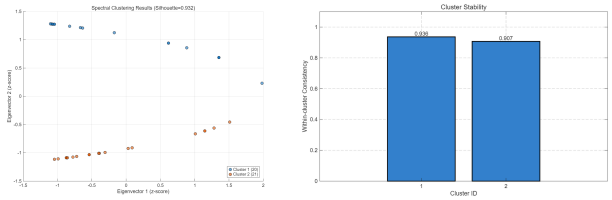

Figure 5. A comparison of five clustering methods.

3.5. The weighted average |ΔCoVaR|

To measure systemic risk in financial markets, we compute the weighted average of |ΔCoVaR| for each cluster. In principle, DebtRank represents the optimal measure for node weighting [60]. The rationale for its use lies in the non-local nature of systemic risk, a critical dimension that traditional indicators fail to capture. DebtRank offers distinct advantages: It integrates both institutional size and

interconnectedness, while also distinguishing between risk contribution and exposure. However, its practical application is hindered by data scarcity on bilateral exposures and sectoral heterogeneity in leverage and capital requirements [61]. Due to the inclusion of securities firms and insurance companies in our analysis, we encountered difficulties in obtaining inter-institutional leverage data. In this study, we adopt normalized PageRank as a feasible alternative that still captures the relative importance of institutions within the network. The PageRank value of 41 financial institutions is calculated; see Figure 6.

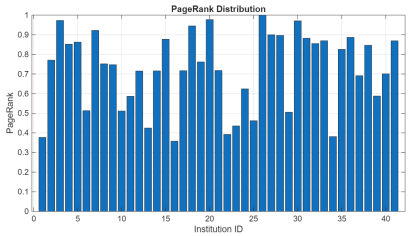

Figure 6. PageRank distribution of 41 financial institutions.

Figure 7 presents the daily weighted average |ΔCoVaR| series for the two clusters from January 2, 2019 to December 30, 2020, illustrating the dynamic evolution of systemic risk within each subgroup during the sample period. Overall, the systemic risk level of Cluster 1 is slightly higher than that of Cluster 2.

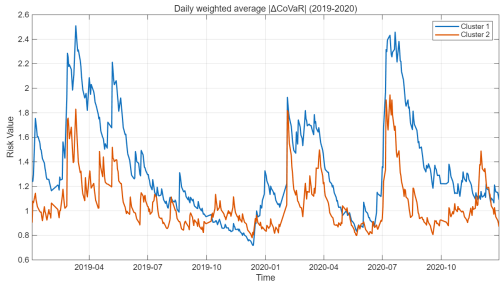

Figure 7. Daily weighted average value of the two clusters from 2019 to 2020.

The daily |ΔCoVaR| series in Figure 7 exhibit irregular and jagged fluctuations. Given the nonlinear and noisy nature of financial time series, it is essential to examine whether these sequences exhibit chaotic behavior. The Lyapunov exponent can be used to test chaos [43, 44]. A time series is considered chaotic if its largest Lyapunov exponent is positive. Among the various methods available for computing the Lyapunov exponent, we adopt the small-data-quantity algorithm [43]. This approach is chosen for its reliability, computational efficiency, and ease of implementation, which makes it well-suited for practical applications. The computed Lyapunov exponents for three |ΔCoVaR| are 0.035 and 0.03, respectively, all exceeding zero. Hence, we conclude that the two series display chaotic dynamics.

Table 4 reports the descriptive statistics of systemic risk for each cluster. Cluster 1 shows a higher mean risk value (1.334), extreme risk (3.252), and considerable volatility. In comparison, Cluster 2 displays a lower mean risk value and a smaller fluctuation amplitude, suggesting that its systemic risk remains relatively stable over the observed period. In general, financial risks propagate from higher-risk institutions to lower-risk entities and further spread to other entities once a spillover threshold is reached [13].

Table 4. Descriptive statistical characteristics of weighted average |ΔCoVaR| value.

Cluster	MAX	MIN	MEAN
Cluster1	3.252	0.685	1.334
Cluster2	2.845	0.737	1.151

Moreover, using the spillover index model proposed by Diebold and Yilmaz in reference [62], we calculate the intensity and direction of risk spillovers. The results are presented in Figure 8. From subgraph "Net Spillover Distribution (Blue=Sender, Red=Receiver)" in Figure 8, the net risk transmitters comprise 19 institutions, 3 insurers and 16 securities firms; while the net risk receivers total 22 institutions, consisting of 16 banks, 5 securities firms, and 1 insurer company. The comparison of the two classifications, as reported in Table 5, reveals that Cluster 1 primarily comprises risk transmitters, and Cluster 2 primarily risk receivers.

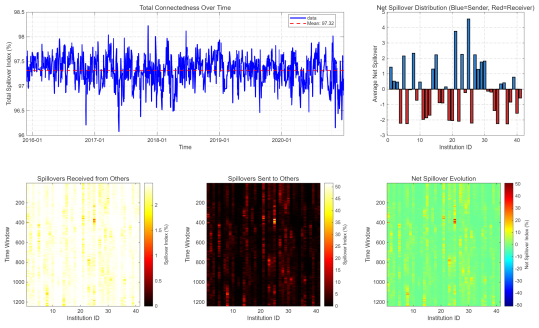

Figure 8. The results of the spillover index model from 2015 to 2020.

Table 5. Descriptive statistical characteristics of weighted average |ΔCoVaR| value.

Cluster	Sender	Receiver
Cluster1	16(securities)	5(securities)
Cluster2	3(insures)	16(banks)1(insure)

4. Prediction and analysis of systemic financial risk

Systemic financial risk is influenced by a multitude of factors, including political, economic, and environmental conditions, making its simulation and prediction of data highly challenging. Conventional signal extraction and differencing methods have limited effectiveness in denoising chaotic |ΔCoVaR| sequences. To address these difficulties, we employ a hybrid approach integrating phase space reconstruction with delay parameterization (RDPM) and a radial basis function neural network (RBFNN) for simulation and forecasting.

4.1. Forecast result

Let the daily weighted average series |ΔCoVaR| for the three clusters be denoted as the observed sequences (y_1, y_2). The dataset is typically partitioned into three subsets: 60% for training, 20% for validation, and the remaining 20% for testing. Following the algorithm described in Section 2, the optimally-trained time delay obtained is

$$\tau^* = (\tau_1, \tau_2) = (49, 25).$$

We first perform a single-step prediction. A comparison between the actual values and the predicted values is presented in Figure 9. The corresponding prediction errors and accuracy, measured by MAE, RMSE, and MRA, for each cluster are reported in Table 6.

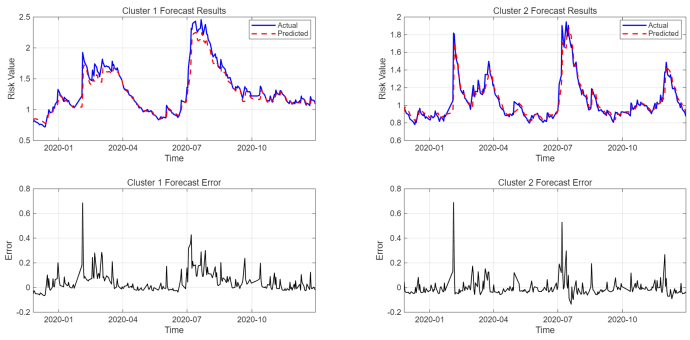

Figure 9. Prediction comparison chart.

Table 6. Prediction error indicators.

Cluster	MAE	RMSE	MRA
Cluster1	0.0476	0.0805	95.24%
Cluster2	0.0481	0.0836	95.19%

Overall, the predicted series largely captures the fluctuation patterns of the actual series, though noticeable deviations occur in certain periods. The prediction error metrics show that the MAE and RMSE values remain below 0.1 for two clusters. The prediction accuracy rates are 96.44% for Cluster 1, and 95.45% for Cluster 2.

In practical risk prevention and control, accurately predicting the direction of financial risk movement provides critical guidance for implementing preventive measures. Following the methodology outlined in [44], we evaluate the accuracy of systemic risk trend forecasts using the Hit ratio metric.

The Hit ratio is defined as

$$\text{Hit ratio} = \frac{1}{n}\sum_{i=1}^{n} D_i, \quad i = 1, 2, \cdots, n \tag{4.1}$$

where n is the number of samples, and $\sum_{i=1}^{n} D_i$ denotes the number of correct forecasts of the direction of financial risk. D_i represents the prediction of the direction change in systemic financial risk. The calculation formula is as follows:

$$D_i = \begin{cases} 1, & (R_{real}(i+1) - R_{real}(i))(R_{forecast}(i+1) - R_{forecast}(i)) > 0 \\ 0, & otherwise \end{cases} \tag{4.2}$$

As shown in Figure 10, the Hit ratio remains stable at 60%–70% as the forecast horizon extends; both clusters exceed 50% for most of the sample. These results suggest that the proposed model provides reliable trend forecasts for timely preventive measures against systemic financial risk.

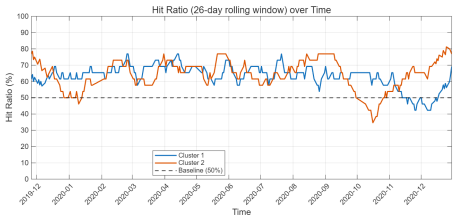

Figure 10. The trend of Hit ratio.

4.2. Comparison with other machine learning methods

Although machine learning offers powerful capabilities for data simulation and prediction, its practical applicability is often constrained by unknown influencing factors and limited interpretability of the underlying predictive mechanisms. In this section, we conduct a comparative analysis of predictions generated by the radial basis function (RBF) neural network, support vector machine (SVM), long short-term memory (LSTM) neural network, and graph neural networks (GNNs). The prediction results are illustrated in Figure 11, and a comparison of error metrics is provided in Table 7.

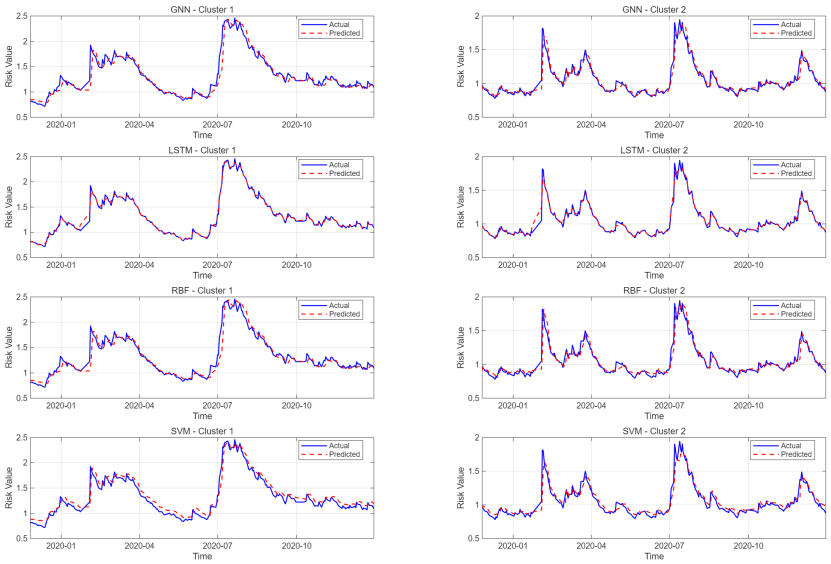

Figure 11. Comparison of machine learning model predictions: GNN, LSTM, RBF, SVM.

The RDPM-RBFNN method demonstrates a high R^2 coupled with a low error rate, indicating strong predictive performance. GNN, LSTM, and RDPM-RBFNN are all capable of capturing the complex nonlinear relationships inherent in financial systems, each demonstrating strong predictive performance from different perspectives. GNN explicitly models inter-institutional risk transmission networks to capture spatial dependencies; however, its effectiveness is contingent on the accuracy of the predefined graph structure and the completeness of the data. LSTMs, as a powerful time series model, excel in large-scale data scenarios, yet its black-box nature limits its applicability in risk interpretation and attribution analysis. By contrast, the RDPM-RBFNN method is grounded in the principles of phase space reconstruction, which stems from dynamical systems theory. It recovers the system's latent dynamical structure from observational data, thereby elucidating nonlinear systemic risk transmission mechanisms. Rather than modeling the two cluster-level time series separately, we jointly estimate their dynamics, allowing each prediction to leverage information from both series. This approach captures interconnected financial risk more comprehensively. Meanwhile, the RDPM-RBFNN offers two distinct advantages: Adaptability to small samples and robustness to non-stationarity. These features make it particularly suitable for financial systems characterized by limited data, structural complexity, and pronounced chaotic dynamics. Nevertheless, we acknowledge that this method faces challenges when handling high-dimensional data. To address this limitation, we first construct a network and apply clustering to reduce dimensionality before prediction. Network construction helps uncover spatial linkages among institutions, providing additional insight into the intrinsic mechanisms governing the financial system.

Table 7. Prediction error of the machine learning model.

Cluster	Model	R^2	MAE	RMSE
Cluster1	RBF	0.9204	0.0657	0.1131
Cluster1	SVM	0.9437	0.0796	0.0951
Cluster1	LSTM	0.9922	0.0199	0.0354
Cluster1	GNN	0.9227	0.0651	0.1114
Cluster1	RDPM-RBFNN	0.9397	0.0476	0.0805
Cluster2	RBF	0.8977	0.045	0.0773
Cluster2	SVM	0.8879	0.0521	0.0809
Cluster2	LSTM	0.9742	0.0217	0.0388
Cluster2	GNN	0.9008	0.0432	0.0761
Cluster2	RDPM-RBFNN	0.9005	0.0481	0.0836

4.3. Robustness test of prediction method

4.3.1. Changing the frequency of data

To assess the general applicability of the RDPM method in predicting systemic financial risk, we perform a robustness test by altering the frequency of the data. As systemic risk evolves over extended periods, forecasting on longer time scales also provides policy-relevant insights. Accordingly, we re-estimate the model using weekly data instead of daily observations. The prediction results based

on weekly degree data are shown in Figure 12, and the corresponding error analysis is summarized in Table 8. Figure 12 and Table 8 indicate that the model achieves the best fit for Cluster 2. The prediction errors remain on the same order of magnitude as those of the daily-frequency analysis, confirming the model's capacity to capture the dynamic risk evolution at weekly frequency.

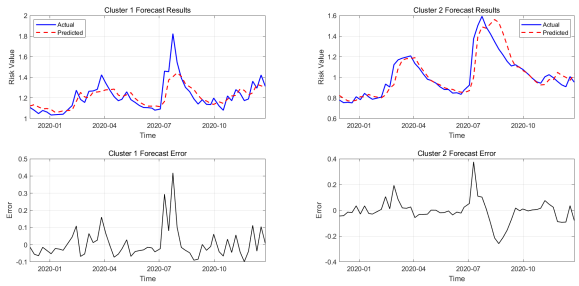

Figure 12. Prediction comparison chart of weekly data.

Table 8. Prediction error of weekly data.

Cluster	MAE	RMSE	MRA
Cluster1	0.0596	0.089	94.04%
Cluster2	0.0585	0.0925	94.15%

Although certain predicted points exhibit noticeable deviations from actual observations, such discrepancies are expected given that lower-frequency data aggregation inevitably loses some high-frequency information, and the reduced sample size may compromise training precision. Nevertheless, the consistency across frequencies underscores the robustness of the proposed method to variations in sampling intervals.

4.3.2. Other sample periods

To further evaluate the robustness of the model, we perform another test using data from a more recent period. Specifically, daily degree data from January 5, 2021, to December 31, 2025, are used to assess the model's temporal stability. The prediction results are shown in Figure 13.

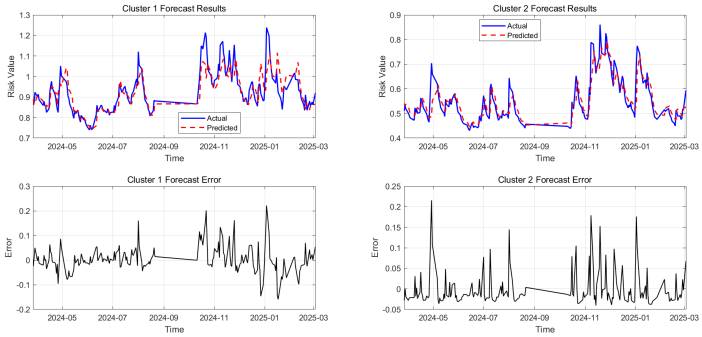

Figure 13. Prediction comparison chart for 2024–2025.

The error analysis is summarized in Table 9. The mean absolute error and mean squared error maintain the same order of magnitude as those obtained from the two previous sample periods, with no anomalous variability observed. These results demonstrate that the model retains its ability to track the dynamic evolution of systemic risk even when applied to a temporally distinct data set. The consistent predictive performance across multiple time intervals reinforces the robustness of the method and supports its applicability to financial risk data from different periods. In implementation, the balance between forecast efficiency and accuracy can be adjusted according to specific practical requirements.

Table 9. Forecast error for 2024–2025.

Cluster	MAE	RMSE	MRA
Cluster1	0.0403	0.0544	95.97%
Cluster2	0.0592	0.0418	94.80%

5. Further study

5.1. The influence of multi-step forecasting days on forecasting accuracy

This section applies the RDPM-RBFNN framework to multi-step recursive forecasting. Under this scheme, the predicted values at each step are fed back into the observed series for subsequent forecasting. Although this approach enhances computational efficiency, it inevitably leads to error accumulation; determining the optimal forecast horizon thus remains a critical challenge. Table 10 indicates that extending the horizon to 5 days yields a systematic increase in error. Beyond 5 days, errors fluctuate irregularly but trend upward overall.

Table 10. Forecast error of multi-step forecasting.

Cluster	R^2	MAE	RMSE
Cluster1	0.7213	0.1326	0.1653
Cluster2	0.6649	0.1213	0.1460

5.2. Validation of risk series: Event responses and external Benchmark comparison

It is well known that external events act as the direct catalyst for the materialization of systemic risk. Under extreme market conditions, the convergence of external shocks with vulnerabilities in the financial system substantially increases systemic risk [63]. The objective of this subsection is simply to validate the effectiveness of the constructed systemic risk series from two perspectives. The analysis that follows is undertaken using only a basic descriptive method, without subjecting it to rigorous empirical examination. This constitutes both a shortcoming of the present study and a focal point for subsequent investigation.

First, we employ an event study approach to examine risk value changes in both clusters around major events. Focusing on four major events (2015 Stock Market Crash (Black Monday), 2016 Circuit Breaker Triggered, 2018 Escalation of US-China Trade Tensions, 2020 COVID-19 Global Market Crash), we observe that the risk values of Cluster 1 and Cluster 2 over a [-5, +5] day window surrounding four major financial events. Cluster 1 exhibits pronounced spikes on the event day (t=0), followed by

partial reversals, whereas Cluster 2 shows relatively muted responses. This pattern corroborates the classification of Cluster 1 as risk transmitters and Cluster 2 as risk receivers; see Figure 14.

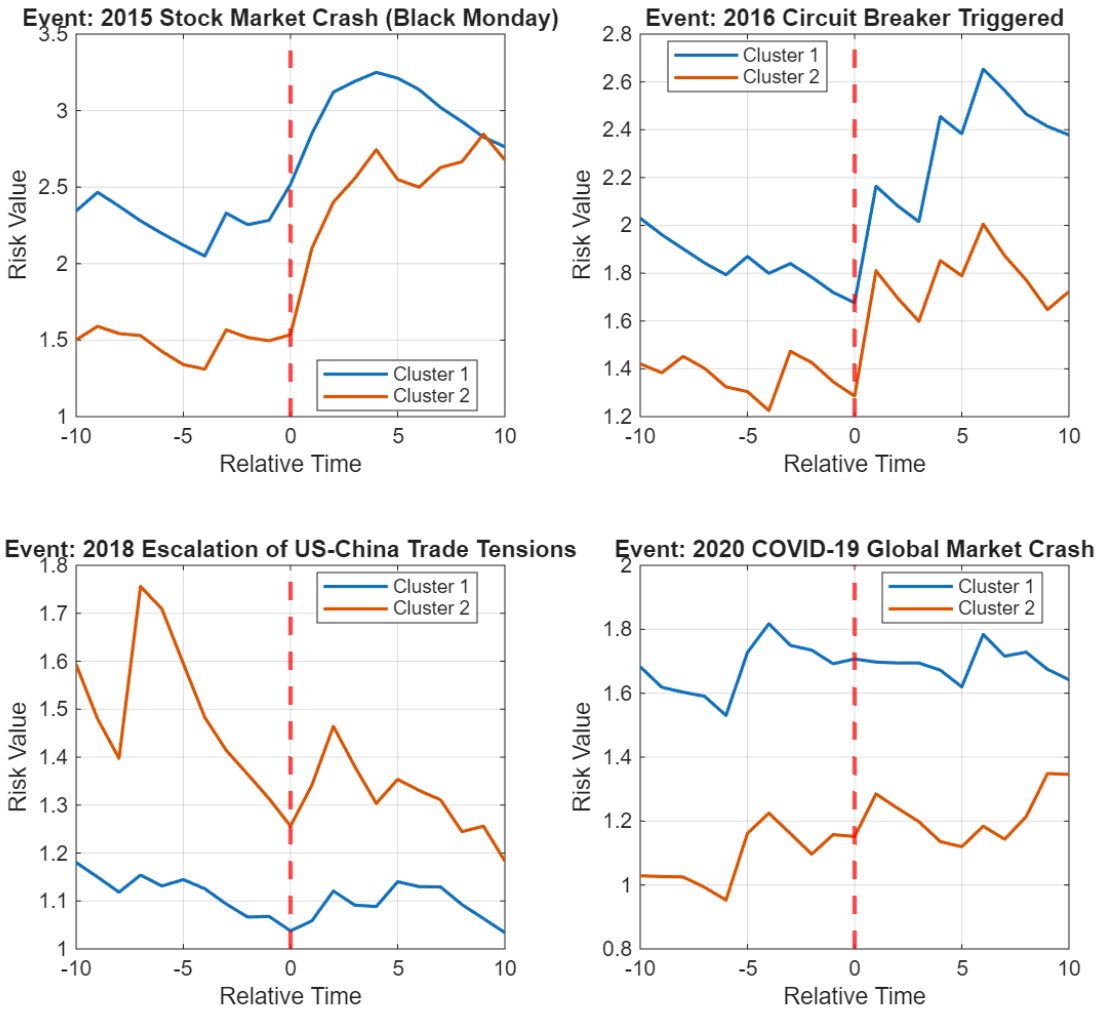

Figure 14. Systemic risk dynamics of the two clusters around major events.

Second, we examine whether external geopolitical shocks transmit to the domestic financial system and whether the transmission channels vary across institution types. We compare our predicted risk series with the geopolitical risk index (GPR), which was introduced in [64]. Figure 15 illustrates the alignment between the high-risk periods identified by our model and GPR peaks (above its 90th percentile) during January–December 2020.

We find that the high-risk periods in Clusters overlap with some instances of high GPR values, but the overall alignment remains suboptimal. The statistical results reveal a negative and statistically insignificant correlation between the two variables ($\varphi \approx -0.07, p > 0.24$ ρ_s ranging from -0.04 to -0.18), which implies that although the high-risk periods identified in each cluster exhibit a temporal correlation with episodes of elevated GPR index, their degree of synchronicity is weak. The result indicates that within

the sample period, the GPR index has limited predictive power for actual high-risk episodes in domestic financial markets. This may stem from the following three reasons: (1) the relatively limited sample size, which restricts statistical power; (2) the likely complex non-linear nature of the relationship between the GPR index and high-risk values; and (3) there may be unaccounted confounding variables that are not controlled for in the analysis. Meanwhile, Cluster 2 displays a greater negative sensitivity than Cluster 1, pointing to differential responses by institution type.

In contrast to the descriptive approach adopted here, future research could pursue rigorous empirical investigation along several dimensions to better capture the complex and nonlinear nature of the interplay between external shocks and systemic risk. On the data side, employing higher-frequency data and incorporating additional geopolitical indicators would enable more granular empirical analysis. On the methodological side, nonlinear models or time-varying frameworks would allow rigorous empirical testing of how external shocks interact with structural vulnerabilities to amplify systemic risk.

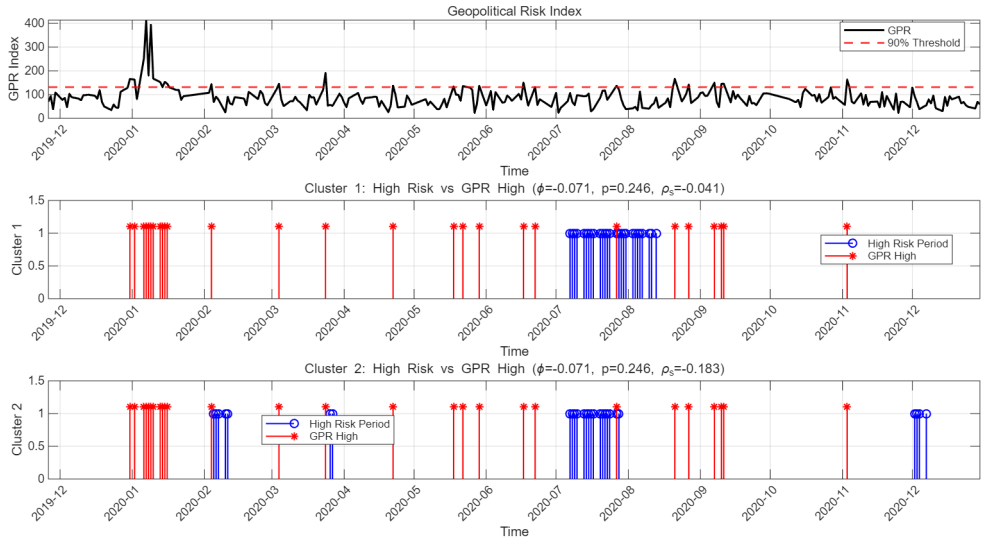

Figure 15. The validity of the model-predicted high-risk periods correspond to the peaks in GPR Index.

6. Conclusions

This paper introduces a useful method for systemic financial risk prediction. We first derive institution-specific |ΔCoVaR| measures from daily stock prices of financial firms and the CSI 300 index, capturing their marginal contribution to system risk. These institutions are then embedded in a complex network, whose higher-order community structure is extracted by spectral clustering, yielding two risk clusters.

The systemic risk at the cluster level is quantified by the weighted mean of |ΔCoVaR| within each group. To forecast its evolution, we design a hybrid model that synergistically combines phase space reconstruction with delay parameterization (RDPM) and a radial basis function neural network (RBFNN). Empirical results indicate that the model achieves more than 94% relative forecast accuracy. Its robustness is rigorously validated through sensitivity checks, frequency-varying tests, and out-of-sample evaluations.

The RDPM-RBFNN approach reconstructs the underlying dynamical system by embedding a univariate time series into a high-dimensional phase space, thereby offering distinct advantages in characterizing the nonlinear dynamics inherent in financial risk. Its principal merit lies in its interpretability: It enables the estimation of measures including the largest Lyapunov exponent and time delay parameters, which quantitatively capture the system's chaotic properties and nonlinear evolutionary patterns. On the other hand, phase space reconstruction requires less data and exhibiting resilience to non-stationarity. It provides a rigorous framework for diagnosing chaotic dynamics in financial systems, thereby yielding deeper insights into the complex mechanisms governing risk propagation.

Although banks typically carry higher leverage as primary lending institutions and are often assumed to be central to systemic risk, our findings show that in the current financial landscape of China, securities and insurance companies constitute the dominant sources of systemic risk, whereas banks show notable stability. This reversal of conventional risk hierarchies illustrates how systemic risk concentration shifts across financial subsectors along with developmental trajectories. The elevated risk contribution of securities and insurance firms during the sample period, driven by their rapid expansion and deepening network integration, calls for a differentiated, activity-based approach to monitoring, with particular emphasis on systemically significant non-bank financial intermediaries.

We think that the proposed framework in this paper is transferable to a broad spectrum of financial and economic systems and can be extended to model systemic interdependence in other complex adaptive systems.

Use of AI tools declaration

The authors declare that no Artificial Intelligence (AI) tools were used in the creation of this article, except for English language editing and polishing.

Acknowledgments

The authors sincerely thank the editors and anonymous reviewers for their valuable comments and suggestions.

This work is supported by the National Natural Science Foundation of China (Grant numbers: 12071302) and the Mentor Academic Guidance Program of Shanghai International Studies University (Grant number: 2022113028).

Conflict of interest

The authors declare there is no conflict of interest.

Author contributions

Methodology, Yulian An and Zhunjun Liu; Data curation, Software, Zhunjun Liu, Zhenkuan Sun and Boan Li; Writing-original draft, Zhunjun Liu, Zhenkuan Sun and Yulian An; Writing-review and editing, Yulian An and Boan Li; Funding acquisition, Yulian An.

References

1. W. Silva, H. Kimura, V. A. Sobreiro, An analysis of the literature on systemic financial risk: A survey, *J. Financ. Stab.*, **17** (2017), 91–114. https://doi.org/10.1016/j.jfs.2016.12.004
2. K. Bluwstein, M. Buckmann, A. Joseph, S. Kapadia, Ö. Şimşek, Credit growth, the yield curve and financial crisis prediction: Evidence from a machine learning approach, *J. Int. Econ.*, **145** (2023), 103773. https://doi.org/10.1016/j.jinteco.2023.103773
3. S. Zedda, G. Cannas, Analysis of banks' systemic risk contribution and contagion determinants through the leave-one-out approach, *J. Banking Finance*, **112** (2020), 105160. https://doi.org/10.1016/j.jbankfin.2017.06.008
4. G. Hale, T. Kapan, C. Minoiu, Shock transmission through cross-border bank lending: Credit and real effects, *Rev. Financ. Stud.*, **33** (2020), 4839–4882. https://doi.org/10.1093/rfs/hhz147
5. M. V. Oordt, C. Zhou, Systemic risk and bank business models, *J. Appl. Econom.*, **34** (2019), 365–384. https://doi.org/10.1002/jae.2666
6. V. Elenev, T. Landvoigt, S. V. Nieuwerburgh, A macroeconomic model with financially constrained producers and intermediaries, *Econometrica*, **89** (2021), 1361–1418. https://doi.org/10.3982/ECTA16438
7. S. Benoit, J. E. Colliard, C. Hurlin, C. Pérignon, Where the risks lie: A survey on systemic risk, *Rev. Finance*, **21** (2017), 109–152. https://doi.org/10.1093/rof/rfw026
8. C. Diks, C. Hommes, J. Wang, Critical slowing down as an early warning signal for financial crises, *Empir. Econ.*, **57** (2019), 1201–1228. https://doi.org/10.1007/s00181-018-1527-3
9. M. Summer, Financial contagion and network analysis, *Annu. Rev. Financ. Econ.*, **5** (2013), 277–297. https://doi.org/10.1146/annurev-financial-110112-120948
10. Z. Du, J. C. Escanciano, Backtesting expected shortfall: Accounting for tail risk, *Manage. Sci.*, **63** (2017), 940–958. https://doi.org/10.1287/mnsc.2015.2342
11. V. V. Acharya, L. H. Pedersen, T. Philippon, M. Richardson, Measuring systemic risk, *Rev. Financ. Stud.*, **30** (2017), 2–47. https://doi.org/10.1093/rfs/hhw088
12. T. Adrian, M. K. Brunnermeier, CoVaR, *Am. Econ. Rev.*, **106** (2016), 1705–1741. http://dx.doi.org/10.1257/aer.20120555
13. E. Grant, J. Yung, The double-edged sword of global integration: Robustness, fragility, and contagion in the international firm network, *J. Appl. Econom.*, **36** (2021), 760–783. https://doi.org/10.1002/jae.2839
14. M. Billio, M. Getmansky, A. W. Lo, L. Pelizzon, Econometric measures of connectedness and systemic risk in the finance and insurance sectors, *J. Financ. Econ.*, **104** (2012), 535–559. https://doi.org/10.1016/j.jfineco.2011.12.010
15. G. Bostanci, K. Yilmaz, How connected is the global sovereign credit risk network, *J. Bank. Finance*, **113** (2020), 105761. https://doi.org/10.1016/j.jbankfin.2020.105761
16. G. J. Wang, L. Wan, Y. Feng, C. Xie, G. S. Uddin, Y. Zhu, Interconnected multilayer networks: Quantifying connectedness among global stock and foreign exchange markets, *Int. Rev. Financ. Anal.*, **86** (2023), 102518. https://doi.org/10.1016/j.irfa.2023.102518

17. Z. Ouyang, X. Zhou, Multilayer networks in the frequency domain: Measuring extreme risk connectedness of Chinese financial institutions, *Res. Int. Bus. Finance*, **65** (2023), 101944. https://doi.org/10.1016/j.ribaf.2023.101944
18. Ö. Akgüller, M. A. Balcı, Detecting financial contagion through higher-order networks: A deep learning approach to emerging market risk, *Comput. Econ.*, (2026), 1–67. https://doi.org/10.1007/s10614-025-11287-3
19. F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, G. Monfardini, The graph neural network model, *IEEE Trans. Neural Networks*, **20** (2009), 61–80. https://doi.org/10.1109/TNN.2008.2005605
20. D. Parnes, A. Gormus, Prescreening bank failures with K-means clustering: Pros and cons, *Int. Rev. Financ. Anal.*, **93** (2024), 103222. https://doi.org/10.1016/j.irfa.2024.103222
21. K. K. Sharma, A. Seal, Spectral embedded generalized mean based k-nearest neighbors clustering with s-distance, *Expert Syst. Appl.*, **169** (2021), 114326. https://doi.org/10.1016/j.eswa.2020.114326
22. A. R. Benson, D. F. Gleich, J. Leskovec, Higher-order organization of complex networks, *Science*, **353** (2016), 163–166. https://doi.org/10.1126/science.aad9029
23. J. Duffy, A. Karadimitropoulou, M. Parravano, Financial contagion in the laboratory: Does network structure matter, *J. Money Credit Bank.*, **51** (2019), 1097–1136. https://doi.org/10.1111/jmcb.12563
24. C. Filippopoulou, E. Galariotis, S. Spyrou, An early warning system for predicting systemic banking crises in the Eurozone: A logit regression approach, *J. Econ. Behav. Organ.*, **172** (2020), 344–363. https://doi.org/10.1016/j.jebo.2019.12.023
25. Z. Gu, Y. Xu, Chaotic dynamics analysis based on financial time series, *Complexity*, (2021), 2373423. https://doi.org/10.1155/2021/2373423
26. M. Pirani, P. Thakka, P. Jivrani, M. H. Bohara, D. Garg, A comparative analysis of ARIMA, GRU, LSTM and BiLSTM on financial time series forecasting, in *2022 IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*, Ballari, India, 2022, 1–6. https://doi.org/10.1109/ICDCECE53908.2022.9793213
27. N. Nazareth, Y. V. R. Reddy, Financial applications of machine learning: A literature review, *Expert Syst. Appl.*, (2023) 119640. https://doi.org/10.1016/j.eswa.2023.119640
28. S. Ahmed, M. M. Alshater, A. El Ammari, H. Hammami, Artificial intelligence and machine learning in finance: A bibliometric review, *Res. Int. Bus. Finance*, **61** (2022), 101646. https://doi.org/10.1016/j.ribaf.2022.101646
29. G. Kou, X. Chao, Y. Peng, F. E. Alsaadi, E. Herrera-Viedma, Machine learning methods for systemic risk analysis in financial sectors, *Technol. Econ. Dev. Econ.*, **25** (2019), 716–742. https://doi.org/10.3846/tede.2019.8740
30. J. Beutel, S. List, G. von Schweinitz, Does machine learning help us predict banking crises, *J. Financ. Stab.*, **45** (2019), 100693. https://doi.org/10.1016/j.jfs.2019.100693
31. V. Balmaseda, M. Coronado, G. C. Santiago, Predicting systemic risk in financial systems using Deep Graph Learning, *Intell. Syst.*, (2023), 200240. https://doi.org/10.1016/j.iswa.2023.200240
32. T. Leonidas, T. Alexandra, B. Aristeidis, Big data in financial risk management: Evidence, advances, and open questions: A systematic review, *Front. Artif. Intell.*, **8** (2025), 1658375. https://doi.org/10.3389/frai.2025.1658375

33. L. J. Vásquez-Serpa, C. Rodríguez, J. R. Pérez-Núñez, C. Navarro, Challenges of Artificial Intelligence for the prevention and identification of bankruptcy risk in financial institutions: A systematic review, *J. Risk Financ. Manage.*, **18** (2025), 26. https://doi.org/10.3390/jrfm18010026
34. A. Wolf, J. B. Swift, H. L. Swinney, J. A. Vastano, Determining Lyapunov exponents from a time series, *Phys. D: Nonlinear Phenom.*, **16** (1985), 285–317. https://doi.org/10.1016/0167-2789(85)90011-9
35. M. D. Johansyah, S. Vaidyanathan, K. Benkouider, A. Sambas, C. Aruna, S. K. Annavarapu, et al., A chaotic butterfly attractor model for economic stability assessment in financial systems, *Mathematics*, **13** (2025), 1633. https://doi.org/10.3390/math13101633
36. K. He, J. Shi, H. Fang, Bifurcation and chaos analysis of a fractional-order delay financial risk system using dynamic system approach and persistent homology, *Math. Comput. Simul.*, **223** (2024), 253–274. https://doi.org/10.1016/j.matcom.2024.04.013
37. X. Yan, H. Wang, Y. An, Forecasting systemic risk of China's banking industry by PDE model, *J. Appl. Anal. Comput.*, **13** (2023), 1–27. https://doi.org/10.11948/20230306
38. L. Yang, T. Gao, Y. Lu, J. Duan, T. Liu, Neural network stochastic differential equation models with applications to financial data forecasting, *Appl. Math. Model.*, **115** (2023), 279–299. https://doi.org/10.1016/j.apm.2022.11.001
39. W. Sun, Y. An, Y. Gao, Systemic risk contagion in China's financial-real estate network: Modeling and forecasting via fractional order PDEs, *Fractal Fract.*, **9** (2025), 557. https://doi.org/10.3390/fractalfract9090557
40. X. Guo, W. Han, J. Ren, Design of a prediction system based on the dynamical feed-forward neural network, *Sci. China Inf. Sci.*, **66** (2023), 1–17. https://doi.org/10.1007/s11432-020-3402-9
41. X. Guo, Y. Sun, J. Ren, Low dimensional mid-term chaotic time series prediction by delay parameterized method, *Inf. Sci.*, **516** (2020), 1–19. https://doi.org/10.1016/j.ins.2019.12.021
42. A. Zhang, Z. Xu, Chaotic time series prediction using phase space reconstruction based conceptor network, *Cogn. Neurodyn.*, **14** (2020), 849–857. https://doi.org/10.1007/s11571-020-09612-7
43. H. Ma, S. Leng, K. Aihara, L. Chen, Randomly distributed embedding making short-term high-dimensional data predictable, *Proc. Natl. Acad. Sci. U.S.A.*, **115** (2018), E9994-E10002. https://doi.org/10.1073/pnas.1802987115
44. C. Chen, R. Li, L. Shu, Z. He, J. Wang, C. Zhang, et al., Predicting future dynamics from short-term time series using an Anticipated Learning Machine, *Natl. Sci. Rev.*, **7** (2020), 1079–1091. https://doi.org/10.1093/nsr/nwaa025
45. J. Hou, H. Ma, D. He, J. Sun, Q. Nie, W. Lin, Harvesting random embedding for high-frequency change-point detection in temporal complex systems, *Natl. Sci. Rev.*, **9** (2022), 92–104. https://doi.org/10.1093/nsr/nwab228
46. M. Lan, Energy control and chaos prediction of a fractional-order financial risk contagion system, *Phys. Scr.*, **100** (2025), 085240. https://doi.org/10.1088/1402-4896/adf3ee
47. Y. Wang, M. F. Ghazali, R. A. Razak, M. A. S. Zaidi, Complex system and PS-LSTM prediction of cryptocurrencies, stocks, bonds, exchange rates and commodities, *Phys. A: Stat. Mech. Appl.*, (2025), 130976. https://doi.org/10.1016/j.physa.2025.130976

48. M. S. Murugan, Large-scale data-driven financial risk management & analysis using machine learning strategies, *Meas.: Sens.*, **27** (2023), 100756. https://doi.org/10.1016/j.measen.2023.100756

49. X. Fan, Y. Wang, D. Wang, Network connectedness and China's systemic financial risk contagion analysis based on big data, *Pac. Basin Finance J.*, **68** (2021), 101322. https://doi.org/10.1016/j.pacfin.2020.101322

50. P. Song, Y. Xiao, Estimating time-varying reproduction number by deep learning techniques, *J. Appl. Anal. Comput.*, **12** (2022), 1077–1089. https://doi.org/10.11948/20220136

51. H. Hu, L. Tang, S. Zhang, H. Wang, Predicting the direction of stock markets using optimized neural networks with Google trends, *Neurocomputing*, **285** (2018), 188–195. https://doi.org/10.1016/j.neucom.2018.01.038

52. I. Choi, W. C. Kim, Practical forecasting of risk boundaries for industrial metals and critical minerals via statistical machine learning techniques, *Int. Rev. Financ. Anal.*, **94** (2024), 103252. https://doi.org/10.1016/j.irfa.2024.103252

53. F. Takens, Detecting strange attractors in turbulence, *Dyn. Syst. Turbul.*, **344** (1981), 366–381. https://doi.org/10.1007/BFb0091924

54. E. R. Deyle, G. Sugihara, Generalized theorems for nonlinear state space reconstruction, *PLoS ONE*, **6** (2011), 18295. https://doi.org/10.1371/journal.pone.0018295

55. L. C. Zheng, X. Q. Huang, X. Y. Lu, Nonbank financial institutions and financial stability: Time series analysis, *Financ. Res. Lett.*, **73** (2025), 106544. https://doi.org/10.1016/j.frl.2024.106544

56. Y. Gao, R. Tan, C. Fu, S. Cai, Revealing stock market risk from information flow based on transfer entropy: The case of Chinese A-shares, *Phys. A: Stat. Mech. Appl.*, (2023), 128982. https://doi.org/10.1016/j.physa.2023.128982

57. L. Barnett, A. B. Barrett, A. K. Seth, Granger causality and transfer entropy are equivalent for Gaussian variables, *Phys. Rev. Lett.*, **103** (2009), 238701. https://doi.org/10.1103/PhysRevLett.103.238701

58. S. Thomas, Measuring information transfer, *Phys. Rev. Lett.*, **85** (2000), 461. https://doi.org/10.1103/PhysRevLett.85.461

59. I. Choi, W. C. Kim, A temporal information transfer network approach considering federal funds rate for an interpretable asset fluctuation prediction framework, *Int. Rev. Econ. Finance*, **96** (2024), 103562. https://doi.org/10.1016/j.iref.2024.103562

60. S. Battiston, M. Puliga, R. Kaushik, P. Tasca, G. Caldarelli, DebtRank: Too central to fail? Financial networks, the FED and systemic risk, *Sci. Rep.*, **2** (2012), 541. https://doi.org/10.1038/srep00541

61. J. Sànchez García, S. Cruz Rambaud, Systemic risk in a macro-multiplex network, *Soft Comput.*, **30** (2026), 1839–1851. https://doi.org/10.1007/s00500-023-09460-7

62. F. X. Diebold, K. Yilmaz, On the network topology of variance decompositions: Measuring the connectedness of financial firms, *J. Econom.*, **182** (2014), 119–134. https://doi.org/10.1016/j.jeconom.2014.04.012

63. H. Tzavellas, A multilayer view of systemic importance and aggregate fluctuations, *Int. Econ. Rev.*, **64** (2023), 1023—1046. https://doi.org/10.1111/iere.12622

64. D. Caldara, M. Iacoviello, Measuring geopolitical risk, *Am. Econ. Rev.*, **112** (2022), 1194–1225. https://doi.org/10.1257/aer.20191823

Supplementary A: 41 listed financial institutions

Table A1. 41 listed financial institutions and their abbreviations.

Institutions	Abbreviation	Institutions	Abbreviation
Founder Securities	FOUND 1	Shanghai Pudong Development Bank	SPDB 22
China Pacific Insurance	CPIC 2	Sealand Securities	SEAL 23
Changjiang Securities	CJSC 3	Industrial Securities	IND 24
Bank of Beijing	BOBJ 4	Guoyuan Securities	GYSC 25
New China Life Insurance	NCI 5	Agricultural Bank of China	ABC 26
China Minsheng Banking	CMBC 6	Dongxing Securities	DXSC 27
Southwest Securities	WST 7	Northeast Securities	NESC 28
Sinolink Securities	SINO 8	Huatai Securities	UTSC 29
Ping An Insurance	PING 9	Soochow Securities	SCSC 30
Guosen Securities	GUOS 10	Ping An Bank	PAB 31
Industrial and Commercial Bank of China	ICBC 11	CITIC Securities	CITIC 32
China Everbright Bank	CEBC 12	Shenwan Hongyuan Group	SWGH 33
China Construction Bank	CCB 13	Bank of Communications	BOCOM 34
China Life Insurance	CLIC 14	China Merchants Securities	CMS 35
Shanxi Securities	SXSC 15	Bank of Ningbo	BONB 36
Haitong Securities	HTSC 16	Bank of China Limited	BOC 37
Bank of Nanjing	BONJ 17	China Merchants Bank	CMB 38
Orient Securities	ORNT 18	Everbright Securities	EBSC 39
Industrial Bank	IBC 19	China CITIC Bank	CNCB 40
Huaxia Bank	HXCB 20	GF Securities	GFSC 41
Western Securities	WST 21		

Supplementary B: A brief description of the Radial Basis Function Neural Network (RBFNN)

1. Baseline configuration: The radial basis function neural network (RBFNN) is trained using MATLAB's newrb function. The input matrix is transposed into the column-vector format required by the Neural Network Toolbox. The fundamental parameters, defined in a configuration file, set the minimum spread at 1, the maximum neuron count at 100, and the target training error at 0.001.

2. Adaptive training strategy: In practice, the network employs an adaptive mechanism to enhance performance. The spread is not static; instead, it is derived from the training data. Up to 200 points are randomly sampled, and the median Euclidean distance between them is calculated. The spread is then set to 0.8 times this median distance, compared against the baseline minimum of 1, with the larger value being selected. Furthermore, the neuron limit is adjusted dynamically to half the size of the training dataset, subject to a minimum of 5 and a maximum of 100.

3. Training execution and output: The newrb function then iteratively adds neurons to the network, continuing until either the target error (0.001) is met or the adjusted maximum neuron count is reached. Should this iterative process (newrb) fail, the function automatically defaults to newrbe to ensure a solution is found. Consequently, the final parameters, including the RBF centers and output weights, are determined internally by the MATLAB algorithms rather than being specified manually.

AIMS Press