

# 频域分析视角下股票指数波动特征研究

郭建平, 秦传清

南京信息工程大学管理工程学院, 江苏 南京

收稿日期: 2023年8月30日; 录用日期: 2023年10月24日; 发布日期: 2023年10月31日

## 摘要

掌握股票指数波动特征是有效管理股市投资风险的重要基础, 选择上证指数数据为样本, 从频域分析视角, 利用希尔伯特-黄变换频域分析方法研究上证指数波动的原因及经济意义。通过分解上证指数, 得到不同频率的特征分量, 分析原序列内在微观构成特征; 通过重构分量, 确定影响上证指数波动的高频部分、低频部分和趋势项部分; 通过希尔伯特变换, 研究其经济意义。实证结果表明, 上证指数的波动是高频部分和低频部分共同作用的结果, 且上证指数会围绕趋势项部分进行波动。在不同时期, 高频部分和低频部分对上证指数波动的贡献程度有所不同, 在2007年牛市中, 低频部分占据上证指数波动的绝大部分; 而在2015年牛市中, 上证指数波动的大部分来自于高频部分。研究为投资者管理股市投资风险提供了新的数量依据。

## 关键词

频域分析, 股票指数, 波动特征, 希尔伯特-黄变换

# Research on Volatility Characteristics of Stock Index from the Perspective of Frequency Domain Analysis

Jianping Guo, Chuanqing Qin

School of Management Engineering, Nanjing University of Information Science & Technology, Nanjing Jiangsu

Received: Aug. 30<sup>th</sup>, 2023; accepted: Oct. 24<sup>th</sup>, 2023; published: Oct. 31<sup>st</sup>, 2023

## Abstract

To master the volatility characteristics of stock index is an important basis for effective management of stock market investment risk. In this paper, Shanghai Stock Index data is selected as a

sample, from the perspective of frequency domain analysis, and the Hilbert-Huang transform frequency domain analysis method is used to study the reasons and economic significance of the volatility of Shanghai Stock index. By decomposing Shanghai Composite index, the characteristic components of different frequencies were obtained, and the intrinsic microscopic composition characteristics of the original sequence were analyzed. By reconstructing component, the high frequency part, low frequency part and trend item part which affect SSE index fluctuation are determined. Through the Hilbert transform, the economic significance is studied. The empirical results show that the fluctuation of the Shanghai Composite Index is the result of the joint action of the high frequency part and the low frequency part, and the Shanghai Composite index will fluctuate around the trend item part. In different periods, the contribution degree of the high frequency part and the low frequency part to the volatility of Shanghai Composite Index varies. In 2007 bull market, the low frequency part occupied the majority of the volatility of Shanghai Composite Index; in the 2015 bull market, most of the volatility of the Shanghai Composite index came from the high-frequency part. This paper provides a new quantitative basis for investors to manage stock market investment risk.

## Keywords

Frequency Domain Analysis, Stock Index, Wave Characteristics, Hilbert-Yellow Transformation

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言及文献综述

股票走势受经济形势、宏观经济政策和投资者情绪变化等多种因素影响,其波动存在极大不确定性,针对股票指数等时间序列数据波动性的研究最早可追溯到上世纪中期对金融资产等时间序列数据分布形态和波动特征的描述性统计分析,之后的学者们通过数量模型进一步研究了市场数据的波动特征,梳理如下。

### (一) 波动时域特征研究

研究发现金融时间序列的方差具有时变特征,依赖于过去的随机误差项,为描述该特性,自回归条件异方差模型(ARCH)被提出,之后,针对 ARCH 模型的不足,广义自回归条件异方差(GARCH)模型被提出。大量研究者基于 GARCH 模型族,就市场中的坏消息对收益率波动的影响、股票市场波动的非对称效应、杠杆效应、不同国家股市波动关联性及溢出效应等问题展开了初步探讨。

国内金融市场等时间序列波动的研究开始较晚,但进展迅速,通过对基本 GARCH 模型进行非线性改进,预测我国股市波动情况,分析股票市场收益率和波动性的相关关系,比较我国股票市场和发达国家股票市场的波动程度,研究沪深 300 股指期货对股票市场波动的平抑作用,探析融资融券制度与我国股市波动率之间的关系,陈海强等分析融资融券制度在一定程度上降低我国股市波动率的机制等研究结论构成了学术界对股票市场波动规律的基本认识[1]。

为进一步明确股指波动特征,新分析方法和分析模型不断被提出,其中有影响力的研究成果包括杨培涛等利用多元向量自回归模型研究我国货币政策影响股价波动的效应[2]、苑莹等利用多分形波动率模型研究我国股市波动[3]、马锋利用马尔科夫状态转换机制修正异质自回归模型进行股市波动的预测研究[4]、卢米雪等利用主成分分析方法研究极端投资者情绪影响股票市场波动的机制[5]、梁琪等利

用高低波动风险溢出模型研究风险溢出方向[6]、文磊等利用疫情防控政策研究影响股价波动的机制[7]以及尹力博等证明了股市波动加剧市场投机的结论[8],上述研究成果加深了对市场波动特征的进一步理解。

## (二) 波动频域特征研究

考虑到从时域的角度研究股票市场波动问题,探究波动形成机理有一定的难度,部分研究开始联合频域的角度研究股票指数等时间序列数据波动问题,在众多时频分析方法中,1998年Huang N. E.提出的包括经验模态分解和希尔伯特谱分析的研究方法(以下简称HHT)应用广泛,成果丰富。

HHT是分析非线性、非平稳时间序列数据的频域方法,包括经验模态分解和希尔伯特变换两个步骤,通过对原始数据施以经验模态分解并对本征模态函数(以下IMF)进行希尔伯特变换,研究IMF序列瞬时特征参数,在工程技术领域应用成果丰富。随着对HHT方法的研究逐渐深入,将HHT方法应用于经济领域,研究经济系统时间序列波动问题不断涌现,取得了令人满意的效果,同时,也为使用频域分析方法研究复杂经济问题提供了有价值的经验。

本文创新之处主要表现在:1)方法应用方面,将频域分析方法应用于经济领域问题研究中,通过时域和频域两个角度识别股票指数的波动特征,分析波动特点,扩展了原有频域分析方法的应用领域;2)将时域框架下影响股指波动的众多因素划分为高频、低频和趋势三类,克服了以往时域分析框架下影响因素无法完全列举的不足,增强了研究结论的说服力;3)研究发现我国股市不同时期股指波动的深层次原因不同,不同时期的牛市是由不同的影响因素造成的。

## 2. 研究方法介绍

### (一) 模态分解和重构基本原理

经验模态分解(以下EMD)是由Huang N.E.等提出的研究非平稳、非线性时间序列分析的方法[9][10],目标是将非线性非平稳的信号最终分解成满足如下两个条件的本征模态分量(IMF):1)在整个数据集中,极值点和穿零点数量必须相等或最多相差1个;2)信号上下包络线关于时间轴局部对称。通过反复筛选,最终把原始信号表达成多个IMF和一个残余项 $r_k(t)$ 之和的形式。

继EMD方法,聚合经验模态分解(Ensemble Empirical Mode Decomposition, EEMD)和完备集合经验模态分解(Complete Ensemble Empirical Mode Decomposition, CEEMD)被提出。有效克服了EMD方法存在的模态混叠问题以及EEMD无法重构原信号和产生虚假分量的问题,也避免了CEEMD方法可能导致分解结果不一致的情形[10]。

在EMD分解过程中,本征模态分量IMF由按照高频到低频到顺序被分解出来,按照以下步骤可以将分解出来的IMF进行重构,根据IMF分量的均值是否显著不同于零,将原序列上证指数重构为高频部分、低频部分和趋势项。具体分析步骤归纳如下:1)将第一个IMF分量指定为 $C_1(t)$ ,将第一个和第二个IMF分量之和指定为 $C_2(t)$ ,前三个IMF分量之和指定为 $C_3(t)$ ,按此规律,将前 $k$ 个IMF分量之和指定为 $C_k(t)$ ;2)对所有 $C_i(t)(i \leq k)$ 进行均值是否显著不为零的 $t$ 检验;3)当指标 $i(i \leq k)$ 的均值显著不为零时,则将第一个IMF分量至第 $i$ 个IMF分量进行叠加,构成原序列的高频部分;将第 $i$ 个IMF分量至第 $k$ 个IMF分量进行叠加,构成原序列的低频分量。

### (二) 希尔伯特变换

已知一个复信号的实部,可以通过希尔伯特变换求它的虚部。对于一个实信号 $Y(t)$ ,其希尔伯特变换为:

$$\widetilde{Y}(t) = H[Y(t)] = Y(t) * \frac{1}{\pi t} = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{Y(\tau)}{t - \tau} d\tau \quad (1)$$

式中,  $\widetilde{Y}(t)$  表示  $Y(t)$  的希尔伯特变换,  $H(\cdot)$  表示希尔伯特变换函数,  $*$  表示卷积运算,  $\tau$  为积分变量。由解析信号的定义,  $Y(t)$  和  $\widetilde{Y}(t)$  代表的解析信号  $Z(t)$  可以表示为:

$$Z(t) = Y(t) + j\widetilde{Y}(t) \quad (2)$$

改写成指数形式为:

$$Z(t) = A(t)e^{j\theta(t)} \quad (3)$$

其中,  $A(t) = \sqrt{(Y(t))^2 + (\widetilde{Y}(t))^2}$  和  $\theta_i(t) = \arctan \widetilde{Y}(t)/Y(t)$ , 分别表示瞬时幅值和瞬时相位。对瞬时相位求导, 可以求出瞬时频率, 瞬时频率为:

$$\omega(t) = d\theta_i(t)/dt \quad (4)$$

通过希尔伯特变换得出的瞬时幅值和瞬时频率都是时间的函数, 通常的做法是在一个三维图形上表示三者之间的关系, 也称为希尔伯特谱, 希尔伯特谱是一种时频谱, 同时从时域和频域两个角度刻画了序列信号的内在特性。根据该谱可以进行希尔伯特频谱分析(Hilbert Spectrum Analysis, HSA)。

根据希尔伯特谱, 可以求出边际谱。计算方法是在希尔伯特谱上对时间轴进行积分, 得到振幅和频率两者之间的关系。边际谱刻画了不同频率下序列幅值的大小, 其计算公式为:

$$h(\omega) = \int_0^T H(\omega, t) dt \quad (5)$$

其中,  $H(\omega, t)$  表示希尔伯特谱,  $t$  表示时间,  $h(\omega)$  表示根据希尔伯特谱计算出来的边际谱。

### 3. 实证分析

#### (一) 数据选择和描述性统计分析

选取上海证券综合指数(以下简称上证指数)作为研究对象。以上证指数 1991/07/09 至 2022/9/30 期间的周度收盘点数作为样本数据, 共计 1583 个观测值。数据来自东方财富 Choice 数据库, 计算工具为 MATLAB。



Figure 1. Shanghai Composite Index weekly closing points chart 1991/07~2022/09

图 1. 上证指数周度收盘点数走势图 1991/07~2022/09

通过观察图 1, 初步可以判断在样本期间内上证指数波动非常大, 前后经历了震荡下跌、暴涨、暴跌和横盘等多种股票走势行情。进一步对序列进行描述性统计分析, 结果如表 1。

**Table 1.** Descriptive statistical analysis of sample data of Shanghai Composite Index  
**表 1.** 上证指数样本数据描述性统计分析

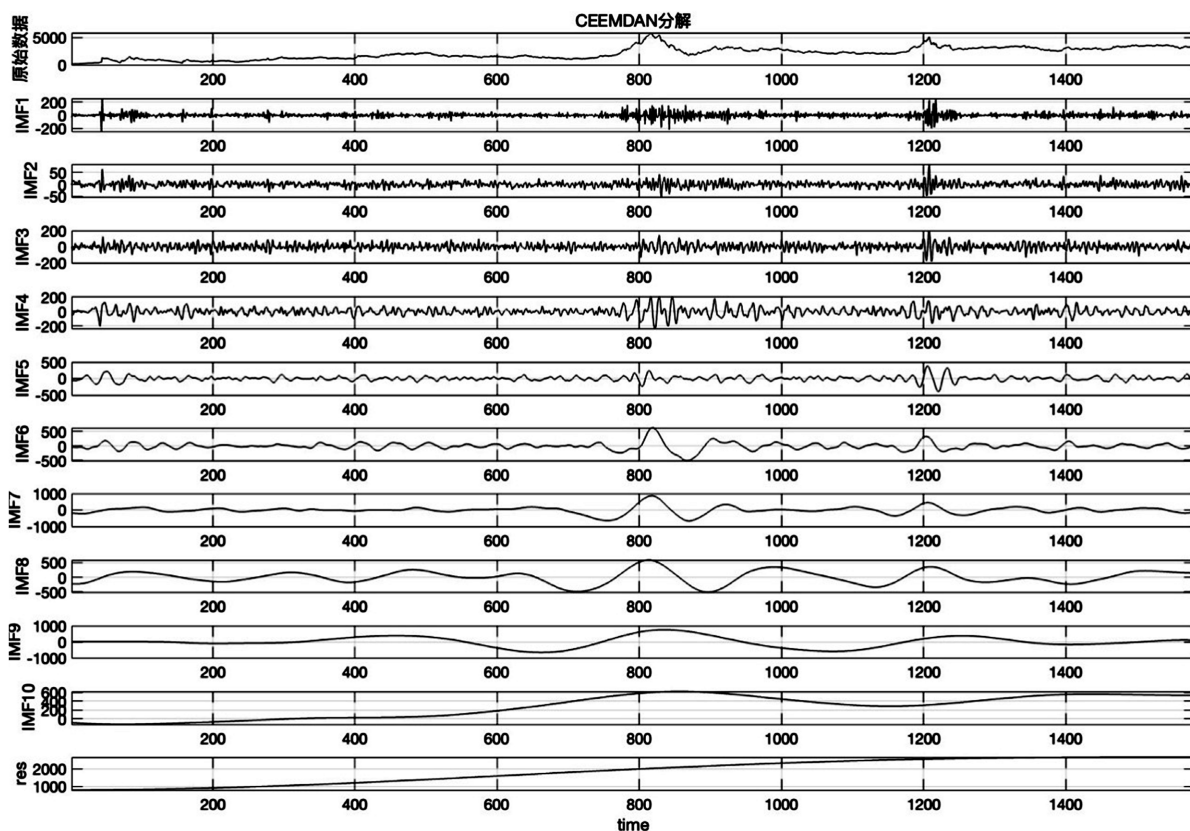
N	Min	Max	kurtosis	skew	Std. Deviation	J-B 统计量	t 统计量	1%临界值
1583	136.7	5903.26	2.75	0.38	1068.41	43.27	-2.90	-3.43

由表 1 可知, 与正态分布相比, 上证指数周度数据的偏度为 0.38, 即表明有更多的观测值出现在右侧, 峰度为 2.75, 具有扁平化的分布特点。Jarque-Bera 统计量为 43.27, 显著异于 0, 有充分理由认为样本数据不满足正态分布假设。对上证指数进行 ADF 单位根检验, ADF 检验统计量的值大于 1%置信水平下的临界值-3.43, 不能拒绝原始序列平稳的假设。综上, 样本数据具有典型金融时间序列数据的基本特征。

(二) 波动性质分析

1、CEEMDAN 分解

添加的白噪声的标准差 $\epsilon$ 设置为 0.2, 将信号集成数量设置为 100, 最大迭代次数设置为 100, 对原序列上证指数进行 CEEMDAN 分解, 得到 10 个 IMF 分量和一个残余项, 如图 2 所示。



**Figure 2.** CEEMDAN decomposition diagram  
**图 2.** CEEMDAN 分解图

图 2 中横轴表示时间向量, 该变量的长度与观测值数据的长度相同, 这里其刻度设置为 10, 纵轴表示频率, 各个 IMF 分量按照频率的大小从上往下依次排列, 第一个表示原始序列的走势, 最后一个表示残余项的走势。图 2 中, signal 表示原序列, IMF 分量按照频率从高到低的顺序进行排列, res 为残余项。

经过 CEEMDAN 分解后, 原序列被分解成 10 平稳的、具有周期性的本征模态分量 IMF 及一个残余项 res, 且第一个 IMF 分量具有周期小、频率大的特点, 而随着分解的不断进行, IMF 分量的周期逐渐



变大, 频率逐渐变小, 第 10 个 IMF 分量的周期最大, 频率最小。IMF 分量中包含的能量, 即振幅, 也具有逐渐增大的特性。残差项 res 代表了原序列的长期趋势, 其图像是一条向上的、单调的曲线, 表明原序列上证指数的长期趋势一直处于不断上升的趋势。

根据对上证指数进行 CEEMDAN 分解所得出的数据, 可以求出 IMF 分量的平均周期、各个 IMF 分量与原序列的 Pearson 相关系数和 Kendall 相关系数、各个统计量的方差、方差占比以及修正方差占比, 如下表 2 所示。

**Table 2.** Features of IMF components and trend items  
**表 2.** IMF 分量和趋势项特征表

	Mean period (week)	Pearson correlation	Kendall Correlation	Variance	Variance as% Of observed	Variance as% of (IMFs + res)
Observed				1,141,492		
IMF1	2.83	0.04	0.01	1617	0.14%	0.20%
IMF2	3.84	0.04	0.02	138	0.01%	0.01%
IMF3	3.49	0.03	0.01	1854	0.16%	0.23%
IMF4	6.62	0.08	0.05	2780	0.24%	0.35%
IMF5	13.65	0.12	0.06	5344	0.47%	0.67%
IMF6	29.87	0.22	0.07	9544	0.84%	1.19%
IMF7	83.32	0.36	0.16	59,218	5.19%	7.39%
IMF8	175.89	0.41	0.23	49,189	4.31%	6.13%
IMF9	395.75	0.35	0.21	89,868	7.87%	11.21%
IMF10	791.50	0.74	0.57	39,407	3.45%	4.91%
res		0.82	0.64	542,895	47.56%	67.71%
sum				801,854	70.25%	1

表 2 中 IMF 分量的平均周期定义为观测值个数除以各个 IMF 分量所包含的极大值点(或极小值点)的个数, 随着分解出来的 IMF 分量频率变小, IMF 分量所包含的极值点变少, 其平均周期相应变大。表中, 第一个 IMF 分量的周期为 2.83 周, 约为 20 天左右, 最后一个 IMF 分量的周期则达到了 791.50 周, 约为 14 年。周期较短的 IMF 分量反应上证指数波动的短期因素, 周期较长的 IMF 分量表现上证指数波动的中期因素, 残差项则表示影响上证指数波动的长期趋势。

Pearson 相关系数用来刻度两个不同时间序列数据之间的线性程度, Kendall 相关系数用来度量两个不同时间序列之间是否具有某种相同的变化趋势。在表 2 中, IMF 分量与原序列上证指数的 Pearson 相关系数随着分解的进行具有逐渐变大的趋势, 其中, IMF8 和 IMF10 的系数分别达到了 0.41 和 0.74, 均超过了 0.4, 这表明, IMF8 和 IMF10 这两个 IMF 分量与上证指数存在着中等以上程度的线性相关性, 残差项 res 与原序列上证指数的 Pearson 相关系数高达 0.82, 超过了 0.8, 表明残差项 res 与原序列上证指数存在着高等以上程度的线性相关性。分解出的 IMF 分量和残差项与原序列上证指数的 Kendall 相关系数均为正, 这表明 IMF 分量和残差项与原序列上证指数具有一致的变化趋势, 且 IMF8、IMF9、IMF10 和残差项的 Kendall 系数相对于其他 IMF 分量的较大, 表明这几个 IMF 分量和残差项与原序列上证指数有着一定相同程度的变化趋势。

上证指数经过 CEEMDAN 分解之后, 各个 IMF 分量和残差项的方差之和小于原序列上证指数的方差, 这导致各要素与原序列上证指数的方差占比之和不为 1, 这里采用修正的方差占比, 即分解出来的

各要素的方差与各要素方差之和的比值。从表 3 中可以看出, IMF7、IMF8、IMF9 和残差项 res 的方差占比较大, 分别占比 7.39%、6.13%、11.21%和 67.71%。

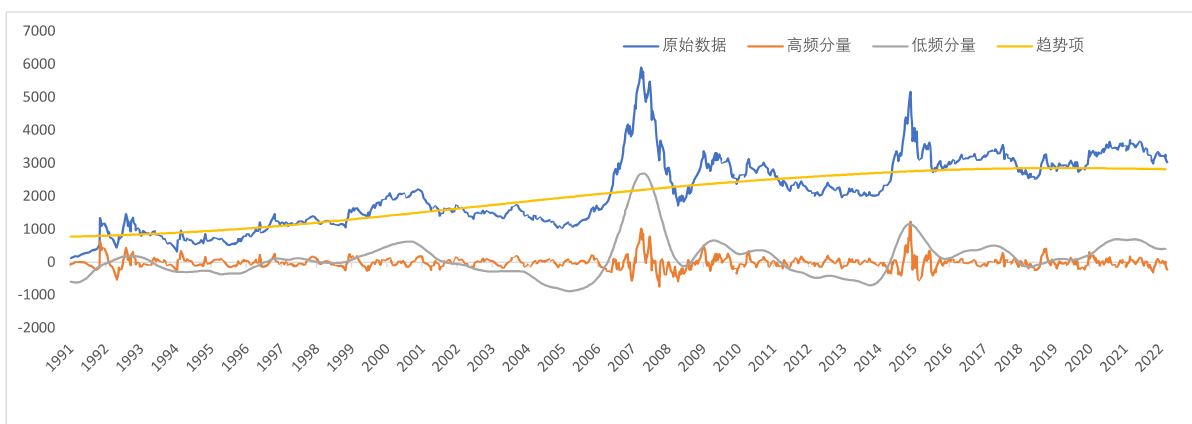
### 2、重构 IMF

重构 IMF 分量, 对原上证指数序列经过 CEEMDAN 分解所产生的 10 个 IMF 进行重构, 确定高频分量和低频分量。按照前述方法, 依次对  $C_1(t) \sim C_i(t)$  的均值进行是否显著为零的  $t$  检验, 在置信水平为 95% 的情况下, 得到分类结果。

**Table 3.** High and low frequency classification of IMF components  
**表 3.** IMF 分量的高低频分类表

	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	IMF10
结果	高频	高频	高频	高频	高频	高频	低频	低频	低频	低频

根据表 3, 将 IMF1~IMF6 进行叠加, 得到原序列上证指数的高频部分; 同理, 将 IMF7~IMF10 进行叠加, 得到原序列上证指数的低频部分。残差项 res 作为原序列上证指数的趋势项。这样, 上证指数就被重构为高频部分、低频部分和趋势项三个子序列之和。



**Figure 3.** Shanghai composite index and its three sub-sequences  
**图 3.** 上证指数和其三个子序列走势图

图 3 是上证指数和其高频部分、低频部分和趋势项三个子序列的走势, 从图中可看出, 与原序列上证指数相比, 三个子序列走势相对平缓、简单, 且具有一定的规律性, 便于对上证指数波动做进一步的分析, 并探究其背后的经济意义。

表 4 给出了上证指数高频部分、低频部分和趋势项的描述性统计情况。观察表中数据, 我们可以看到高频部分的周期为 4.07 周, 约为一个月, 低频部分的周期为 87.94 周, 约为 1 年 6 个月。高频部分的

**Table 4.** Descriptive statistics of high frequency part, low frequency part and trend items  
**表 4.** 高频部分、低频部分和趋势项描述性统计表

	Mean period (week)	Min	Max	Mean	Std.Deviation
High frequency component	4.07	-734.18	1239.09	-4.40	165.91
Low frequency component	87.94	-867.64	2706.85	120.58	580.59
Trend		781.15	2862.09	2000.80	736.81

极差为 1973.27 点, 趋势项点极差为 2080.94 点, 低频部分的极差为 3574.49 点, 超过了高频部分和趋势项。从均值上看, 趋势项的均值为最大, 为 2000.80 点, 远超过高频部分和低频部分。

**Table 5.** Sub-sequence characteristics of SSE index

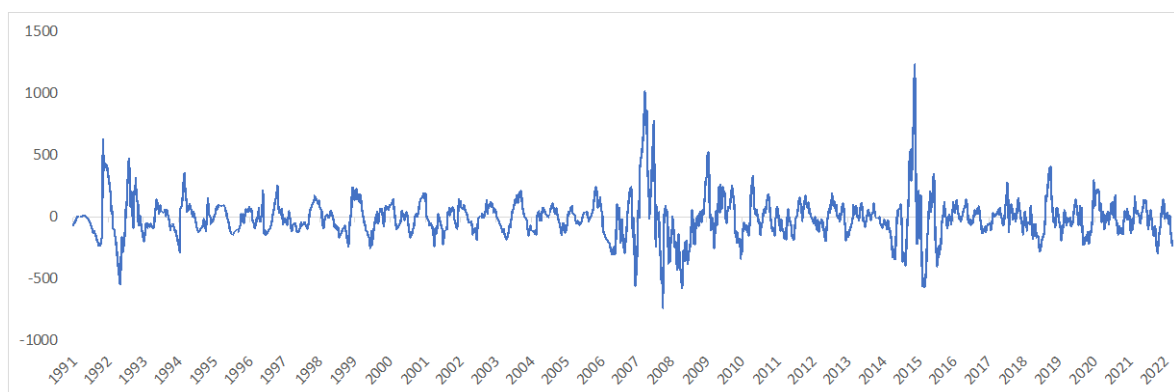
**表 5.** 上证指数子序列特征表

	Pearson correlation	Kendall Correlation	Variance	Variance as% of observed	Variance as% of components
Observed			1,141,492		
High frequency component	0.23	0.11	27527	2.41%	3.03%
Low frequency component	0.73	0.51	337,089	29.53%	37.14%
Trend	0.82	0.64	542,895	47.56%	59.83%
Sum			907,511	79.50%	1

表 5 是上证指数三个子序列的特征情况表, 包含了各个子序列和原序列上证指数的 Pearson 和 Kendall 相关系数、各个子序列及原序列的方差和相对应的方差占比。从 Pearson 相关系数来看, 低频部分与原序列上证指数的相关系数达到了 0.73, 超过了 0.4, 表明趋势项和原序列具有中等程度以上的相关性。趋势项和原序列上证指数的相关系数达到了 0.82, 超过了 0.8, 表明低频部分和原序列具有高度相关性。

在 Kendall 相关系数中, 高频部分、低频部分和趋势项与原序列上证指数的相关系数都为正数, 这说明上证指数的三个子序列与上证指数具有相同的变化趋势。其中, 与高频部分相比, 低频部分和趋势项与原序列的 Kendall 相关系数最大, 说明低频部分和趋势项与原序列的变化趋势的同步程度更高。从方差占比上看, 低频部分和趋势项的方差较大, 达到了原序列的 29.53% 和 47.56%, 修正方差占比达到了 37.14% 和 59.83%。综上表明, 低频部分和趋势项与原序列上证指数具有较高的相似性和方差占比, 一定程度上表现了上证指数长期的特点, 构成了上证指数的绝大部分, 高频部分则具有更短的周期, 更小的振幅(能量), 表现了上证指数短期内的波动。

上证指数的波动受多种因素的影响, 通过上述步骤将其分解成三个子序列, 把其影响因素按照时间的标准划分为三类。第一类, 高频部分所代表的短期波动, 如下图 4 所示, 高频部分的走势有着较频繁的波动, 它的振幅较小, 它是指一些短期突发事件对上证指数造成的影响。第二类, 低频部分所代表的中期波动, 它对应能影响上证指数波动一些具有影响力的重大事件, 其走势如图 5 所示, 相比于高频部分, 它的频率较低、周期和振幅较大。第三类, 趋势项代表对长期趋势, 长期趋势是指一些长期重大事件对上证指数造成的影响。如图 6 所示, 趋势项的走势一条平滑的、不断上升的曲线。



**Figure 4.** Trend chart of the high frequency part of Shanghai Composite Index

**图 4.** 上证指数高频部分走势图



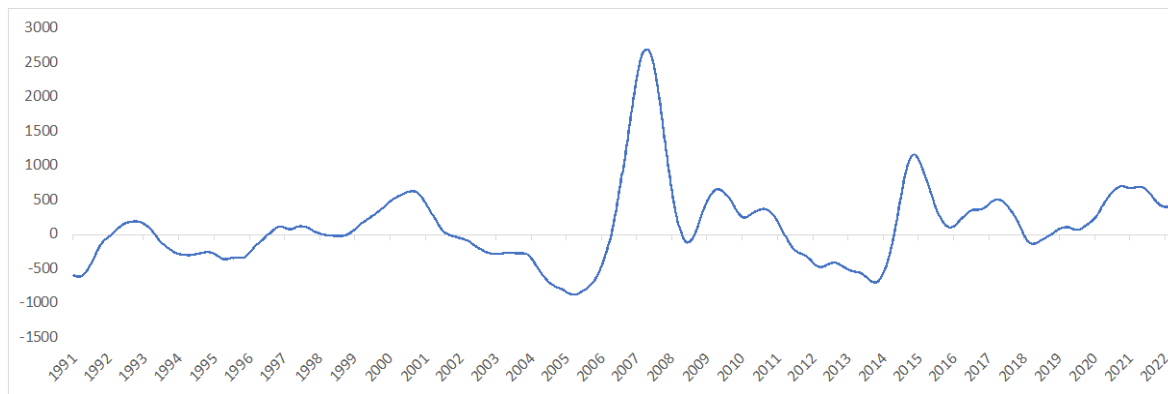


Figure 5. Trend chart of the low frequency part of Shanghai Composite Index  
图 5. 上证指数低频部分走势图

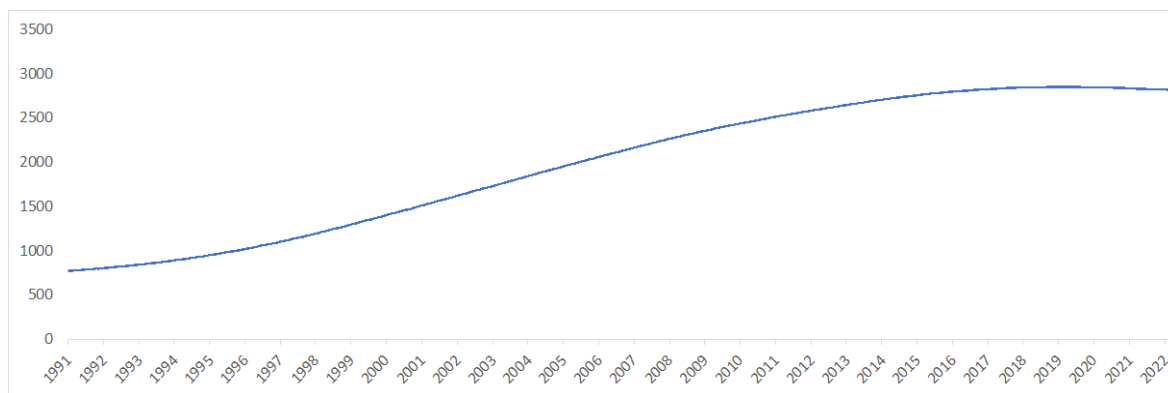


Figure 6. Trend chart of Shanghai Composite Index  
图 6. 上证指数趋势项走势图

### 1) 趋势项

趋势项代表了上证指数波动的长期趋势，由表 5 知，趋势项与原序列上证指数的线性相关程度达到了 82%，在三个子序列中最高，达到了高等程度线性相关水平。从图 6 中，随着时间的推移，尽管上证指数在建模区间内经历了暴涨、暴跌的走势，但趋势项呈现着不断上升的趋势，这种不断上升的长期趋势可以对应于我国在此期间内经济发展水平、国民收入的提高。

然而，可以看到，最近 10 年，趋势项的增长趋势放缓，这与我国近些年来国民经济增速放缓相对应。在图 6 中，通过对比趋势项和原序列上证指数的走势，我们可以看到，无论上证指数因何种事件如何波动，只要影响上证指数的事件随着时间的推移而结束，上证指数的走势终归会与趋势项的走势趋于一致。例如，金融危机期间，上证指数先是经历了暴涨，从 2000 点左右涨到了 6000 点以上，短时间内与趋势项有很大的差距。随着金融危机的影响逐渐接近尾声，上证指数又跌回 2000 点附近，与趋势项几乎处于同一水平上。

### 2) 低频部分

上证指数的低频部分，是指造成上证指数波动的中期事件。这些中期事件具有两个特点，一是对上证指数的影响大，而是影响时间较长。由表 6 知，上证指数低频部分与原序列上证指数的线性相关程度达到了 73%，超过了中等程度的线性相关水平，且低频部分点方差占比达到了 37.14%，占据了大部分。这表明低频部分是上证指数波动来源的一个重要部分。低频部分包含的 IMF 分量具有较高的能量(振幅较大)，如图 5 所示，低频部分在某些时刻对上证指数的贡献点数超过了 2000 点。

此外，通过前述分析知，低频部分由本征模态分量 IMF7~IMF10 构成，由表 2 知，这些 IMF 分量中，

平均周期最大的是 791.50 周, 平均周期最小的是 83.32 周, 最少都超过了一年, 这表明这些影响上证指数波动的中期事件在短时间内很难消除, 都具有持续性。由于上证指数的趋势项变动缓慢, 高频部分的振幅较小, 所以, 上证指数的波动大多数都来自于其低频部分的波动。

通过对上证指数分解出来低频部分进行分析, 可以研究具体中期事件对上证指数走势波动的影响有多大, 会持续多长时间, 预测未来重大事件的发生会给上证指数造成多大的影响。例如, 在 2015 年牛市期间, 低频部分从 2014 年 5 月份开始快速上涨, 一直到 2016 年 5 月份才逐渐趋于平稳, 在此期间, 上证指数从最低点 2011.13 点涨到了 5178.19, 涨幅为 3167.06 点, 然后又跌回到 3000 点以下, 而同期低频部分从低点-693.12 点涨到了 1167.34 点, 涨幅为 1860.22 点, 在上证指数的涨幅中, 低频部分的涨幅占比高达 58.73%。

### 3) 高频部分

除了长期趋势和低频部分的影响, 上证指数的波动还受其高频部分的影响。高频部分具有频率高、周期和振幅小的特点。高频部分包括的事件一般是指突发事件, 在短时间内引起上证指数的波动。例如, 经济政策的出台、CPI 和 PPI 指数等一些重要宏观经济数据的公布和投资者情绪的变化等影响短期股票市场供需平衡的事件。

根据以上分析, 高频部分由本征模态分量 IMF1~IMF6 构成, 在这些 IMF 分量中, 平均周期最小的为 2.83 周, 最大的为 29.87 周, 均不超过一年, 且这些 IMF 分量的振幅较小, 所携带能量(振幅)较小, 其叠加之和构成的高频部分对上证指数的波动贡献较小, 远低于趋势项和低频部分。除特殊点之外, 高频部分对上证指数波动的贡献大约处于-200~200 点之间。虽然由市场短期不平衡构成高频部分对上证指数波动的冲击较小, 但它发生的频率很高, 造成了上证指数短期的起伏。在中长期的分析和预测中, 我们可以将它忽略不计, 只需考虑趋势项和低频部分, 但是在短期研究和预测中, 我们必须将它考虑在内, 因为它确实是造成上证指数短期波动的决定性因素。

根据上述分析, 上证指数在 2022 年 9 月 30 号为 3024.39 点, 其中, 高频部分贡献了-230.41 点, 占比 7.62%; 低频部分贡献了 427.45 点, 占比 14.13%; 作为长期趋势的趋势项贡献了 2827.35 点, 占比高达 93.48%。但在 2015 年牛市时, 上证指数最高点为 5166.35 点(周度数据), 其中, 趋势项部分贡献了 2763.29 点, 占比仅为 53.49%; 低频部分贡献了 1163.97 点, 占比为 22.53%; 而高频部分贡献了 1239.09 点, 占比高达 23.98%。这进一步说明了趋势项表现了上证指数的长期趋势, 来自于中短期事件的冲击是上证指数发生波动的主要原因。

### 3、希尔伯特变换

通过上述分析, 上证指数走势中短期的波动均来自高频分量和低频分量, 为了更好地对上证指数波动的特征进行分析, 这里对经过 CEEMDAN 分解重构后得到的高频分量和低频分量分别施以希尔伯特变换。基于希尔伯特变换可以得到希尔伯特谱, 希尔伯特谱反映了振幅随瞬时频率和信号时间变化的内在关系, 该图谱可以用来研究序列中的子序列随时间变化的规律, 用来识别局部特征。如图 7 所示, 图中纵轴分别表示频率和幅值, 横轴表示观测时间。图中不同颜色的点表示在某一时刻下子序列所对应的瞬时频率和瞬时幅值, 颜色越深、点数越密集的地方表明在该时间段内序列所携带能量越高。

为进一步分析上证指数的波动特征, 在希尔伯特谱的基础上得到边际谱, 计算方法是希尔伯特谱在时间轴上进行积分, 使之从幅值-时间-频率三者间的关系转变为幅值-频率两者间的关系, 描述的是幅值(或能量)在频率轴上的分布。

图 8 分别表示高频分量和低频分量的边际谱, 图中横轴表示频率, 纵轴表示不同频率下所对应的累计幅值。从图中可以看出, 高频分量和低频分量的边际谱图像具有相似性, 都是在低频率时, 对应的累计幅值更大, 具有更大的能量; 在高频率, 能量较小。这一结论符合前面分析内容, 即低频具有更大的振幅, 包含着更多的能量, 也表明中长期因素才是决定股票市场走势的关键因素。

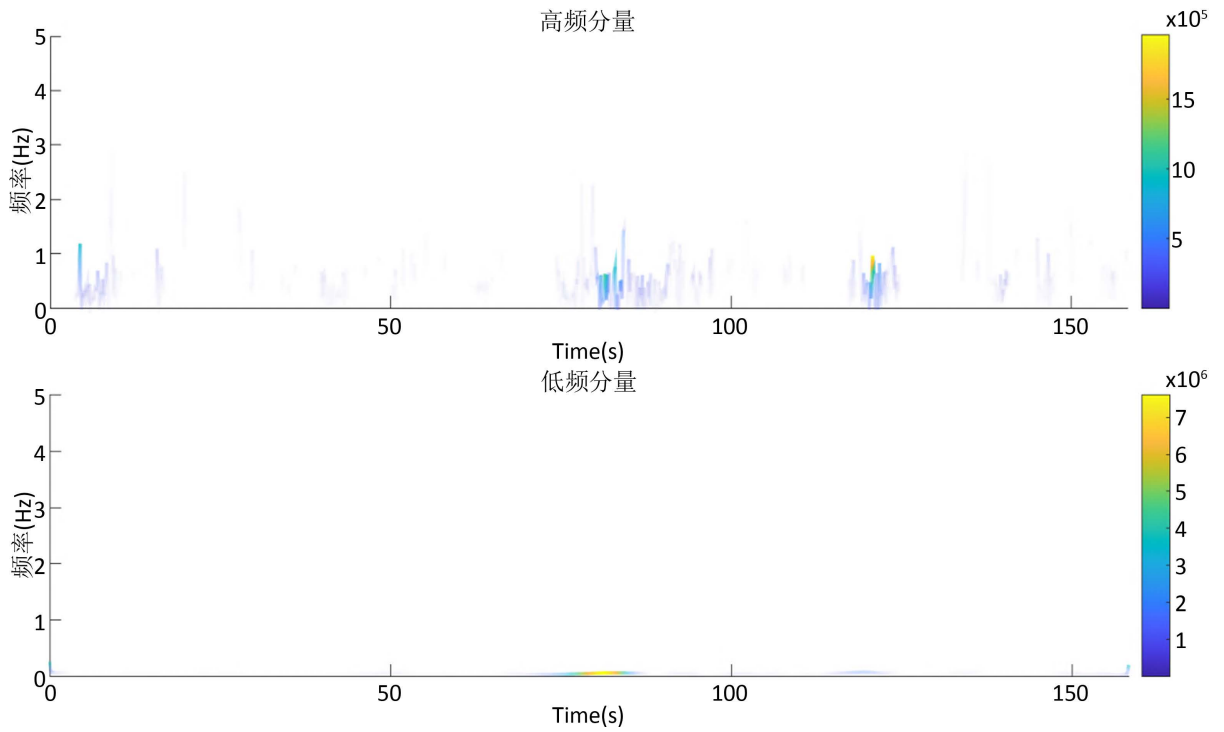


Figure 7. Hilbert spectra of high frequency and low frequency components  
图 7. 高频分量、低频分量 Hilbert 谱

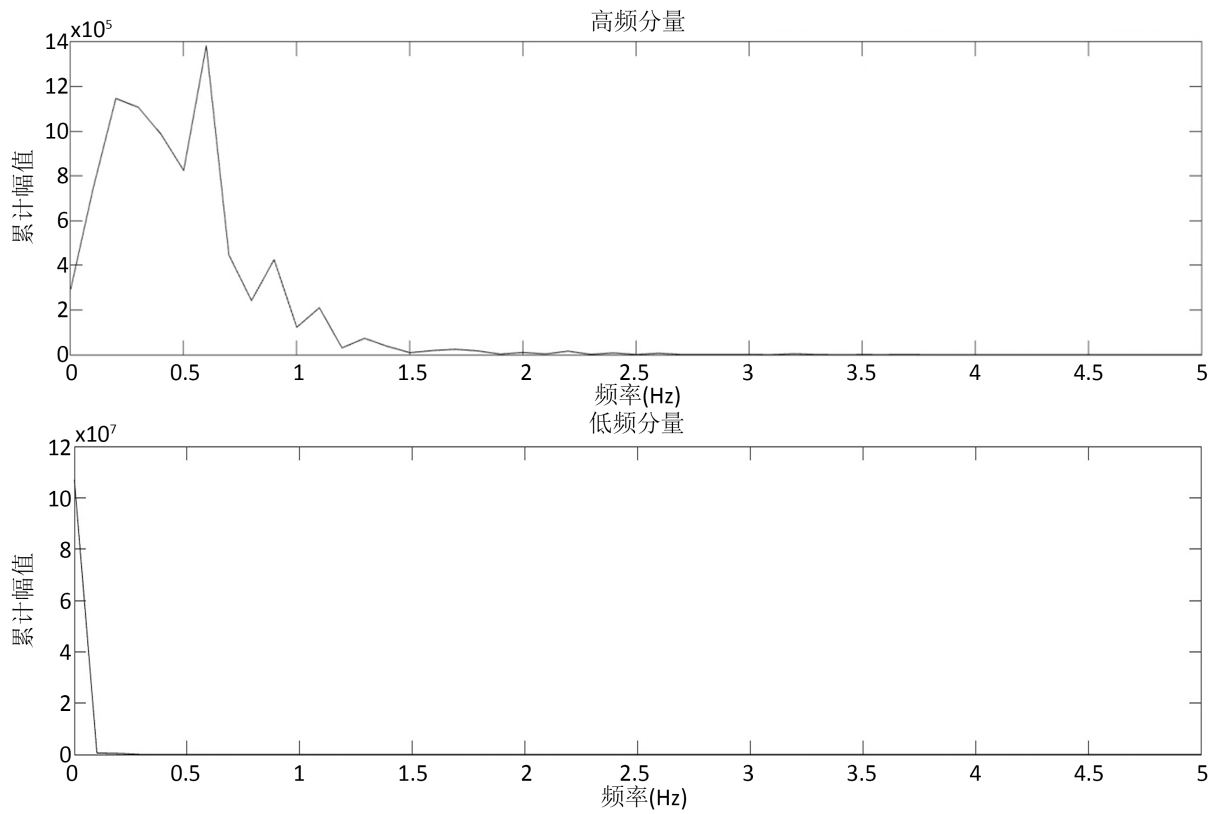
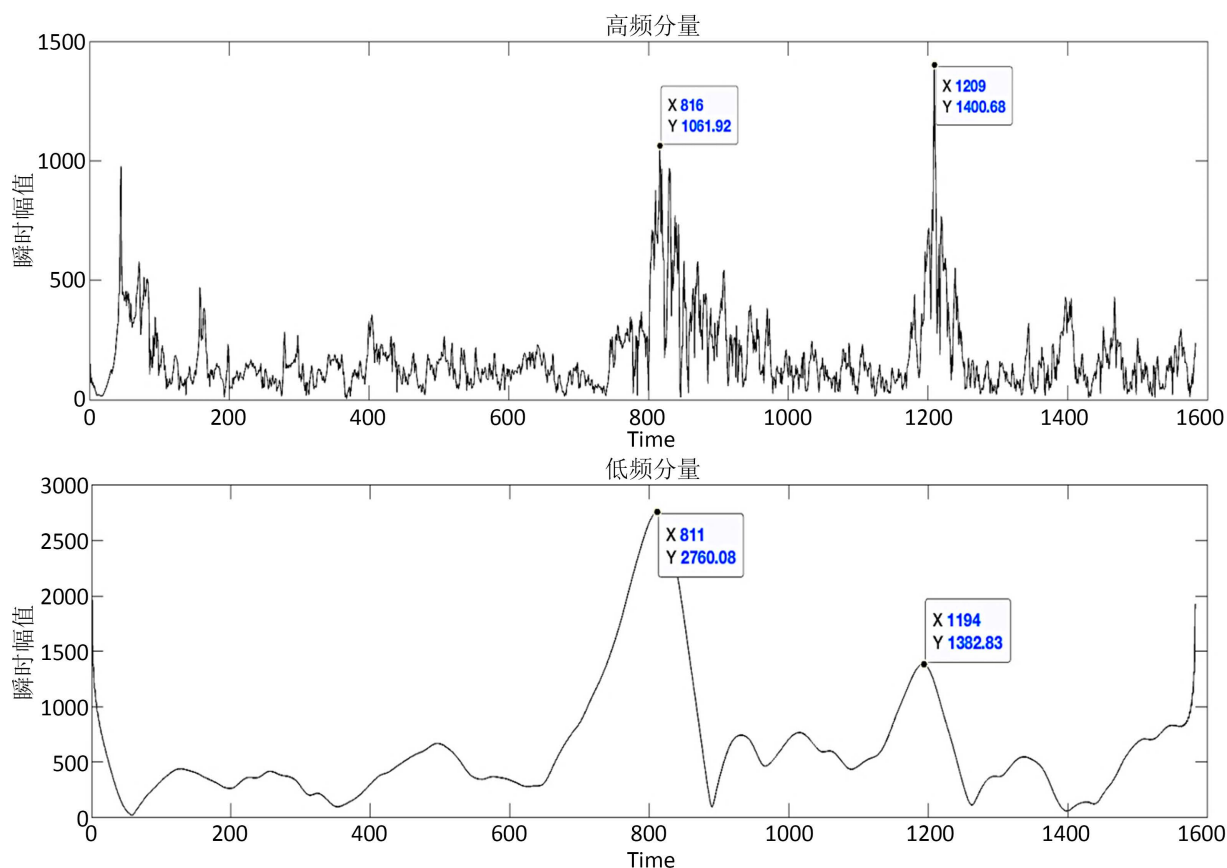


Figure 8. Marginal spectra of high-frequency and low-frequency components  
图 8. 高频分量、低频分量边际谱

不妨将其波动过程分为短期震荡(如短期题材利好、游资投机炒作、宏观经济数据的公布等对指数未来走势产生短暂影响、很快就会恢复正常的波动信号)和长期波动(如股权分置改革、四万亿救市措施等引起长期、深远影响的波动过程),那么短期震荡具有时间短、变化快和频次高的特点,对股票指数未来走势的影响一般都小于长期波动,原因是频率较低的信号其波形更平稳,走势往往更难以改变,一旦长期波动过程的波形改变,必然是股市大环境的突变造成的,这些突变很可能导致指数的剧烈震荡,导致股票指数暴涨暴跌,并最终引发大牛市或金融危机。但两者的差异也比较明显,随着频率的增大,高频分量累计幅值衰减的速度远远小于低频分量。图中,高频分量的边际谱有一定的波动,随着频率的增大逐步下降;而低频分量的累计幅值则是迅速下降。这表明,低频分量的波动随着频率的增大衰减的很快,其走势波动程度远远小于高频分量。



**Figure 9.** Instantaneous amplitudes of high-frequency and low-frequency components

**图 9.** 高频分量、低频分量瞬时幅值图

图 9 是高频分量和低频分量的瞬时幅值图,改图表示在各个不同的时刻,子序列的瞬时幅值的大小。图中,低频分量的瞬时幅值的数量级明显高于低频分量,表明低频分量普遍比高频分量携带更多的能量,是上证指数波动的主要原因,这与我们前述分析一致。但是,也有特殊的情况。例如图中标识所示,2007 年牛市时期,高频分量的瞬时幅值达到了 1061.92 点,低频分量的瞬时幅值达到了 2760.08 点,占据了波动绝大部分,而在 2015 年牛市时期,低频分量的瞬时幅值为 1382.83 点,高频分量的瞬时幅值为 1400.68 点,超过了低频分量的瞬时幅值,表明短期因素对这次上证指数波动的影响更大。

对于 2007 年大牛市,目前普遍认为是股权分置改革的完成、对外进一步资本开放和基金规模的扩大等中长期因素叠加造成了这一轮大牛市,图 9 低频分量在这一时刻的瞬时幅值占据波动的绝大部分证明

了这一点。在 2015 年牛市中, 银行降息降准等宏观利好政策等实施, 场外配资政策的迅速增长等短期因素是造成市场波动的主要原因, 据统计, 在 2015 年牛市中, 股票市场通过场外配资获得了超过 4 万亿的资金。图中, 我们可以看到, 2015 年牛市时期, 高频分量的瞬时幅值罕见的超过了低频分量, 说明进一步说明了此轮上证指数的波动更多的来自于短期因素。

#### 4. 研究结论

本文通过引入 CEEMDAN 方法, 将上证指数分解成一系列按照本征模态分量 IMF 和一个残余项, 这些 IMF 分量相互独立, 且按照频率的高低进行排序, 表现了不同时间尺度下上证指数的波动特点, 有助于从微观结构上研究上证指数的变化特征。利用重构原理, 将分解出的 IMF 分量分类为高频 IMF 分量和低频 IMF 分量, 将高频 IMF 分量进行叠加构成原序列上证指数的高频部分, 将低频 IMF 分量进行叠加构成原序列上证指数的低频部分, 残余项则代表了上证指数的长期趋势。这样, 上证指数就由其高频部分、低频部分和长期趋势这三部分构成, 其中, 长期趋势指的是在长期中能影响上证指数波动的事件, 例如, 国家经济增长率, 国民收入水平等; 低频部分是指在中期中, 能造成上证指数大幅波动的事件, 这些事件的影响时间都在一年以上, 例如, 金融危机、中美贸易战和新冠疫情等。高频部分是指在短期内能造成上证指数波动的事件, 这些事件的持续时间都低于一年且发生频率较高, 例如, CPI、GDP 等宏观数据的公布、经济政策、产业政策的发布和投资者情绪的变化等。通过分析, 得出以下结论。

第一, 对上证指数日常的波动起决定作用的是其低频部分和高频部分。其中, 与高频部分相比, 低频部分具有振幅大、周期长的特点, 对上证指数走势的影响时间长, 影响程度大, 是上证指数波动的一个主要来源。

第二, 趋势项决定了上证指数的长期趋势。无论上证指数是因为某些事件的影响而出现多大的波动, 只要事件的影响结束, 上证指数会和趋势项处于同一水平上, 这说明了长期投资的合理性。

第三, 不同时期低频部分和高频部分对上证指数波动的影响有所不同, 在 2007 年牛市期间, 低频部分是造成这次波动的主要原因; 在 2015 年牛市期间, 造成上证指数波动的因素更多的来自高频部分。

第四, 在中长期内, 上证指数的波动主要来自于低频部分, 高频部分可以忽略不计, 但在短期内, 高频部分是造成上证指数波动的一个重要因素, 高频部分波动的频率较大, 振幅较小, 是上证指数短期波动的主要来源。

与传统的分析方法不同, 本文通过对 CEEMDAN 分解出来的 IMF 分量进行重构, 确定了影响上证指数波动的三个子序列, 并指出了每个子序列背后的经济意义, 使研究结论更具有现实意义, 并对造成上证指数波动的主要序列高频分量和低频分量进行希尔伯特变换和希尔伯特谱分析, 从时域和频域两个角度分析上证指数波动的内在特征。本文提出的思路, 丰富了股票指数等时间序列数据的研究方法, 为股票指数等时间序列数据波动的分析提供了新的视角。

#### 基金项目

教育部人文社会科学研究规划基金: 频域分析视角下金融系统极值风险防范与化解机制研究(19YJA630023)。

#### 参考文献

- [1] 陈海强, 范云菲. 融资融券交易制度对中国股市波动率的影响——基于面板数据政策评估方法的分析[J]. 金融研究, 2015(6): 159-172.
- [2] 杨培涛, 王帅, 朱玉林. 我国货币政策对股价波动的效应测度[J]. 统计与决策, 2020, 36 (24): 125-128.
- [3] 苑莹, 张同辉, 庄新田. 中国股市多分形波动率建模及预测研究[J]. 系统工程理论与实践, 2020, 40(9):



---

2269-2281.

- [4] 马锋, 王继谦, 郭杨莉, 陆菲. 基于跳跃、跳跃强度和机制转换的股票市场波动建模及其预测研究[J/OL]. 系统工程理论与实践: 1-15, 2022-12-02. <http://kns.cnki.net/kcms/detail/11.2267.N.20221202.1053.001.html>
- [5] 卢米雪. 投资者情绪的测量及其对股市波动率的影响效应研究[J]. 宏观经济研究, 2022(9): 106-119.
- [6] 梁琪, 常姝雅. 全球股票市场系统性风险的预警与防范——基于高低波动风险溢出网络的分析[J]. 国际金融研究, 2022(9): 67-76.
- [7] 文磊, 李宏兵. 疫情防控与股价波动: 来自中国上市公司的经验证据[J]. 管理评论, 2022, 34(8): 54-64.
- [8] 尹力博, 韦亚. 股市波动助长了市场投机性吗? [J]. 管理评论, 2021, 33(12): 15-29.
- [9] Huang, N.E., Shen, Z., Long, S.R., *et al.* (1998) The Empirical Mode Decomposition and the Hilbert Spectrum for Non-Linear and Non-Stationary Time Series Analysis. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, **454**, 903-995. <https://doi.org/10.1098/rspa.1998.0193>
- [10] Wu, Z. and Huang, N.E. (2009) Ensemble Empirical Mode Decomposition: A Noise Assisted Data Analysis Method. *Advances in Adaptive Data Analysis*, **1**, 1-41. <https://doi.org/10.1142/S1793536909000047>