[1]Shuangqin Cheng
[2]Qiyi Zhang

SD2S: Multiscale Granger causality analysis based on serial decomposition state space models

Abstract: - Granger causality, widely recognized for its ease of use, intuitive interpretability, and applicability to complex multivariate systems, facilitates the inference of causal connections between variables through observational data and elucidates their dynamic interactions. With the advancement of the significant data era, an increasing multiscale characteristic of data is evident, presenting dynamics across multiple temporal scales. Current research in this domain typically relies on vector autoregressive models and wavelet transformations, which are susceptible to noise and dependent on substantial prior knowledge for selecting basis functions. To more accurately interpret and analyze Granger causality at multiple scales, this paper employs state space models and Empirical Mode Decomposition, introducing a novel multiscale Granger causality analysis approach based on serial decomposition state space models (SD2S). Experiments on simulated and real datasets confirm that (1) the integration of Empirical Mode Decomposition with state space models enhances the analysis of multiscale Granger causality; (2) serial decomposition state space models can improve the accuracy of multiscale analysis and effectively reduce computation time; (3) the proposed method successfully identifies dynamic causal relationships that vary with time scales in real-world data.

Keywords: granger causality; serial empirical mode decomposition; state space model; multiscale analysis.

I. INTRODUCTION

In disciplines such as physics, biology, and many others, complex time-series data often exhibit dynamics that span multiple timescales, known as multiscale data. Understanding the multiscale characteristics of these time series is crucial for uncovering the internal mechanisms of the sequences because they help reveal underlying physical laws, biological processes, and mechanisms of information transmission. Previous research [1–5] has deeply explored these issues by resampling original measurement data at various timescales to assess the dynamics of rescaled series, thus investigating their multiscale properties. This approach has achieved notable success in quantifying the multiscale behaviors of individual dynamic processes and has provided powerful tools for research in related fields. Granger causality (GC), initially proposed by Norbert Wiener [6] and further developed by Clive Granger [7–9], is a method used to determine if there is a causal relationship between two events. If using historical information about Y significantly improves predictions about future changes in X, then Y can be considered a Granger cause of X. Over time, Granger causality has been increasingly applied in fields such as economics, engineering, and neuroscience [10,11].

As data exhibits multiscale characteristics [12], there is a need to understand more profoundly the challenges and impacts of multiscale data on Granger causality analysis to mine Granger causality relationships at multiple scales effectively. Although empirical studies [13,14] have attempted to explore this area, the information transmission process involves complex interactions and influences between scales. Thus, accurately describing and computing the complexity of causality within a multiscale framework remains an urgent problem to solve. Most current research on Granger causality uses vector autoregressive (VAR) models for modeling [15], which are easily affected by noise, leading to inaccurate estimates. Additionally, the commonly used wavelet transform methods for extracting multiscale features require extensive prior knowledge to select appropriate wavelet bases[16].

Numerous scholars have pursued improvements in addressing the deficiencies of vector autoregressive models, particularly in handling noise and downsampling. In 1991, Aoki M and others introduced the state-space model for time series modeling, and Solo V emphasized that state-space models could maintain robust Granger causality without the influence of noise and downsampling, advocating for the use of state-space models over VAR models [17]. Barnett L and colleagues substantiated through comparative experiments that state-space models enhance statistical power and reduce bias in estimating Granger causality [18]. Hence, this study considers employing state space models to estimate Granger causality.

[1] * Department of Computer Science, College of Information Science and Technology, Jinan University, Guangzhou, 510632, China;
[2] Department of Mathematics, College of Information Science and Technology, Jinan University, Guangzhou 510632, China;
Correspondence: sqcheng@stu2021.jnu.edu.cn;

In light of the extensive prior knowledge required by traditional wavelet transformations, researchers have developed Empirical Mode Decomposition (EMD), an adaptive decomposition method that leverages the local characteristics of signals without predetermined basis functions. EMD iteratively decomposes signals to extract intrinsic mode functions (IMFs), representing local oscillatory patterns of the signal, and it has been proven to have excellent local feature extraction capabilities. Improved EMD approaches, introducing noise or adjusting iterations, have mitigated the problem of mode mixing, thereby enhancing the accuracy and stability of decomposition [19]. Variational Empirical Mode Decomposition (VMD), a signal decomposition method grounded in optimization theory, extracts finer local features by minimizing the discrepancy between the signal and its IMFs [20]. Beyond improving prediction accuracy [21,22], EMD has also been applied for multiscale analysis, with He K and others using it to reveal underlying driving factors [23] more clearly.

However, with increasing signal dimensions, traditional EMD algorithms face significant computational load challenges. To address this issue, Zhang et al proposed Serial-EMD, an innovative rapid EMD method that significantly reduces the computational complexity of multidimensional signal decomposition through serialization techniques [24].

This paper aims to explore multiscale Granger causality by introducing EMD and serial-EMD methods to capture signal characteristics across various timescales. Combined with state space models, it estimates multiscale Granger causality between variables, enhancing the accuracy and efficiency of computations.

The paper is structured as follows: Section 2 introduces the proposed SD2S method in detail, providing definitions for empirical mode decomposition, multiscale Granger causality, and state space models. Section 3 outlines the evaluation metrics utilized in this study, presents comparative experiments against vector autoregression-based methods, and includes an application on a real GDP dataset, analyzing Granger causality relationships across four different timescales for ten countries. Finally, Section 4 concludes the research and suggests future work directions.

II. Materials and Methods

A. Overview of the methodology

To enhance the analysis of multiscale Granger causality relationships, this paper introduces a novel analysis method termed Serial Decomposition State Space Models (SD2S) for Multiscale Granger Causality Analysis. This approach is founded on state space models and incorporates serial complete ensemble empirical mode decomposition with adaptive noise (SCEEMDAN) technique to investigate Granger causality across multiple scales.

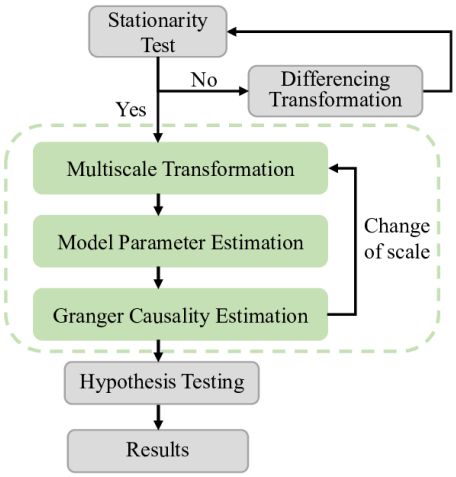


Figure 1 Method Flowchart

Initially, the framework for multiscale Granger causality analysis based on the SD2S method is presented, with the steps of the SD2S method delineated within a green dashed box in Figure 1. An augmented Dickey-Fuller test is initially conducted to assess data stationarity within this framework. If the data is stationary, the analysis proceeds directly to the Granger causality analysis. Conversely, if the data exhibits non-stationary characteristics, it must undergo differencing transformations until stationarity is achieved, after which the SD2S method is applied.

The core of the SD2S method lies in multiscale transformation, which includes two critical steps: multiscale decomposition using SCEEMDAN and the replacement of driver variable data.

The SD2S is applied at each scale s to unveil potential causal links at different temporal scales. Subsequently, the scale value s is altered, and the multiscale transformation and causality estimation process is iterated to determine

Granger causality relationships across all scales. After completing the SD2S method and hypothesis testing, the final results of the multiscale Granger causality analysis are obtained.

B. Serial Empirical Mode Decomposition

1) Empirical Mode Decomposition and Class Empirical Mode Decomposition

Empirical Mode Decomposition (EMD) is an adaptive signal processing method introduced in 1998 by American scientist N. E. Huanget al. [25], primarily designed for analyzing nonlinear and non-stationary signals. This method iteratively identifies local extrema in the signal, interpolates these extrema to estimate the upper and lower envelopes, and then removes the mean of these envelopes to obtain a low-pass trend line, thereby isolating the high-frequency oscillatory components as mode signals. By repeating this process, it is possible to extract Intrinsic Mode Functions (IMFs), each representing a local oscillatory mode of the signal, which helps to reveal more about the signal's intrinsic structure. However, EMD has limitations such as boundary effects, sensitivity to noise, and potential mode mixing in the presence of high-frequency noise [25].

Ensemble Empirical Mode Decomposition (EEMD) is an improved method over EMD, which enhances the original technique by adding white noise to the signal multiple times and performing multiple EMD decompositions. By averaging these decomposition results, artificially added random noise is eliminated, and the multi-frequency components of the signal are appropriately projected onto different scales. This method effectively reduces the mode mixing issues found in EMD and enhances the stability and accuracy of the decomposition.

Each xi(t) in the EEMD algorithm is decomposed independently under different added white noise and each decomposition may produce different modes and residuals of ri(t). To address the shortcomings of EEMD, M.E. Torres et al [19] proposed the Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN). Compared to EEMD, the CEEMDAN method introduces an additional signal-to-noise ratio to control the noise level during each decomposition process and can resolve the issue of mode mixing by decomposing the original data into IMFs of different frequencies. Although CEEMDAN resolves the problems of large reconstruction errors and inconsistent mode numbers due to different noise realizations seen in EEMD, it still presents some notable issues. For example, the modes may contain some noise involved, and signal information may appear later than in the EEMD method, resulting in some spurious modes in the early stages of decomposition [26].

Variational Mode Decomposition (VMD) is a signal decomposition technique based on variational optimization, introduced by N. Dragomiretskiy and D. Zosso [20], primarily designed to decompose complex nonlinear and non-stationary signals into a series of IMFs with varying frequencies and amplitudes. In VMD, each IMF is constructed as an analytic signal through the Hilbert transform to compute its unilateral spectrum. The spectral contents are then shifted to the baseband using the shifting properties of the Fourier transform, and bandwidths are estimated using H1 Gaussian smoothness. VMD aims to decompose the original signal into multiple smooth, decreasing-frequency modes by minimizing the total spectral bandwidth of all IMFs. This enhancement improves the practicality and accuracy of VMD in signal processing.

Furthermore, Rehman and Aftab expanded on VMD to develop Multivariate Variational Mode Decomposition (MVMD) for handling multichannel or multivariate data[27]. The core of MVMD lies in identifying a series of common multivariate modulated oscillations that share a common frequency component across all input data channels. This method extracts a set of multivariate modes with minimal collective bandwidth through a variational optimization problem, ensuring these modes can fully reconstruct the input signal. MVMD solves the optimization problem directly in the frequency domain using the method of multipliers and alternating directions without the need for additional user-defined parameters. This technique achieves mode alignment in multichannel data processing, ensuring oscillations with similar frequencies across multiple channels are assigned to the same mode. Consequently, MVMD performs well in handling multichannel signals, providing richer and more accurate signal features for fields such as biomedical signal analysis.

2) Serial Empirical Mode Decomposition

With the advent of the big data era, the surge in data volume has escalated the demand for real-time signal analysis, posing a challenge to traditional Empirical Mode Decomposition (EMD) and its variants in balancing data dimensionality growth with analysis speed. Zhang Jin et al [24] introduced a Serial Empirical Mode Decomposition method (SEMD) based on serialization technology to accelerate the decomposition of multidimensional signals. This method serializes multivariate or multidimensional signals into one-dimensional signals, allowing for the decomposition of multidimensional signals using the one-dimensional EMD algorithm. Such serialization not only

significantly reduces computation time but also offers a new perspective for optimizing current EMD algorithms, which is to transform the structure of the input signal rather than developing new envelope identification techniques or multi-EMD algorithms.

EMD often confronts the challenge of mode mixing, where components of signals at different scales may be amalgamated within the same IMF, or similar scale signals maybe scattered across different IMFs. The CEEMDAN method substantially reduces this mode mixing phenomenon by introducing ensemble averaging, albeit at the expense of significantly increased computation time. While the SEMD method reduces the computation time of traditional EMD, it does not effectively resolve the mode mixing problem. However, Serial-CEEMDAN (SCEEMDAN) leverages its serial processing attribute to effectively lower mode mixing issues and significantly shorten computation time, enhancing overall performance.

Therefore, the SD2S method proposed in this paper employs SCEEMDAN for multi-scale decomposition. The main steps of the SCEEMDAN method include signal serialization, signal decomposition, and IMF deserialization. A multidimensional signal, initially considered a composite multivariate signal composed of multiple one-dimensional signals, is serialized into a one-dimensional signal through a specific transformation function for further analysis. This conversion retains the original signal information and creates a smooth transition zone via linear interpolation between the ends of consecutive signals, ensuring continuity. Notably, the length of the transition zone is crucial; it not only influences the smoothness of signal transitions and preservation of original signal characteristics but also significantly affects computational load and efficiency. Therefore, in some instances, choosing an appropriate transition length is vital for correctly separating different modes within the signal.

CEEMDAN can decompose the serialized one-dimensional signal during the decomposition phase to obtain Intrinsic Mode Functions (IMFs). For a given signal *x(t)*, let $\widetilde{IMF_i}(t)$ denote the *i*-th IMF obtained by the CEEMDAN method. Let EMD *j(·)* denote the *j*-th IMF obtained by EMD decomposition.

Step 1: Add white noise with a signal-to-noise ratio of ε_0 to *x(t)* for a total of *I* times;

$$x_i(t) = x(t) + \varepsilon_0 w_i(t), i = 1,2,\dots,I$$

Step 2: Perform EMD decomposition on $x_i(t)$ quentially, taking the first IMF from each decomposition as $\mathrm{IMF}_{i,1}$. Compute the first IMF for CEEMDAN, $\widetilde{IMF_i}(t)$, and the corresponding residue $r_1(t)$ using $\mathrm{IMF}_{i,1}$;

$$\widetilde{IMF}_1(t) = \frac{1}{I}\sum_{i=1}^{I} \mathrm{IMF}_{i,1}(t) = \widetilde{\mathrm{IMF}}_1(t)$$

$$r_1(t) = x(t) - \widetilde{\mathrm{MF}}_1(t)$$

Step 3: Add an adaptive noise term $\varepsilon_1 \mathrm{EMD}_1(\mathrm{w}_i(\mathrm{t}))$ to $\mathrm{r}_1(\mathrm{t})$. Here, ε_1 is the signal-to-noise ratio, and is the $\mathrm{EMD}_1(\mathrm{w}_i(\mathrm{t}))$ first IMF obtained from the EMD decomposition of $\mathrm{w}_i(\mathrm{t})$. Then perform EMD decomposition *I* times, obtaining the average of the Idecompo- sitions as the new IMF;

$$\widetilde{\mathrm{IMF}}_2(t) = \frac{1}{I}\sum_{i=1}^{I} \mathrm{EMD}_1\left(r_1(t) + \varepsilon_1 \mathrm{EMD}_1\left(w_i(t)\right)\right)$$

$$r_2(t) = r_1(t) - \widetilde{\mathrm{IMF}}_2(t)$$

Step 4: To $\mathrm{r}_k(\mathrm{t})$, $k = 2, \dots, K$, add an adaptive noise term, which is the first IMF obtained from the EMD decomposition of the white noise $\mathrm{w}_i(\mathrm{t})$ with a signal-to-noise ratio ε_k, and obtain the new IMF;

$$\widetilde{\mathrm{IMF}}_k(t) = \frac{1}{I}\sum_{i=1}^{I} \mathrm{EMD}_1\left(r_{k-1}(t) + \varepsilon_{k-1}\mathrm{EMD}_{k-1}\left(w_i(t)\right)\right)$$

$$r_k(t) = r_{k-1}(t) - \widetilde{\mathrm{IMF}}_k(t)$$

Step 5: Repeat step 4 until the residue can no longer be decomposed, that is, when the number of extrema in the residue does not exceed two. At this point, the final residue is:

$$R(t) = x(t) - \sum_{k=1}^{K} \widetilde{\mathrm{IMF}}_k$$

Therefore, the original signal can be reconstructed as follows:

$$x(t) = R(t) + \sum_{k=1}^{K} \widetilde{\mathrm{IMF}}_k$$

During the signal decomposition process, the transition interval of the serialized signal is also decomposed. To extract the IMFs of the original signal, the IMF deserialization is done by reshaping and slicing to eliminate the transition interval, thus extracting the original signal's IMFs from the serialized IMFs.

SCEEMDAN, by serializing multidimensional signals into one-dimensional signals, reduces the computational load of finding local extrema and decomposing signals in high-dimensional spaces, thereby reducing mode mixing effects and significantly increasing decomposition speed. This provides a more flexible way to handle signal structures. In summary, SEMD adopts an effective method for multidimensional signal processing, facilitating rapid decomposition and real-time analysis of multidimensional signals.

C. State space model

Let's examine a typical constant State Space (SS) model,

$$\begin{aligned} X_{t+1} &= AX_t + W_t \\ Y_t &= CX_t + V_t \end{aligned} \tag{1}$$

in which W_t and V_t represent noise components, possessing semi-positive definite covariance matrices expressed as var $\begin{pmatrix} W_t \\ V_t \end{pmatrix} = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix}$ The defining parameters of the model (1) are (A, C, Q, R, S).

Projecting the state variable onto a subspace influenced by the historical values of the observed variable yields $Z_t = \mathbb{E}\{X_t|Y_{t-1}^-\}$, where $Y_{t-1}^- = [Y_{t-1}^T, Y_{t-2}^T \ldots]^T$ forms an infinite column vector. Consequently, Z_t is the updated state variable and linear regression residuals on y_t using its infinite past.

Thus, the innovation model of the state space (ISS) for Z_t, mirroring the SS model, is delineated as:

$$\begin{aligned} Z_{t+1} &= AZ_t + KE_t \\ Y_t &= CZ_t + E_t \end{aligned} \tag{2}$$

Here, the innovation Et is defined as $E_t = Y_t - \mathbb{E}\{Y_t|Y_{t-1}^-\}$ and its covariance matrix is denoted by $\sum = \mathbb{E}\{E_t E_t^T\}$. The parameters for the model (2) include $(A, C, K, \sum)$.

Regarding the ISS model, the error from a comprehensive regression is captured as $\lambda_{full} = \sum(j,j)$. Regarding the ISS model, the error from a comprehensive regression is captured as:

$$\begin{aligned} Z_{t+1} &= AZ_t + K\varepsilon_t \\ Y_t^{(jk)} &= C^{(jk)}Z_t + \varepsilon_t^{(jk)} \end{aligned} \tag{3}$$

where (*jk*) indicates selecting the *j*, *k* rows of the matrix. Consequently, the parameters of this adjusted state space model (3) are listed as $(A, C^{(jk)}, K\sum K^T, \sum(jk,jk), K\sum(:,jk))$. Notably, the ISS and SS models are interchangeable [18]. SS model parameters are transposed into ISS model parameters via solving the DARE equation, hence obtaining the error $\lambda_{reduced} = \sum^R(j,j)$ for the restricted model.

Ultimately, Granger causality is computed as follows:

$$F_{i\to j} = ln\frac{\sum^R(j,j)}{\sum(j,j)}$$

D. Multiscale Granger causality

Begin with n smooth complex processes $Y_n = [y_{1,n}, \cdots, y_{M,n}]^T$, each being a zero-mean scalar process. Let y_i represent the target variable and y_i the driver variable, while the rest, constituting $M-2$ processes of Y_n, form Y_k, where k represents the complementa set in $\{1,2,\ldots,M\}$ excluding $\{i,j\}$. The notation Y_n^- indicates the historical data of $Y_{n,}$ and Y_n^- denotes the past information of y_n. Granger causality from y_i to y_j assesses how much including the historical data of y_i, denoted as $y_{i,n}^-$, enhances the prediction of $y_{j,n}$ using the past data $y_{j,n}^-$ and $y_{k,n}^-$.

To predict the current state of the target variable using all past processes, referred to as the full regression, the prediction error is calculated as follows:

$$\varepsilon_n = y_{j,n} - \mathbb{E}[y_{j,n}|Y_n^-]$$

Here, E represents the expectation operator.

The error from the restricted regression, which utilizes the historical data of all variables except the driver variable for predicting the current state of the target, is given by:

$$\varepsilon_n^R = y_{j,n} - \mathbb{E}\left[y_{j,n}\middle|y_{j,n}^-, Y_k^-\right] \quad (4)$$

where the superscript *R* indicates reduced regression.

By evaluating the log-likelihood of the errors from both the full and restricted regressions, which are $\lambda = \mathbb{E}[\varepsilon_n \varepsilon_n^T]$ and $\lambda^R = \mathbb{E}[\varepsilon_n^R (\varepsilon_n^R)^T]$ respectively (with T representing transpose), the Granger causality measure from y_i to y_j is determined as:

$$F_{i\to j} = \ln \frac{\lambda^R}{\lambda}$$

To compute the multiscale Granger causality estimates, Empirical Mode Decomposition employed to derive intrinsic mode functions c_s at various scales, and the driver variable y_i is replaced with the corresponding scale's c_s to extract signal information better. As the reduced regression does not utilize the driver variable y_i data, the error in reduced regression remains unaffected by the multiscale transformation. It is still as depicted in equation (4) . At this point, the prediction error for the full regression is given as $\varepsilon_s = y_j - \mathbb{E}\left[y_j\middle|y_j^-, Y_k^-, c_s^-\right]$, thus $\Sigma_s^R(j,j) = \mathbb{E}[\varepsilon_s \varepsilon_s^T]$.

Hence, at scale $s(s = 1, \ldots, N)$, the Granger causality value from y_i to y_j is expressed as

$$F_{i\to j} = \ln \frac{\sum_s^R (j,j)}{\sum (j,j)}$$

III. Simulation Study and Results

A. Evaluation Indicator

Several evaluation metrics are utilized to facilitate the experiments, and their computation methods are subsequently detailed.

Initially, a reject rate (RR) is established based on the significance test outcomes as follows:

$$\mathrm{RR} = \frac{NRS}{TNS}$$

Here, NRS is the count of rejections from significance tests, and TNS is the total number of these tests conducted. A low reject rate under true causality suggests the algorithm's effectiveness in achieving accurate estimations before hypothesis testing, thus minimally depending on such testing methods. Conversely, a high reject rate when absent causality suggests that hypothesis testing can help mitigate estimation errors.

To encapsulate the effects comprehensively, this study amalgamates the reject rates from both scenarios into a single metric termed the overall reject rate (ORR), defined as:

$$\mathrm{ORR} = \mathrm{RR}_E + 1 - \mathrm{RR}_N$$

Here, RRE denotes the RR when Granger causality is truly present, and RRN is the rate at which Granger causality does not exist.

This study incorporates the metrics of Accuracy and Precision, defined as follows:

$$\mathrm{Accuracy} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FN} + \mathrm{FP} + \mathrm{TN}}$$

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$$

Here, TP represents the count of items correctly identified as causally related both in prediction and reality. FN counts the items wrongly predicted as non-existent, though present; TN is the count of items correctly identified as

non-existent both in prediction and reality, and FP is the count of items wrongly identified as present, though they are absent.

B. Simulation Model

1) Data Description

We employ a VAR(3) model for simulation and theoretical analysis. The equations of the VAR model are as follows:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2r_1\cos(\theta_1) & 0 & 0 \\ -0.356 & 2r_2\cos(\theta_2) & 0 \\ 0 & -0.3098 & 2r_3\cos(\theta_3) \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \\ x_{3,t-1} \end{bmatrix} + \begin{bmatrix} -r_1^2 & 0 & 0 \\ 0.7136 & -r_2^2 & 0 \\ 0 & 0.5 & r_3^2 \end{bmatrix} \begin{bmatrix} x_{1,t-2} \\ x_{2,t-2} \\ x_{3,t-2} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ -0.356 & 0 & 0 \\ 0 & -0.3098 & 0 \end{bmatrix} \begin{bmatrix} x_{1,t-3} \\ x_{2,t-3} \\ x_{3,t-3} \end{bmatrix} + \begin{bmatrix} \omega_{1,t} \\ \omega_{2,t} \\ \omega_{3,t} \end{bmatrix} \quad (5)$$

Here, $r_i = \{0.9,0.7,0.8\}$, $\theta_i = f_i \Delta t 2\pi, i = \{1,2,3\}$, $\frac{1}{\Delta} = 120Hz$, and considering the sample rate is $\frac{1}{\Delta} = 120Hz$, with $f_i = \{40,10,50\}$Hz. This setup facilitates modeling the causality from x1to x2 and from x2 to x3.

2) Validation of state space model validity

The following experiments were conducted to verify the influence of the state space model on the validity of the empirical modal decomposition. Firstly, based on the VAR(3) model (5), we generated experimental data with a length of 512 and 10 trials. Under four modes and a significance level of α = 0.05, empirical mode decomposition was performed using state-space and VAR-based methods to analyze multiscale Granger causality, and various evaluation metrics was calculated. All subsequent experiments were conducted using this dataset and conditions.

Table 1 Multiscale Estimation Effects of the SD2S Method

%	EMD	EEMD		CEEMDAN			VMD	MVMD		
	VAR	SSVAR		SSVAR		SS	VAR	SSVAR		SS
Accuracy	71.25	75.42	69.17	77.08	72.50	77.92	32.50	55.42	39.17	65.83
ORR	31.29	24.58	31.58	22.92	30.55	22.08	44.67	44.58	42.77	34.17
Precision	60.38	88.89	54.84	87.88	60.00	84.62	32.48	41.82	34.58	49.29

The table shows that the SD2S method outperforms the VAR method, particularly in the VMD and MVMD approaches, with SD2S improving accuracy by up to 20%. This underscores the effectiveness of the SD2S method in multiscale Granger causality analysis, especially when using state space models, significantly enhancing the accuracy of multiscale GC estimates under empirical mode decomposition. Furthermore, EMD's decomposition based on local extrema enables better capture of signal local features to identify causal relationships between variables. In contrast, VMD and similar methods perform poorly, likely due to their global optimization approach to decomposition, which may not be sensitive enough to signal local features. Additionally, the table data indicate that the MVMD method significantly improves over VMD in all metrics, likely due to its attribute of mode alignment in decomposition, which is crucial for multivariate data analysis.

3) Validation of SD2S validity

Next, we employ VAR and SD2S methods based on serial empirical mode decomposition to estimate multiscale Granger causality. The results are displayed in Table 2, where SVMD, SEMD, SEEMD, and SCEEMDAN represent the serial VMD, EMD, EEMD, and CEEMDAN decomposition methods.

Table 2 Estimation Results Using Serial Decomposition Methods

%	VAR				SS			SD2S
	SVMD	SEMD	SEEMD	SCEEMDAN	SVMD	SEMD	SEEMD	
Accuracy	38.75	72.50	67.08	71.25	58.75	75.83	78.75	81.67
ORR	45.17	30.18	31.82	31.22	41.25	24.17	21.25	18.33
Precision	35.11	62.96	50.75	57.14	43.45	95.83	89.19	90.91

Notably, lower is generally considered better according to the rejection rate definition. In the table, most methods show a decrease in rejection rates compared to the original data, while only a few show an increase, but none exceed a 3% increase. Therefore, from the perspective of rejection rates, it can be said that the serial methods have effectively improved the accuracy of estimating multiscale Granger causality relations.

Overall, using the serial method, almost all evaluation metrics have improved. Par- particularly, the State Space (SS) method performs better than the original across all metrics. Meanwhile, our proposed SD2S method outperforms other methods in both Accuracy and Overall Rejection Rate (ORR), and ranks just below the SEMD method in Precision. This suggests that the SD2S method, compared to the SEMD method, demonstrates better generalization effects overall. Although the SEMD method excels at identifying variable combinations where causal relationships exist, its overall performance does not match that of the SD2S method.

Additionally, the main advantage of the serial method is its significant reduction in the decomposition process runtime. We estimated the runtime and calculated the mean and standard deviation across ten trials to verify this.

Table 3 Runtime Using Original and Serial Methods

Method	EMD	EEMD	VMD	CEEMDAN
Original	0.192±0.011	47.211±0.517	0.463±0.150	30.762±0.391
Serial	0.193±0.020	28.585±0.018	0.371±0.244	25.281±0.443

The data in Table 3 displays the runtime in seconds for the original methods EMD, VMD, EEMD, CEEMDAN, and their serial counterparts. Aside from EMD, which already had a relatively short runtime, the VMD, EEMD, and CEEMDAN methods showed time reductions of 19.87%, 39.45%, and 17.82%, respectively, when using the serial versions. Furthermore, the runtime difference between SEMD and EMD is negligible. Thus, it can be concluded that integrating the serial method effectively reduces the runtime of decomposition methods, especially as the number of variables and signal length increase, where the advantages of the serial method become even more apparent.

C. Practical Application

For a long time, fluctuations in economic growth cycles have been a topic of significant interest among economists. In the context of globalization, the dependence and interconnection between national economies have continuously strengthened. Fluctuations in one country's economy often affect others and trigger a chain reaction. Therefore, in-depth study of macroeconomic data, especially the relationship between economic fluctuations in different countries, is crucial for grasping global economic dynamics, predicting economic trends, and formulating effective economic policies.

In 2003, Lee et al. proposed using the VAR model to analyze how the business cycles of the United States and Japan affect the Australian economy [28]; other scholars have used wavelet analysis to study the relationship between energy consumption and economic growth within the same country at different timescales [29,30]. However, studies on the impact of economic fluctuations between different countries are relatively scarce.

This study selects GDP data from 1970 to 2022 for ten countries from the World Bank's World Development Indicators database [31]. All countries' data are based on the year 2015, and the rate of GDP change is calculated using the following formula, where Sn represents the GDP data for year n:

GDP is chosen as the main indicator because it is the core measure of a country's economic size and total economic activity, and its fluctuations directly reflect the stability and growth potential of the nation's economy. Additionally, this study will use Granger causality analysis to deeply explore the potential connections and mutual influence mechanisms between economic fluctuations of different countries, thus providing theoretical support for international economic cooperation and coordination.

Through in-depth analysis of macroeconomic data, we can reveal the intrinsic connections between economic fluctuations in different countries and explore the underlying economic logic and transmission mechanisms. This not only helps us better understand the operational laws of the global economy but also provides important decision-making bases for policymakers, thereby promoting the stability and sustainable development of the global economy.

1) Data Description

Before conducting a multiscale Granger causality analysis, we first calculated the GDP change rates of various countries and plotted them, as shown in Figure 2.

From the figure, it can be observed that over time, each country's GDP exhibits varying degrees of fluctuation. For example, the fluctuations in China's GDP time series are more pronounced than in other countries, while similar fluctuation patterns are evident in other countries during different periods.

Furthermore, Table 4lists the basic descriptive statistics for ten countries, including Brazil, the United States, and France, covering mean, standard deviation, minimum, median, maximum, skewness, and kurtosis, to help us understand the overall trends and fluctuations of the data.

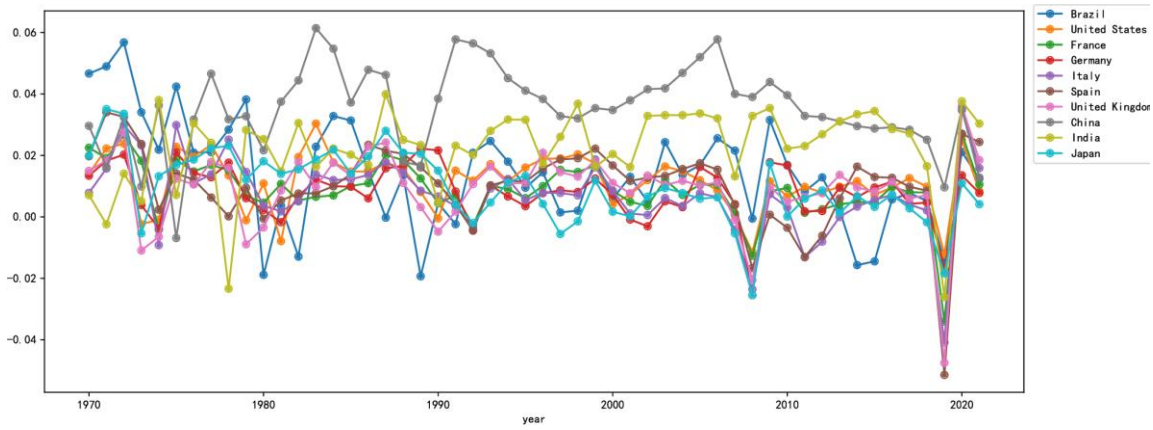


Figure 2 GDP Change Rates of Various Countries

As shown in Table 4, the mean values of GDP change rates for countries like the United States, France, and Germany are lower than their medians, and the skewness is less than zero, indicating that the sample's GDP change rates are mostly left-skewed with most changes being minimal. All countries have positive kurtosis, indicating that the distribution of GDP change rates is more concentrated than a normal distribution. Spain and the UK especially have higher kurtosis, suggesting that their change rates are more concentrated within a specific range.

Table 4 Descriptive Statistics of GDP Change Rates for Various Countries

	Mean	Standard Deviation	Minimum	Median	Maximum	Skewness	Kurtosis
Brazil	0.0143	0.0172	-0.0193	0.0141	0.0568	0.1205	0.0720
USA	0.0117	0.0089	-0.0122	0.0119	0.0303	-0.7341	0.7632
France	0.0089	0.0095	-0.0340	0.0092	0.0271	-1.7430	7.5644
Germany	0.0080	0.0089	-0.0255	0.0082	0.0222	-1.2803	3.4343
Italy	0.0070	0.0123	-0.0408	0.0073	0.0347	-1.0577	4.2834
Spain	0.0103	0.0134	-0.0514	0.0120	0.0340	-2.0279	8.1184
UK	0.0091	0.0126	-0.0475	0.0110	0.0361	-1.9905	7.7244
China	0.0355	0.0135	-0.0069	0.0358	0.0614	-0.5997	1.0258
India	0.0224	0.0135	-0.0261	0.0252	0.0399	-1.7800	4.1829
Japan	0.0095	0.0114	-0.0255	0.0090	0.0351	-0.4013	1.1677

To preliminarily understand the correlations among different countries, this paper presents a heatmap of Pearson correlation coefficients between ten countries, as shown in Figure 3. The results indicate that most countries have correlation coefficients above 0.9, with a few around 0.7, suggesting strong positive correlations among these nations. It is important to note that correlation coefficients reflect the similarity between variables and provide a preliminary basis for causality research.

2) Multiscale Granger Causality Analysis

This section uses the Multiscale Granger Causality Analysis Method based on the Serial Decomposition State Space Model (SD2S) to explore the causal relationships between GDP growth rates among different countries.

Western economists have defined and classified economic cycles based on their durations: short-term economic cycles, usually triggered by changes in inventory levels, last 3 to 5 years and are known as "Kitchen cycles"; medium-term economic cycles, caused by changes in fixed investments in industrial and financial sectors, last 7 to 11 years and are called "Juglar cycles"; medium-long-term economic cycles, initiated by changes in construction

investments, last 15 to 20 years and are known as "Kuznets cycles"; long-term economic cycles, driven by technological innovations, last 50 to 60 years and are referred to as "Kondratieff cycles" [32]. We set the maximum decomposition scale to 4 and used the SCEEMDAN method to decompose the data for further analysis.

According to the experimental results, we present the Granger causality relationships among ten countries at different scales, displayed as directed graphs, as shown in Figure 4. The figure shows the causality-directed graphs for the first three scales. According to the results, no causality was found at the fourth scale. This may be due to the gradual weakening of economic trade and interactions among countries over time. Conversely, the data sample covering only 1970-2022 may be insufficient to detect long-term relationships. This indicates that short-term economic activities and interactions deserve more attention and are likely the primary factors influencing international relations.

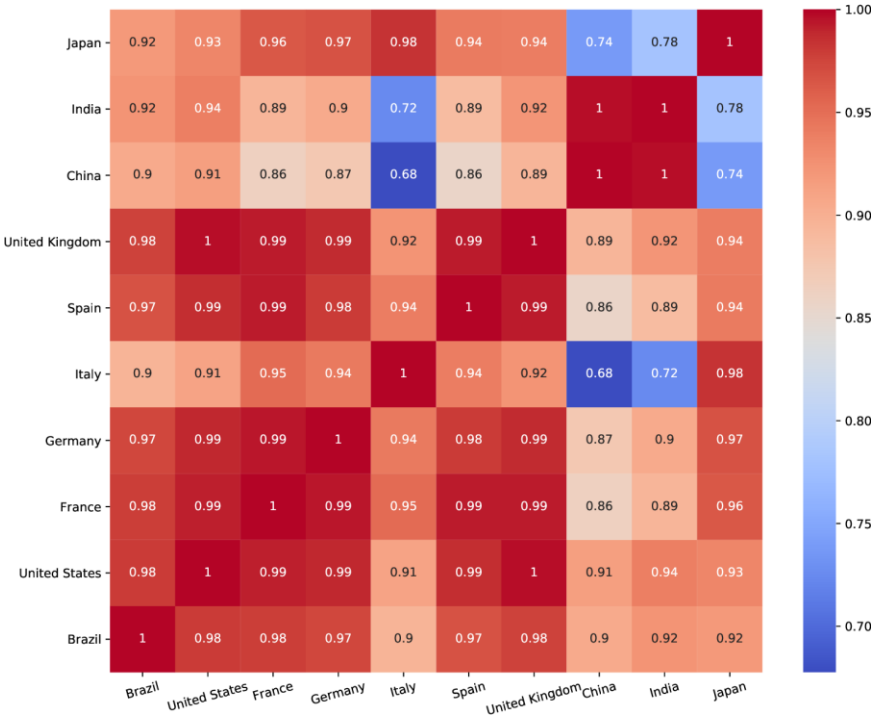
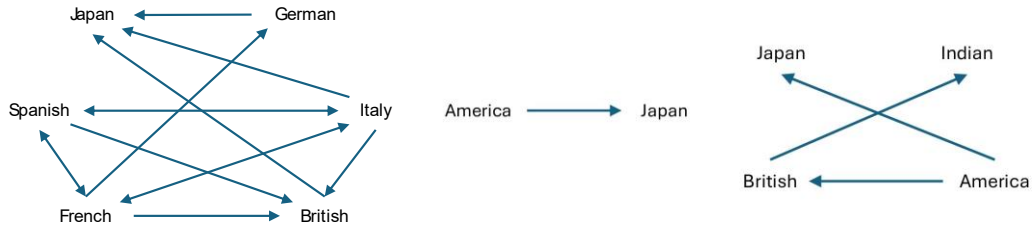


Figure 3 Heatmap of GDP Correlation Coefficients Among Countries

From Figure 4a, it is observed that compared to other countries, Japan and the UK receive more arrows, suggesting that they are more susceptible to influences from other nations. Conversely, Italy and France appear to be dominant in driving economic activities to some extent. Additionally, the causal relationships among Spain, Italy, and France indicate that these geographically adjacent EU member countries have close economic ties and trade relations, all using the euro and influenced by the same monetary policy.

From Figure 4b, it is evident that the United States has a significant impact on Japan during the medium-term economic cycle. The Plaza Accord signed in 1985 with Japan severely affected Japan's export trade, impacting the stock market and leading to a bubble burst and prolonged economic slump, which slowed down GDP growth [33].



(a) Scale 1 (b) Scale 2 (c) Scale 3

Figure 4 Multiscale Granger Causality Relationships of GDP Among Countries

As seen in Figure 4c, the UK significantly influences India. As a former colonial ruler, the UK has profoundly impacted India's political, legal, educational, and economic aspects, with these influences continuing to this day.

Over the past few decades, investment relations between the UK and India have grown rapidly, with the UK being the largest investor from the G20 countries in India [34]. Also, it is a major export destination for India [35].
These observations reveal that the results of multiscale Granger causality analysis differ from those of correlation coefficient analysis. This discrepancy is partly because correlation coefficients only reflect the similarity between datasets, whereas Granger causality analysis investigates how historical data affects current data to establish causality. Moreover, the analysis method adopted in this paper focuses on Granger causality at different scales, identifying the inherent multiscale complexities that are not captured by correlation coefficients and traditional Granger causality methods.

IV. Conclusions

With the advent of the big data era, data increasingly exhibits multiscale and multi-dimensional characteristics, posing new challenges for analyzing Granger causality. On the one hand, noise in multidimensional data may lead to significant biases when utilizing vector autoregressive models. On the other hand, traditional methods for multiscale feature extraction rely heavily on prior knowledge and do not fully utilize the inherent features of the data itself. To address these challenges, this paper proposes a new SD2S method that integrates SS models with the SCEEMDAN to enhance the accuracy of model estimates and the effectiveness and efficiency of multiscale decomposition. This study initially compares the traditional vector autoregressive models with empirical mode decomposition methods combined with state space models, revealing that the state space models improved across all evaluation metrics, with the highest accuracy increase of 18.8%, thus proving their effectiveness. Furthermore, the paper compares five different empirical mode decomposition methods, whether serial or not, demonstrating that the SCEEMDAN method reduced computation time by 18.82% compared to traditional CEEMDAN, with an improved accuracy rate. In summary, the SD2S method proposed in this paper effectively enhances the estimation of multiscale Granger causality and computational efficiency.
Nevertheless, there is still room for improvement in this method. On the one hand, although this paper has compared state space models with vector autoregressive models, given that existing research has attempted to combine Granger causality with neural networks [36], future studies could consider integrating with more neural networks or deep learning algorithms for a more comprehensive exploration of Granger causality. On the other hand, the multiscale analysis framework in this paper involves model order selection based on information criteria and hypothesis testing based on statistical testing, where the inconsistency of mathematical foundations between these two methods could affect the consistency of results [37]. Future research could use methods such as description length to unify model complexity and error terms, simplify the model selection process, and further optimize the multiscale Granger causality analysis method.

References

[1] Ivanov, P.C.; Amaral, L.A.N.; Goldberger, A.L.; Havlin, S.; Rosenblum, M.G.; Struzik, Z.R.; Stanley, H.E. Multifractalityin human heartbeat dynamics. Nature 1999, 399, 461–465.
[2] Kang, X.; Jia, X.; Geocadin, R.G.; Thakor, N.V.; Maybhate, A. Multiscale entropy analysis of EEG for assessment of post-cardiac arrest neurological recovery under hypothermia in rats. IEEE Transactions on Biomedical Engineering 2009, 56, 1023–1031.
[3] Valencia, J.F.; Porta, A.; Vallverdu, M.; Claria, F.; Baranowski, R.; Orlowska-Baranowska, E.; Caminal, P. Refined Multiscale Entropy: Application to 24-h Holter Recordings of Heart Period Variability in Healthy and Aortic Stenosis Subjects. IEEE Transactions on Biomedical Engineering 2009, 56, 2202–2213.
[4] Chou, C.M. Wavelet-based multi-scale entropy analysis of complex rainfall time series. Entropy 2011, 13, 241–253.
[5] Wang,J.; Shang, P.; Zhao, X.; Xia,J. Multiscale entropy analysis of traffic time series. International Journal of Modern Physics C 2013, 24, 1350006.
[6] Wiener, N.; Masani, P. The prediction theory of multivariate stochastic processes. Acta Mathematica 1957, 98, 111–150.
[7] Granger, C.W.J. Economic processes involving feedback. Information and control 1963, 6, 28–48.
[8] Granger, C.W.J. Investigating causal relations by econometric models and cross-spectral methods. Econometrica: journal of the Econometric Society 1969, pp. 424–438.
[9] Zhang,Q.; Zhang, C.; Cheng, S. Wavelet Multiscale Granger Causality Analysis Based on State Space Models. Symmetry 2023, 15, 1286. https://doi.org/10.3390/SYM15061286.
[10] Granger, C.W.J. Some properties of time series data and their use in econometric model specification. Journal of econometrics 1981, 16, 121–130.

[11] Coronado, S.; Martinez,J.N.; Gualajara, V.; Romero-Meza, R.; Rojas, O. Time-Varying Granger Causality of COVID-19 News on Emerging Financial Markets: The Latin American Case. Mathematics 2023, 11.

[12] Leung, T.; Zhao, T. A Noisy Fractional Brownian Motion Model for Multiscale Correlation Analysis of High-Frequency Prices. Mathematics 2024, 12.

[13] Peng, Y.; Chen, W.; Wei, P.; Yu, G. Spillover effect and Granger causality investigation between China's stock market and international oil market: A dynamic multiscale approach. Journal of Computational and Applied Mathematics 2020, 367, 112460.

[14] Jammazi, R.; Ferrer, R.; Jareño, F.; Shahzad, S.J.H. Time-varying causality between crude oil and stock markets: What can we learn from a multiscale perspective? International Review of Economics & Finance 2017, 49, 453–483.

[15] Amaral, A.; Dyhoum, T.E.; Abdou, H.A.; Aljohani, H.M. Modeling for the Relationship between Monetary Policy and GDP in the USA Using Statistical Methods. Mathematics 2022, 10.

[16] Saadatmorad, M.; Talookolaei, R.A.J.; Pashaei, M.H.; Khatir, S.; Wahab, M.A. Pearson correlation and discrete wavelet transform for crack identification in steel beams. Mathematics 2022, 10, 2689.

[17] Solo, V. On causality I: sampling and noise. In Proceedings of the 2007 46th IEEE Conference on Decision and Control. IEEE, 2007,pp. 3634–3639.

[18] Barnett, L.; Seth, A.K. Granger causality for state-space models. Physical Review E 2015, 91, 040101.

[19] Torres, M.E.; Colominas, M.A.; Schlotthauer, G.; Flandrin, P. A complete ensemble empirical mode decomposition with adaptive noise. In Proceedings of the 2011 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, 2011, pp. 4144–4147.

[20] Dragomiretskiy, K.; Zosso, D. Variational mode decomposition. IEEE transactions on signal processing 2013, 62, 531–544.

[21] Lin, C.S.; Chiu, S.H.; Lin, T.Y. Empirical mode decomposition–based least squares support vector regression for foreign exchange rate forecasting. Economic Modelling 2012, 29, 2583–2590.

[22] Plakandaras, V.; Gupta, R.; Gogas, P.; Papadimitriou, T. Forecasting the US real house price index. Economic Modelling 2015, 45, 259–267.

[23] He, K.; Zha, R.; Wu, J.; Lai, K.K. Multivariate EMD-based modeling and forecasting of crude oil price. Sustainability 2016, 8, 387.

[24] Zhang,J.; Feng, F.; Marti-Puig, P.; Caiafa, C.F.; Sun, Z.; Duan, F.; Solé-Casals,J. Serial-EMD: Fast empirical mode decomposition method for multi-dimensional signals based on serialization. Information Sciences 2021, 581, 215–232.

[25] Barbosh, M.; Singh, P.; Sadhu, A. Empirical mode decomposition and its variants: A review with applications in structural health monitoring. Smart Materials and Structures 2020, 29, 093001.

[26] Colominas, M.A.; Schlotthauer, G.; Torres, M.E. Improved complete ensemble EMD: A suitable tool for biomedical signal processing. Biomedical Signal Processing and Control 2014, 14, 19–29.

[27] Ur Rehman, N.; Aftab, H. Multivariate variational mode decomposition. IEEE Transactions on signal processing 2019, 67, 6039–6052.

[28] Lee, H.H.; Huh, H.S.; Harris, D. The relative impact of the US and Japanese business cycles on the Australian economy. Japan and the world economy 2003, 15, 111–129.

[29] Aslan, A.; Apergis, N.; Yildirim, S. Causality between energy consumption and GDP in the US: evidence from wavelet analysis. Frontiers in Energy 2014, 8, 1–8.

[30] Ha,J.; Tan, P.P.; Goh, K.L. Linear and nonlinear causal relationship between energy consumption and economic growth in China: New evidence based on wavelet analysis. PloS one 2018, 13, e0197785.

[31] Bank, W. World Development Indicators. https://data.worldbank.org/indicator/EN.ATM.CO2E.PC, 2024. Accessed on 21 Feb.2024.

[32] Mao, X.; Yang, A.C.; Peng, C.K.; Shang, P. Analysis of economic growth fluctuations based on EEMD and causal decomposition. Physica A: Statistical Mechanics and its Applications 2020, 553, 124661.

[33] Lee, Y. Economic interdependence and peace: A case comparison between the US-China and US-Japan trade disputes. East Asia 2018, 35, 215–232.

[34] Matharoo, R.; Singh, G. India and UK Trade. International Journal of Trade and Commerce-IIARTC 2017, 6, 444–449.

[35] Ranjan, P. India-UK Investment Relations and 'Global Britain' . King's Law Journal 2023, 34, 120–144.

[36] Rosoł, M.; Mły ´nczak, M.; Cybulski, G. Granger causality test with nonlinear neural-network-based methods: Python package and simulation study. Computer Methods and Programs in Biomedicine 2022, 216, 106669.

[37] Hu, Z.; Li, F.; Cheng, M.; Lin,Q. Investigating Large-Scale Network with Unified Granger Causality Analysis. Computational and Mathematical Methods in Medicine 2022, 2022, 1–15.