Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.5 (2026)

Dynamic Monitoring of Banking Systemic Risk: An Analysis Based on Smooth Quantile-LSTM Model

Chun Hou, Shuanghua Luo*

School of Science, Xi'an Polytechnic University, Xi'an, Shaanxi, 710048, China

Corresponding author: Shuanghua Luo, Shuanghualuo@xpu.edu.cn

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: Preventing systemic risk is a core task of macroprudential supervision, and accurately measuring the marginal contribution of individual banks to systemic risk remains a key challenge. To address the estimation bias of traditional parametric models in characterizing extreme tail risk, this paper constructs a Smooth Quantile Long Short-Term Memory (SQ-LSTM) model and conducts an empirical analysis using daily return data of 14 European Global Systemically Important Banks identified by the Financial Stability Board from 2008 to 2021. The model introduces the Horowitz smooth kernel to reformulate the traditional quantile loss function so that it is differentiable over the entire domain, employs LSTM gating mechanisms to capture long-term dependencies and regime-switching features of financial time series, and uses a multi-output neural network to simultaneously estimate VaR and CoVaR for all institutions. The empirical results show that SQ-LSTM is highly responsive to extreme events, with ΔCoVaR peaks closely coinciding with major shocks such as the 2008 financial crisis, the 2011 European debt crisis, Brexit in 2016, and the COVID-19 pandemic in 2020; risk rankings remain highly stable throughout the sample period, with Spearman correlation coefficients consistently above 0.80; and out-of-sample rolling forecasts with Diebold–Mariano tests confirm that, at the 1% extreme quantile level, SQ-LSTM achieves significantly higher prediction accuracy than the benchmark Q-LSTM model. This study provides a distribution-free, robust, and sensitive quantitative framework for identifying systemically important institutions and dynamically monitoring systemic risk.

Keywords: Systemic Risk; ΔCoVaR; Smooth Quantile Regression; LSTM; Macroprudential Regulation

Published: Nov 4, 2026

DOI: https://doi.org/10.62177/apemr.v3i5.1846

1. Introduction

Systemic financial risk is characterized by strong destructiveness, wide contagion, and long duration. Its transmission is not a linear additive process but spreads through complex higher-order connectedness structures in financial networks [1][2][3][4]. Historical events such as the 2008 global financial crisis, the 2011 European sovereign debt crisis, and the COVID-19 shock in 2020 have repeatedly confirmed that traditional regulatory indicators based on linear assumptions struggle to capture the accumulation process and nonlinear abrupt turning points of systemic risk. More flexible nonparametric measurement tools are therefore urgently needed.

Conditional Value at Risk (CoVaR) and its difference ΔCoVaR have become mainstream tools in macroprudential supervision [5][6]. However, existing measurements mostly rely on GARCH-type parametric models, which have three fundamental limitations: normal/t-distribution assumptions underestimate tail risk [7]; linear structures cannot characterize

nonlinear abrupt changes during crises [8][9]; and short-memory specifications lead to accumulating forecasting errors for long-lasting shocks [10]. Although improvements such as Copula [11][12] and DCC-GARCH [13][14] have been proposed, they have not escaped the parametric framework.

To break through the inherent constraints of parametric models, deep learning techniques have gradually been introduced into financial time series analysis. Long Short-Term Memory (LSTM) networks, with their gating structure, can effectively extract long-term dependencies and nonstationary dynamic features from sequential data [15]. Meanwhile, quantile regression neural networks (QRNNs) can directly estimate conditional quantiles without imposing any prior assumptions on the return distribution [16], and have shown good applicability in forecasting scenarios such as meteorology [17], electricity [18], and carbon prices [19], indicating that the integration of quantile regression and deep learning has broad prospects. On this basis, Aprea et al. [20] proposed the Q-LSTM model, which embeds quantile regression into LSTM to construct an end-to-end ΔCoVaR measurement framework, effectively avoiding error transmission between two-stage models [21]. Their empirical results show that it significantly outperforms traditional benchmarks such as GARCH-family models and quantile regression in characterizing nonlinear dynamics of tail risk and long memory in time series.

However, Q-LSTM still has an unavoidable technical shortcoming: the traditional quantile loss function is nondifferentiable at the origin, which makes gradient-based backpropagation prone to oscillation and even instability during training. This problem is particularly prominent at extreme quantile levels (e.g.,). To address this limitation, this paper proposes the Smooth Quantile Long Short-Term Memory (SQ-LSTM) network. The core contributions of this model are threefold:

(1) The Horowitz smooth kernel is introduced to reconstruct the quantile loss function, replacing the nondifferentiable indicator function with a continuously differentiable cumulative kernel function, thereby ensuring the stability of gradient propagation at the algorithmic level .

(2) The smooth quantile loss is jointly trained with LSTM in an end-to-end manner, enabling simultaneous estimation of VaR and CoVaR without any preset distributional assumptions.

(3) A multi-output neural network architecture is adopted to estimate the VaR and CoVaR of multiple institutions in parallel, structurally avoiding the quantile crossing problem and eliminating error accumulation. Without relying on distributional assumptions, this framework balances the stability of ranking structure during normal periods and the sensitivity of risk measurement around crisis events, providing a directly deployable quantitative tool for macroprudential supervision.

2. Theoretical Basis and Model Architecture

2.1 Measurement of Systemic Risk

Let $R_t^{(i)}$ and $R_t^{(sys)}$ denote the logarithmic returns of financial institutionand the financial system at time, respectively. Given a quantile level $p \in [0,1]$ (representing the institution's distress state, typically 1%), the institution's Value-at-Risk (VaR) $\mathrm{VaR}_{p,t}^{(i)}$ satisfies:

$$P\left(R_t^{(i)} \le \mathrm{VaR}_{p,t}^{(i)}\right) = p \ .$$

Given a tail quantile level q, the financial system's $\mathrm{CoVaR}_{q,p,t}^{(sys|i)}$ is the Value-at-Risk of the financial system conditional on bank*i*being in distress. Here, distress means that bank *i*'s market logarithmic return at time equals its $\mathrm{VaR}_{p,t}^{(i)}$:

$$P\left(R_t^{(sys)} \le \mathrm{CoVaR}_{q,p,t}^{(sys|i)} \middle| R_t^{(i)} = \mathrm{VaR}_{p,t}^{(i)}\right) = q \ .$$

To simplify notation, let $\Delta\mathrm{CoVaR}_{q,p,t}^{(i)}$ denote institution*i*'s marginal contribution to systemic risk, defined as:

$$\Delta\mathrm{CoVaR}_{q,p,t}^{(i)} = \mathrm{CoVaR}_{q,p,t}^{(i)} - \mathrm{CoVaR}_{q,50,t}^{(i)} \ ,$$

where the 50% quantile represents the financial system's normal (median) state, and the quantile level $p = 0.01$ represents the crisis (tail) state. Accurate estimation of $\Delta\mathrm{CoVaR}$ relies on precise characterization of the tail features of financial institutions' return distributions, providing the theoretical foundation for introducing quantile regression methods.

2.2 Smooth Quantile Regression

Quantile regression, introduced by Koenker and Bassett , overcomes key limitations of ordinary least squares by estimating conditional quantiles through minimizing a weighted sum of absolute errors. For a given quantile level $\tau \in (0,1)$, the linear conditional quantile function takes the form:

$$Q_\tau(Y|X) = X^T\beta(\tau) \ .$$

The corresponding estimator is obtained by solving the following optimization problem:

$$\widehat{\beta}(\tau) = arg\min_{\beta \in R^d} \sum_{i=1}^{n} \rho_\tau\left(y_i - x_i^T \beta\right) ,$$

where $\rho_\tau(u) = u\big(\tau - I(u<0)\big) = \begin{cases} (\tau-1)u, & u<0 \\ \tau u, & u \geq 0 \end{cases}$ denotes the pinball loss function.

While quantile regression aligns naturally with the estimation requirements of $\Delta\mathrm{CoVaR}$, its standard formulation poses a critical computational bottleneck: the pinball loss $\rho_\tau(u)$ is non-differentiable at the origin $u = 0$. This non-differentiability causes gradient instability during backpropagation, particularly at extreme quantiles such as $\tau = 0.01$, where accurate estimation is most challenging.

To overcome this limitation, this study adopts the smooth quantile regression framework developed by Horowitz [22], which replaces the non-differentiable indicator function with a continuously differentiable kernel function. Let $K(\cdot)$ denote a bounded kernel function with compact support on $[-1,\ 1]$ satisfying $\int_{-1}^{1} K(u)du = 1$. Define its cumulative distribution function as $G(x) = \int_{-\infty}^{x} K(u)du$. For the seventh-degree polynomial kernel proposed by Horowitz, the explicit form of $G(x)$ is given by:

$$G(x) = \begin{cases} 0 & , \ x \leq -1 \\ 0.5 - \dfrac{105}{64}\left(x - \dfrac{5}{3}x^3 + \dfrac{7}{5}x^5 - \dfrac{3}{7}x^7\right) & , \ |x| \leq 1 \\ 1 & , \quad x \geq 1 \end{cases} .$$

The scaled version is $G_h(x) = G(x/h)$, where h is a positive bandwidth parameter. The smooth quantile loss function is then obtained by replacing the indicator function $I(u<0)$ in the traditional pinball loss with $G_h(u)$. yielding:

$$\tilde{\rho}_\tau(u) = u(\tau - G_h(u)) + h\int_{-\infty}^{u/h} K(v)dv ,$$

where $u = y - \hat{y}$ is the prediction residual .

This smoothing ensures continuous differentiability across all *u*, enabling stable gradient-based optimization while preserving the asymptotic properties of quantile regression. The bandwidth parameter *h* governs the trade-off between approximation accuracy and numerical stability: smaller values yield closer approximation to the original pinball loss but may retain some instability, while larger values produce smoother gradients at the cost of increased bias. In our implementation, *h* is selected adaptively based on the scale of the data, as detailed in Section 2.4.

2.3 Long Short-Term Memory Network

Recurrent neural networks (RNNs) have the ability to process sequential data by allowing unit outputs to be fed back as inputs in subsequent time steps through recurrent connections. Let $x_t \in R^{h_0}, 0 < t \leq T$, be a multivariate time series. An RNN layer with $h_1 \in N$ units can be expressed as:

$$\begin{aligned} &z_t(x) = \varphi\left(\omega_0 + Wx_t + Uz_{t-1}^{(1)}(x)\right), \ \ t = 1,\ldots,T, \\ &z_0(x) = 0, \end{aligned}$$

where $U \in R^{h_1 \times h_1}$ is the weight related to the output of the previous time step.

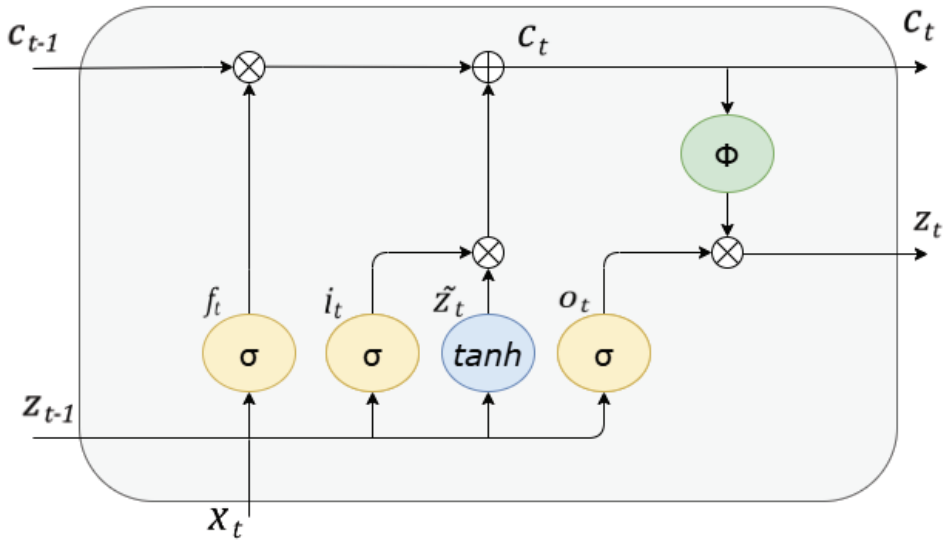
However, standard RNNs often suffer from the vanishing gradient problem during training. Long Short-Term Memory (LSTM) networks alleviate this deficiency by introducing memory cells and gate control mechanisms, which can selectively store or release long-term information. Let $W^{(j)} \in R^{h_1 \times h_0}$, $U^{(j)} \in R^{h_1 \times h_1}$, $b^{(j)} \in R^{h_1}$ be the relevant weights and biases for each subnetwork $j \in \{i, o, f, z\}$. The mechanism of the LSTM layer for $t = 1,\ldots,T$ can be described as follows:

$$\begin{aligned} i_t &= \sigma\left(\omega_0^{(i)} + W^{(i)}x_t + U^{(i)}z_{t-1}\right), \\ o_t &= \sigma\left(\omega_0^{(o)} + W^{(o)}x_t + U^{(o)}z_{t-1}\right), \\ f_t &= \sigma\left(\omega_0^{(f)} + W^{(f)}x_t + U^{(f)}z_{t-1}\right), \\ \tilde{z}_t &= tanh\left(\omega_0^{(z)} + W^{(z)}x_t + U^{(z)}z_{t-1}\right), \\ c_t &= c_{t-1} \odot f_t + i_t \odot \tilde{z}_t, \end{aligned}$$

$$z_t = \varphi(c_t) \odot o_t,$$
$$z_0 = 0,$$
$$c_0 = 0,$$

where $\sigma(\cdot): R \mapsto (0,1)$ and $tanh\,(\cdot): R \mapsto (-1,1)$ are the sigmoid and hyperbolic tangent activation functions, respectively, and c_t is the memory cell state at time. This structure enables LSTM to selectively retain or forget information over longer time horizons, thereby effectively capturing long-term dependencies and nonstationary dynamics in financial time series.

Figure 1: LSTM neural network structure diagram



2.4 SQ-LSTM Model

The SQ-LSTM model proposed in this paper consists of a multi-output neural network that can simultaneously provide estimates of $CoVaR^{(i)}_{q,p,t}$ and $CoVaR^{(i)}_{q,50,t}$ for different financial institutions, and then calculate the ΔCoVaR of each institution. The network contains two LSTM layers and multiple fully connected layers (FCNs), and is trained end-to-end by minimizing a custom loss function based on the smooth quantile loss.

Let I be the set of banks considered. The model processes the multivariate time series of logarithmic returns of financial institutions, $R_t = \left(R_t^{(i)}\right)_{i\in I}$ and simultaneously estimates the VaR and system CoVaR of all institutions.

Choose a lookback period $L \in N$,the first LSTM layer of size $S_{50} \in N$ processes the time lags $R_{t-1}, l = 1, \ldots, L,$. The activation value of the LSTM at the last time step is denoted as $z_{50,t} \in R^{S_{50}}$, which can be regarded as a set of features summarizing the information contained in the input data. Through an Idimensional fully connected layer, $z_{50,t}$ is processed to obtain the vector of median-level Value at Risk of financial institutions, $VaR^{(i)}_{50,t}$, $i \in I$:

$$\mathrm{VaR}^{(i)}_{50,t} = \varphi_{50}^{(\mathrm{VaR})}\left(b_{50}^{(\mathrm{VaR})} + W_{50}^{(\mathrm{VaR})} z_{50,t}\right),$$

where $b_{50}^{(VaR)} \in R^{|I|}$ and $W_{50}^{(VaR)} \in R^{|I|\times S_{50}}$ are network coefficients.

The second LSTM layer of size S_p is used to extract features describing the quantiles of logarithmic returns at level $p \in [0,0.5]$ (e.g., 0.01). The output $z_{p,t} \in R^{q_p}$ has the same interpretation as $z_{50,t}$ However, the features $z_{50,t}$ summarize information used to describe $VaR^{(i)}_{50,t}$, while $z_{p,t}$ selects useful information for describing $VaR^{(i)}_{p,t}$.

The latter is calculated by subtracting from the output of an FCN layer from $VaR^{(i)}_{50,t}$. In this case, $VaR^{(i)}_{p,t}$ is computed as:

$$\widetilde{\mathrm{VaR}}^{(i)}_{p,t} = \varphi_p^{(\mathrm{VaR})}\left(b_p^{(\mathrm{VaR})} + W_p^{(\mathrm{VaR})} z_{p,t}\right),$$
$$\mathrm{VaR}^{(i)}_{p,t} = \mathrm{VaR}^{(i)}_{50,t} - \widetilde{\mathrm{VaR}}^{(i)}_{p,t},$$

which solves the quantile crossing problem.

Finally, $CoVaR^{(i)}_{q,50,t}$ for different institutions is calculated by processing $CoVaR^{(i)}_{50,t}$ with an FCN layer:

$$CoVaR^{(i)}_{q,50,t} = \varphi_p^{(CoVaR)}\left(b_p^{(CoVaR)} + \omega_p^{(CoVaR)} VaR^{(i)}_{50,t}\right),$$

where $b_p^{(CoVaR)}, \omega_p^{(CoVaR)} \in R$ are network coefficients.

The matrix elements and bias vectors of different neural network layers need to be properly calibrated. Let θ denote the vector containing all network parameters. The training process can be regarded as an unconstrained optimization problem, in which an appropriate loss function $L(\theta)$ is selected. Neural network training is usually conducted using the backpropagation (BP) algorithm, in which weights are updated based on the gradient of the loss function $L(\theta)$. The weights are iteratively adjusted to reduce the error between network outputs and certain reference values.

In the model of this paper, the loss function used for training is the sum of three parts:

$$\begin{aligned} L(\theta) &= L_{50}^{(\mathrm{VaR})} + L_p^{(\mathrm{VaR})} + L_p^{(\mathrm{CoVaR})} \\ &= \sum_{i,t} \tilde{\rho}_{50}\left(r_t^{(i)} - \widehat{\mathrm{VaR}}_{50,t}^{(i)}\right) + \sum_{i,t} \tilde{\rho}_p\left(r_t^{(i)} - \widehat{\mathrm{VaR}}_{p,t}^{(i)}\right) + \sum_{i,t} \tilde{\rho}_p\left(r_t^{(sys)} - \widehat{\mathrm{CoVaR}}_{q,50,t}^{(i)}\right), \end{aligned}$$

where $\tilde{\rho}_\tau$ is the smooth quantile loss function. The first term is the error in estimating the median logarithmic return; it imposes the same penalty on positive and negative residuals. The second term relates to the estimation of $VaR_{p,t}^{(i)}$, using the pinball loss at level $p \in]0,0.5[$, which penalizes positive residuals more heavily than negative residuals. The third term relates to the estimation of $CoVaR_{t,q,50}^{(i)}$; in this case, the pinball loss with parameter $\tau = q$ is applied to estimate the conditional quantile of the financial system relative to different financial institutions.

The bandwidth parameter h is selected using an adaptive data-driven method based on optimal bandwidth theory in nonparametric statistics. Its formula is:

$$h = \hat{\sigma} \times n^{-1/5},$$

where $\hat{\sigma}$ is the standard deviation of prediction residuals, used to measure the magnitude of data fluctuations; *n* is the batch sample size; and the exponent −1/5 comes from the optimal convergence rate of kernel density estimation under Horowitz's smooth quantile regression framework . This adaptive mechanism ensures that the smooth kernel function can be dynamically adjusted according to the scale of the data and sample size, adapting to volatility changes in financial time series while maintaining asymptotic theoretical properties. The use of the smooth loss function together with the Nadam optimization algorithm makes the model converge more robustly at extreme quantiles such as $\tau = 0.01$.

3. Experimental Design and Evaluation Metrics

3.1 Data Sources

This study selects the Global Systemically Important Banks (G-SIBs) published by the Financial Stability Board (FSB), covering 14 banks from Italy, Germany, France, Spain, Belgium, and the Netherlands. The original daily stock price and market capitalization data are obtained from the Refinitiv Eikon financial database. The sample period is from January 2008 to December 2021, covering major events such as the subprime mortgage crisis, the European debt crisis, Brexit, and the COVID-19 pandemic. After logarithmic differencing, daily logarithmic return series are obtained. Each series contains 3,653 observations, and data completeness is good.

3.2 Operating Environment

Data processing and analysis are completed in the R language environment, using RStudio for programming, graphing, and analysis. The operating system is Windows 11, and the CPU is an Intel(R) Core(TM) i7-10750H CPU @ 2.60 GHz. The experimental platform is Rx64 4.4.2 and RStudio.

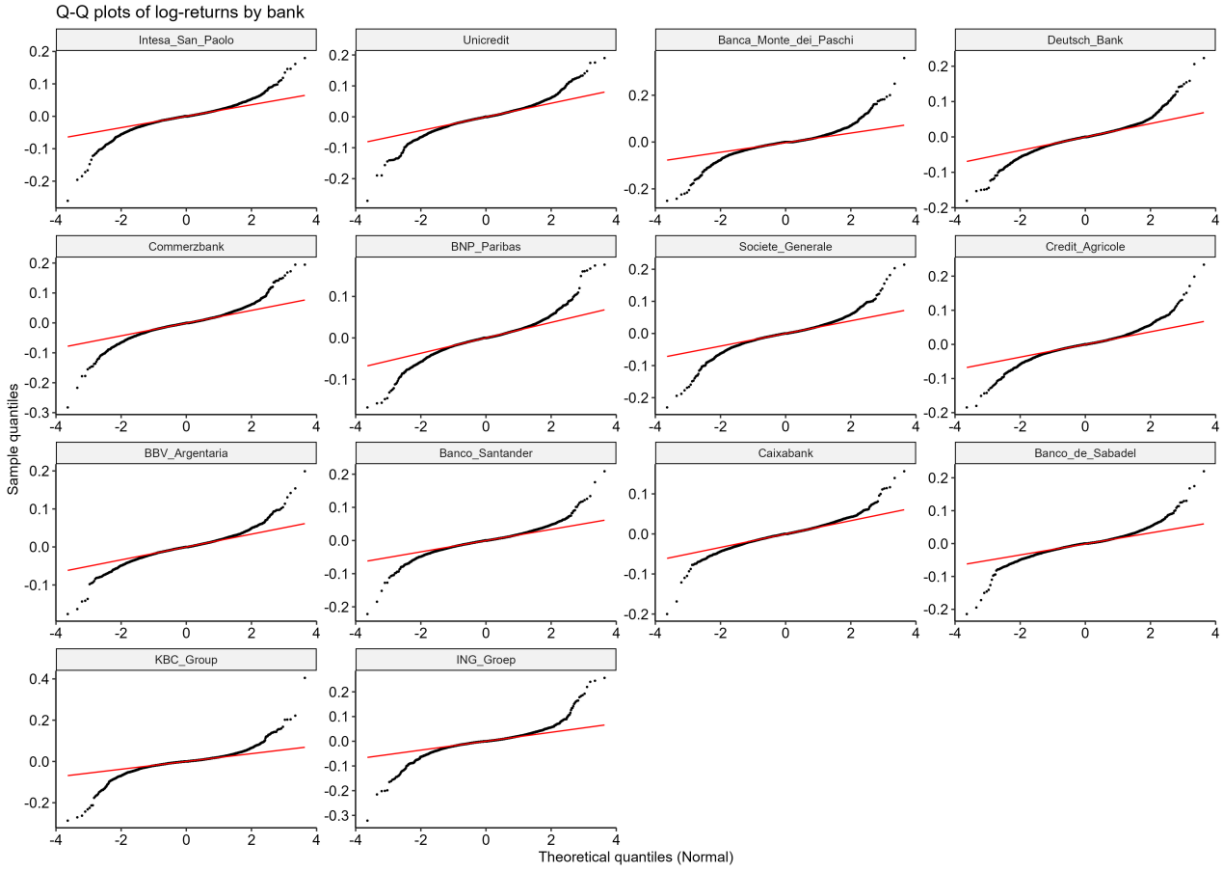
This section conducts a preliminary exploration of the dataset. Table 1 reports the descriptive statistics of the logarithmic return series of each institution. The average daily returns of all banks are negative, reflecting the overall profitability pressure on the European banking sector during the sample period. The standard deviation of returns ranges from 0.0215 to 0.0349, higher than the average level of nonfinancial industries in mature markets, highlighting the high volatility of the banking sector. In terms of skewness, 8 banks exhibit negative skewness and 6 exhibit positive skewness; the kurtosis of all banks is significantly greater than 3, indicating obvious fat-tail characteristics in return distributions. The p-values of the ADF unit root test are all less than 0.01, indicating that the series are stationary and satisfy the modeling prerequisites.

Table 1: Sample Banks and Descriptive Statistics

Bank	Sample Size	Minimum	Maximum	Mean	Std. Dev.
Intesa San Paolo	3653	-0.260594	0.179636	-0.000218	0.026311
Societe Generale	3653	-0.23034	0.214255	-0.000292	0.029256
Unicredit	3653	-0.271652	0.190067	-0.000652	0.030739
KBC Group	3653	-0.286635	0.404843	-0.000066	0.033119
Caixabank	3653	-0.200029	0.156447	-0.000138	0.021456
Commerzbank	3653	-0.282482	0.194573	-0.000863	0.03057
ING Groep	3653	-0.321361	0.256527	-0.000142	0.030659
Banco Santander	3653	-0.221726	0.208773	-0.000283	0.023588
BBV Argentaria	3653	-0.176492	0.199073	-0.000243	0.023332
BNP Paribas	3653	-0.167572	0.176624	-0.000049	0.026155
Credit Agricole	3653	-0.18469	0.233615	-0.000145	0.027299
Banco de Sabadel	3653	-0.214339	0.219867	-0.000568	0.024662
Banca Monte dei Paschi	3653	-0.250988	0.358634	-0.002111	0.034948
Deutsche Bank	3653	-0.180711	0.223032	-0.000503	0.027103
Intesa San Paolo	-0.546739	11.667074	8.667074	-14.450345	0.01
Societe Generale	-0.225112	10.474597	7.474597	-15.090548	0.01
Unicredit	-0.212477	9.048984	6.048984	-15.265291	0.01
KBC Group	-0.15128	20.416118	17.416118	-15.632977	0.01
Caixabank	-0.141425	9.394128	6.394128	-14.534408	0.01
Commerzbank	-0.124558	10.36427	7.36427	-14.580012	0.01
ING Groep	-0.107982	17.103629	14.103629	-14.491876	0.01
Banco Santander	-0.099421	11.798564	8.798564	-15.71936	0.01
BBV Argentaria	0.043895	9.655382	6.655382	-14.931367	0.01
BNP Paribas	0.070931	9.369236	6.369236	-15.957839	0.01
Credit Agricole	0.085687	9.840405	6.840405	-15.359134	0.01
Banco de Sabadel	0.095957	12.04786	9.04786	-15.857803	0.01
Banca Monte dei Paschi	0.197991	13.710434	10.710434	-15.550672	0.01
Deutsche Bank	0.22422	10.095487	7.095487	-15.316432	0.01

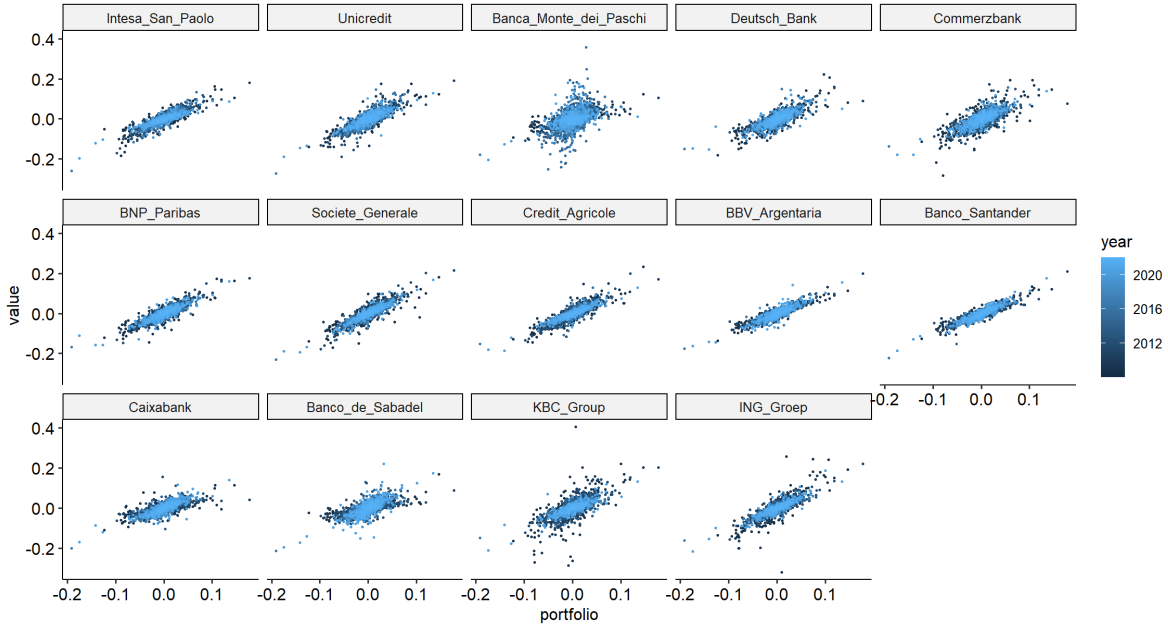
To further verify the non-normality of return distributions, this study plots Q-Q plots of the logarithmic returns of each bank. The results show that the sample quantiles of all institutions deviate significantly from the normal reference line at both tails, confirming fat-tail characteristics.

Figure 2: Q-Q plots of logarithmic returns by bank



The scatter plot (Figure 3) shows that the dependence between institutions and the system strengthens significantly in the extreme negative return region of the left tail, and that correlation coefficients exhibit time-varying characteristics, supporting the use of a flexible model without preset dependence structures.

Figure 3: Scatter plots of individual bank logarithmic returns and system portfolio logarithmic returns



3.3 SQ-LSTM Model Calibration

The SQ-LSTM model requires selection of hyperparameters defining the network architecture. Table 2 reports the parameter values used. The dimensions of the fully connected layers and logic layers used to calculate VaR and CoVaR are set to the number of banks $|L|$.

Table 2: Hyperparameter values of the SQ-LSTM architecture

Hyperparameter	Value
Hidden units	8
lookback period	8
Dropout rate	20%
Epochs	1000
Batch size	16
Activation function	tanh

To ensure fairness of comparison, all comparison models (including SQ-RNN and the unsmoothed Q-LSTM) adopt exactly the same calibration procedure, hyperparameters, and optimization algorithm (Nadam). Since this study aims to comprehensively evaluate the dynamics of systemic risk over the entire observation period rather than focusing only on point forecasts, full-sample fitting is used as the main approach, while out-of-sample testing is conducted separately through rolling windows.

4. Empirical Analysis

4.1 Model Fitting Ability

Figure 4 shows the fitting performance of SQ-LSTM for the daily logarithmic returns of 14 banks. The median forecast (red) closely follows the central trend of actual returns; the 1% quantile forecast (yellow) shifts downward significantly during the 2008 and 2020 crises, accurately reflecting the amplification of tail risk. At all time points, the 1% quantile is strictly below the median, indicating that the model effectively avoids the quantile crossing problem.

Figure 4: Fitted returns and 0.01 and 0.50 quantile forecasts of the SQ-LSTM model

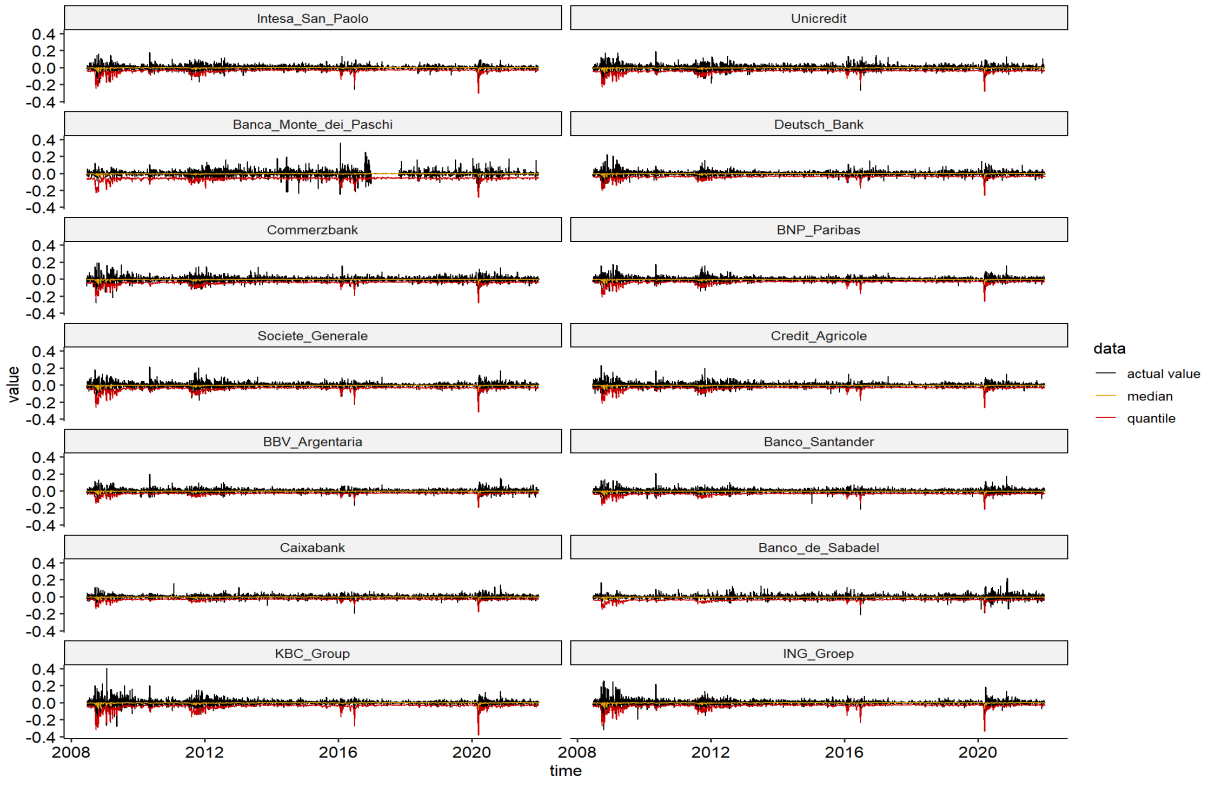


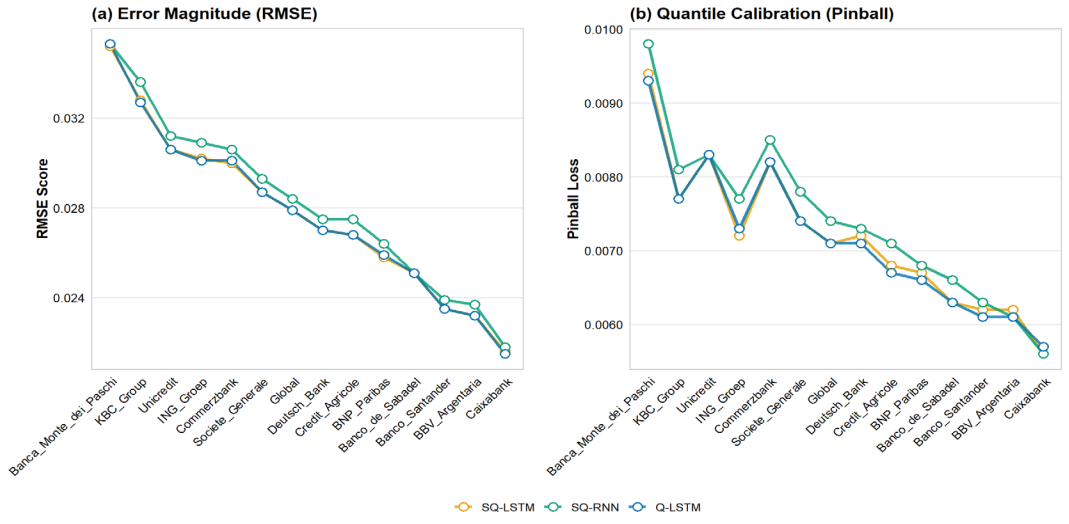
Table 3 quantitatively evaluates the goodness of fit of the model. The median RMSE ranges from 0.0216 to 0.0352, with an overall mean of 0.0279; the 1% quantile pinball loss ranges from 0.0056 to 0.0094, with an overall mean of 0.0071. The low RMSE and pinball loss indicate that the model fits the observed data well.

Table 3: RMSE and loss values of median forecasts for each bank

Bank	RMSE	Pinball
Intesa_San_Paolo	0.0263	0.0068
Unicredit	0.0306	0.0083
Banca_Monte_dei_Paschi	0.0352	0.0094
Deutsch_Bank	0.027	0.0072
Commerzbank	0.03	0.0082
BNP_Paribas	0.0258	0.0067
Societe_Generale	0.0287	0.0074
Credit_Agricole	0.0268	0.0068
BBV_Argentaria	0.0232	0.0062
Banco_Santander	0.0235	0.0062
Caixabank	0.0216	0.0056
Banco_de_Sabadel	0.0251	0.0063
KBC_Group	0.0328	0.0077
ING_Groep	0.0302	0.0072
Global	0.0279	0.0071

To further objectively evaluate the goodness of fit of SQ-LSTM, this paper systematically compares it with Q-LSTM and SQ-RNN within a unified framework (Figure 5). The results show that SQ-LSTM achieves the lowest RMSE and quantile loss for the vast majority of banks.

Figure 5: Comparison of RMSE and quantile loss of three models for different banks



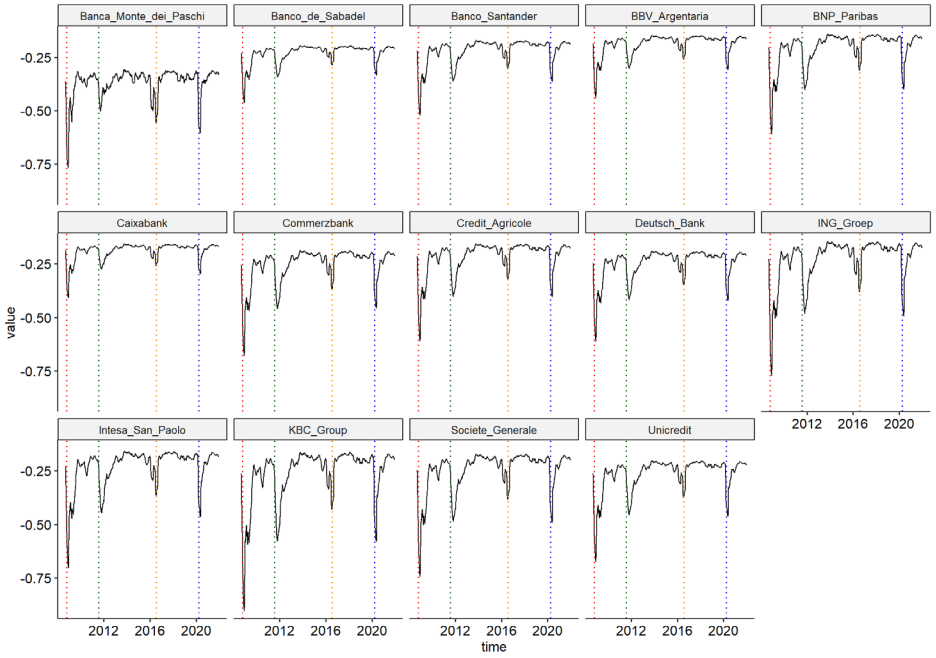
Existing literature has fully verified that Q-LSTM significantly outperforms traditional GARCH-family and quantile regression methods in capturing long-term dependencies in financial time series, characterizing nonlinear dynamics of tail risk, and ensuring the stability and persistence of ΔCoVaR rankings. On this basis, this paper further confirms that the smoothing strategy not only improves gradient stability but also indirectly enhances the model's fitting accuracy for tail risk. Moreover, compared with SQ-RNN, which also belongs to the smooth framework, LSTM performs better in dynamic risk fitting due to its advantage in capturing long-memory features.

4.2 ΔCoVaR Estimation

This paper further estimates the daily ΔCoVaR of 14 institutions and calculates their 60-day moving averages to observe trend changes. Figure 6 presents the time-series trajectories of ΔCoVaR for all institutions. The ΔCoVaR series of all banks exhibit four significant common peaks, corresponding respectively to the bankruptcy of Lehman Brothers in September 2008, the spread of the European debt crisis to Italy and Spain in July 2011, Brexit and the Banca Monte dei Paschi crisis in June

2016, and the COVID-19 shock in March 2020. The model responds sensitively to these abrupt changes, fully demonstrating LSTM's rapid adaptability to market regime shifts.

Figure 6: Time-series trajectories of ΔCoVaR for all banks



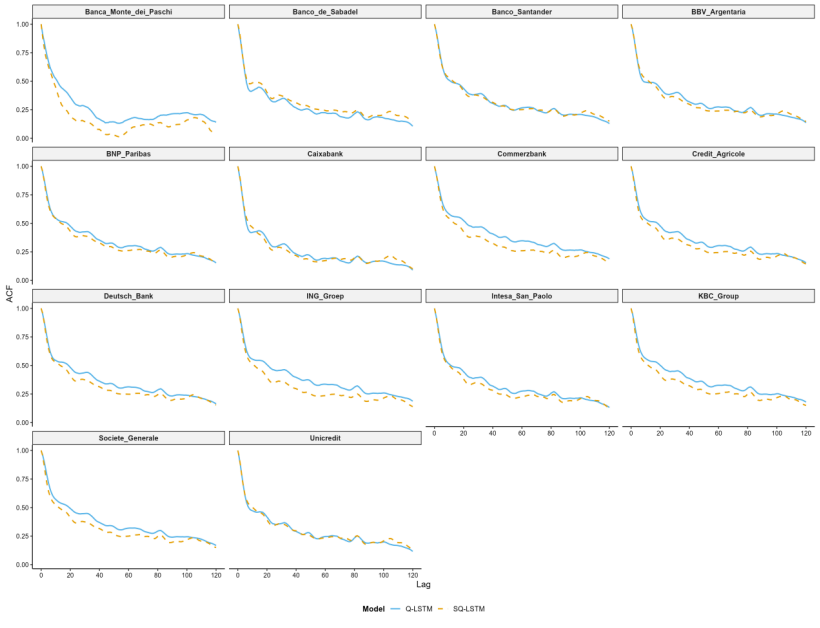
At the same time, heterogeneity among institutions is also significant: the ΔCoVaR of Banca Monte dei Paschi and Deutsche Bank remains high for a long time, reflecting their strong risk spillover effects; Santander, BBVA, and others fluctuate relatively smoothly and are less sensitive to systemic shocks.

4.3 Ranking Stability and Persistence

Two core attributes of systemic risk measurement tools are ranking persistence and stability. The former measures the degree of autocorrelation of risk estimates over time, while the latter measures the consistency of cross-period rankings.

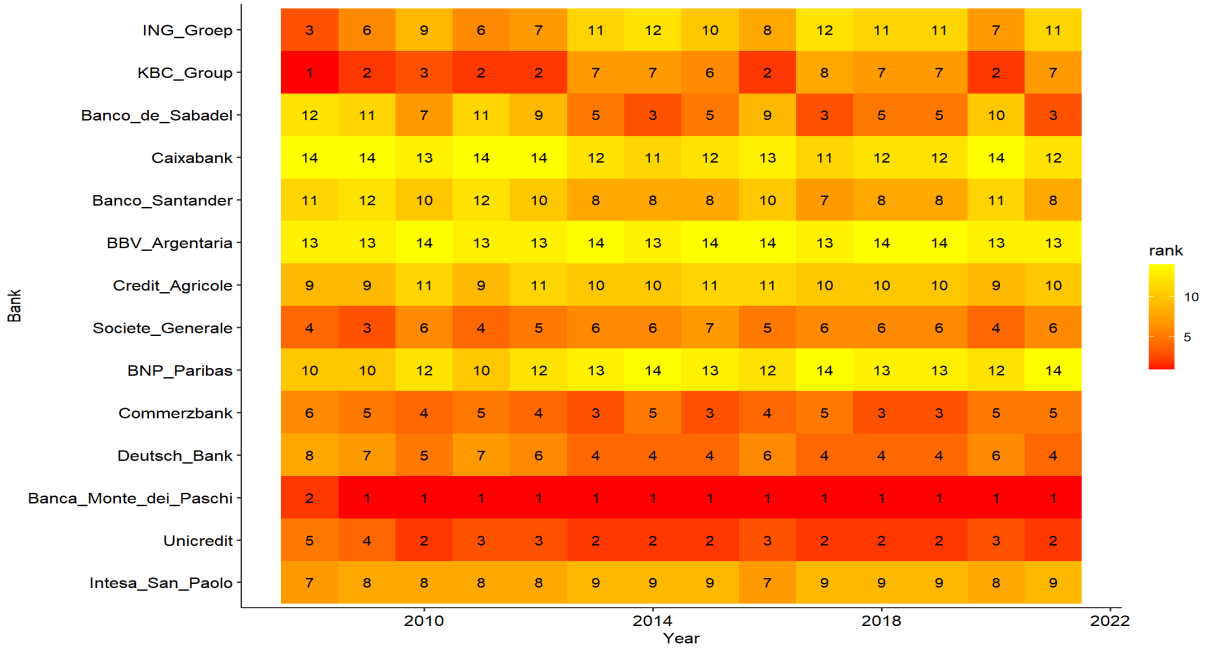
Persistence: Figure 7 compares the autocorrelation coefficients of ΔCoVaR between SQ-LSTM and Q-LSTM. The difference between the two is very small (mostly within 0.05), and SQ-LSTM fully inherits the excellent persistence of Q-LSTM. The autocorrelation coefficients of SQ-LSTM are slightly lower for most banks, indicating that smoothing reduces the interference of extreme values on long-term dependence structures and makes estimates more robust.

Figure 7: Comparison of autocorrelation coefficients of ΔCoVaR estimates between the two models



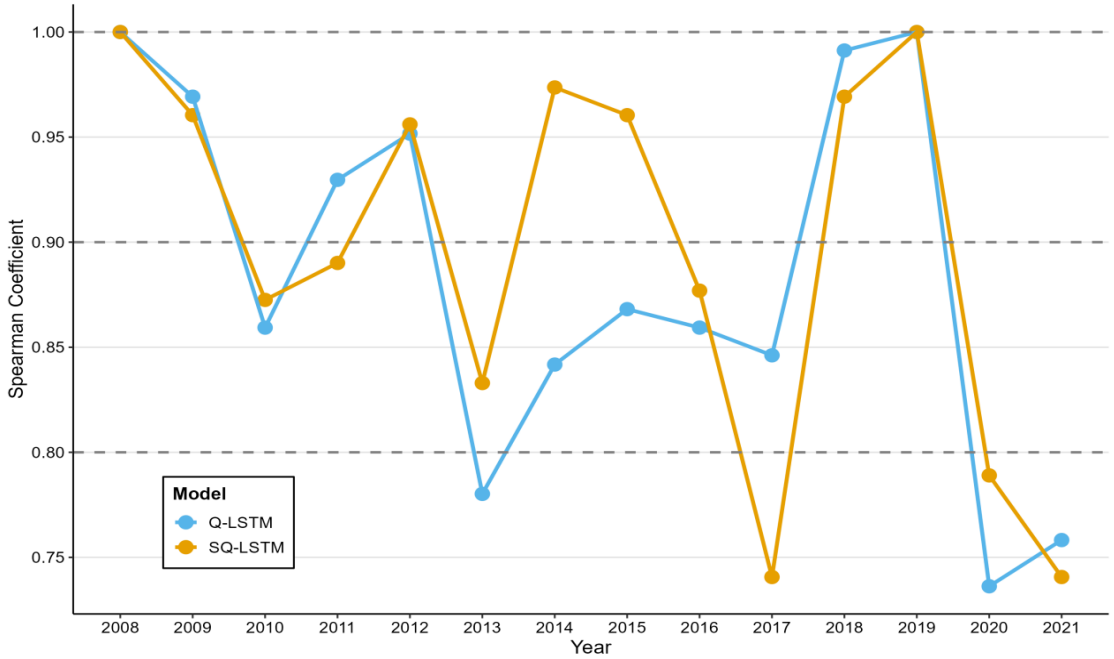
Stability: Based on annual mean rankings (Figure 8), high-risk institutions (Banca Monte dei Paschi, Deutsche Bank) consistently rank at the top throughout the sample period, while low-risk institutions (Caixabank, BBVA) always rank at the bottom. Ranking dynamics also capture changes in the macro environment: KBC and ING ranked high in 2010; the systemic contribution of Italian banks increased in 2014; and the post-2020 period reflects risk repricing in the post-pandemic era.

Figure 8: Annual systemic risk rankings of banks estimated by the SQ-LSTM model



To further quantify ranking stability, Spearman correlation coefficients between consecutive years (Figure 9) are all above 0.80, confirming the temporal coherence of rankings. The coefficients of SQ-LSTM during stress periods (2014, 2015, 2020) are significantly higher than those of Q-LSTM, indicating that the smooth loss enhances estimation robustness under extreme events.

Figure 9: Spearman correlation coefficients of the two models by year



In addition, computational efficiency is also a key consideration in practical deployment. Under the same hardware conditions, the training time of the SQ-LSTM model is 1.530 hours, while that of the Q-LSTM model is 1.645 hours. The shorter training time of SQ-LSTM benefits from faster convergence brought by smooth gradients.

4.4 Out-of-Sample Forecasting and Statistical Significance Tests

To verify the generalization ability of SQ-LSTM and avoid concerns about overfitting, this paper uses a time-series rolling window method for out-of-sample forecasting evaluation. The sample is divided chronologically into the first 80% as the training set and the last 20% as the test set. On the test set, model parameters are kept unchanged (i.e., test set information is not used for retraining), one-step-ahead forecasts are made daily, and the pinball loss of the 1% quantile forecast (τ=0.01) is calculated for each institution.

Table 4: Out-of-sample pinball loss of the SQ-LSTM model (τ=0.01)

Bank	Average out-of-sample pinball loss
Intesa San Paolo	0.038320
Unicredit	0.041512
Banca Monte dei Paschi	0.056865
Deutsch Bank	0.039337
Commerzbank	0.041313
BNP Paribas	0.032700
Societe Generale	0.037856
Credit Agricole	0.036652
BBV Argentaria	0.036499
Banco Santander	0.032822
Caixabank	0.031638
Banco de Sabadel	0.040161
KBC Group	0.043259
ING Groep	0.036182
Mean	0.038701

Table 5 shows that the out-of-sample losses of the 14 institutions range from 0.0316 to 0.0569, with a mean of 0.0387. Compared with the in-sample loss (0.0071), it rises somewhat but remains overall controllable and does not diverge, indicating that the model has a certain ability to generalize in characterizing tail risk.

To further determine whether the above differences are statistically significant, this paper uses the Diebold–Mariano (DM) test. The null hypothesis is that there is no difference in forecasting accuracy between the two methods, and the alternative hypothesis is that the pinball loss of SQ-LSTM is significantly lower than that of Q-LSTM. Newey–West HAC standard errors are used to correct for residual autocorrelation.

Table 5: Diebold–Mariano test results (SQ-LSTM vs. Q-LSTM)

Bank	DM statistic	p-value	Conclusion
Intesa San Paolo	−13.9961	0.000	SQ-LSTM significantly better
Unicredit	−23.0831	0.000	SQ-LSTM significantly better
Banca Monte dei Paschi	−13.9643	0.000	SQ-LSTM significantly better
Deutsch Bank	−40.4122	0.000	SQ-LSTM significantly better
Commerzbank	−28.0227	0.000	SQ-LSTM significantly better
BNP Paribas	−33.5852	0.000	SQ-LSTM significantly better
Societe Generale	−24.1483	0.000	SQ-LSTM significantly better
Credit Agricole	−22.4939	0.000	SQ-LSTM significantly better
BBV Argentaria	−28.1823	0.000	SQ-LSTM significantly better

Bank	DM statistic	p-value	Conclusion
Banco Santander	−33.1501	0.000	SQ-LSTM significantly better
Caixabank	−34.3676	0.000	SQ-LSTM significantly better
Banco de Sabadel	−38.6877	0.000	SQ-LSTM significantly better
KBC Group	−17.3569	0.000	SQ-LSTM significantly better
ING Groep	−27.0367	0.000	SQ-LSTM significantly better

The DM test results show highly consistent statistical significance: the DM statistics of all 14 sample institutions are negative (ranging from -13.96 to -40.41), and the corresponding p-values are all close to 0. The null hypothesis is rejected at the 1% significance level. This indicates that the SQ-LSTM model significantly outperforms the Q-LSTM model using the standard quantile loss in out-of-sample prediction accuracy for 1% VaR, and that this advantage is not accidental random fluctuation but a structural improvement brought by the smooth quantile loss function.

From an economic perspective, this result indicates that the introduction of the smoothing mechanism not only improves training stability and convergence efficiency by enhancing the differentiability of the loss function, but also fundamentally strengthens the model's ability to generalize in characterizing extreme tail risk. The nondifferentiable loss function of Q-LSTM may lead to unstable gradients and excessively high prediction errors on the test set; in contrast, SQ-LSTM makes gradient propagation smoother through the smooth kernel function, avoids oscillations in the optimization process, and thereby learns a more robust tail dependence structure. Unlike the "lazy" benchmark based on zero forecasts in existing literature, this paper uses a standard quantile model with the same architecture, data, and training configuration as the reference, excluding interference from factors such as network structure complexity, and directly verifying the net contribution of the smoothing mechanism to prediction accuracy. This provides reliable causal evidence for smoothing strategies in the integration of quantile regression and deep learning.

Overall, the results of out-of-sample evaluation and DM tests establish the advantage of the SQ-LSTM model over the traditional Q-LSTM model at the level of statistical inference, providing a solid econometric foundation for applying the model to systemic risk monitoring and identification of systemically important institutions.

5. Conclusions and Recommendations

5.1 Conclusions

This paper focuses on measuring the marginal contribution of Global Systemically Important Banks to systemic risk. To address the three limitations of traditional parametric models in distributional assumptions, linear dependence, and long-memory characterization, it constructs the SQ-LSTM model, which integrates smooth quantile regression and LSTM. Through a systematic empirical study of daily data from 14 European G-SIBs over 14 years, the following conclusions are drawn:

(1) The SQ-LSTM model can achieve robust estimation of ΔCoVaR without presetting any distributional form. The smooth quantile loss effectively resolves the gradient oscillation caused by the nondifferentiability of the traditional pinball loss. The multi-output neural network structure simultaneously estimates the VaR of all institutions and the system CoVaR, structurally ensuring quantile monotonicity and eliminating two-stage error accumulation.

(2) SQ-LSTM exhibits high sensitivity to the four major extreme events during the sample period, and the ΔCoVaR peaks closely match the timing of these events. Moreover, the model performs well in ranking stability and persistence, with Spearman correlation coefficients remaining above 0.80 for a long time, indicating that it combines crisis sensitivity with normal-period stability.

(3) Strict out-of-sample rolling forecasts and DM tests confirm that SQ-LSTM is significantly superior to the benchmark Q-LSTM model in 1% extreme quantile prediction accuracy (all institutions have p-values of approximately 0.000). The introduction of the smoothing mechanism brings statistically reliable and economically significant gains.

5.2 Recommendations

Based on the above conclusions, this paper proposes the following recommendations and prospects:

(1) At the data level, this study only uses daily stock market return data and has not yet incorporated micro-level risk information such as balance sheets, interbank lending, and derivative exposures. Future research can explore the integration of textual data (e.g., financial report disclosures and news sentiment) with structured data to construct a multidimensional risk measurement framework.

(2) At the methodological level, although SQ-LSTM can effectively measure risk levels and capture dynamic changes, it still lacks structural explanation of the internal mechanisms of risk transmission. Future research can combine interpretability tools such as SHAP values to decouple LSTM gate activation values and reveal the temporal driving paths of risk spillovers.

(3) At the application level, the sample of this paper is concentrated in developed European markets, and the applicability of the model needs to be further verified in emerging economies and other types of financial institutions (e.g., insurance companies and securities firms). In addition, the adaptive selection strategy of the bandwidth parameter can be further optimized to improve the robustness of the model under different market conditions.

Funding

This work was supported by the Natural Science Foundation of Shaanxi Province of China (2024JC-YBMS-007).

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Kupiec, P., & Nickerson, D. (2004). Assessing systemic risk exposure GSEs under alternative approaches to capital regulation. The Journal of Real Estate Finance and Economics, 28(2), 123–145.

[2] Zhu, B., Li, L., Zhou, Y., & Wang, H. (2019). How does information disclosure affect bank systemic risk in the presence of a deposit insurance system? Emerging Markets Finance and Trade, 55(11), 2497–2522.

[3] Billio, M., Getmansky, M., Lo, A. W., & Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. Journal of Financial Economics, 104(3), 535–559.

[4] Akgüller, Ö., & Balcı, M. A. (2026). Detecting financial contagion through higher-order networks: a deep learning approach to emerging market risk. Computational Economics. https://doi.org/10.1007/s10614-025-11287-3

[5] Caporin, M., Costola, M., & Garibal, J. C. (2022). Systemic risk and severe economic downturns: A targeted and sparse analysis. Journal of Banking and Finance, 134, 106339.

[6] Acharya, V. V., Pedersen, L. H., Philippon, T., & Richardson, M. (2017). Measuring systemic risk. The Review of Financial Studies, 30(1), 2–47.

[7] Leonardo, M. B., Giovanni, L. D., & Giorgia, R. (2023). Non-Gaussian models for CoVaR estimation. International Journal of Forecasting, 39(1), 391–404.

[8] Yao, C. Z., & Li, M. J. (2023). GARCH-MIDAS-GAS-copula model for CoVaR and risk spillover in stock markets. North American Journal of Economics and Finance, 66, 101910.

[9] Giglio, S., Kelly, B., & Pruitt, S. (2016). Systemic risk and the macroeconomy: An empirical evaluation. Journal of Financial Economics, 119(3), 457–471.

[10] Ding, Z. X., Granger, C. W. J., & Engle, R. F. (2001). A long memory property of stock market returns and a new model. In C. W. J. Granger (Ed.), Essays in econometrics: Collected papers of Clive W. J. Granger (pp. 349–372). Cambridge University Press.

[11] Zhao, S. J. (2022). Systemic risk measurement: A limiting threshold copula approach to CoVaR. Computers & Industrial Engineering, 171, 108464.

[12] Yang, L., Yang, L., Ho, K., & Wang, Y. (2020). Dependence structures and risk spillover in China's credit bond market: A copula and CoVaR approach. Journal of Asian Economics, 68, 101200.

[13] Zhang, X., & Wei, C. (2023). Systemic risk of Chinese financial institutions and asset price bubbles. North American Journal of Economics and Finance, 64, 101880.

[14] Chen, X., Zhou, L., Wang, L., & Liu, H. (2023). Risk spillover in China's real estate industry chain: a DCC-EGARCH-

ΔCoVaR model. Humanities and Social Sciences Communications, 10(1), 1–16.

[15] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.

[16] Taylor, W. J. (2000). A quantile regression neural network approach to estimating the conditional density of multiperiod returns. Journal of Forecasting, 19(4), 299–311.

[17] Cannon, A. J. (2011). Quantile regression neural networks: Implementation in R and application to precipitation downscaling. Computers & Geosciences, 37(9), 1277–1284.

[18] He, Y., Qin, Y., Wang, S., & Li, J. (2019). Electricity consumption probability density forecasting method based on LASSO-Quantile Regression Neural Network. Applied Energy, 233, 565–575.

[19] Zhu, B., Wei, Y., Zhang, L., & Li, Y. (2020). Forecasting carbon price using empirical mode decomposition and quantile regression neural network. Journal of Cleaner Production, 275, 122842.

[20] Aprea, I. L., Scognamiglio, S., & Zanetti, P. (2024). Systemic risk measurement: A Quantile Long Short-Term Memory network approach. Applied Soft Computing, 152, 111224.

[21] Horowitz, J. L. (1998). Bootstrap methods for median regression models. Econometrica, 66(6), 1327–1351.

[22] Koenker, R., & Bassett, J. G. (1978). Regression quantiles. Econometrica, 46(1), 33–50.