Time Series Analysis: Examples of Real World Applications

MSDM 5058

Prepared by S.P. Li

Example One:
Financial Network Analysis

The Shanghai Stock Exchange Composite Index

"Analysis of Network Clustering Behavior of the Chinese Stock Market", with Y. Mai and H. Chen, Physica A414(2014)360-7

Table 1. China Industrial Classification Standard of CSI300

Industry	Code	Number of stocks	Closing price average	Closing price SD
Industrials	*IN*	64	14.19	29.90
Financials	*FI*	52	13.08	15.15
Materials	*MA*	51	16.52	36.60
Materials	*MA*	51	16.52	36.60
Consumer discretionary	*CD*	27	15.67	22.19
Energy	*EN*	23	20.13	32.50
Health care	*HC*	16	26.70	44.93
Consumer staples	*CS*	14	33.25	81.87
Utilities	*UT*	13	8.49	7.36
Information technology	*IT*	8	14.28	16.48
Telecommunication services	*TS*	3	14.28	29.44

** During the period October 9, 2007 through March 29, 2013, we choose 214 stocks, each has 1334 observations.

Analysis of structure of the China Security Index 300 (CSI300) sectors

Define the logarithmic price return $R_i(t)$ of stock *i* on trading day *t* as

$$R_i(t) = \ln p_i(t) - \ln p_i(t-1),$$

where $p_i(t)$ is the closing price of the i^{th} stock on trading day *t*. To compare the result of different datasets, we further normalize the logarithmic price return as follows

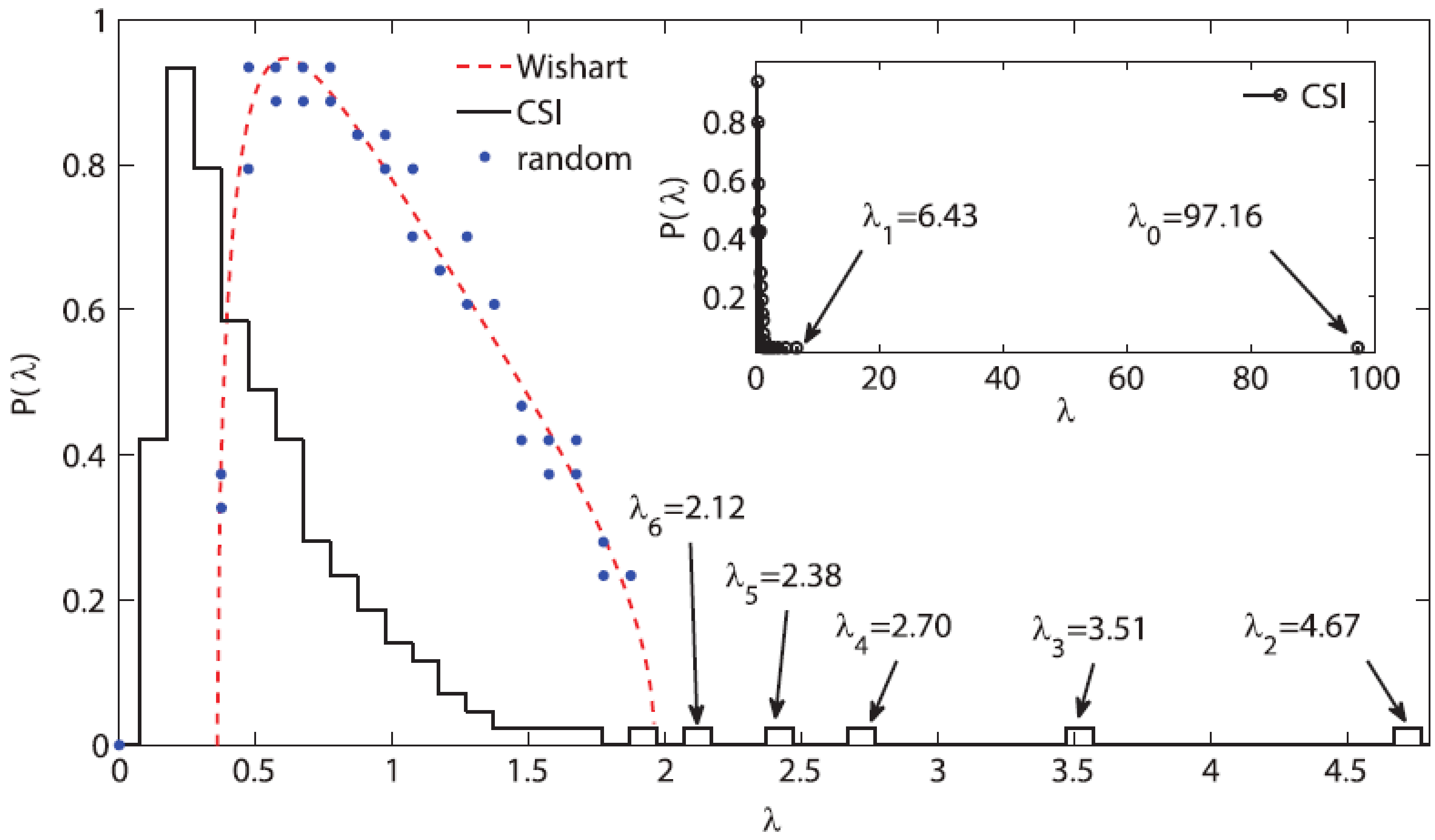
$$r_i(t) = \frac{R_i(t) - <R_i(t)>}{\sqrt{<R_i(t)^2> - <R_i(t)>^2}}$$

where <*X*> stands for the average of *X* during the trading period.

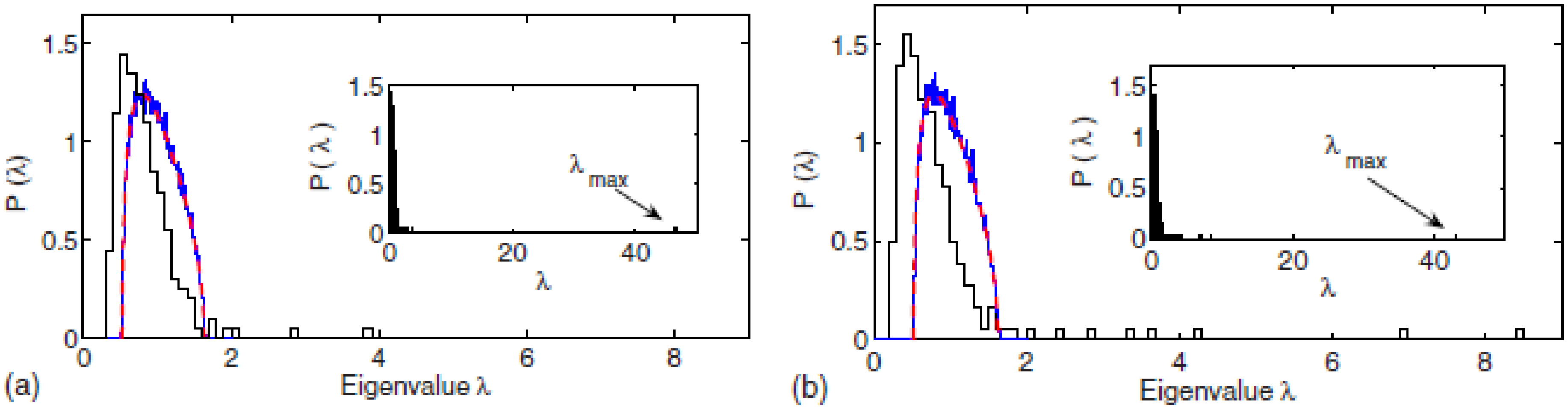
We define the cross-correlation of stocks *i* and *j* as

$$C_{ij} = < r_i(t) r_j(t) >$$

C_{ij} is a symmetric matrix with values between -1 and 1.

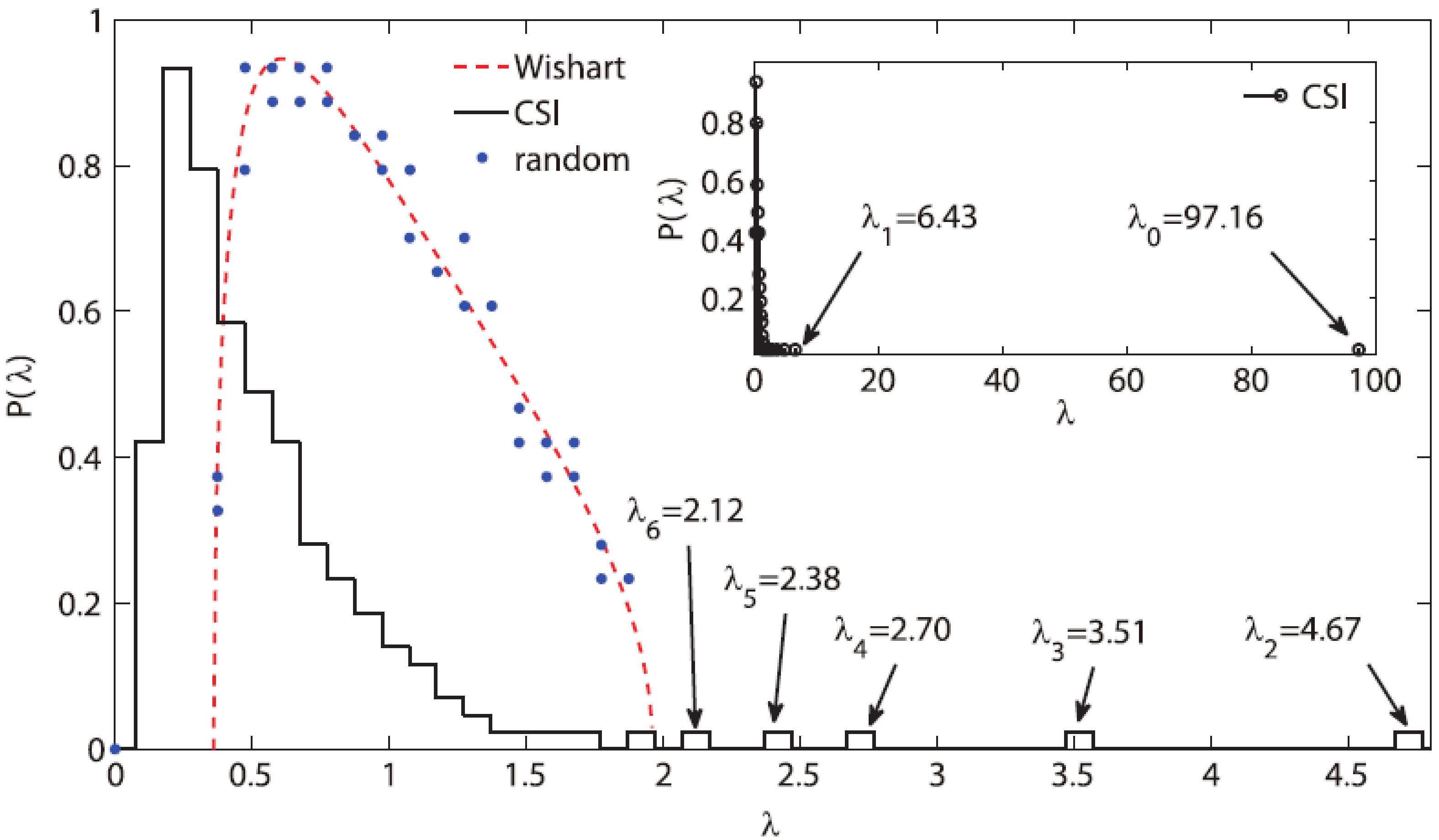


The probability distribution of the eigenvalues of the correlation matrix C for CSI300. (The red dotted curve is the distribution of the Wishart Matrix. The black curve is the probability distribution of the eigenvalues of the correlation matrix C. The blue data points are the distribution of the eigenvalues of the correlation matrix C for the reshuffled time series.)

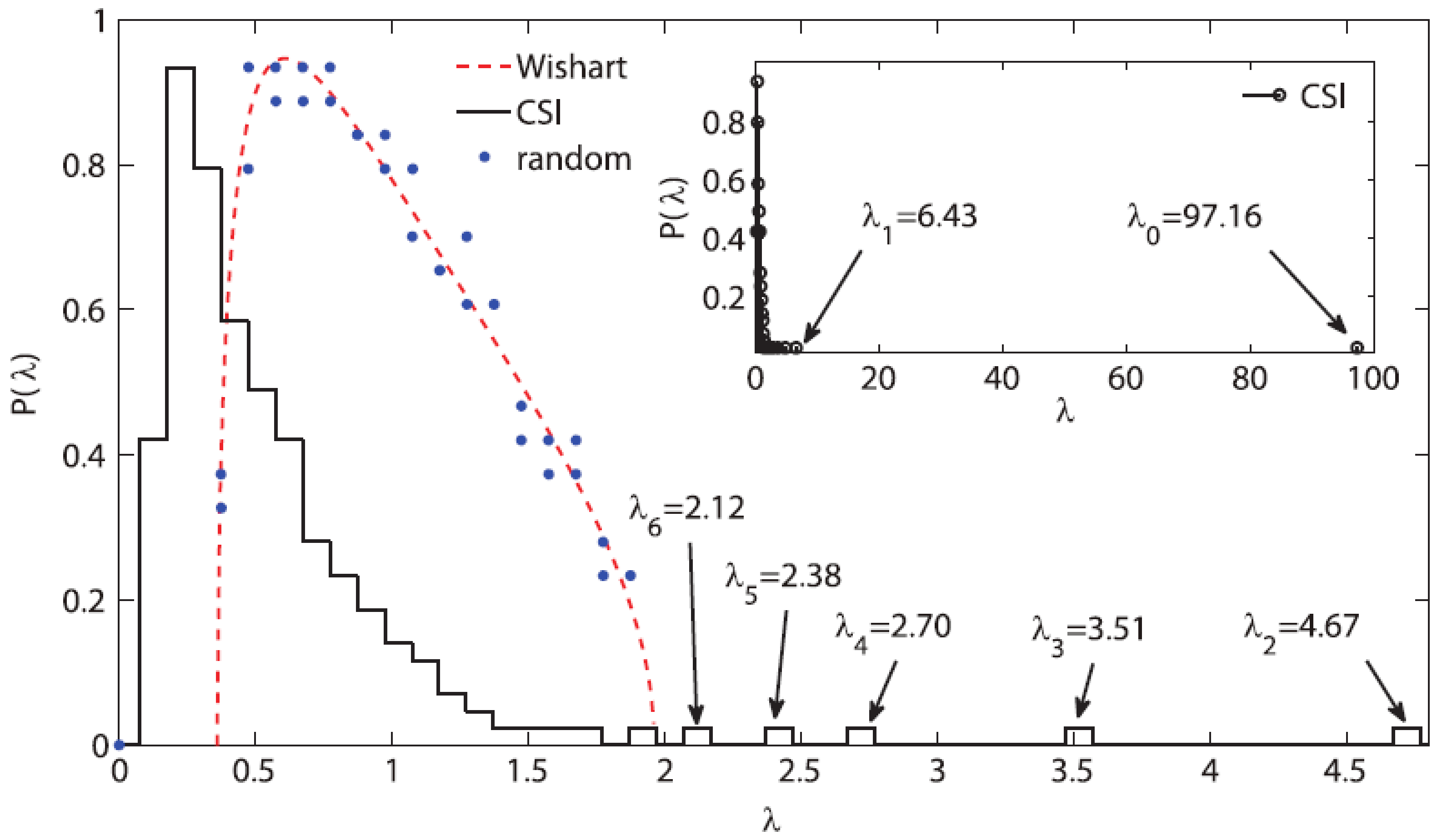


The probability density function of the eigenvalues of the correlation matrix C for (a) NSE (India) and (b) NYSE. For comparison, the theoretical distribution predicted by Wishart Matrix is shown using broken curves, which overlaps with the distribution obtained from the surrogate correlation matrix generated by randomly shuffling each time series.

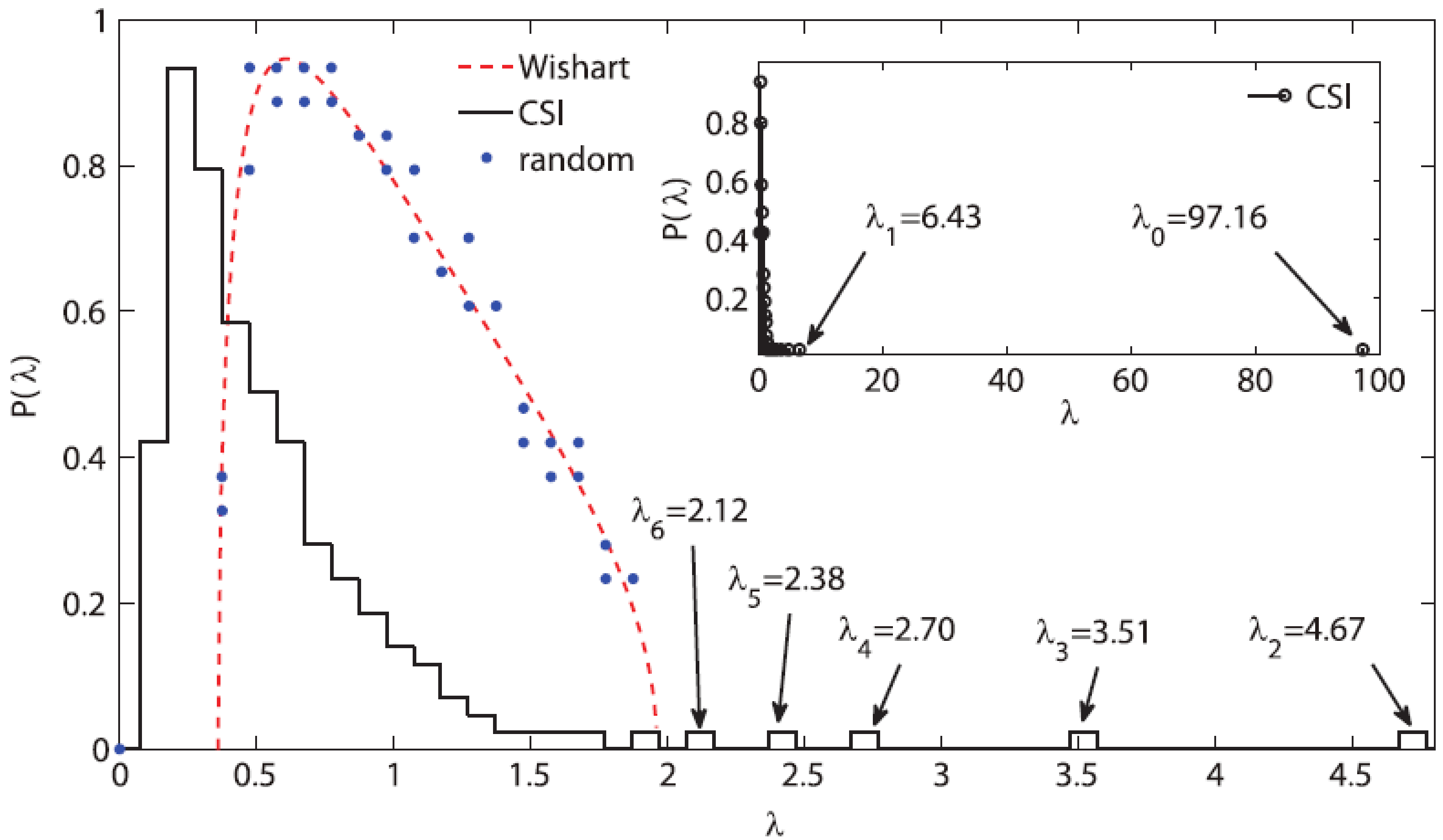
** R.K. Pan and S. Sinha, Phys. Rev. E 76 (2007) 046116.



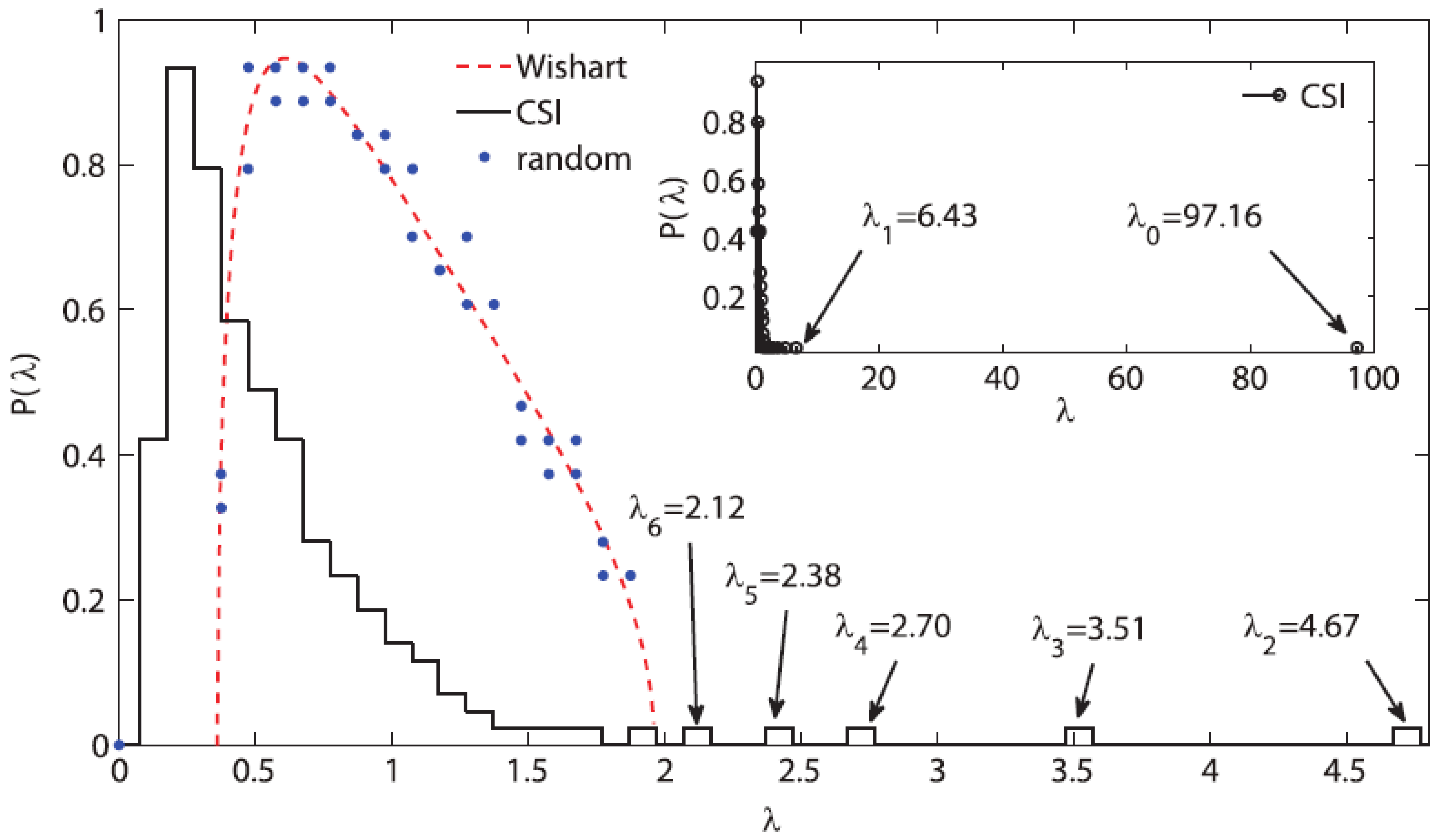
The bulk of the cross-correlation matrix eigenvalue spectrum $P(\lambda)$ is similar to $P_{rm}(\lambda)$, but some large eigenvalues exceed the upper bound of λ_{max}^{ran}, suggesting non-random interactions. The largest eigenvalue λ_0 is associated with the market mode. The other large eigenvalues (λ_i, $\lambda_i > \lambda_{i+1}$, i=1,…) are more related to the dynamics of different sectors of the market.



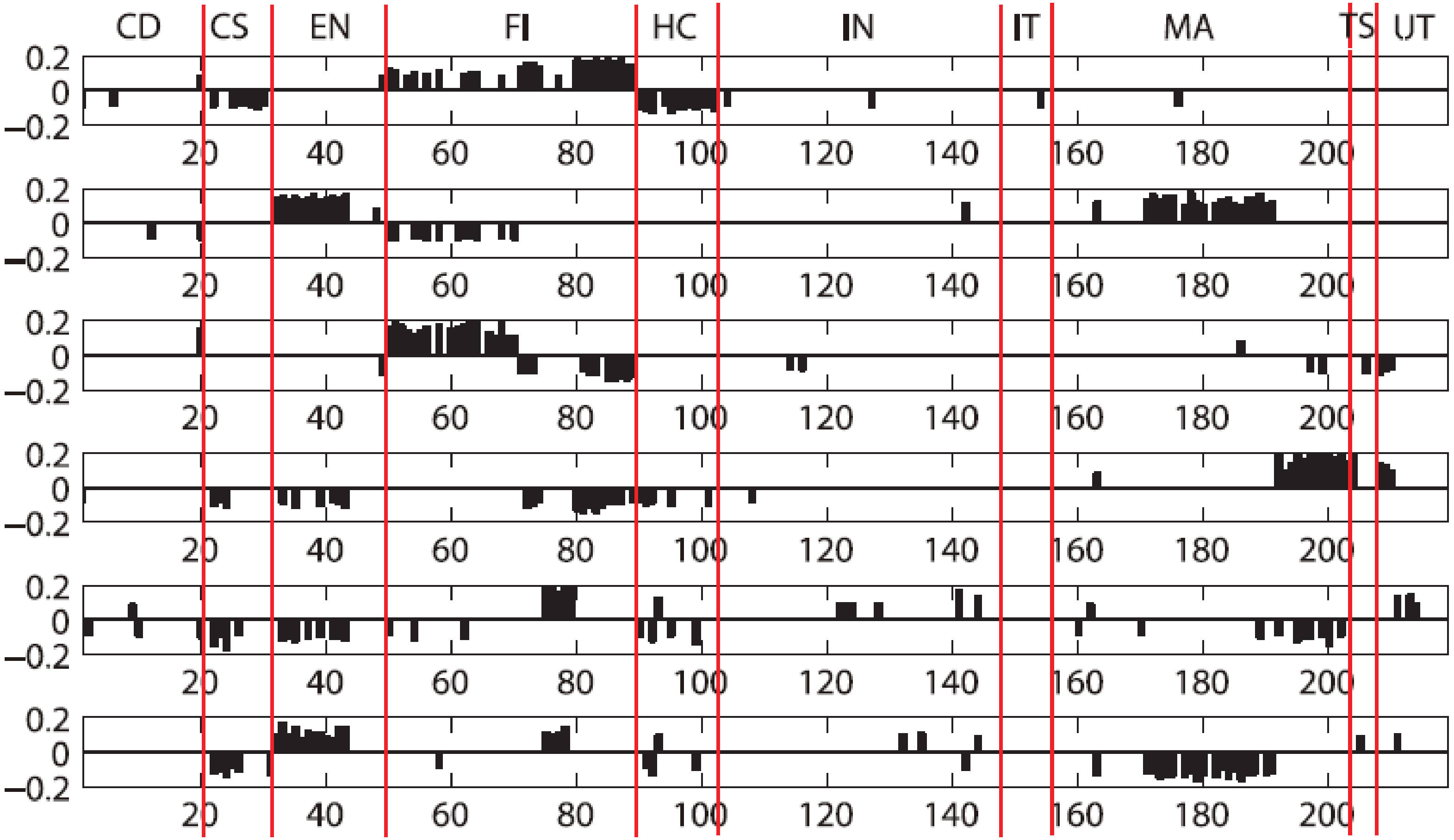
The largest eigenvalue λ_0 is about 97, more than 49 times the maximum value predicted by RMT and also much larger than the values of New York Stock Exchange (NYSE) and Indian National Stock Exchange (NSE), suggesting that the Chinese market is more vulnerable to external stimulation such as global financial crisis and austerity measures taken by the Chinese government.



We use $u_\alpha(i)$ to represent the *i-th* component of the eigenvector corresponding to eigenvalue λ_α . To identify the sector corresponding to a large eigenvalue λ_α, we introduce thresholds $u^+{}_c$ and $u^-{}_c$, and separate the sector into two subsectors, positive subsector with $u_\alpha(i) \geq u^+{}_c$ and negative subsector with $u_\alpha(i) \leq u^-{}_c$. We can study the anti-correlation property of the two subsectors.



As an example, we use $u^{\pm}{}_{c} = \pm u_c$ with $u_c \geq 1/\sqrt{N}$ (we use 0.08 in our analysis here).



Eigenvectors $u_{1,\ldots,6}$ of the six largest eigenvalues $\lambda_{1,\ldots,6}$ of the correlation matrix C. Each stock i in the eigenvector satisfies $|u_\alpha(i) \geq 0.08|$. The codes at the top correspond to the different sectors as listed in Table 1.

Table 2. Subsector structure of the CSI300 for the period from 2007 through 2013

	λ_1		λ_2		λ_3		λ_4	
Sign	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$
Dominant industry	CB&RE	PH&DV	MI	RE	RE	CB	ST	\
$u_c^{\pm} = 0.08$	25/29	22/30	38/40	22/24	19/27	13/23	12/24	\
$u_c^{\pm} = 0.09$	23/24	18/24	36/38	20/21	18/21	12/15	12/20	\
$u_c^{\pm} = 0.10$	19/20	12/15	25/26	16/16	14/15	12/13	12/16	\

** “Sign” here refers to the sign of their components of the corresponding eigenvectors. The denominator represents the total number of stocks within the subsector and the numerator represents the number of stocks of the dominant industry.

The cross-correlation $\boldsymbol{C}$ of the stocks can now be represented as

$$\boldsymbol{C} = \sum_{i=0}^{N-1} \lambda_i \boldsymbol{u}_i \boldsymbol{u}_i^T$$

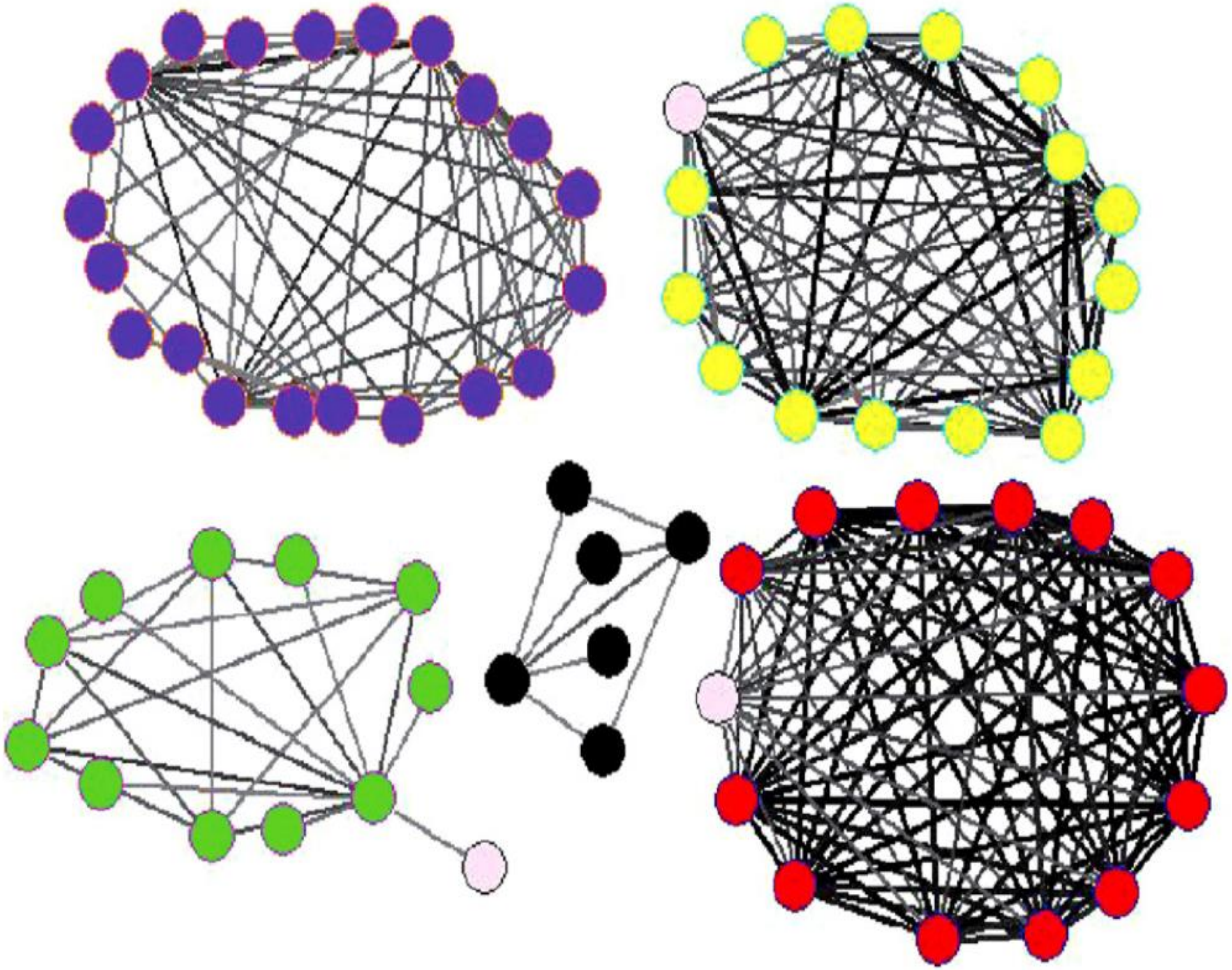
where N is the total number of eigenvalues. One can decompose the correlation matrix into three components: the market component $\boldsymbol{C}^m$, the group component $\boldsymbol{C}^g$ and the random component $\boldsymbol{C}^r$.

$$\boldsymbol{C} = \boldsymbol{C}^m + \boldsymbol{C}^g + \boldsymbol{C}^r = \lambda_0 \boldsymbol{u}_0 \boldsymbol{u}_0^T + \sum_{i=1}^{n} \lambda_i \boldsymbol{u}_i \boldsymbol{u}_i^T + \sum_{i=n+1}^{N-1} \lambda_i \boldsymbol{u}_i \boldsymbol{u}_i^T$$

$\boldsymbol{C}^g \left(= \sum_{i=1}^{n} \lambda_i \boldsymbol{u}_i \boldsymbol{u}_i^T\right)$ can be used to construct an interaction network for the stocks to understand the interaction nature of stocks within sectors and to compare with the results shown in Table 2.

Define A to be the linking matrix of the stock interaction network.

If $C^g{}_{ij} > c_{th}$, where c_{th} is a preset threshold value, then $A_{ij} = C^g{}_{ij}$, i.e., there is a link between stock i and stock j, otherwise, $A_{ij} = 0$.



The interaction network when the threshold c_{th}=0.16. (The yellow, red, green, purple and black nodes denote stocks in the RE, CB, PH&DV, MI, and ST sectors respectively. The pink nodes denote stocks from other sectors.

Dynamical Behavior of CSI300 Sectors

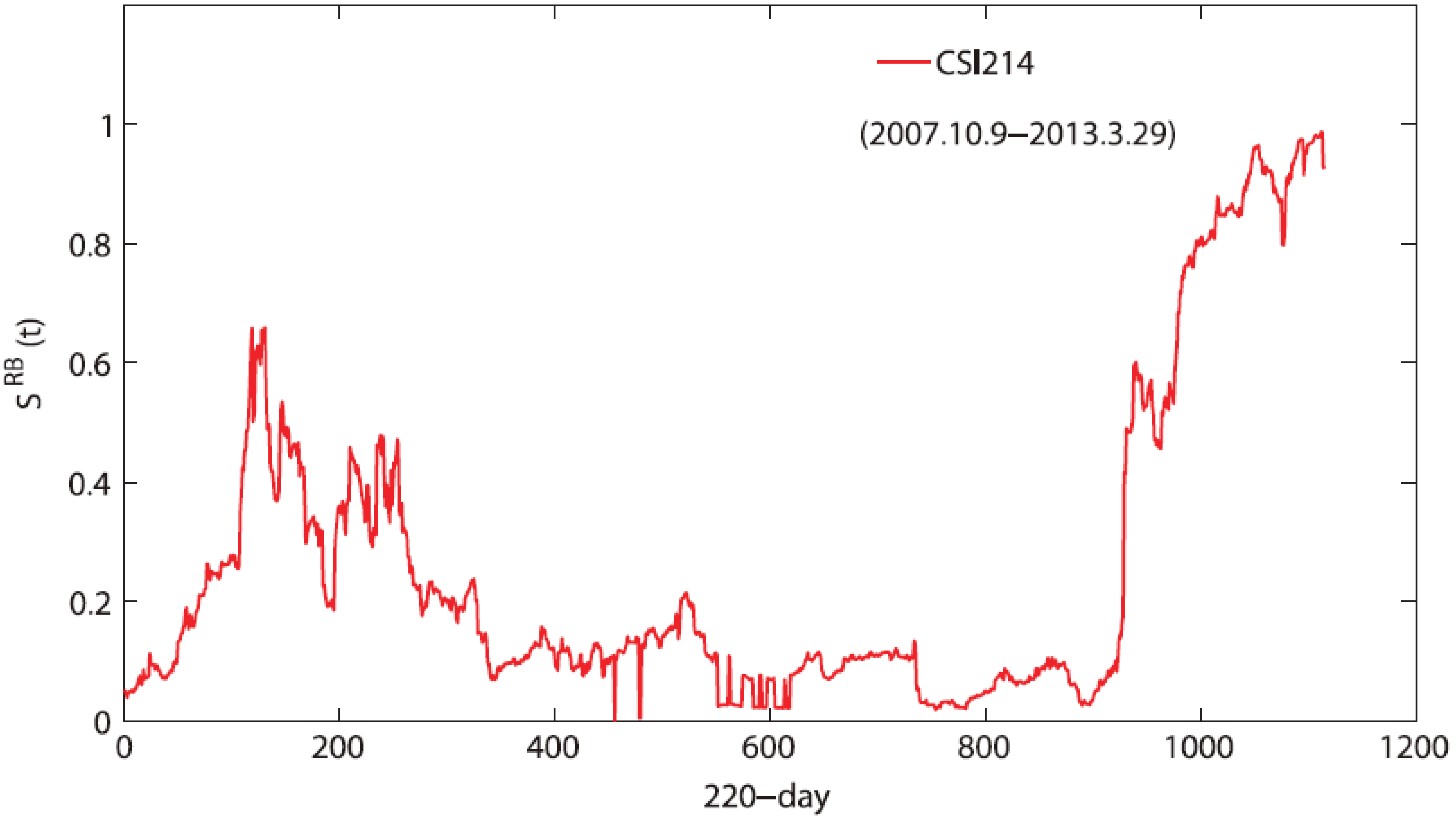
The sector structures of the 214 stocks of the CSI300, in 5 different time periods each having 265 days at $c_{th} = 0.1$.

Time period	n	λ_1		λ_2		λ_3	
		+	−	+	−	+	−
2007.10.9–2008.11.5	3	CB 13/17	PH&DV 7/7	MI 19/19	RE 8/9	\	\
2008.11.6–2009.12.4	3	CB&RE 15/16	PH&DV 10/13	MI 26/27	RE 6/8	RE 16/18	CB 8/16
2009.12.7–2010.12.13	3	CB&RE 13/13	PH&DV 15/24	MI 25/25	RE 13/14	CB 13/19	RE 11/14
2011.1.4–2012.2.14	3	CB&RE 18/19	\	MI 26/27	PH&DV 12/13	\	ST 9/12
2012.2.15–2013.3.18	4	PH&DV 13/15	CB 13/14	\	RE 16/19	\	MI 11/11

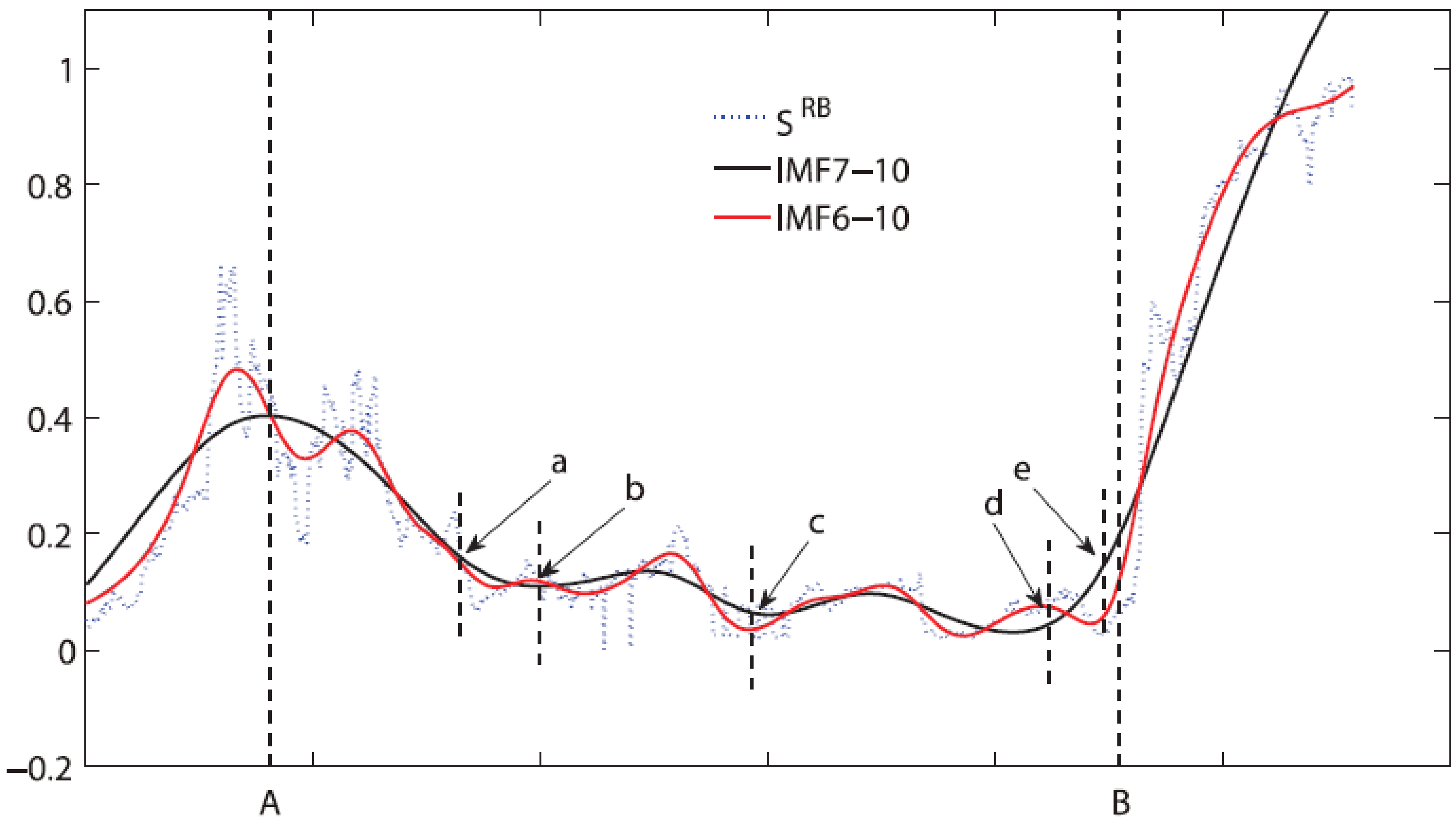
Define S^{RB} to be the anti-correlation function of CB and RE

$$S^{\mathrm{RB}}(t) = \frac{\sum_{\alpha=1}^{n} \sum_{i \in \mathrm{RE}, j \in \mathrm{CB}} \lambda_\alpha u_{i\alpha} u_{j\alpha} (\mathrm{sign}(u_{i\alpha} u_{j\alpha}) - 1)}{2 \sum_{\alpha=1}^{n} \sum_{i \in \mathrm{RE}, j \in \mathrm{CB}} \lambda_\alpha \left| u_{i\alpha} u_{j\alpha} \right|}$$

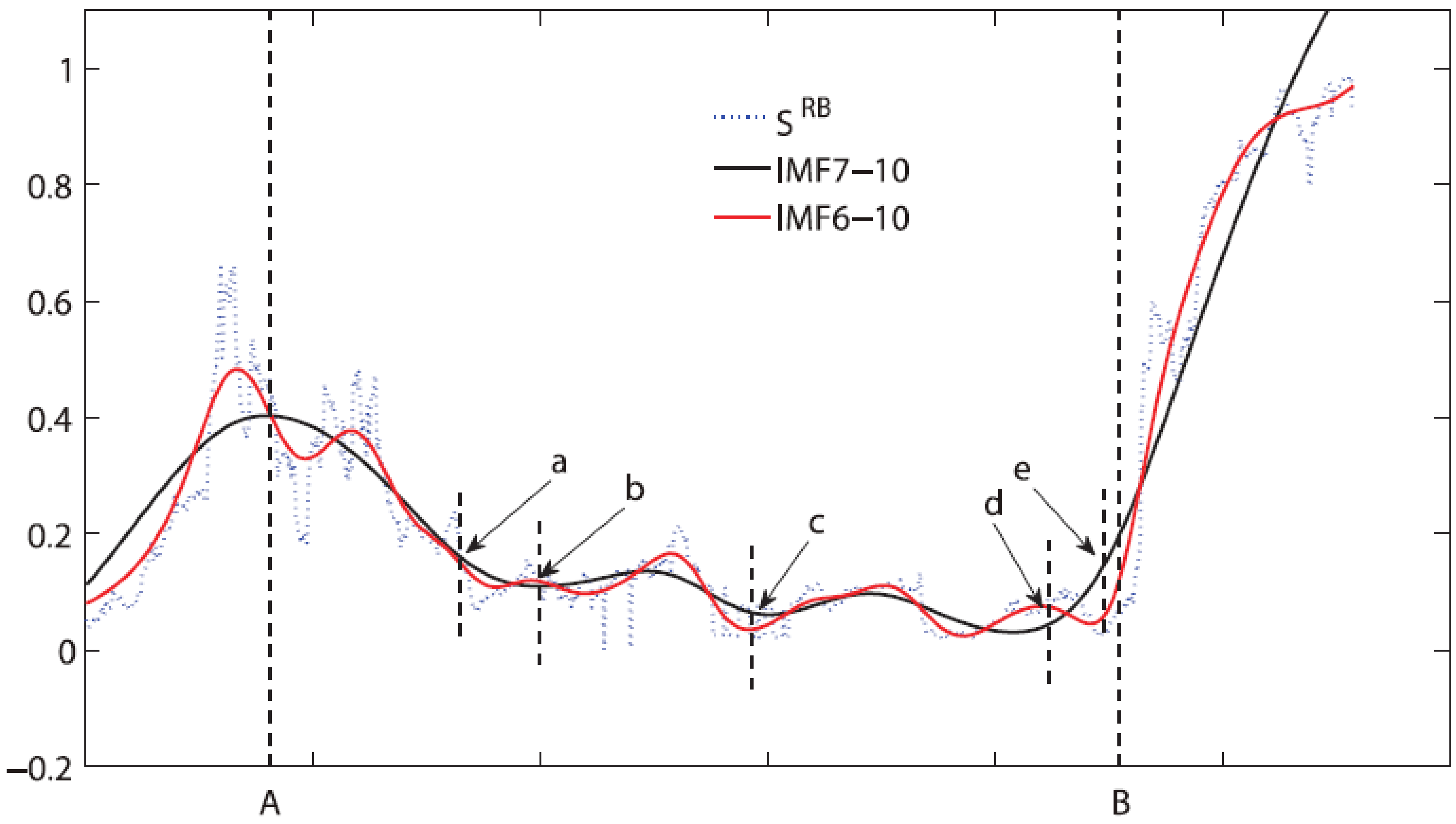
where $i \in$ RE and $j \in$ CB are the stocks corresponding to RE and CB sectors respectively, $u_{i\alpha}$ and $u_{j\alpha}$ are the eigenvector components of RE stock i and CB stock j of the eigenvalue λ_α respectively.



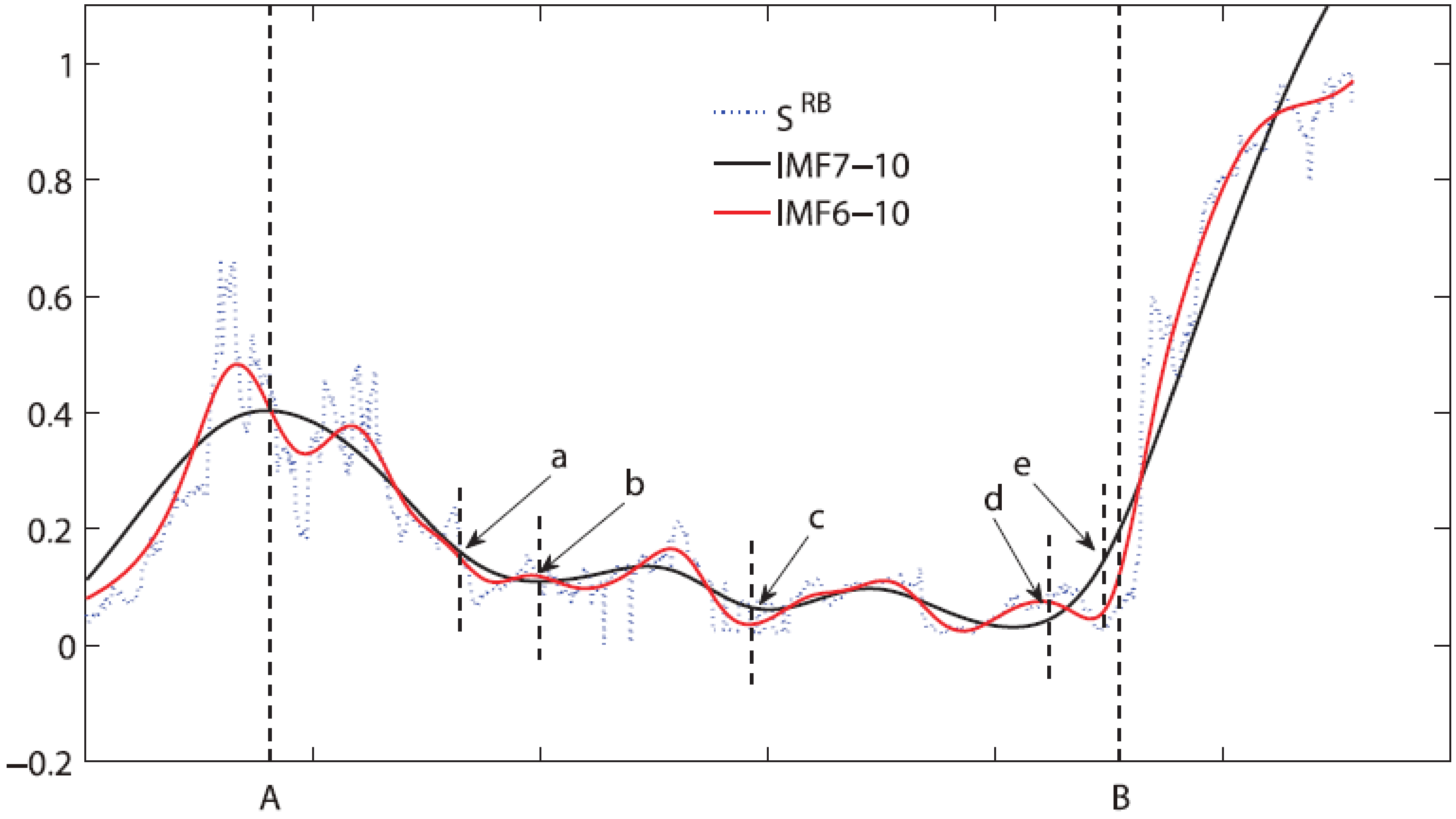
The anti-correlation fluctuations of the real estate and banking sectors in the CSI300 for the period between October 2007 and March 2013 with a moving window of size 220 days.



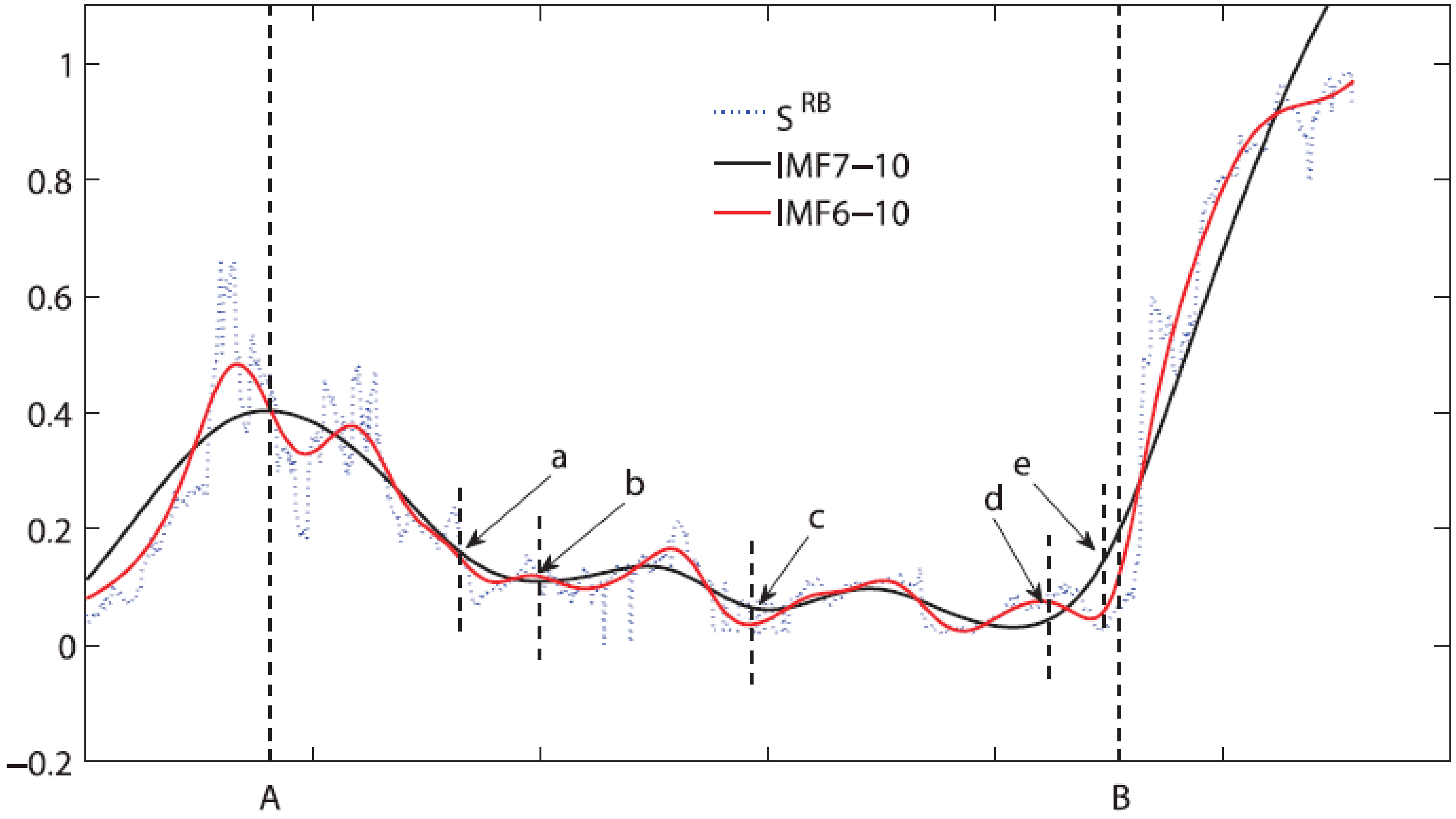
EEMD decomposition of anti-correlation fluctuations of the real estate and banking sectors in the CSI300 for the period corresponding to the above figure.



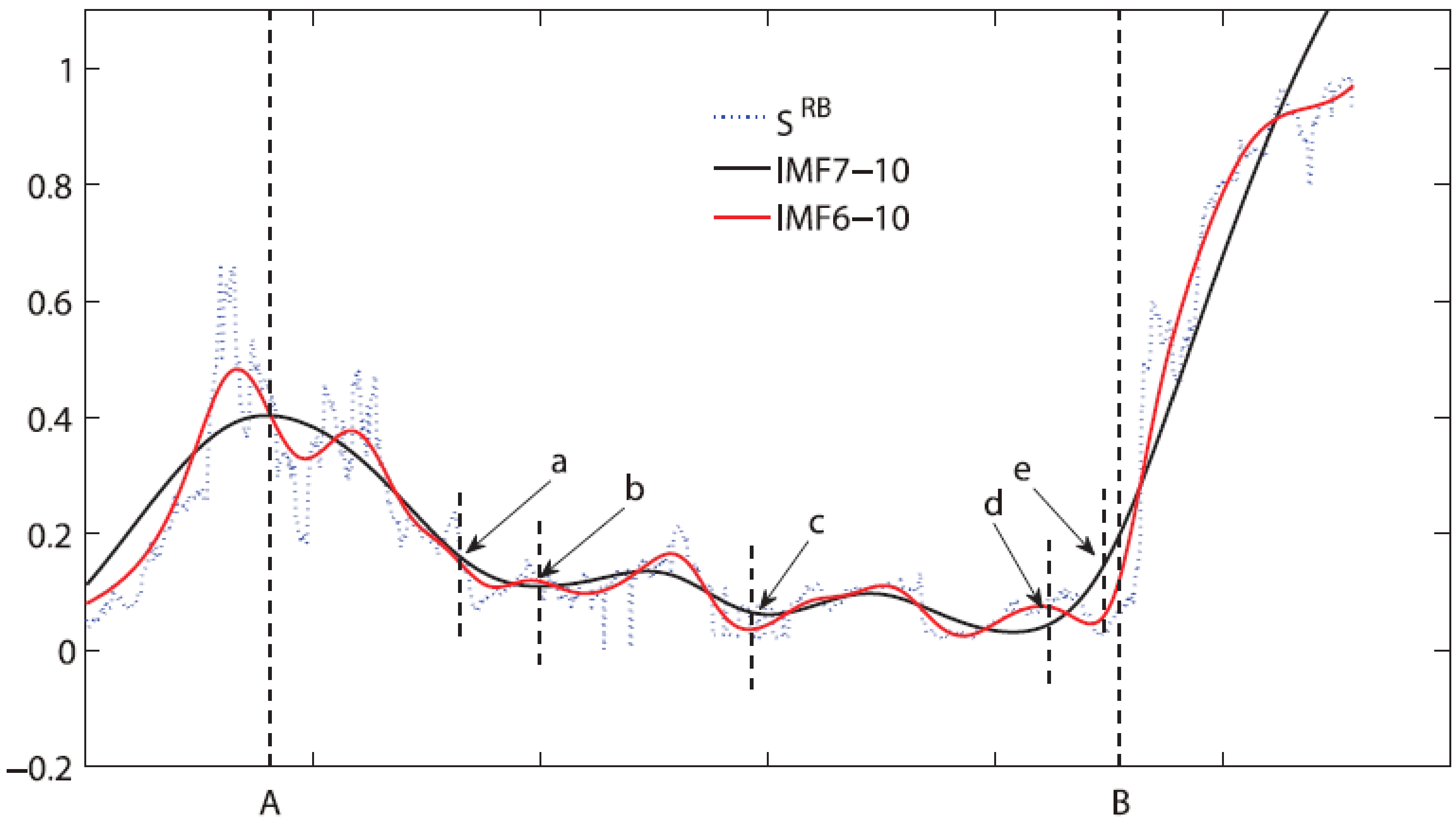
Point A corresponds to the 220-day period from June 3, 2008 till April 28, 2009 while point B corresponds to the 220-day period from June 30, 2011 till May 29, 2012.



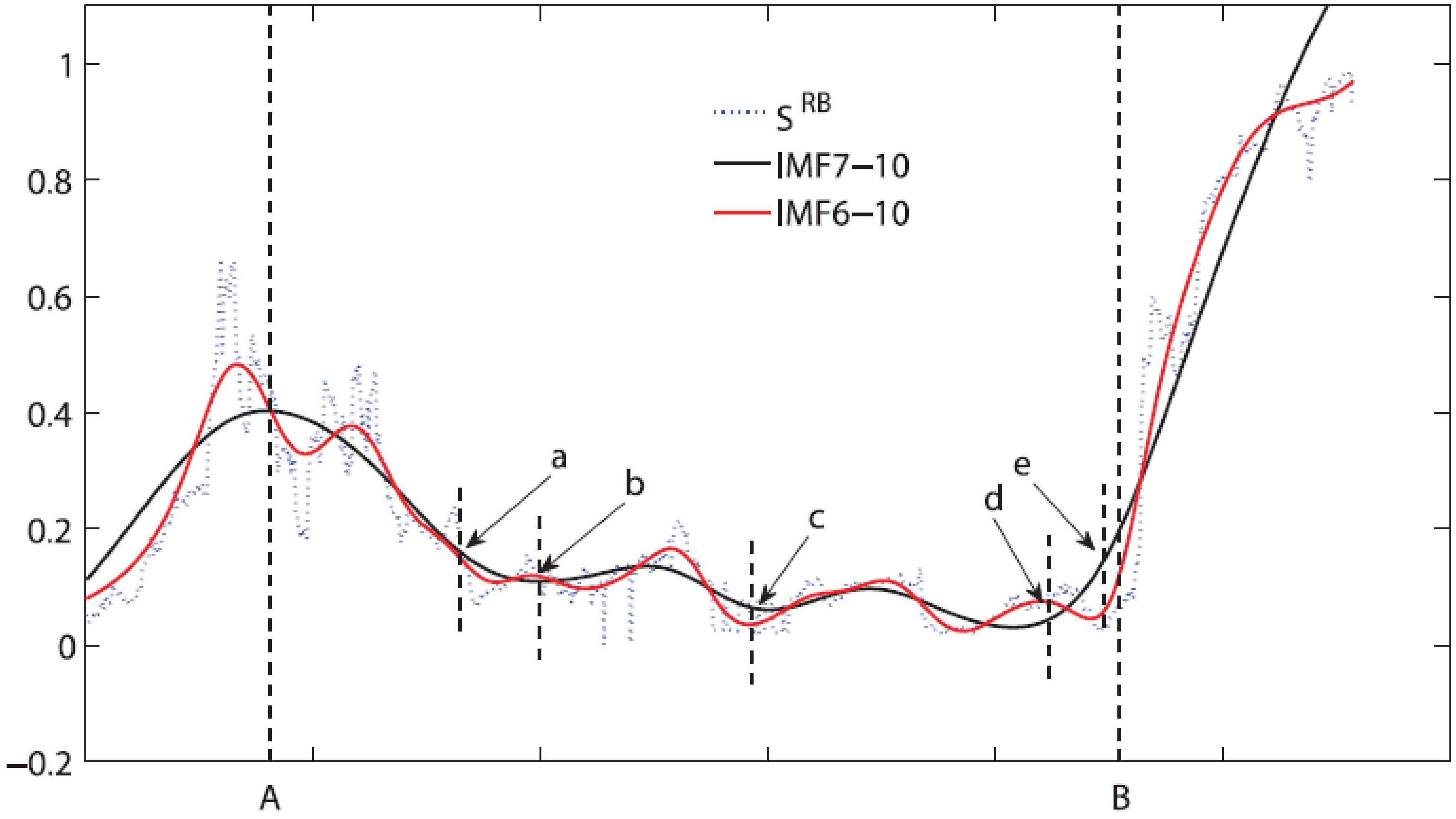
In May, 2009 the Chinese government decided to impose the austerity policy in order to cool down investments in real estate (point A). By the end of 2009, the Chinese government lifted all the previous real estate investment measures (point a).



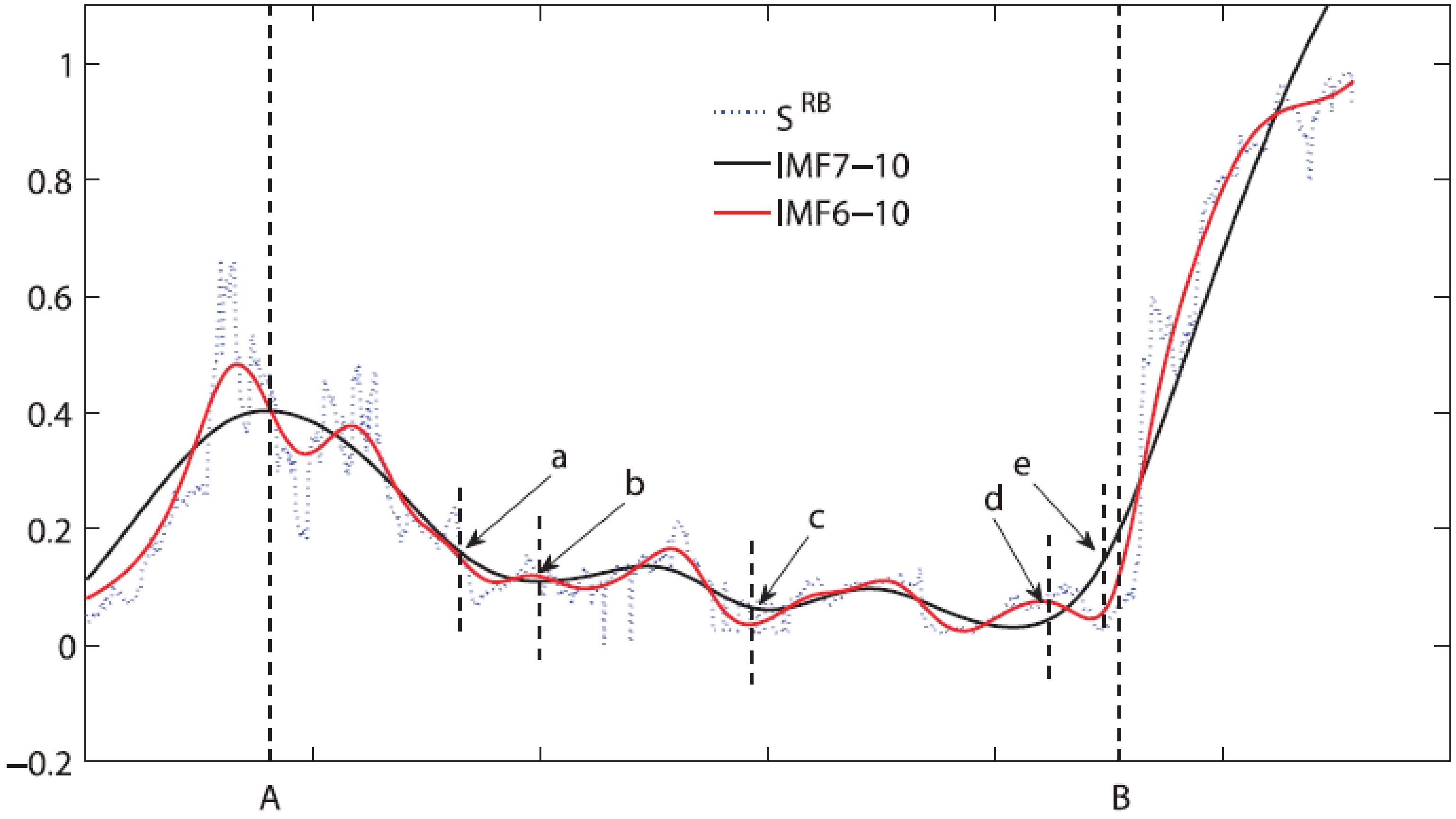
On April 17, 2010, the Chinese government announced the Ten New Measures on investments in real estate (point b).



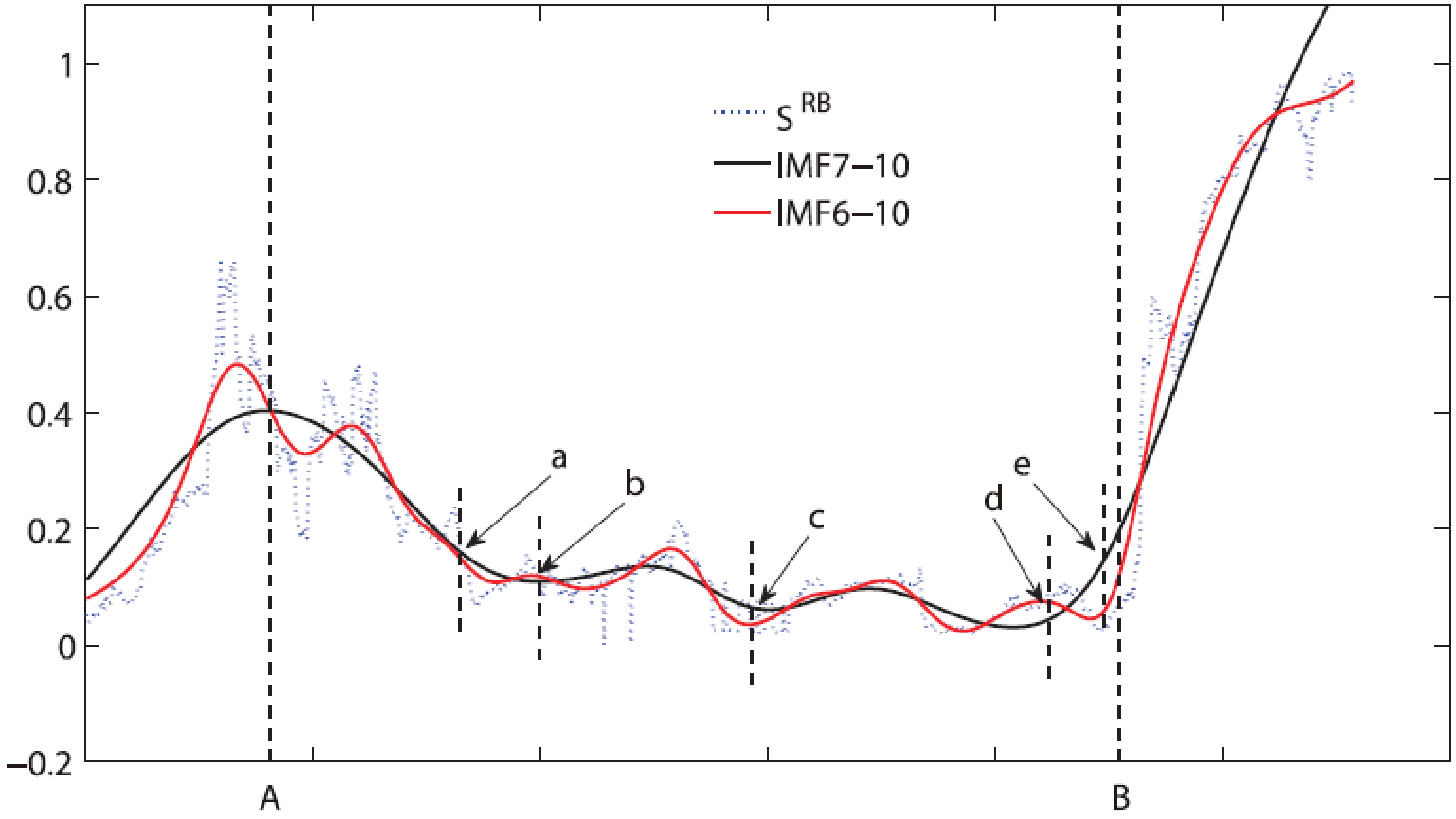
The government added 8 more new measures on January 26, 2011 (point c) to further cool down investments in real estate. As a result of the series of measures, the anti-correlation between the real estate and commercial bank sectors weakened (from point A through point c).



In order to guarantee the targeted economic growth, the Chinese government carried out the quantitative easing policy to stimulate the market in February, 2012. On February 24, 2012 the Central Bank of China lowered the required deposit reserve ratio (point d).



The required deposit reserve ratio was then lowered again on May 18, 2012 by 0.5% (point e) and on June 8, 2012 by 0.25% (point B).



This series of measures taken by the Central Bank helped the commercial banks in their bank credit business because of the ample capital flow in the market. Many enterprises got out of their financial predicaments and subsequently helped the economic growth of the nation. The result is the drastic rise in the stock prices of the CB sector. On the other hand, the performance of the stocks in the real estate sector is still hampered by the measures taken earlier by the government, resulted in a substantial increase of the anti-correlation behavior between the two sectors RE and CB starting from point B.

Example Two:
Time Series Analysis on Cryptocurrency Market

Overview

- Blockchain and cryptocurrency are popular now, with a market cap of around $1.1 trillion dollars
- Bitcoin is the biggest cryptocurrency with over $550 billion market cap and has over $18 billion 24-hour trading volume in Binance
- Recent studies suggest that the cryptocurrency market is more and more correlated with traditional financial markets
- Ethereum implements the merge, changes its consensus mechanism from Proof-of-Work to Proof-of-Stake

Objectives

To study the correlations between the cryptocurrency market and the traditional financial market

To observe the impact of the merge on the cryptocurrency market

To use the models to forecast future cryptocurrency prices

Data (Yahoo Finance Website)

- BTC: Bitcoin
- ETH: Ethereum
- LTC: Litecoin
- XRP: XRP (Ripple)
- NASDAQ: Nasdaq Composite Index

Time period: 2017-11-09 – 2023-03-23

Time Series Models Used

ARMA
Autoregressive Moving Average Model

ARIMA
Autoregressive Integrated Moving Average Model

VARMA
Vector Autoregressive Moving Average Model

LSTM
Long Short-Term Memory Network Model

Denote the value of the market variable at the end of day i to be S_i. Define the daily return by using the popular log-return as the daily return of the market variable. The variable u_i is defined as the return during day i given by,

$$u_i = \ln\left(\frac{S_i}{S_{i-1}}\right)$$

In the ARMA(p, q) model

$$x_t = a_0 + \sum_{i=1}^{p} a_i\, x_{t-i} + \sum_{j=1}^{q} \beta_j \varepsilon_{t-j} + \varepsilon_t$$
$$= a_0 + a_1 x_{t-1} + \cdots + a_p x_{t-p} + \beta_1 \varepsilon_{t-1} + \cdots + \beta_q \varepsilon_{t-q} + \varepsilon_t$$

ARMA(p, 0) = AR(p) , Autoregressive Model of order p.
ARMA(0, q) = MA(q) , Moving Average Model of order q.

For example, AR(1) with $a_0 = 0;\ a_1 = 1$ takes the form, $x_t = x_{t-1} + \varepsilon_t$, which corresponds to a random walk. By differencing x_t, i.e. $x_t - x_{t-1} = \nabla x_t$, we get $\nabla x_t = \varepsilon_t$, which is a stationary time series. One can perform differencing to get an ARMA process, known as the integrated ARMA or ARIMA. A process x_t is then said to be ARIMA(p, d, q) if

$$\nabla^d x_t = (1 - B)^d x_t$$

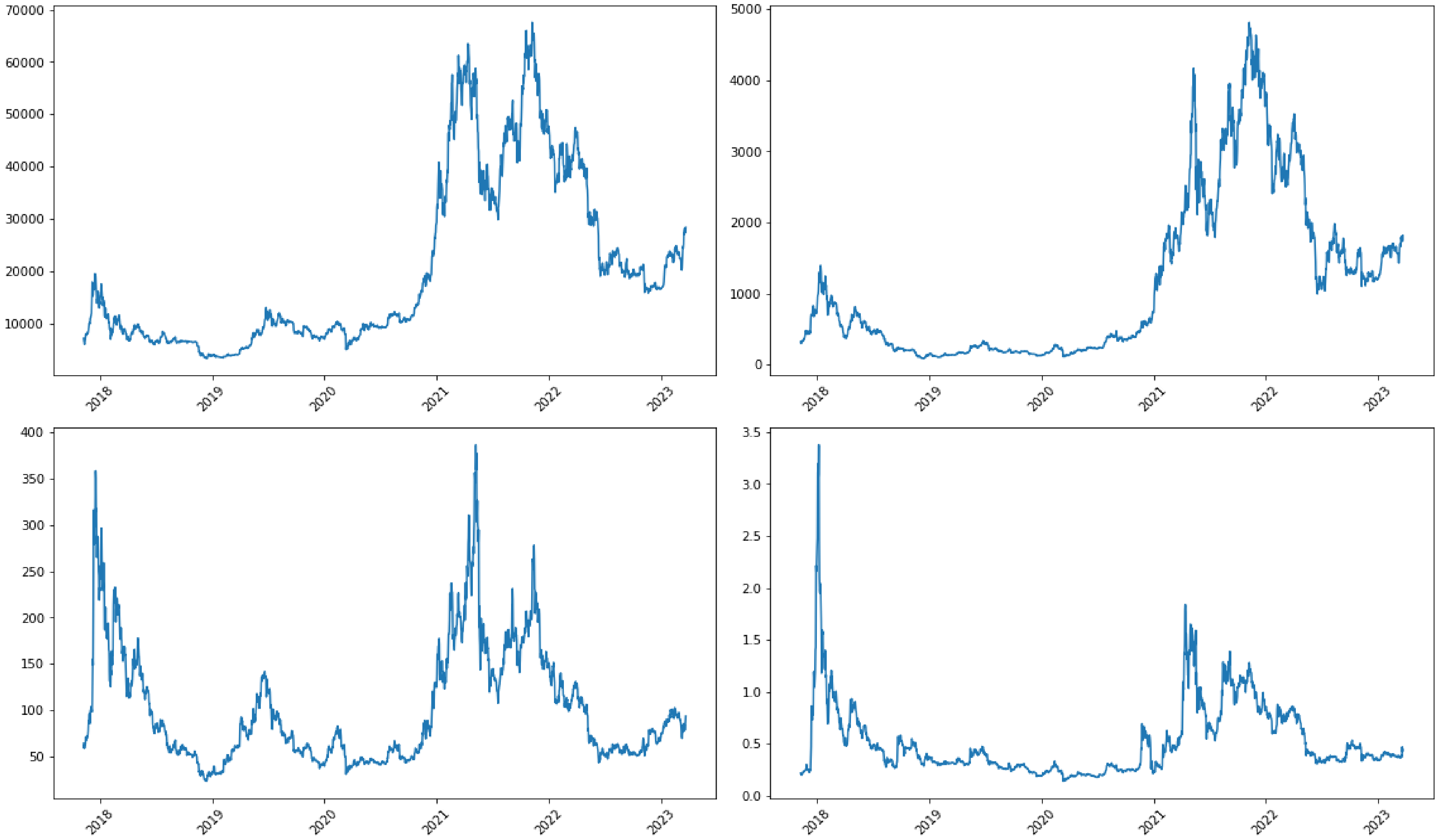
is ARMA(p, q). B here refers to the lag operator. For example, $B x_t = x_{t-1}$, $B^2 x_t = x_{t-2}$, etc.

VARMA models generalize **ARMA** models and take the form

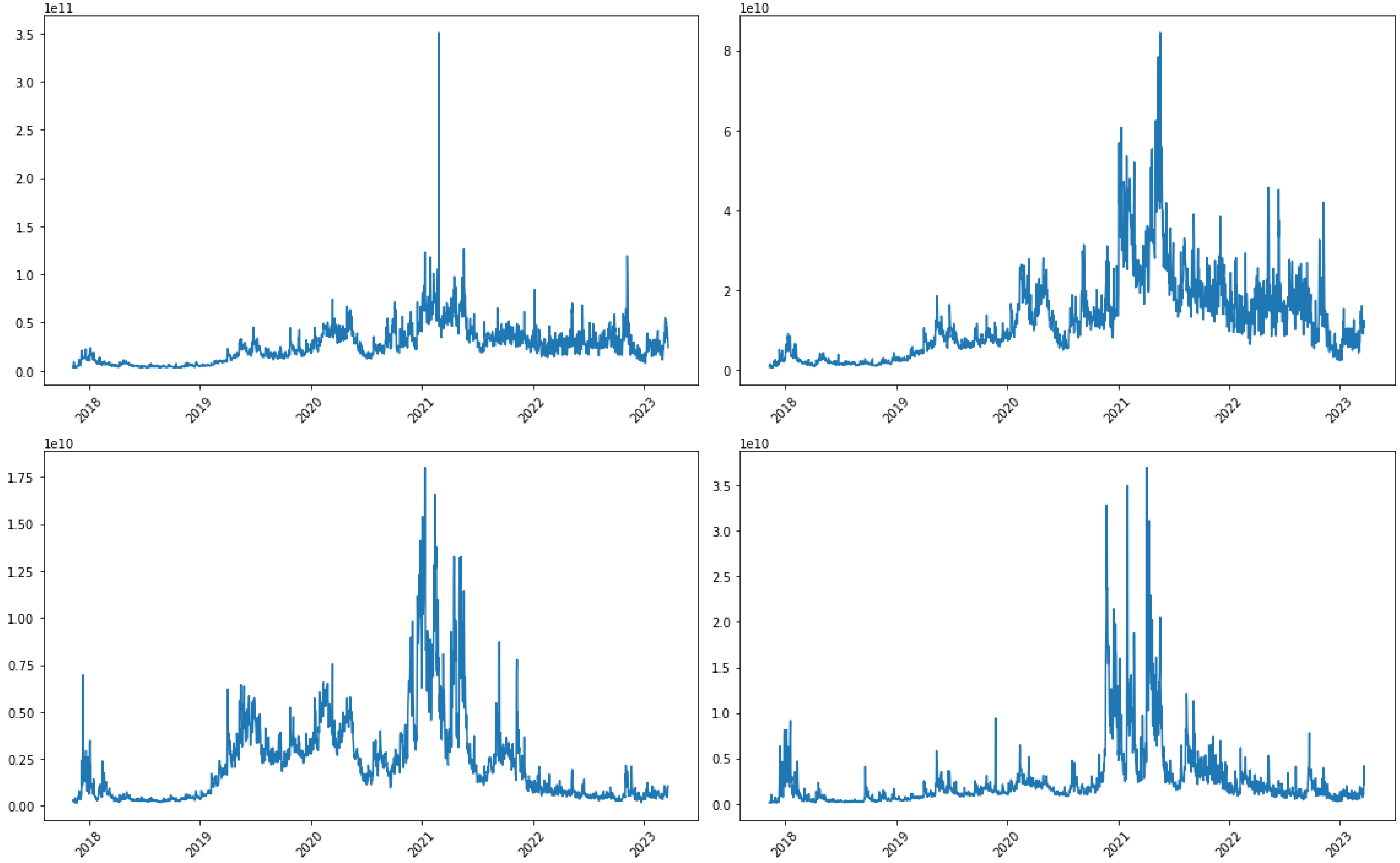
$$\boldsymbol{x_t} = \boldsymbol{\Phi_0} + \sum_{\boldsymbol{j=1}}^{\boldsymbol{p}} \boldsymbol{\Phi_j}\, \boldsymbol{x_{t-j}} + \sum_{\boldsymbol{k=1}}^{\boldsymbol{q}} \boldsymbol{\Theta_k} \boldsymbol{\varepsilon_{t-k}} + \boldsymbol{\varepsilon_t} \,.$$

where the $\boldsymbol{\Phi_0} \in \boldsymbol{R^N}$ is an $N \times 1$ vector, $\boldsymbol{\Phi_j} \in \boldsymbol{R^{N \times N}}$ is an $N \times N$ matrix, $\boldsymbol{\varepsilon_t}$ is a white noise series with zero mean and constant covariance matrix S_e .
For example, VARMA(1,0) with two time random variables takes the form

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \Phi_{x,0} \\ \Phi_{y,0} \end{bmatrix} + \begin{bmatrix} \Phi_{xx,1} & \Phi_{xy,1} \\ \Phi_{yx,1} & \Phi_{yy,1} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$$

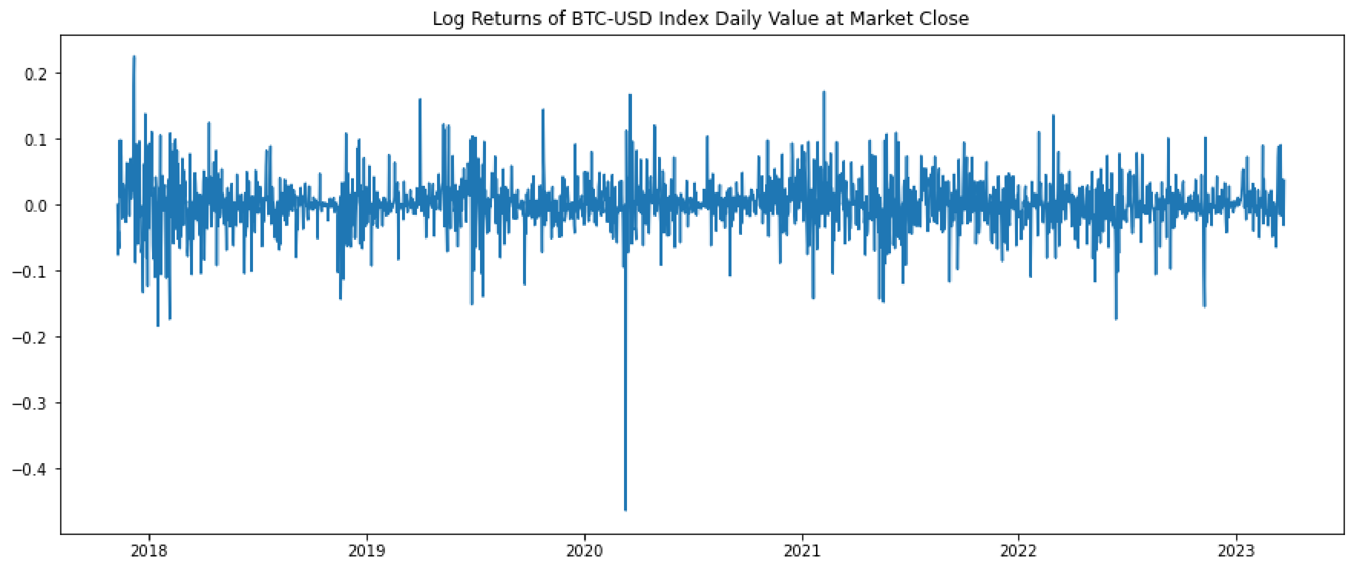


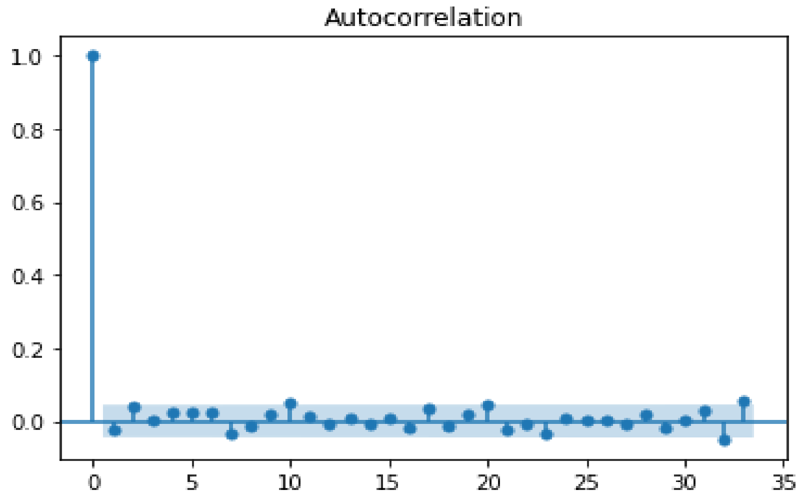
Daily closing prices of the four cryptocurrency in our study. Upper left: BTC; Upper right: ETH; Lower left: LTC; Lower right: XRP.



Daily trading volumes of the four cryptocurrency in our study. Upper left: BTC; Upper right: ETH; Lower left: LTC; Lower right: XRP.

Upper: BTC daily Log return values; Lower: Autocorrelation function.





Result of ADF test for Bitcoin:

ADF Statistic: -30.459909990493856 p-value: 0.0 used lag: 1

critical values: {'1%': -3.433692465214729, '5%': -2.8630165042063163, '10%': -2.567556030574193}

VARMAX(2,0) for the four cryptocurrency time series (2017-11-09 – 2023-03-18)

$$\begin{bmatrix} x_{1,t} \\ x_{2,t} \\ x_{3,t} \\ x_{4,t} \end{bmatrix} = \begin{bmatrix} -0.000026 \\ -0.000036 \\ -0.000034 \\ -0.000037 \end{bmatrix} + \begin{bmatrix} 0.0943 & -0.0644 & -0.0122 & -0.0433 \\ -0.0473 & 0.0168 & 0.0418 & -0.0770 \\ 0.0408 & -0.1290 & 0.0958 & -0.0622 \\ -0.0644 & -0.0074 & -0.0474 & 0.0588 \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \\ x_{3,t-1} \\ x_{4,t-1} \end{bmatrix} + \begin{bmatrix} -0.0140 & 0.0497 & -0.0161 & 0.0174 \\ 0.0213 & 0.0465 & -0.0038 & -0.0025 \\ 0.1176 & -0.0316 & -0.0037 & -0.0306 \\ 0.0063 & -0.0543 & 0.1399 & -0.0199 \end{bmatrix} \begin{bmatrix} x_{1,t-2} \\ x_{2,t-2} \\ x_{3,t-2} \\ x_{4,t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \\ \varepsilon_{4,t} \end{bmatrix}$$

VARMAX(2,0) for the four cryptocurrency plus NASDAQ index time series (2017-11-09 – 2023-03-18)

$$\begin{bmatrix} x_{1,t} \\ x_{2,t} \\ x_{3,t} \\ x_{4,t} \\ x_{5,t} \end{bmatrix} = \begin{bmatrix} -0.000031 \\ -0.000040 \\ -0.000038 \\ -0.000040 \\ -0.000010 \end{bmatrix} + \begin{bmatrix} 0.0985 & -0.0605 & -0.0122 & -0.0434 & -0.0822 \\ -0.0435 & 0.0189 & 0.0408 & -0.0762 & -0.0502 \\ 0.0426 & -0.1290 & 0.0946 & -0.0611 & -0.0071 \\ -0.0554 & 0.0008 & -0.0475 & 0.0587 & -0.1720 \\ 0.0014 & 0.0092 & -0.0137 & 0.0034 & -0.1330 \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \\ x_{3,t-1} \\ x_{4,t-1} \\ x_{5,t-1} \end{bmatrix} + \begin{bmatrix} -0.0153 & 0.0494 & -0.0172 & 0.0181 & 0.0124 \\ 0.0166 & 0.0424 & -0.0044 & -0.0020 & 0.0884 \\ 0.1126 & -0.0364 & -0.0036 & -0.0305 & 0.0997 \\ 0.0032 & -0.0553 & 0.1376 & -0.0182 & 0.0369 \\ 0.0027 & -0.0079 & 0.0023 & 0.0030 & 0.0068 \end{bmatrix} \begin{bmatrix} x_{1,t-2} \\ x_{2,t-2} \\ x_{3,t-2} \\ x_{4,t-2} \\ x_{5,t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \\ \varepsilon_{4,t} \\ \varepsilon_{5,t} \end{bmatrix}$$

Cryptocurrency Daily Closing Prices

Before the merge (2022-03-16 to 2022-09-14):

$$\begin{bmatrix} x_{1,t} \\ x_{2,t} \\ x_{3,t} \\ x_{4,t} \end{bmatrix} = \begin{bmatrix} -0.0048 \\ -0.0040 \\ -0.0041 \\ -0.0054 \end{bmatrix} + \begin{bmatrix} -0.0859 & 0.1755 & -0.1121 & -0.0499 \\ -0.2767 & 0.1751 & -0.0002 & -0.0195 \\ -0.3179 & 0.0859 & -0.1035 & 0.1320 \\ -0.1654 & 0.1593 & -0.0983 & -0.0429 \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \\ x_{3,t-1} \\ x_{4,t-1} \end{bmatrix} + \begin{bmatrix} -0.2770 & 0.1737 & 0.1702 & -0.0366 \\ -0.1928 & 0.1323 & 0.1673 & -0.0283 \\ -0.2300 & 0.0418 & 0.1562 & 0.0419 \\ -0.1054 & -0.0485 & 0.2555 & -0.0602 \end{bmatrix} \begin{bmatrix} x_{1,t-2} \\ x_{2,t-2} \\ x_{3,t-2} \\ x_{4,t-2} \end{bmatrix} + \boldsymbol{\varepsilon_t}$$

After the merge (2022-09-15 to 2023-03-16):

$$\begin{bmatrix} x_{1,t} \\ x_{2,t} \\ x_{3,t} \\ x_{4,t} \end{bmatrix} = \begin{bmatrix} 0.0005 \\ -0.0009 \\ 0.0008 \\ -0.0011 \end{bmatrix} + \begin{bmatrix} 0.2967 & -0.1195 & 0.1196 & -0.1767 \\ 0.1993 & -0.0605 & 0.1552 & -0.2727 \\ 0.1924 & -0.2779 & 0.2839 & -0.1823 \\ 0.3686 & -0.4435 & 0.0870 & -0.0983 \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \\ x_{3,t-1} \\ x_{4,t-1} \end{bmatrix} + \begin{bmatrix} 0.2814 & -0.1413 & -0.1438 & -0.0543 \\ 0.1560 & -0.2232 & -0.0328 & -0.0945 \\ 0.0482 & -0.2556 & -0.0175 & -0.0907 \\ 0.4812 & -0.6944 & 0.1625 & -0.0518 \end{bmatrix} \begin{bmatrix} x_{1,t-2} \\ x_{2,t-2} \\ x_{3,t-2} \\ x_{4,t-2} \end{bmatrix} + \boldsymbol{\varepsilon_t}$$

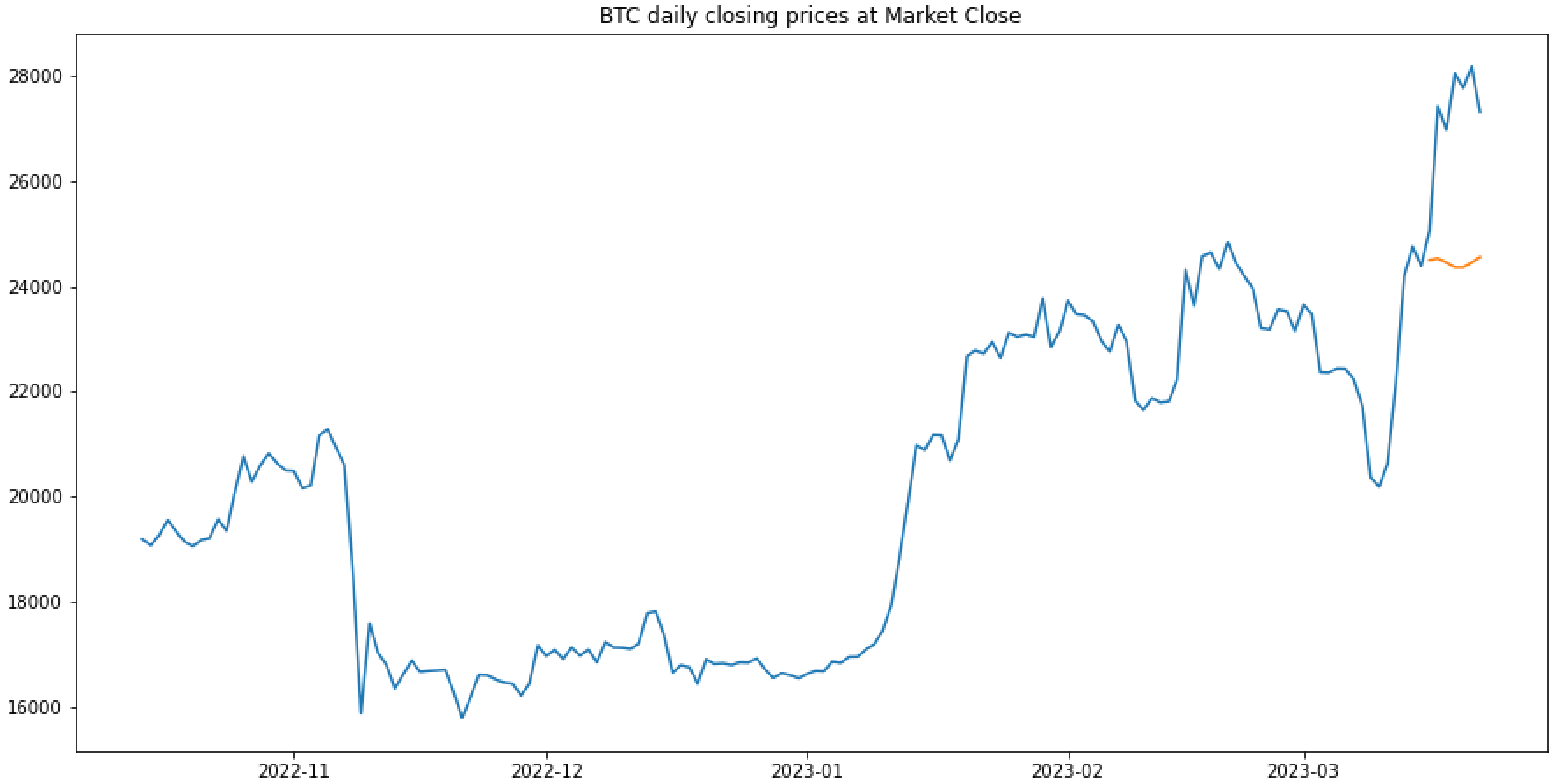
Cryptocurrency Daily Trading Volume

Before the merge (2022-03-16 to 2022-09-14):

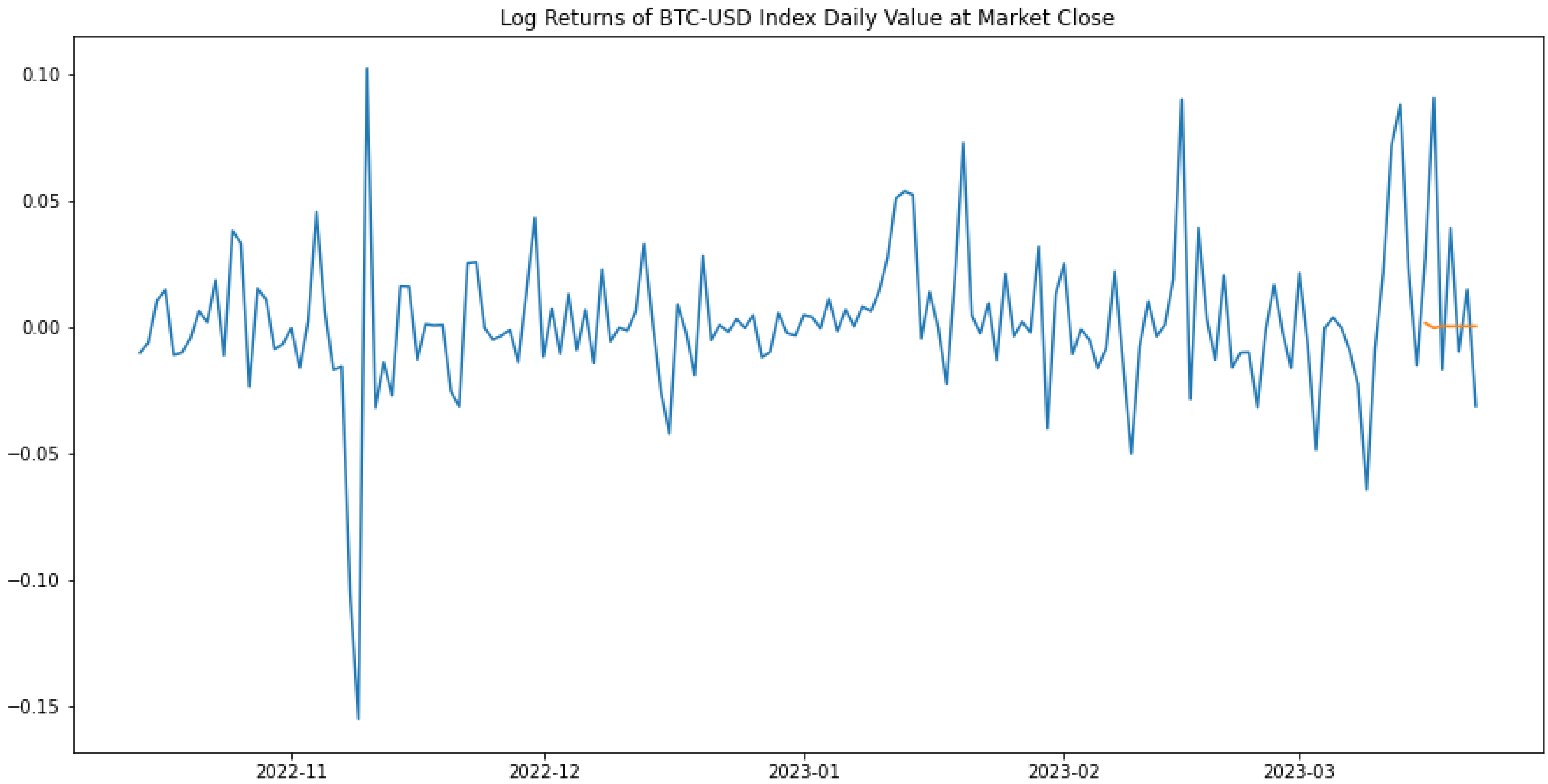
$$\begin{bmatrix} v_{1,t} \\ v_{2,t} \\ v_{3,t} \\ v_{4,t} \end{bmatrix} = \begin{bmatrix} 0.0020 \\ 0.0007 \\ -0.0018 \\ -0.0056 \end{bmatrix} + \begin{bmatrix} -0.3229 & 0.0545 & -0.0951 & 0.0857 \\ -0.2054 & -0.0243 & -0.1074 & 0.1220 \\ 0.0701 & 0.0148 & -0.4461 & 0.1059 \\ 0.0474 & 0.0947 & 0.0368 & -0.2857 \end{bmatrix} \begin{bmatrix} v_{1,t-1} \\ v_{2,t-1} \\ v_{3,t-1} \\ v_{4,t-1} \end{bmatrix} + \begin{bmatrix} -0.2260 & -0.0746 & 0.0794 & -0.0449 \\ -0.0058 & -0.3210 & -0.0216 & 0.0209 \\ -0.0175 & 0.0578 & -0.2082 & -0.0373 \\ 0.0370 & -0.1655 & 0.0926 & -0.2833 \end{bmatrix} \begin{bmatrix} v_{1,t-2} \\ v_{2,t-2} \\ v_{3,t-2} \\ v_{4,t-2} \end{bmatrix} + \boldsymbol{\varepsilon_t}$$

After the merge (2022-09-15 to 2023-03-16):

$$\begin{bmatrix} v_{1,t} \\ v_{2,t} \\ v_{3,t} \\ v_{4,t} \end{bmatrix} = \begin{bmatrix} -0.0021 \\ -0.0078 \\ -0.0023 \\ -0.0063 \end{bmatrix} + \begin{bmatrix} 0.1178 & -0.2123 & -0.1377 & 0.0798 \\ 0.4694 & -0.5752 & -0.1504 & 0.0492 \\ 0.4271 & -0.3597 & -0.2960 & 0.0259 \\ 0.4198 & -0.2579 & -0.1405 & -0.1981 \end{bmatrix} \begin{bmatrix} v_{1,t-1} \\ v_{2,t-1} \\ v_{3,t-1} \\ v_{4,t-1} \end{bmatrix} + \begin{bmatrix} -0.2821 & 0.0656 & 0.0796 & -0.0742 \\ 0.0417 & -0.1567 & 0.0531 & -0.0739 \\ 0.0685 & -0.0869 & -0.0631 & -0.0921 \\ -0.3063 & 0.0104 & 0.4021 & -0.3852 \end{bmatrix} \begin{bmatrix} v_{1,t-2} \\ v_{2,t-2} \\ v_{3,t-2} \\ v_{4,t-2} \end{bmatrix} + \boldsymbol{\varepsilon_t}$$

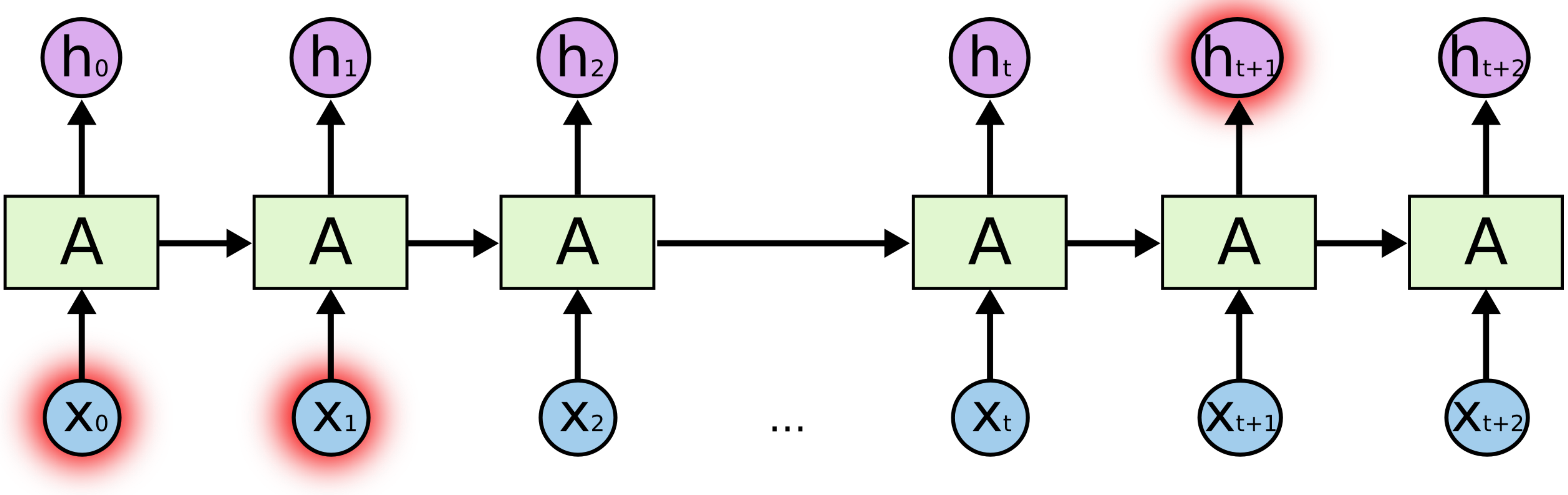


Prediction of the daily closing prices of Bitcoin by using the ARIMA model. The blue curve shows the Bitcoin daily closing prices from 2022-10-14 to 2023-03-23. The orange line is the prediction of the last 5 days (2023-03-19 – 2023-03-23) of the dataset by using the previous daily closing price values from 2017-11-09 to 2023-03-18.

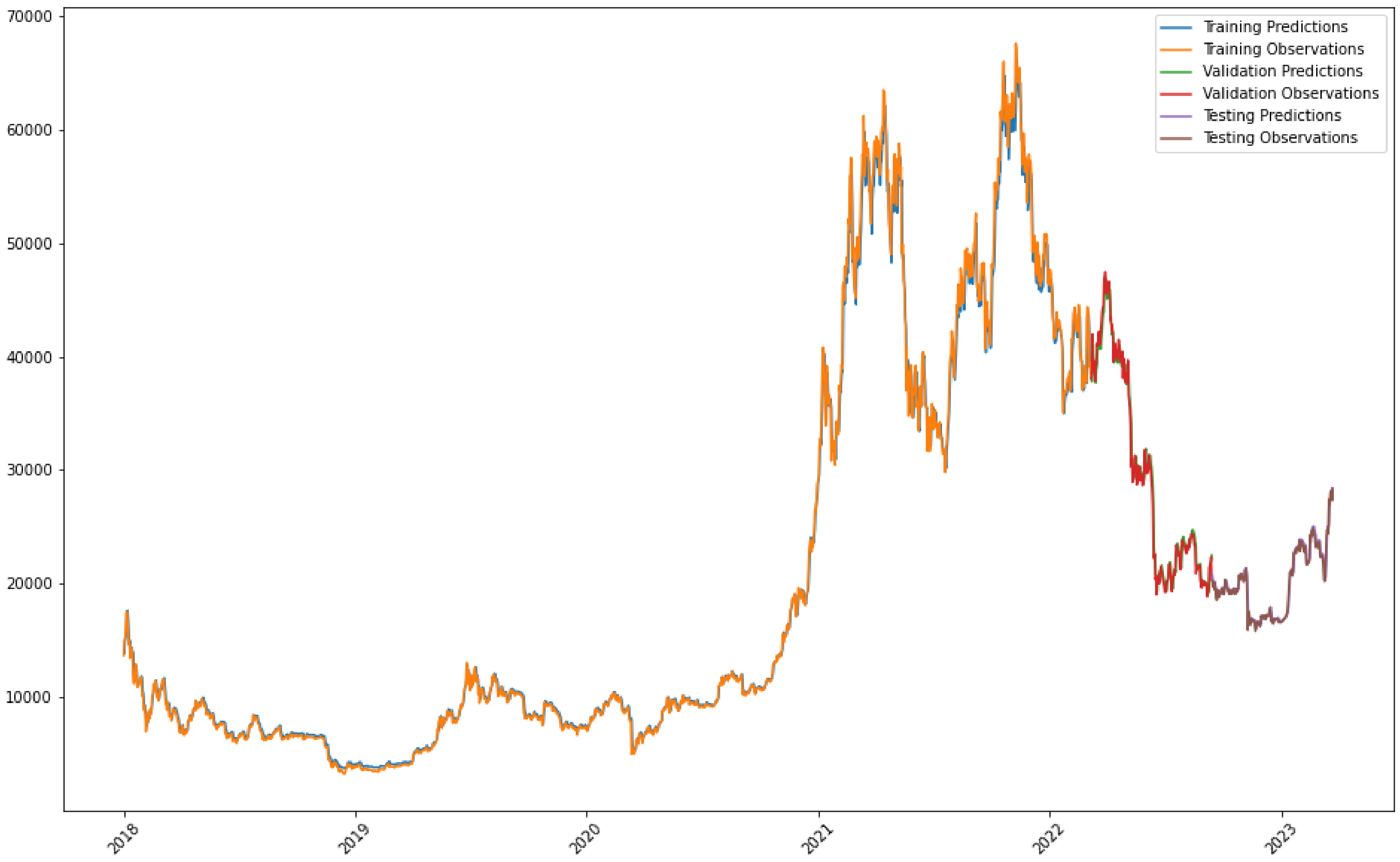


Prediction of the daily Log return of Bitcoin by using the ARIMA model. The blue curve shows the Bitcoin daily Log return from 2023-10-14 to 2023-03-23. The orange line is the prediction of the last 5 days (2023-03-19 – 2023-03-23) of the dataset by using the previous daily Log return values from 2017-11-09 to 2023-03-18.

A Long short-term memory (LSTM) Network



Bitcoin daily closing prices. (a) The dataset is split into 3 subsets, namely training, validation and testing sets; (b) Predictions and observations of the training set; (c) Predictions and observations of the validation set; (d) Predictions and observations of the testing set.



LSTM result for BTC. We show the training, validation and testing predictions and observations in the same graph for comparison.

Technical Indicators

- Pricing data is not enough to learn patterns
- Add more features (technical indicators) to further improve the LSTM network model predictions.

Technical Indicators	Types	Description
Simple Moving Average 7 days (SMA 7)	Trend	Simple moving averages (SMA) calculate the average of range of prices by the number of periods within that range. SMA can help to determine the trend of an asset price.
Simple Moving Average 30 days (SMA 30)	Trend	
Simple Moving Average 90 days (SMA 90)	Trend	
Bollinger Channel Middle Band (B.B.M)	Volatility	Bollinger Bands are a type of chart indicator and are composed of three lines. It is used to determine overbought and oversold market conditions. The width of the band can be an indicator of an asset's volatility.
Bollinger Channel High Band (B.B.H)	Volatility	
Bollinger Channel Low Band (B.B.L)	Volatility	
Parabolic Stop and Reverse (PSAR)	Trend	Parabolic SAR indicator is used to determine trend direction and potential reversals in price.
Moving Average Convergence Divergence (MACD)	Trend	MACD shows the relationship between two moving averages and the momentum of price movement.
Average True Range (ATR)	Volatility	ATR measures the degree of price volatility
Relative Strength Index (RSI)	Momentum	RSI measures the magnitude of price changers to evaluate overbought or oversold conditions
On-balance volume (OBV)	Volume	OBV uses volume to predict changes in stock price.

Conclusion:

- Correlations between the cryptocurrency market and the traditional financial market are not as strong as recent studies have suggested
- There is a significant impact of the merge on the cryptocurrency market, which can be observed by comparing the correlations of the daily closing prices and trading volumes of the cryptocurrencies before and after the merge
- Modern deep learning algorithms provide much better forecasting of future cryptocurrency prices

LSTM Network Analysis

Platform: TensorFlow

API (application programming interface): Keras

Input Layer: (5, 1)

LSTM Layer: 64 neurons (units)

Dense Layer 1: 32 neurons; activation='relu‘

Dense Layer 2: 32 neurons; activation='relu'

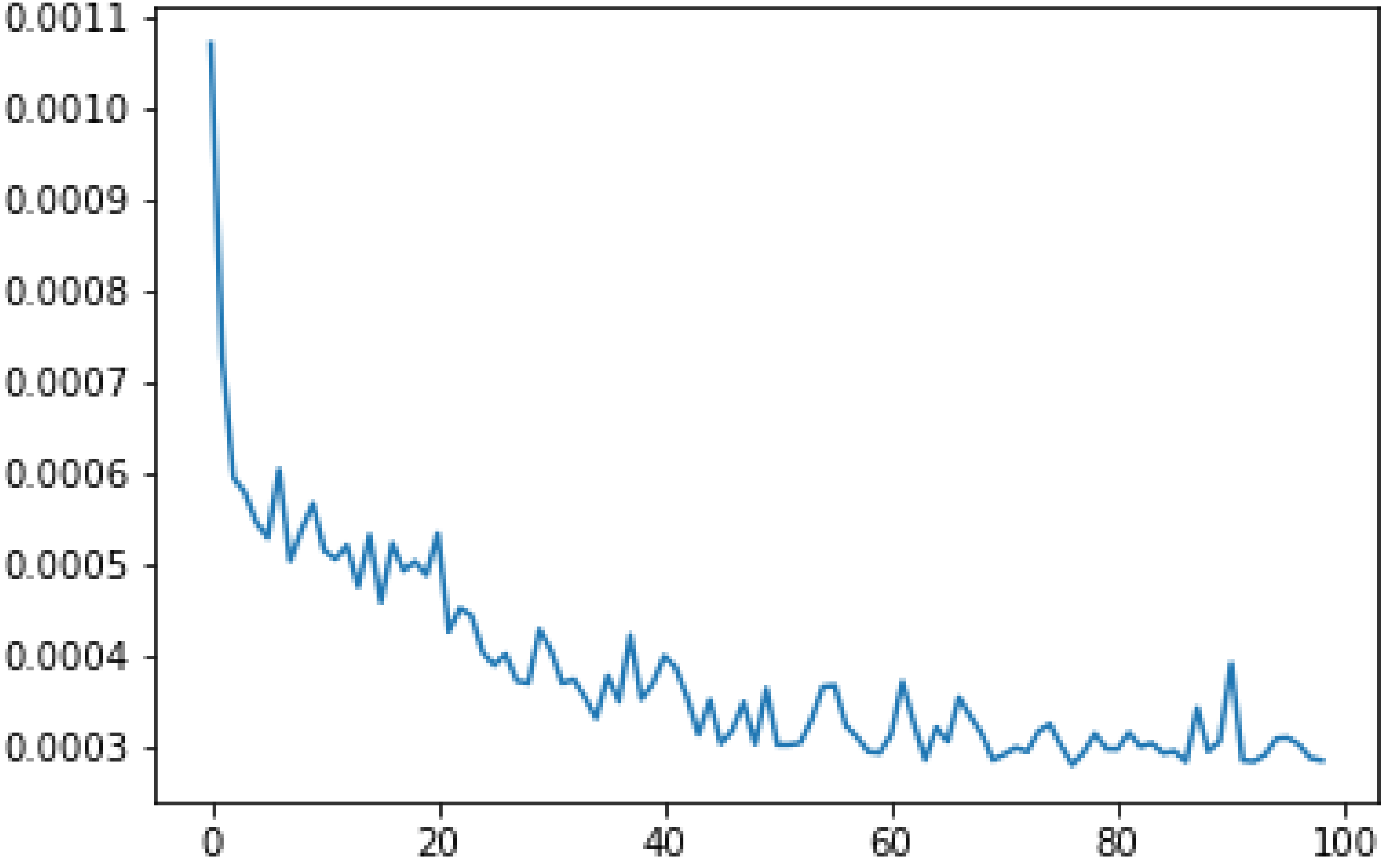
Dense Layer 3: 1 neuron

Optimizer: Adam (Learning rate=0.001)

Activation Function used: Rectified Linear Unit (ReLU)

Adam: Adaptive Moment Estimation

LSTM Network Analysis



Loss per epoch