

特征提取方法及其在经营管理中的应用

中国科学院数学与系统科学研究院

中国科学院大学经济与管理学院

中国科学院预测科学研究中心

汪寿阳

Email: sywang@amss.ac.cn

OUTLINE

- ◎ 一个真实故事
 - ◎ 经验模态分解 (EMD) 方法
 - ◎ 一个新的原油定价理论
 - ◎ 一个基于数据分解的事件分析方法
 - ◎ 结论与启示
-
- ◎ 直接滤波方法DFA
 - ◎ 奇异谱方法SSA
 - ◎ 季节调整方法
 - ◎ 程序与实现

一个真实的故事

A REAL STORY

- ❑ **Goldman Sachs** predicted in March 2008 that the crude oil price would be over US\$ 200 per Barrel by the end of 2008.
- ❑ **Our team** forecasted in a press conference on March 14, 2008 that the crude oil price would drop below US\$ 80 per Barrel by the end of 2008 (the price on that day was US\$ 109 per Barrel).
- ❑ **The price of crude oil on the last transaction day of 2008 was US\$ 35.82 per Barrel!**
- ❑ All other institutions predicted that the price would be up (not down)
- ❑ **Some good comments, such as “the unique institution in the world made a correct prediction of crude oil price”** appeared in some newspapers, including The Wall Street Journal, Financial Times, etc.

IMPLICATIONS

- **The methodology** is the most important in economic forecasting;
- **Quantitative Modeling** is **extremely crucial** in any good methodology applied to economic forecasting and analysis, but a “good” methodology might be more important;
- Our **main advantage** in comparing with those top economic forecasting institutions (such as World Bank, Goldman Sachs, Lancaster Center for Forecasting) is that we have a number of outstanding scientists in the several fields of mathematics, economics, computer science, management and systems science in our Center.

Official Think-Tank of Chinese Government

NEWS CENTER
Tuesday, February 09, 09:28 am IST

NEWS MARKETS MUTUAL FUND
Budget Business General News

Moneycontrol » News Center » General News

Want more about
General News to
land in your
mailbox?
Set an alert here



Official think-tank

Published on Sun, Feb 07, 2010 at 20:32 |
Email Print Blish This

Sustainable Economy
Siemens has answers for sustainable



An official think-tank around 10% in 2010, agency reported. Strong first-quarter expansion dipping Science at the C

weekend.

Growth in the third and fourth quarters

- Listed Company Directory
- Business News
- IPO News
- Listing IPO
- Columnists
- Analysts
- Standard Money
- Announcement
- Circulars
- Annual Report/Interim Report
- Company Supplement
- Membership
- Columnist



Business News

Top think-tank

08/02/2010

China will return to the first quarter

The Center for Forecasting Science will report growth for the second quarter respectively.

It also estimated that economic recovery will jump 5.22 percent

Investments will see annual investor tank's forecasts more evidence of

The investment bank from December's

Bloomberg BusinessWeek

HOME

REUTERS

EDITION: U.S.

News & Markets

Sectors & Industries

Analysis & Opinion

Login Register

Search

SEARCH

THE VISA BLACK CARD.
LEARN MORE

ARTICLE

COMMENTS (0)

SAVE EMAIL PRINT REPRINTS



CONSUMER STAPLES SECTOR | FOOD PRODUCTS IND

NESTLE SA-REG (N)

LAST SFr.49.42 CHF CHANGE TODAY +1.07 2.3

As of 11:31 AM 02/08/10 All times are local (Market data)

Snapshot News Charts Financials Es

PR Newswire 02/08/2010 9:55 AM ET

Zhongpin Reports Production a Higher Market Prices

CHANGGE and BEIJING, China, Feb. 8 /F Asia-FirstCall/ -- Zhongpin Inc. ("Zhongpin HOGS"), a leading meat and food process in the People's Republic of China, today n smooth production at its new Tianjin plant and foreign trade and exports higher market prices

MORE REUTERS RESULTS FOR:
""center for forecasting science"
february 2010"

Login or register

Latest from My Topics

0 Saved Items

5 Suggested for You

Home Run and knocking frozen pizza sales out of the park

Official think-tank sees China 2010 GDP up 10 pct

BEIJING, Feb 7 (Reuters) - An official think tank has forecast China's gross domestic product (GDP) will expand around 10 percent in 2010, driven by investment and domestic consumption, the official Xinhua agency reported on Sunday.

Sun Feb 7, 2010 6:16am EST

Strong first-quarter growth of 11 percent will slow slightly for the rest of the year, with the pace of expansion dipping just below 10 percent in the second half, the Center for Forecasting Science at the Chinese Academy of Sciences said in the report, which was issued at the weekend.

Growth in the third and fourth quarters will be 9.5 percent and 9.8 percent respectively, it said.

Foreign trade will rebound as the global economy improves, but overall net exports will be a drag

EMPIRICAL MODE DECOMPOSITION (EMD)方法

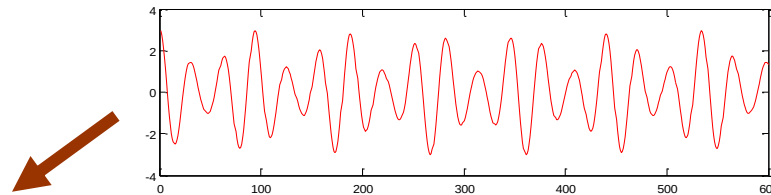
EMPIRICAL MODE DECOMPOSITION (EMD)



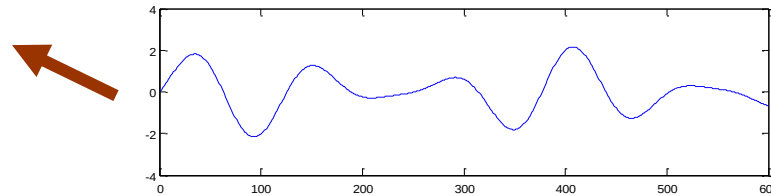
Norden E Huang



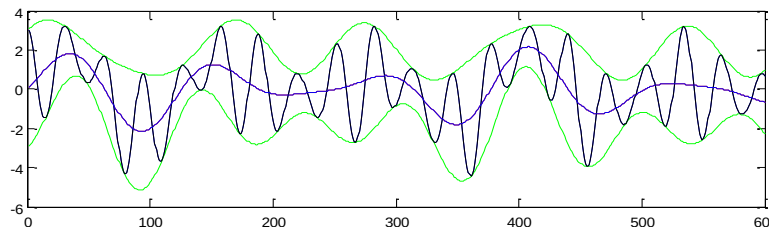
Empirical Mode Decomposition (EMD) is an empirical, intuitive, direct and self-adaptive data processing method which is proposed especially for nonlinear and non-stationary data. (Huang et al, 1998)



signal
source 1



signal
source 2



**Overall
Signal**

THE PRINCIPLE OF EMD

- ◉ EMD assumes that the data, depending on its complexity, may have many different coexisting modes of oscillations at the same time.
- ◉ EMD can extract these intrinsic modes from the original time series, based on the local characteristic scale of data itself, and represent each intrinsic mode as an intrinsic mode function (IMF), which meets the following two conditions:
 - have the same numbers of extrema and zero-crossings or differ at most by one;
 - are symmetric with respect to local zero mean.
- ◉ An IMF is a nearly periodic function and the mean is set to zero. IMF is a harmonic-like function, but with variable amplitude and frequency at different times.

EMD ALGORITHM

- 1) Identify all the maxima and minima of time series $x(t)$;
- 2) Generate its upper and lower envelopes, ($e_{\min}(t)$ and $e_{\max}(t)$), with cubic spline interpolation.
- 3) Calculate the point-by-point mean ($m(t)$) from upper and lower envelopes:

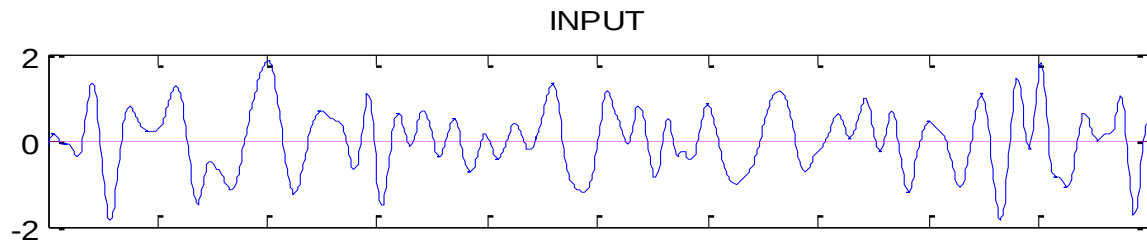
$$m(t) = (e_{\min}(t) + e_{\max}(t)) / 2 \quad (1)$$

- 4) Extract the mean from the time series and define the difference of $x(t)$ and $m(t)$ as $d(t)$

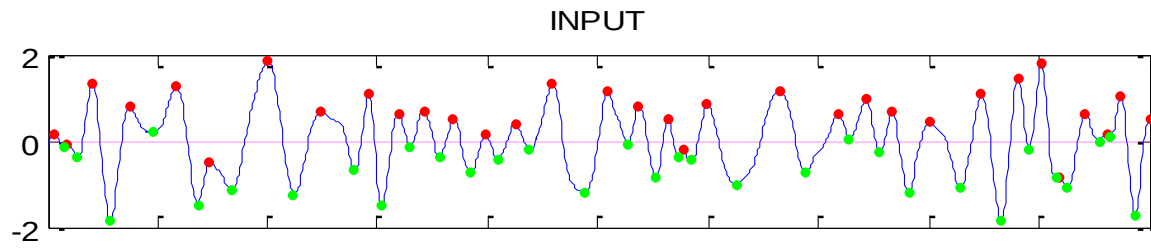
$$d(t) = x(t) - m(t) \quad (2)$$

- 5) Check the properties of $d(t)$:
 - If it is an IMF, denoted $d(t)$ as the i^{th} IMF and replace $x(t)$ with the residual $r(t) = x(t) - d(t)$; The $d(t)$ is often denoted as $c_i(t)$;
 - If it is not, replace $x(t)$ with $d(t)$;
- 6) Repeat steps 1)–5) until the residual satisfies some stopping criterion.

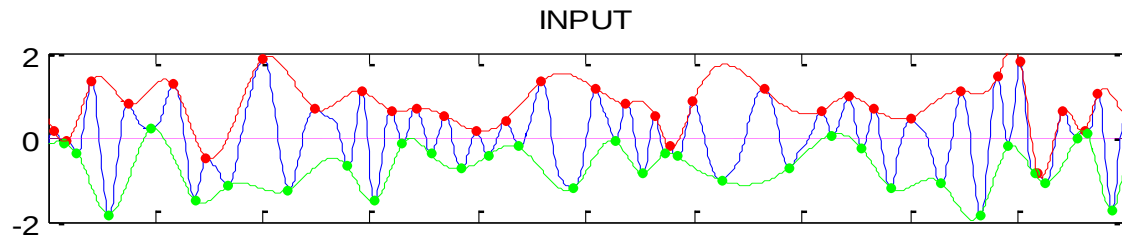
SIFTING



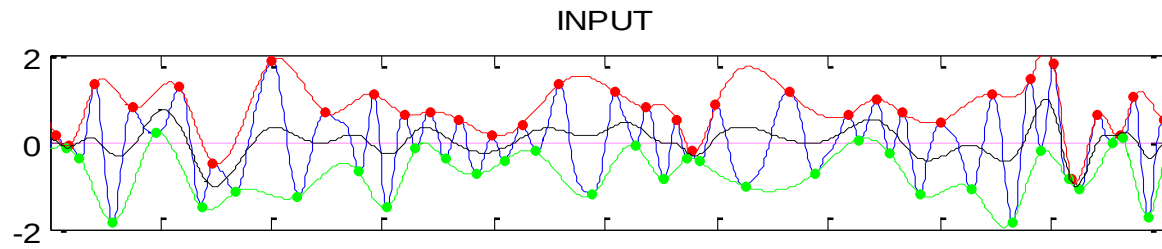
SIFTING



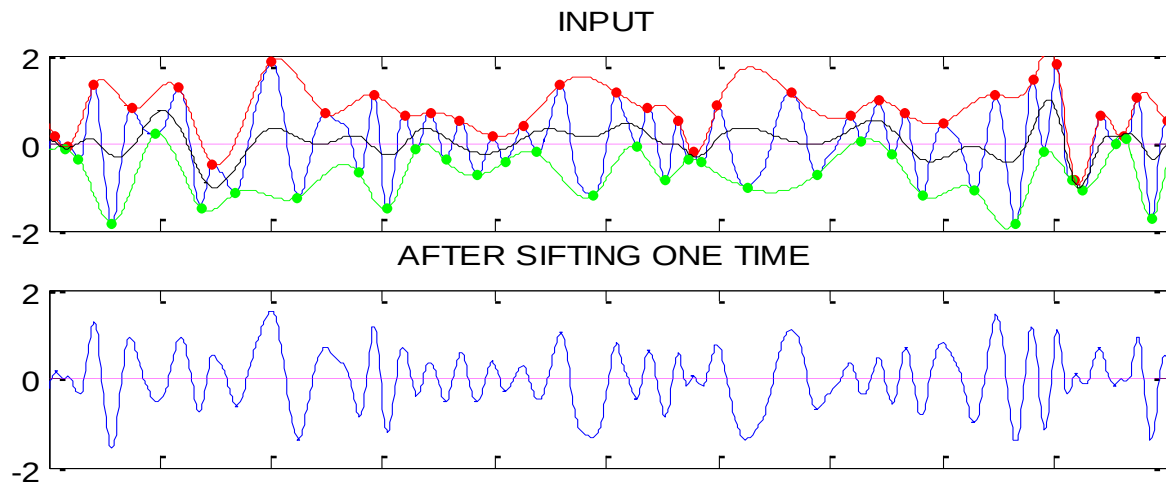
SIFTING



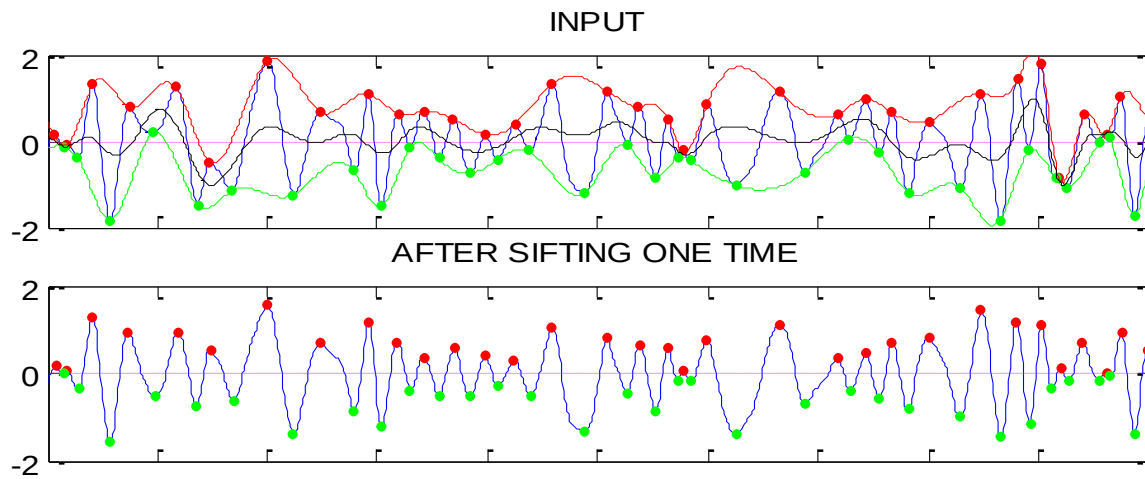
SIFTING



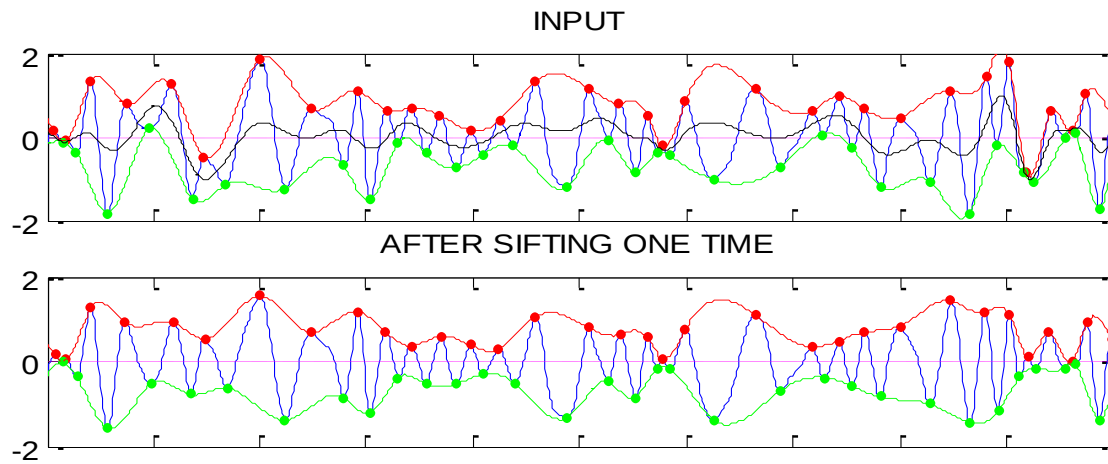
SIFTING



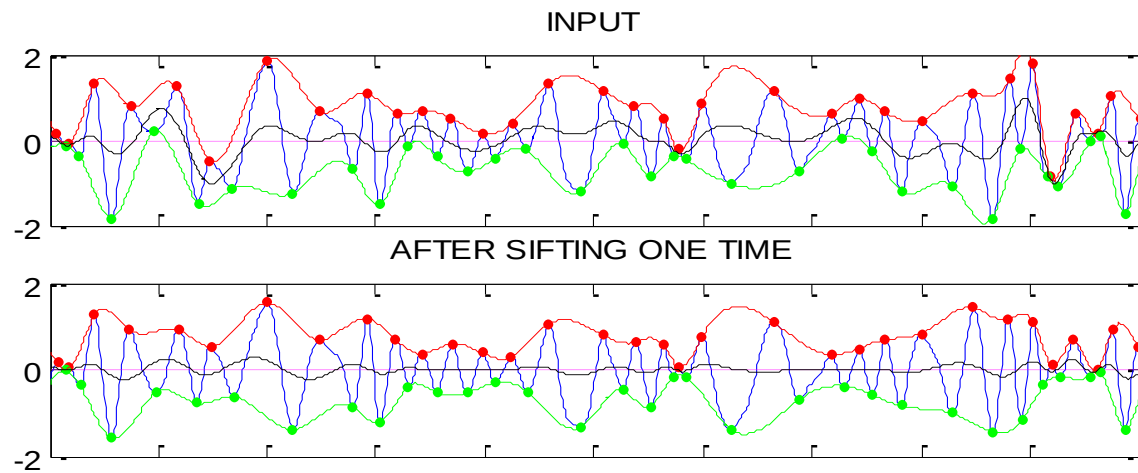
SIFTING



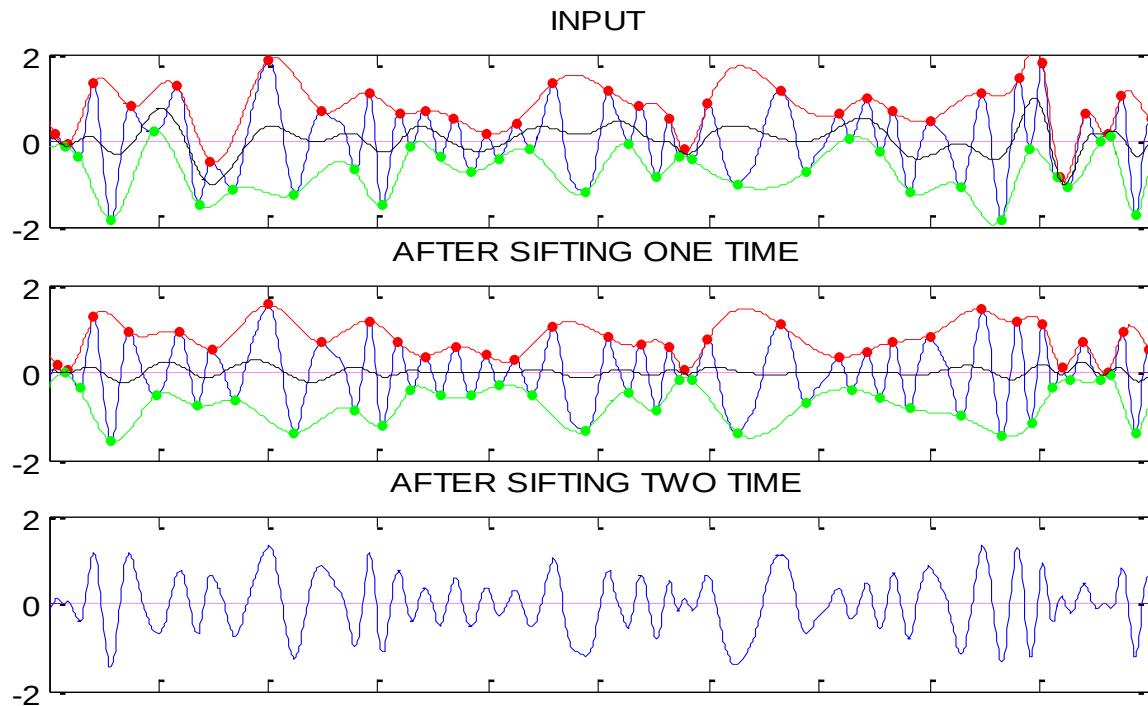
SIFTING



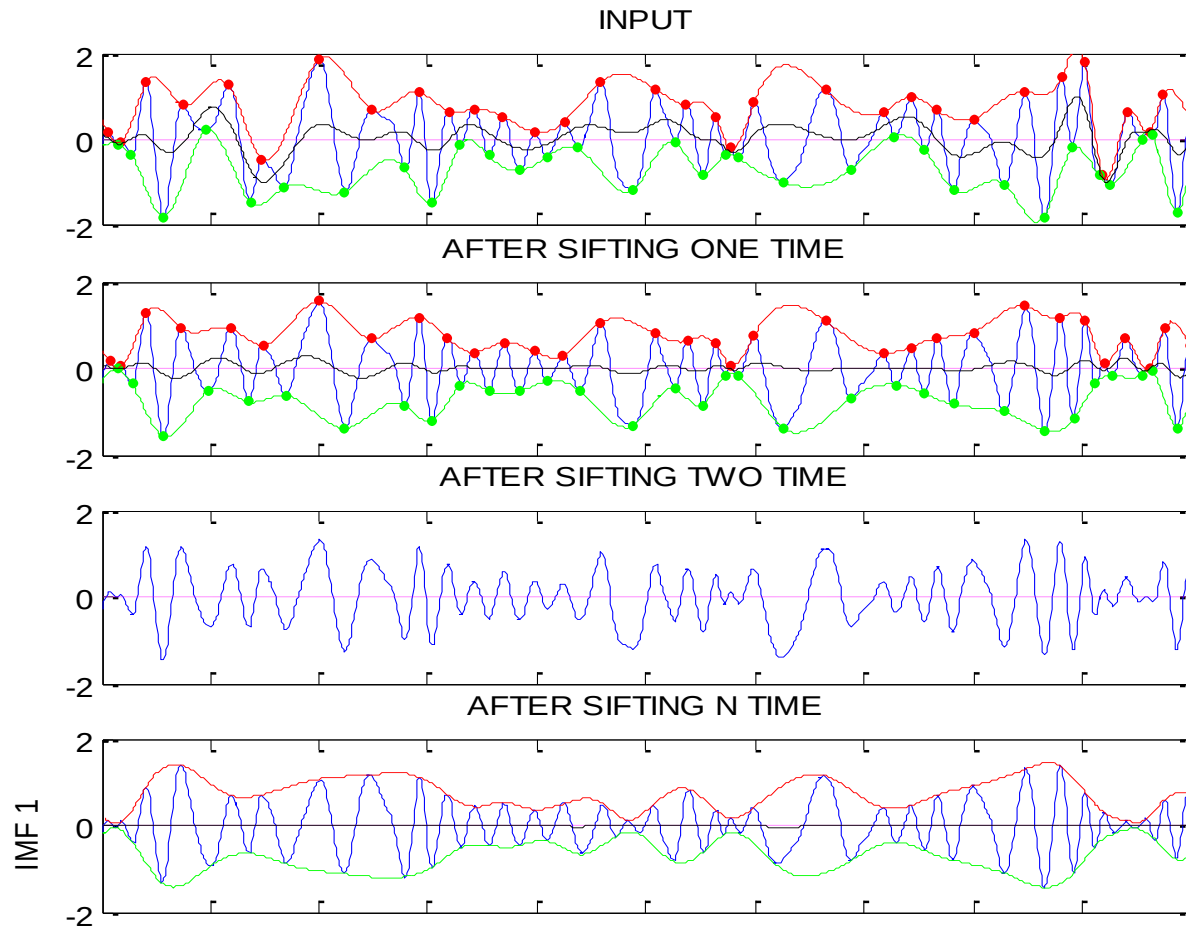
SIFTING



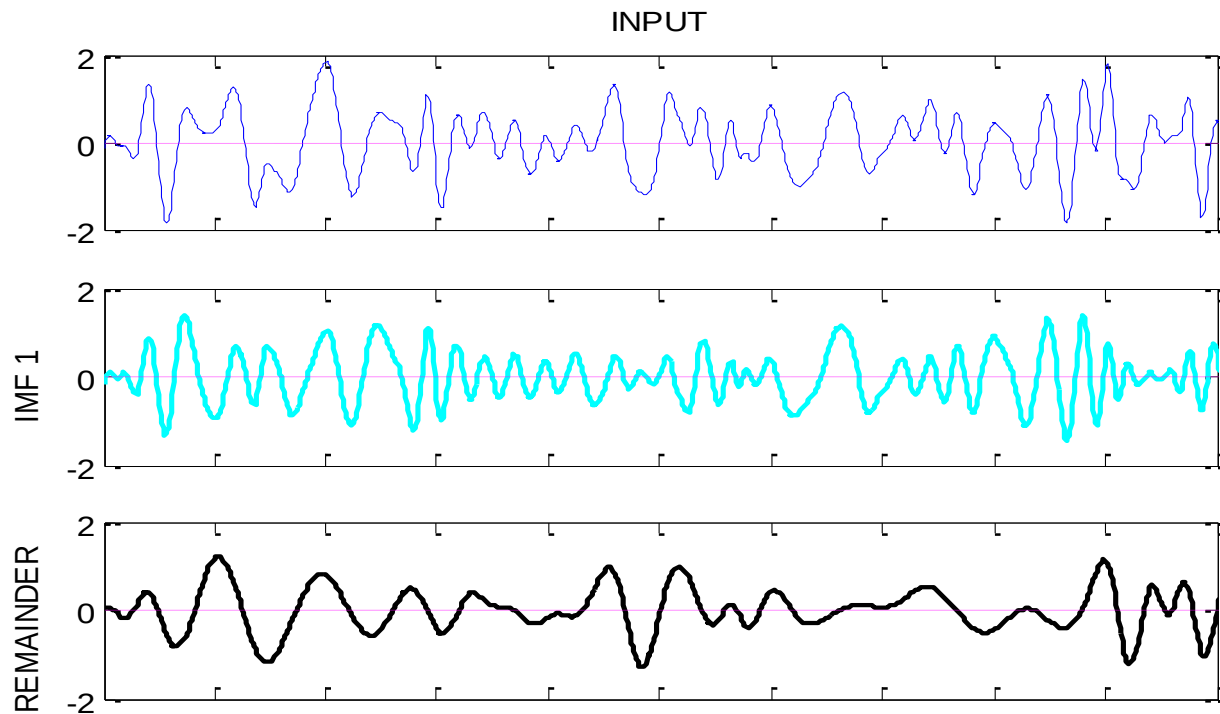
SIFTING



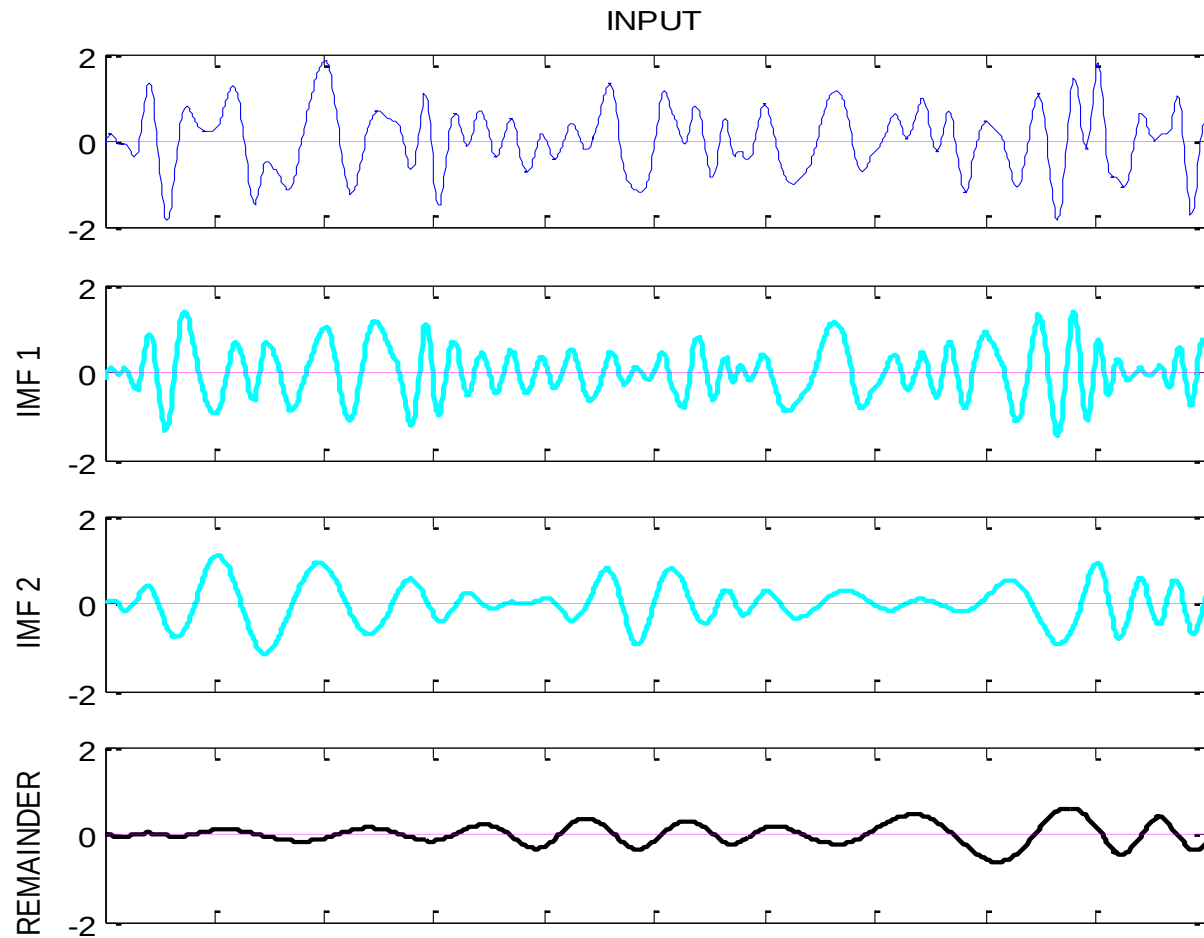
SIFTING



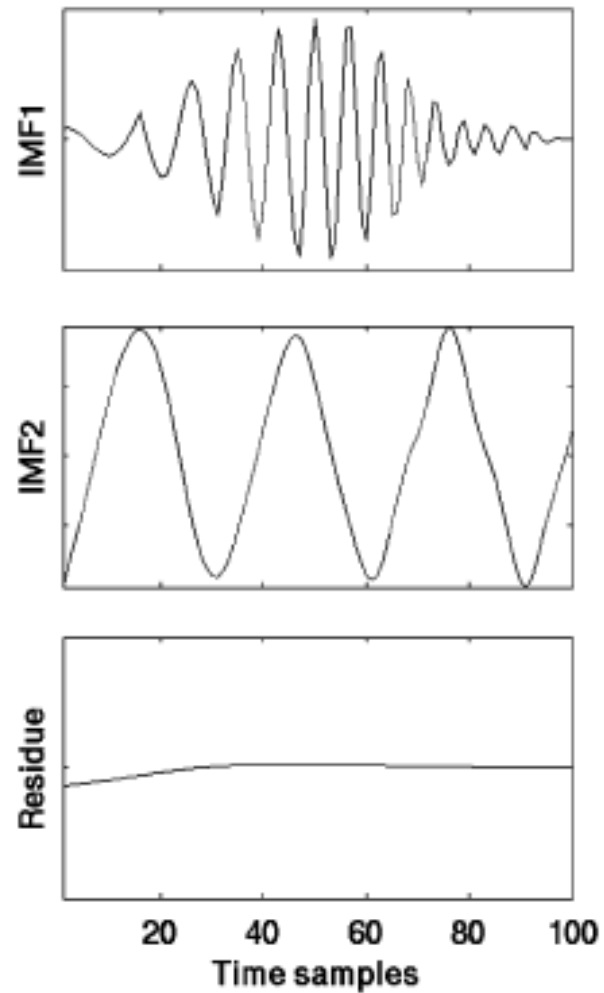
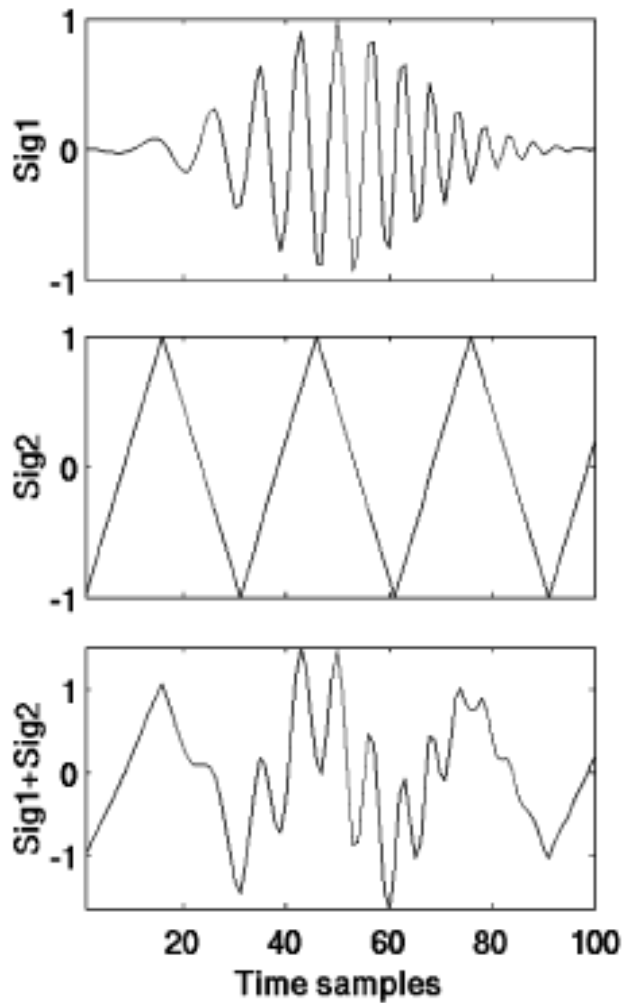
SIFTING



SIFTING



TOY EXAMPLE



COMPARISONS BETWEEN FOURIER, WAVELET AND EMD

Table 1. Comparisons between Fourier, Wavelet and Hilbert-Huang Transform in Data analysis.

	Fourier	Wavelet	HHT
basis	<i>a priori</i>	<i>a priori</i>	adaptive
frequency	deconvolution: global, uncertainty	deconvolution: regional, uncertainty	differentiation: local, certainty
presentation	energy- frequency	energy-time- frequency	energy-time- frequency
nonlinear	no	no	yes
non-stationary	no	no	yes
feature extraction	no	discrete: no continuous: yes	yes
theoretical base	complete	complete	empirical

ADVANTAGES OF EMD

- ◉ EMD can reduce any data, from non-stationary and nonlinear processes, into simple independent intrinsic mode functions;
- ◉ Since the decomposition is based on the local characteristic time scale of the data and only extrema are used in the sifting process, it is local, self-adaptive, concretely implicational and highly efficient;
- ◉ The IMFs have a clear definition of instantaneous frequency as the derivative of the phase function, so Hilbert transformation can be applied to the IMFs, allowing us to analyze the data in a time-frequency space.

ADVANTAGES OF EMD

- ◉ EMD can reduce any data, from non-stationary and nonlinear processes, into simple independent intrinsic mode functions;
- ◉ Since the decomposition is based on the local characteristic time scale of the data and only extrema are used in the sifting process, it is local, self-adaptive, concretely implicational and highly efficient;
- ◉ The IMFs have a clear definition of instantaneous frequency as the derivative of the phase function, so Hilbert transformation can be applied to the IMFs, allowing us to analyze the data in a time-frequency space.

一个新的原油定价理论

- ◎ A New Approach for Crude Oil Price Analysis Based on Empirical Mode Decomposition

Xun Zhang, Shouyang Wang and K.K. Lai

Energy Economics, Vol 29 (2008)

能源经济学的公开问题一

今天的原油价格由什么决定？

能源经济学的公开问题二

如果原油价格不全由供求关系决定，那供求关系对价格的形成占多大比重？

APPLICATIONS OF EMD IN SOCIAL SCIENCE BEFORE 2005

- ◉ Limited to natural and engineering fields
 - ◉ Non-destructive evaluation for structural health monitoring
 - ◉ Vibration, speech, and acoustic signal analyses
 - ◉ Earthquake engineering
 - ◉ Bio-medical applications
 - ◉ Cosmological gravitational wave and planets hunting
- ◉ Only two significant applications in social sciences
 - ◉ Financial market data analysis
 - examine the changeability of the markets by Huang et al. (2003).
 - ◉ Epidemics transmission
 - show the existence of a spatial-temporal traveling wave in the incidence of dengue hemorrhagic fever in Thailand

DATA: WTI OIL PRICE

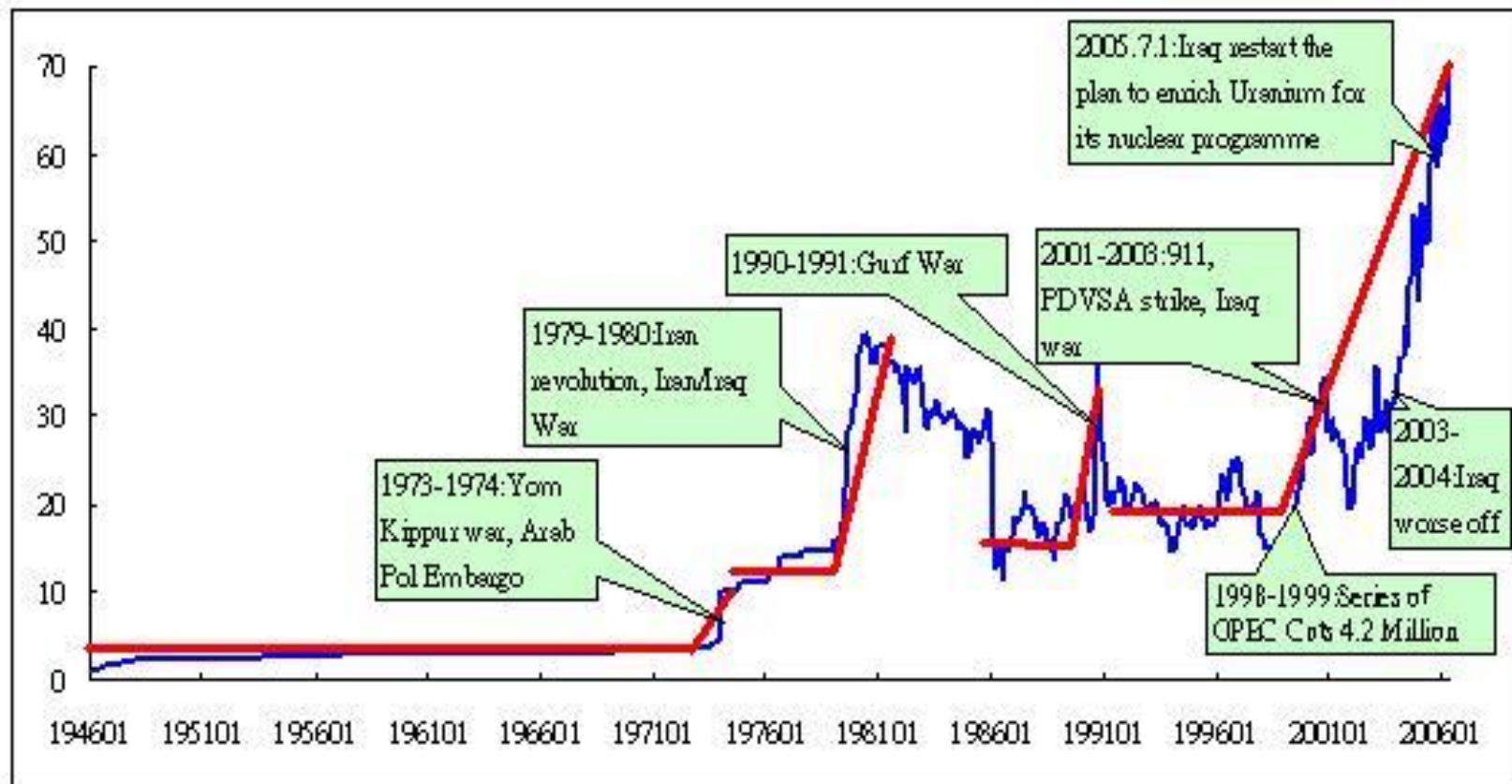
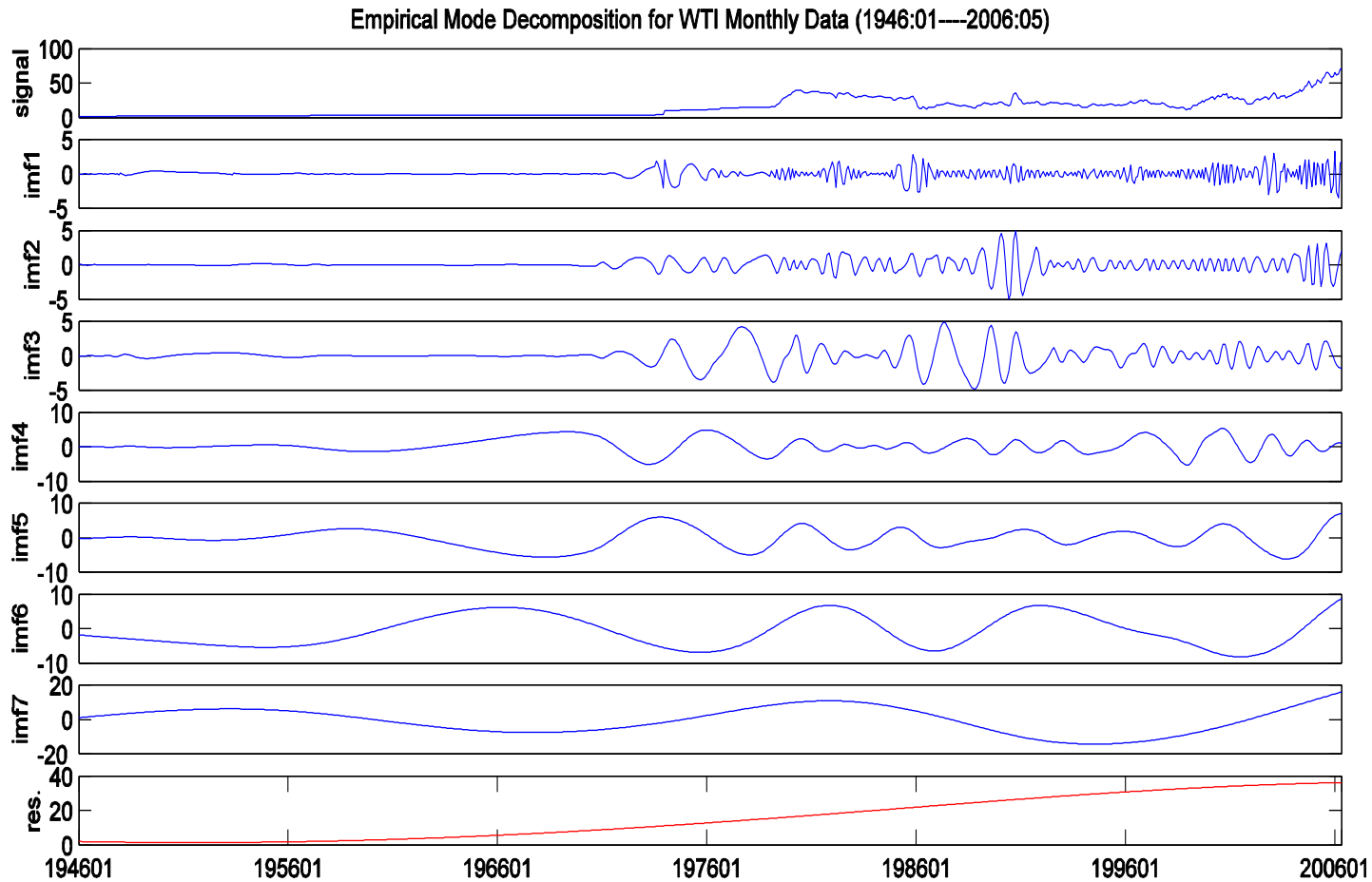


Figure 1 The observed time series of WTI oil price from Jan.1946 to May, 2006

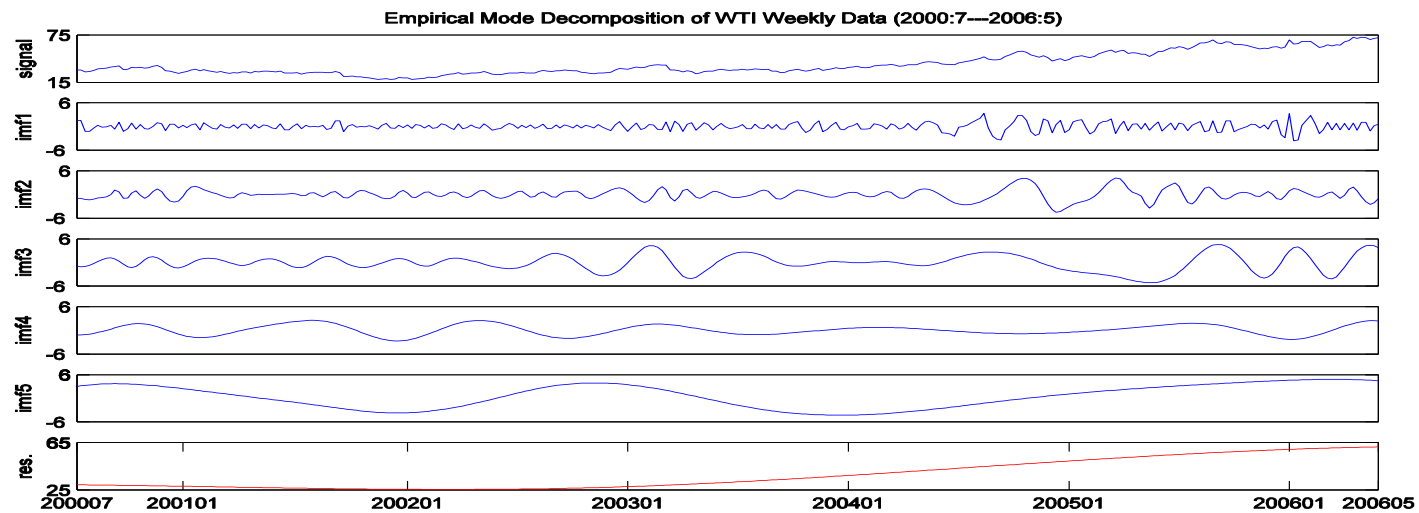
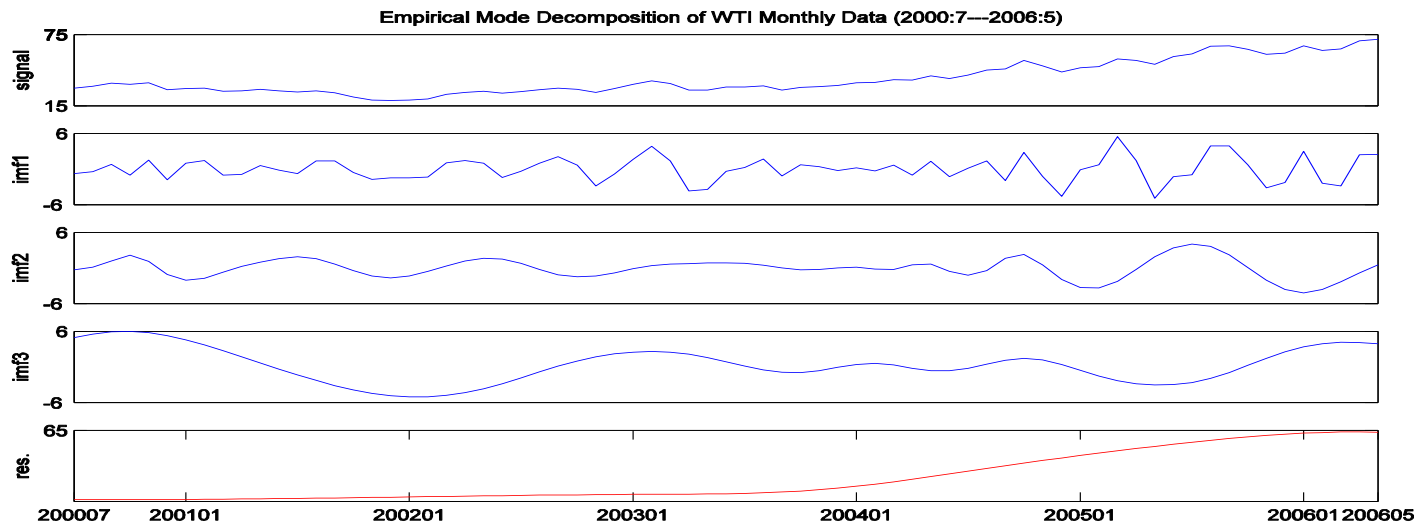
DATA SETS

- ◉ Three data sets are used.
- ◉ Data set 1: WTI spot oil price, monthly data, 1946.1–2006.5
- ◉ Data set 2: WTI spot oil price, monthly data, 2000.7–2006.5
- ◉ Data set 3: WTI spot oil price, weekly data, 2000.7–2006.5
- ◉ For brevity, I will only present the results for data set 1 in details today.
- ◉ In fact, the results derived from the three data sets are very consistent. Except for that the high frequency data cover more information about high frequency fluctuations in the original time series, whereas the data with longer range enable components with much longer periods be extracted.

WTI OIL PRICE DECOMPOSITION



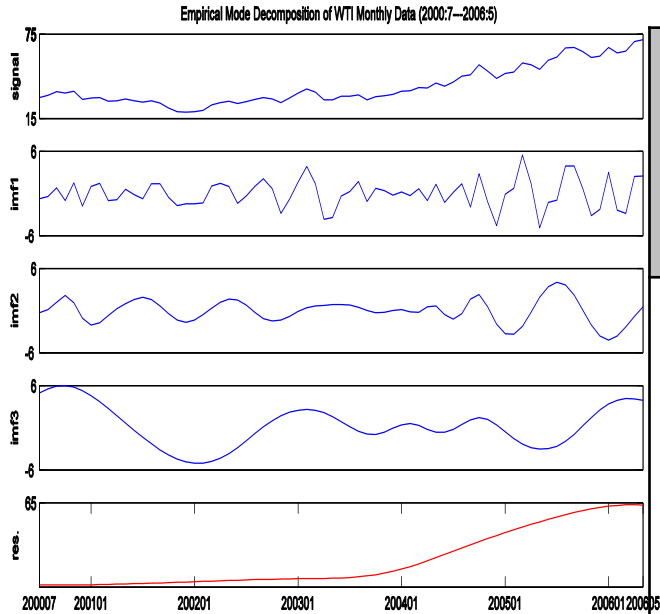
DECOMPOSITION



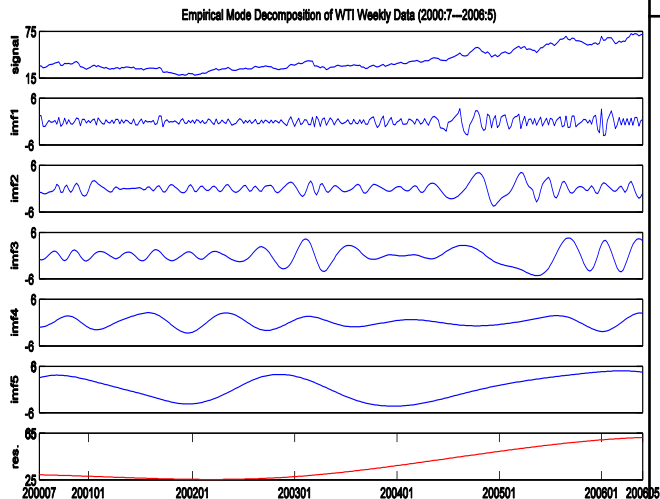
IMF STATISTICS

	Mean Period(Mon th)	Pearson Correlatio n	Kendall Correlat ion	Varianc e	Variance as a % of Observed	Variance as a % of (Σ IMFs +residua l)
Observed				181.40		
IMF1	5.41	0.10**	0.06**	0.54	0.30%	0.26%
IMF2	11.51	0.07	0.06**	0.65	0.36%	0.32%
IMF3	20.71	0.11**	0.08**	0.79	0.44%	0.39%
IMF4	42.65	0.15**	-0.01	2.21	1.22%	1.08%
IMF5	80.56	0.25**	0.10**	3.17	1.75%	1.55%
IMF6	241.67	0.24**	0.04	7.38	4.07%	3.62%
IMF7	362.50	0.32**	-0.03	39.08	21.55%	19.16%
Residue		0.81**	0.75**	150.18	82.79%	73.61%
Sum					112.47%	100.00%

IMF STATISTICS



	Mean Per iod	Pearso n Cor relat ion	Kendal l Cor relat ion	Varianc e	Variance as a % of Obse rved	Variance as a % of (Σ IMF s+resi dual)
Observed				143.24		
IMF1	3.94	0.2	0.16	3.76	2.62%	2.43%
IMF2	8.88	-0.07	-0.33	2.44	1.70%	1.58%
IMF3	14.2	0.23	0.23**	8.08	5.64%	5.22%
Residue	71	0.96**	0.66**	140.3	97.95%	90.76%
Sum					107.91 %	100.00%



Observed				143.65		
IMF1	3.5	0.09	0.076*	0.93	0.65%	0.68%
IMF2	8.32	0.12	0.059	1.78	1.24%	1.30%
IMF3	19.25	-0.04	0.091*	3.87	2.70%	2.83%
IMF4	44	0.12*	0.066	1.56	1.09%	1.14%
IMF5	102.6 7	0.45**	0.355* *	7.88	5.48%	5.76%
Residue		0.95**	0.794* *	120.74	84.05%	88.28%
Sum					95.21%	100.00%

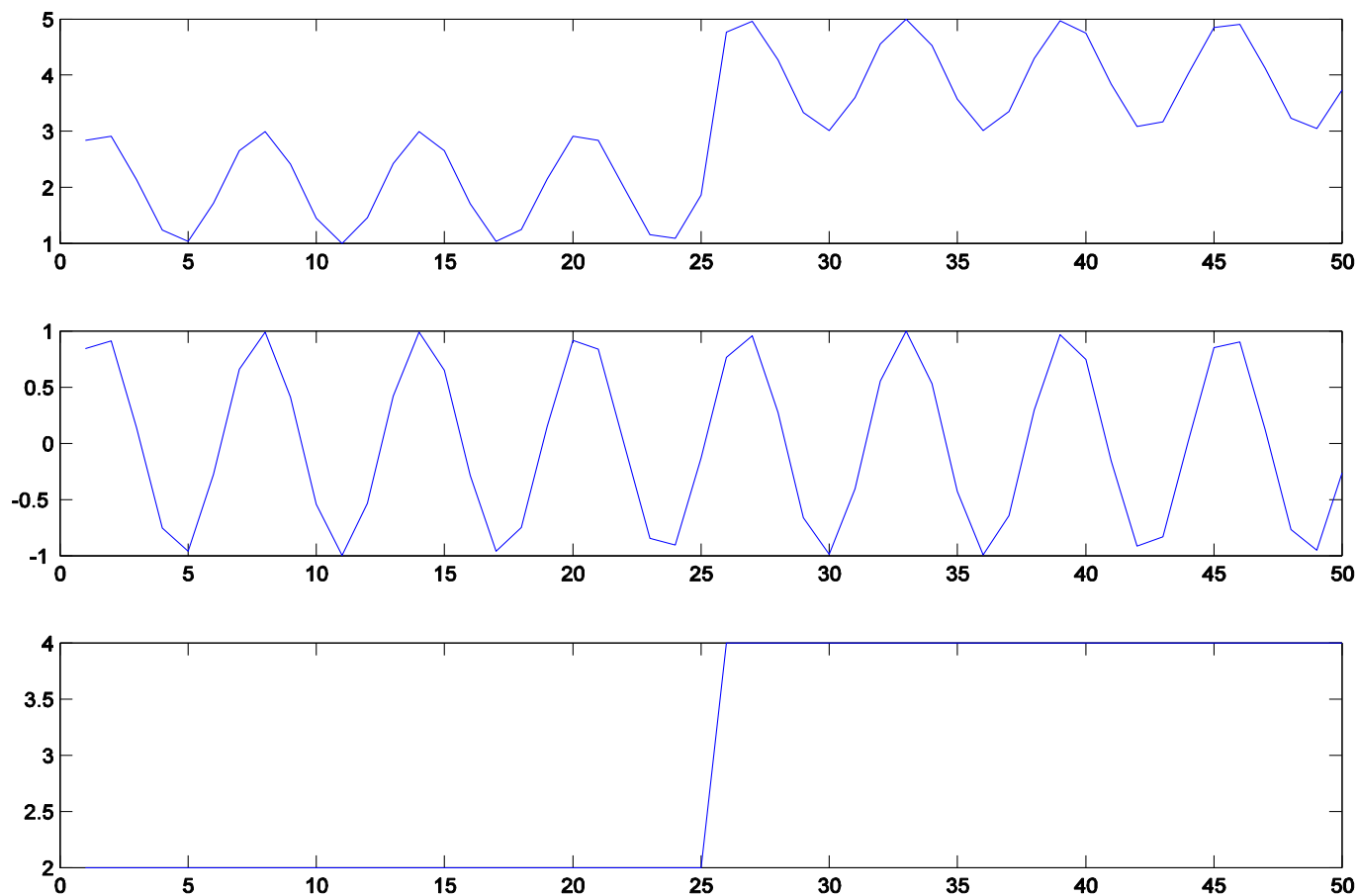
FEATURES

- ◉ The dominant mode of the observed data is not any IMF but the residue.
 - Both Pearson and Kendall coefficients: more than 0.75 in Table 1, 0.66 in Table 2 and 0.79 in Table 3.
 - Variances of the residue: account for more than 73% of the total variability. The highest one is even more than 90%.
- ◉ Although the increasing trends are very similar for the three residues, their values may differ at the same points of time due to different ranges of the original data series.
- ◉ One difference between weekly and monthly data

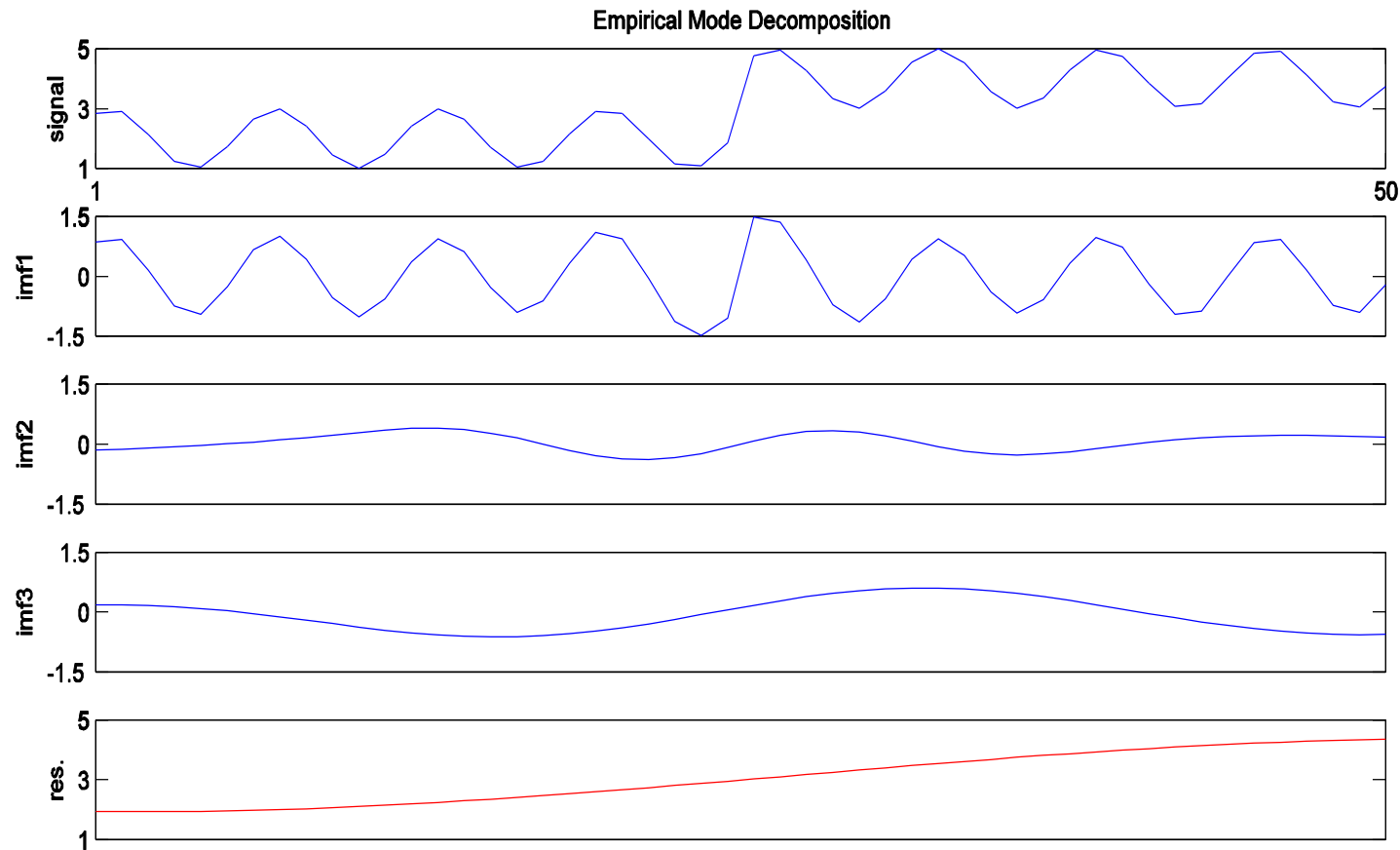
WHY COMPOSITION?

- ◉ Scale-mixing in Decomposition
- ◉ Help focus on components of different characteristics

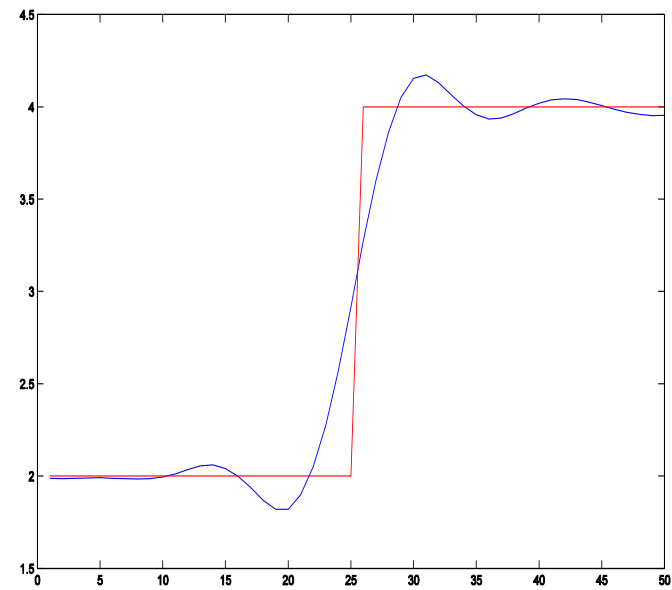
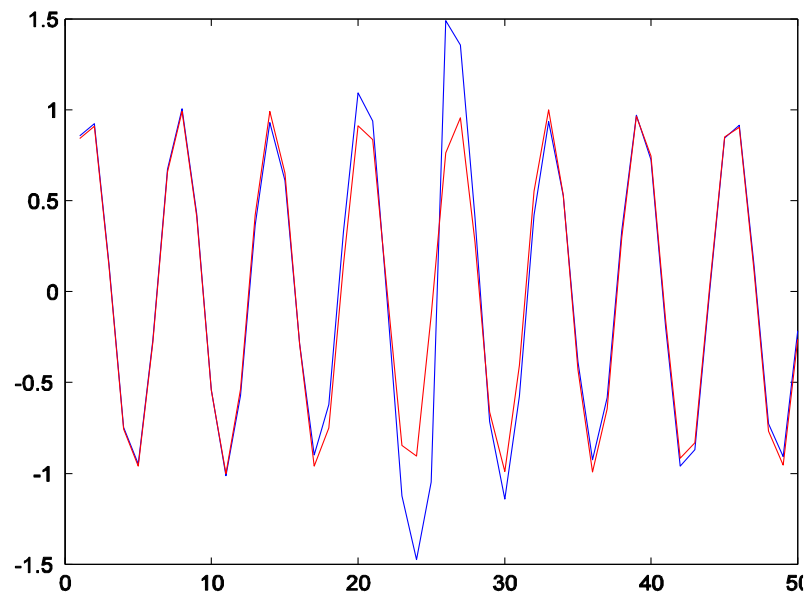
EMD WITH INTERMISSION(1): ORIGINAL SIGNAL



EMD WITH INTERMISSION(2): DECOMPOSITION RESULTS



EMD WITH INTERMISSION(3):COMPARISON



COMPOSITION ALGORITHM- EMD BASED FILTER

- ◉ EMD can be used as a filter to separate high frequency (fluctuating process) and low frequency (slowing varying component) modes.
- ◉ Algorithm:
 - Computing the mean of the sum of c_1 to c_i for each component (except for the residue);
 - Using t-test to identify for which i the mean significantly departs from zero;
 - Once i is identified as a significant change point, partial reconstruction with IMFs from this to the end, is identified as the slow-varying mode and the partial reconstruction with other IMFs is identified as the fluctuating process.

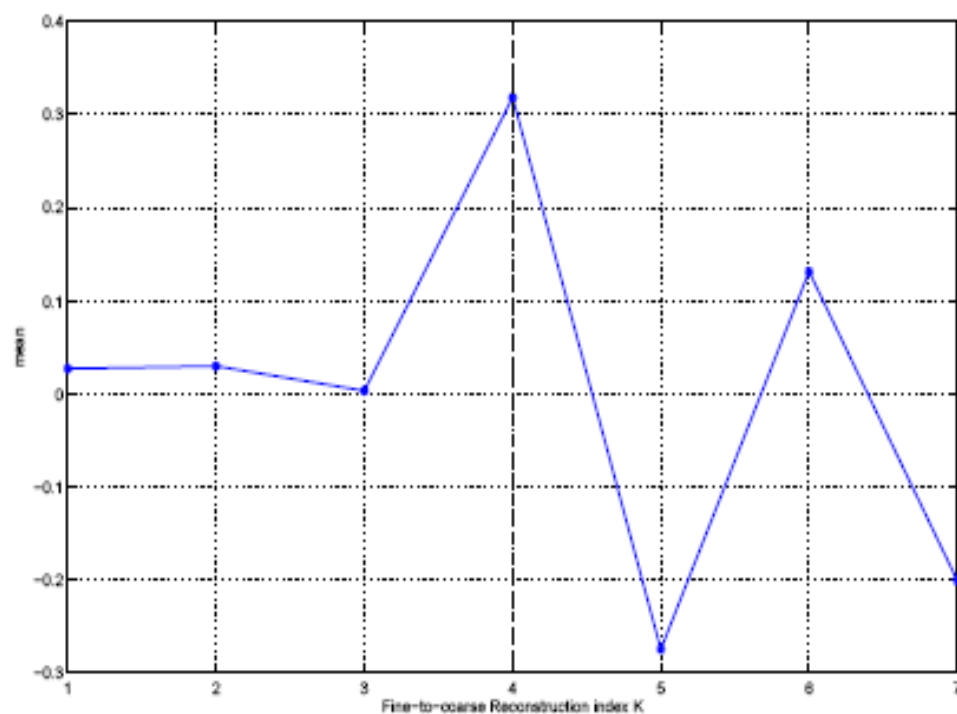
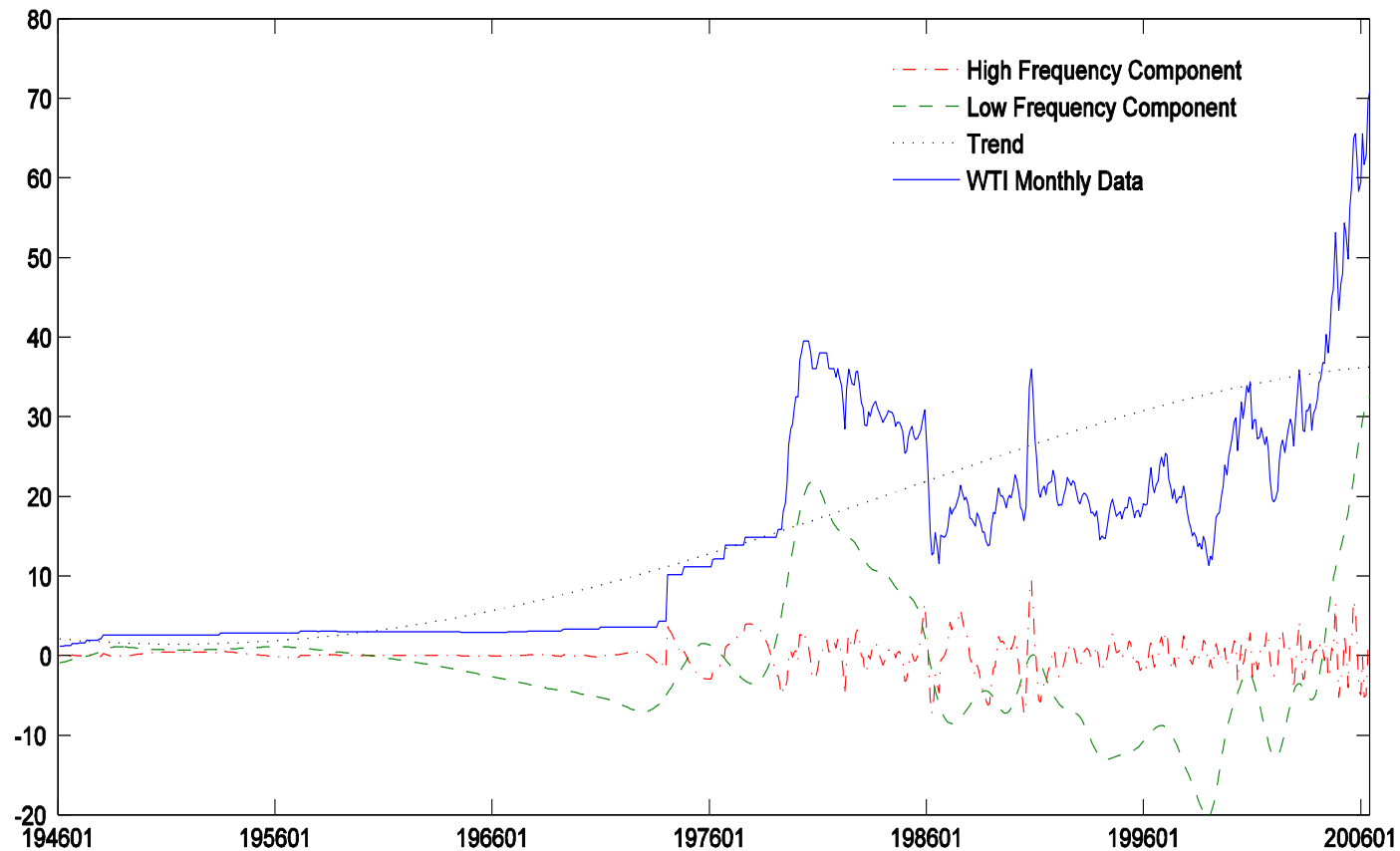


Figure 5: The mean of the fine-to-coarse reconstruction as a function of index K. The vertical dash-line at $K = 4$ indicates that the mean departs significantly from zero ($p < 0.01$).

THREE COMPONENTS



TREND: DETERMINISTIC FORCE IN THE LONG RUN

- The deterministic force for oil price evolution in the long run.
 - holds a high correlation with the original price
 - accounts for more than 70% percent of variability,
- The long term trend of crude oil price may be determined by global economic development
 - The continuing increasing trend is consistent with the economic development of the world.

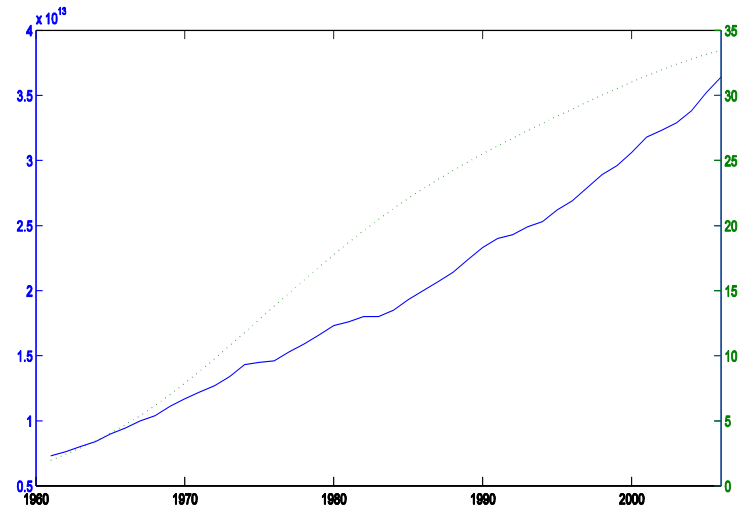
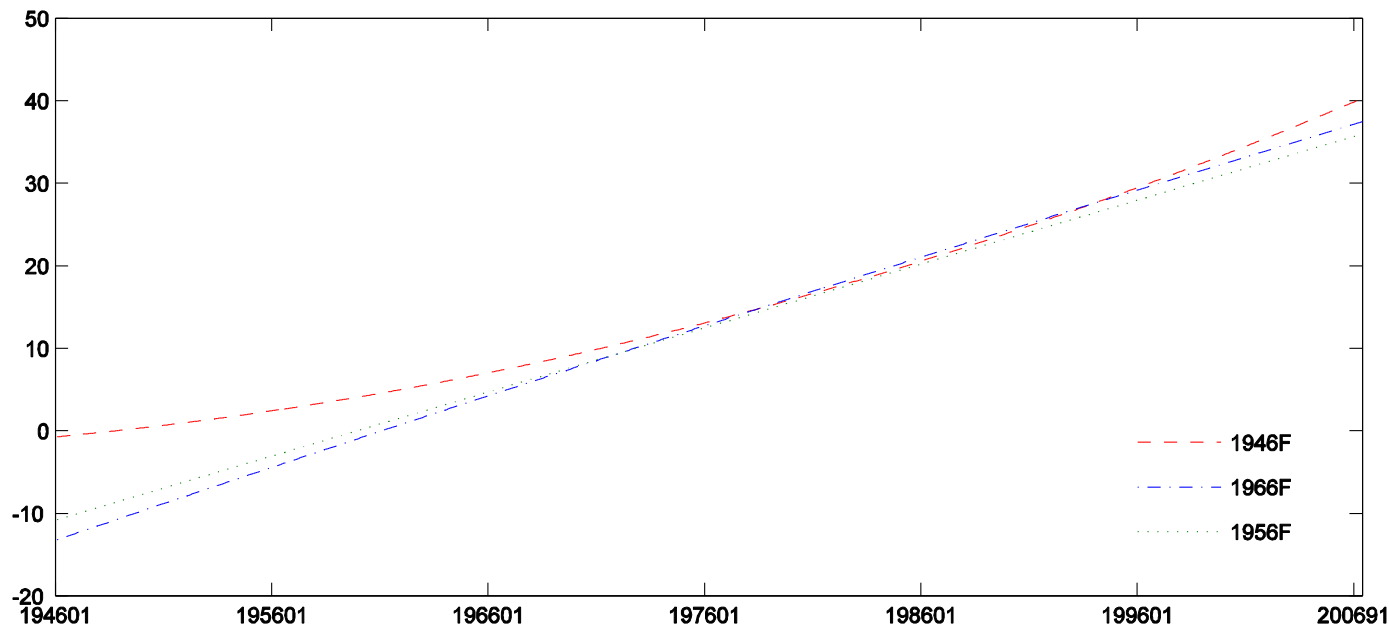


Figure World GDP and the trends in oil price. The solid line is the annual world GDP measured by constant 2000 US\$ from 1960 to 2005. The data is from the World Economic Indicators Database. The dotted line is the trend from 1960 to 2005. The yearly trend value is defined as the average of monthly residue in the same year

TREND: IS THERE A SUDDEN STRUCTURE CHANGE IN 1973?

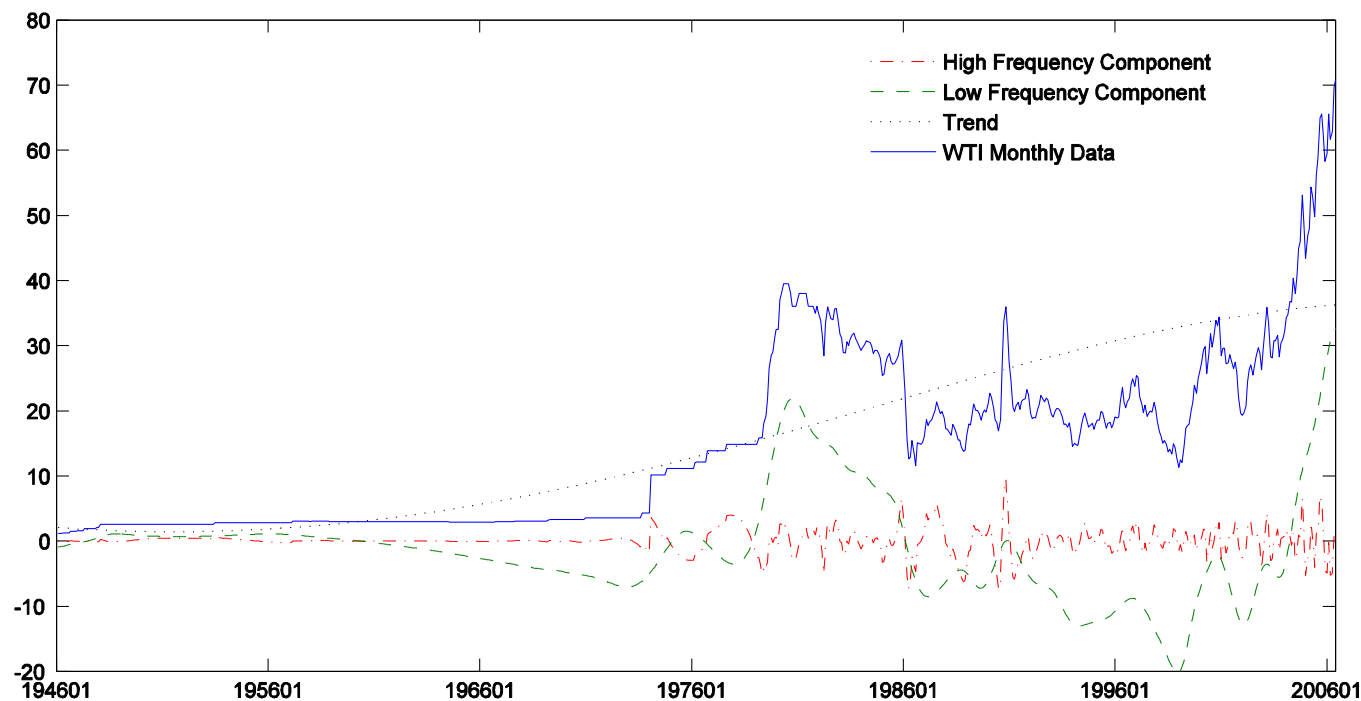
- The trend line can be fitted accurately with a quadratic function. But it will drift while changing the sample range due to the end effect in decomposition. This finding is similar to what Pindyck (1999) found for real crude oil price.
- The long term trend of oil price changes continuously with time.



EFFECTS OF SIGNIFICANT EVENTS: THE CAUSES OF LARGE FLUCTUATIONS IN MEDIUM TERM

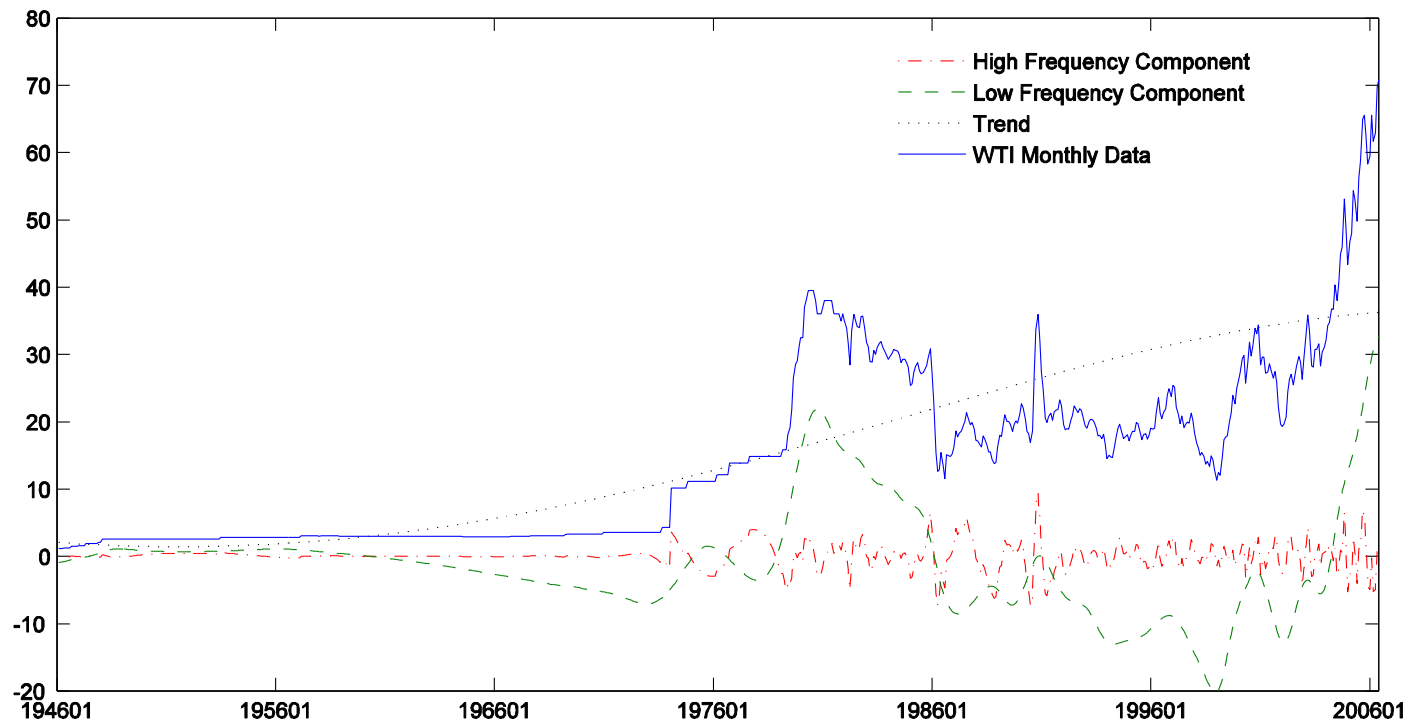
- Described by IMF4 to IMF7

- The mean periods range from 3.5 years to 30 years
- The amplitudes at some data points could be more than \$10 or even higher, suggesting the effects of some significant events on oil price may be very serious. .



NORMAL MARKET DISEQUILIBRIUM: IMPORTANT FOR SHORT TERM FORECASTING

- Have no serious impact on oil price -generally within \$5.
- But these events are becoming more and more frequent and have become the fundamental impetus for oil prices going up recently.



ASSUMING CRUDE OIL PRICE IS COMPOSED OF THE THREE COMPONENTS IS REASONABLE

■ residue

- It is also described as the “trend” in EMD, could represent the major trend of oil price in the long run.

■ low frequency component

- It can be treated as the effect of significant events. It is the main reason for the dramatic oil price variability in the medium term.

■ high frequency component

- It should be explained as normal market fluctuations or events which only bring a short term impact to crude oil price.

the price of \$70.84/barrel in May 2006 =

a trend price (\$36.26)+

a significant event price (\$32.63)+

a normal fluctuation of \$1.21.

A NEW THEORY OF CRUDE OIL PRICING

Crude Oil Price

■ Composition

- a long term trend price

- a shock from significant events

- a short term fluctuation

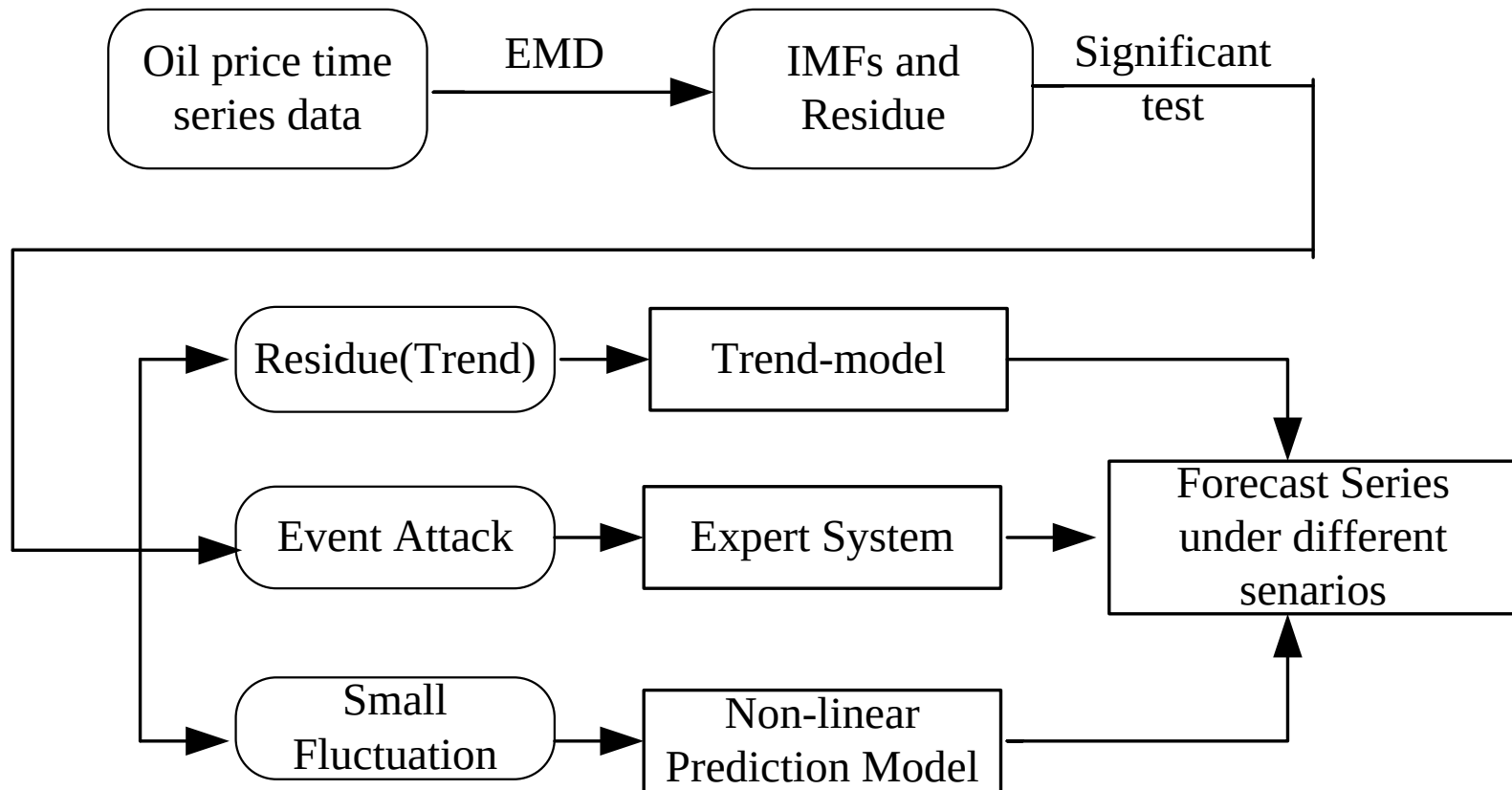
■ The features of crude oil price in different terms

Oil price in the long run is basically determined by the trend, which changes continuously and keeps around the long term mean.

The sharp downs or ups in oil prices are triggered by unpredictable and significant events, the impact of which may endure for several years.

The small fluctuations in the short term are mainly driven by normal market activities or some small events which can not have a serious influence on oil markets.

A PROPOSED FRAMEWORK FOR CRUDE OIL PRICE PREDICTION BASED ON EMD



一个基于EMD的 事件分析方法

ESTIMATING THE IMPACT OF EXTREME EVENTS ON CRUDE OIL PRICE: AN EMD-BASED EVENT ANALYSIS METHOD

Xun Zhang, Lean Yu and Shouyang Wang
Energy Economics, Vol. 31 (2009)

OUTLINE



Introduction

Methodology Formulation

Case Study

Conclusions

INTRODUCTION

- ◉ The forecasting of future prices is one of the most fundamental one among many issues related to crude oil markets.
- ◉ However, practical performances of most crude oil price forecasting models did not show significant improvements over the simple random walk model according to forecast evaluations.

INTRODUCTION

- ◉ Previous studies have summarized some reasons that lead to failures of predictions.
- ◉ An important issue is **the evident impacts of extreme and irregular events on crude oil markets.**
- ◉ Some studies suggested to incorporate an event evaluation mechanism in forecasting models to improve the performance of forecasting. (Wang et al. (2005), Yu et al. (2007a, 2007b) and Zhang et al. (2008)).

INTRODUCTION

- ◉ Traditional event evaluation methods
 - Event study
 - Intervention analysis
- ◉ Both the two traditional methods consider a time series as the sum of a normal evolution part and an exceptional shock, and then they implement the “divide and conquer” strategy to separately modeling the two components.
- ◉ But as addressed by Huang et al. (1998), although a complex time series can be treated as sums of several simple oscillation patterns, generally **the number of these patterns is more than two.**

INTRODUCTION

- ◉ Furthermore, an extreme event, as **an exceptional shock**, may make different types of impulses to these patterns.
- ◉ Therefore, a feasible decomposition method which can extract the inherent oscillations from the time series will be more useful for exploring extreme event's impact.
- ◉ We apply a competitive decomposition algorithm, empirical mode decomposition (EMD, Huang et al., 1998), to event analysis and formulate **an innovative EMD-based extreme event analysis approach**.

OUTLINE



Introduction



Methodology Formulation



Case Study



Discussions and conclusions

EMD

- A promising nonlinear, non-stationary data processing method proposed by Huang et al. (1998).
- It considers the real time series as fast oscillations superimposed on slow oscillations. Those oscillations in data are extracted based on the principle of local scale separation, and are approximated by “intrinsic mode functions” (IMF).
- Only extrema are needed in the sifting process, so it is local and adaptive;
- The inductive basis is posterior rather than prior and, therefore, it is physically meaningful.
- An improvement of EMD--- **Ensemble EMD** is used in this study.

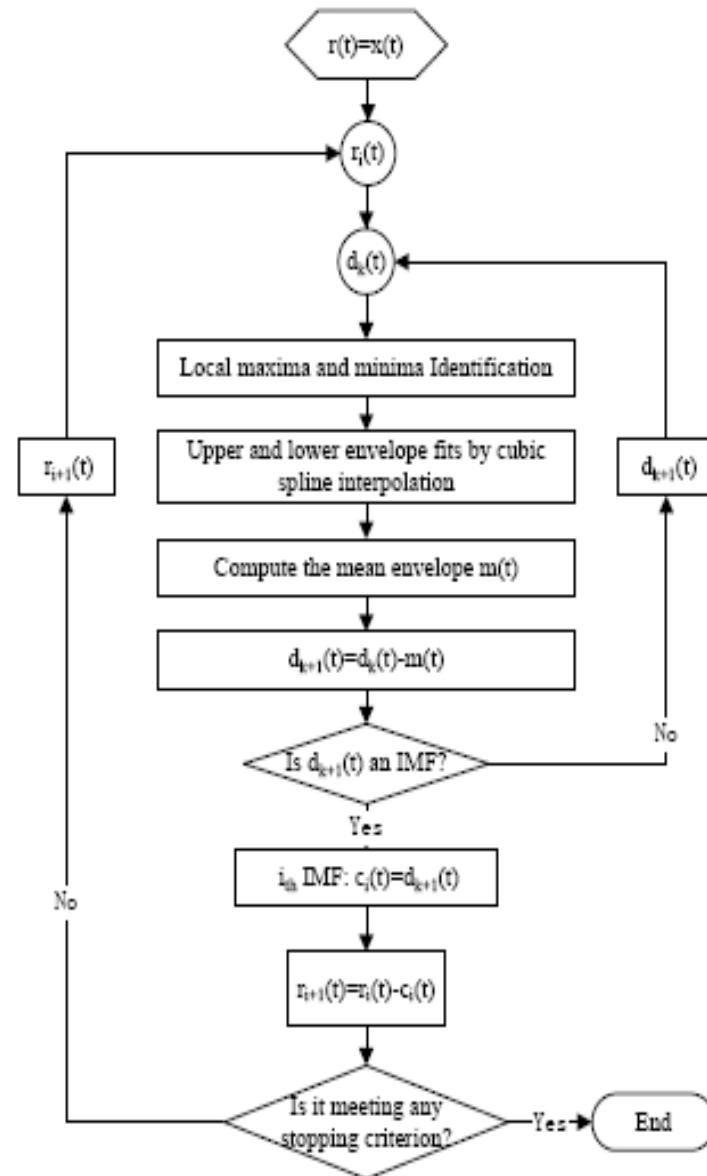
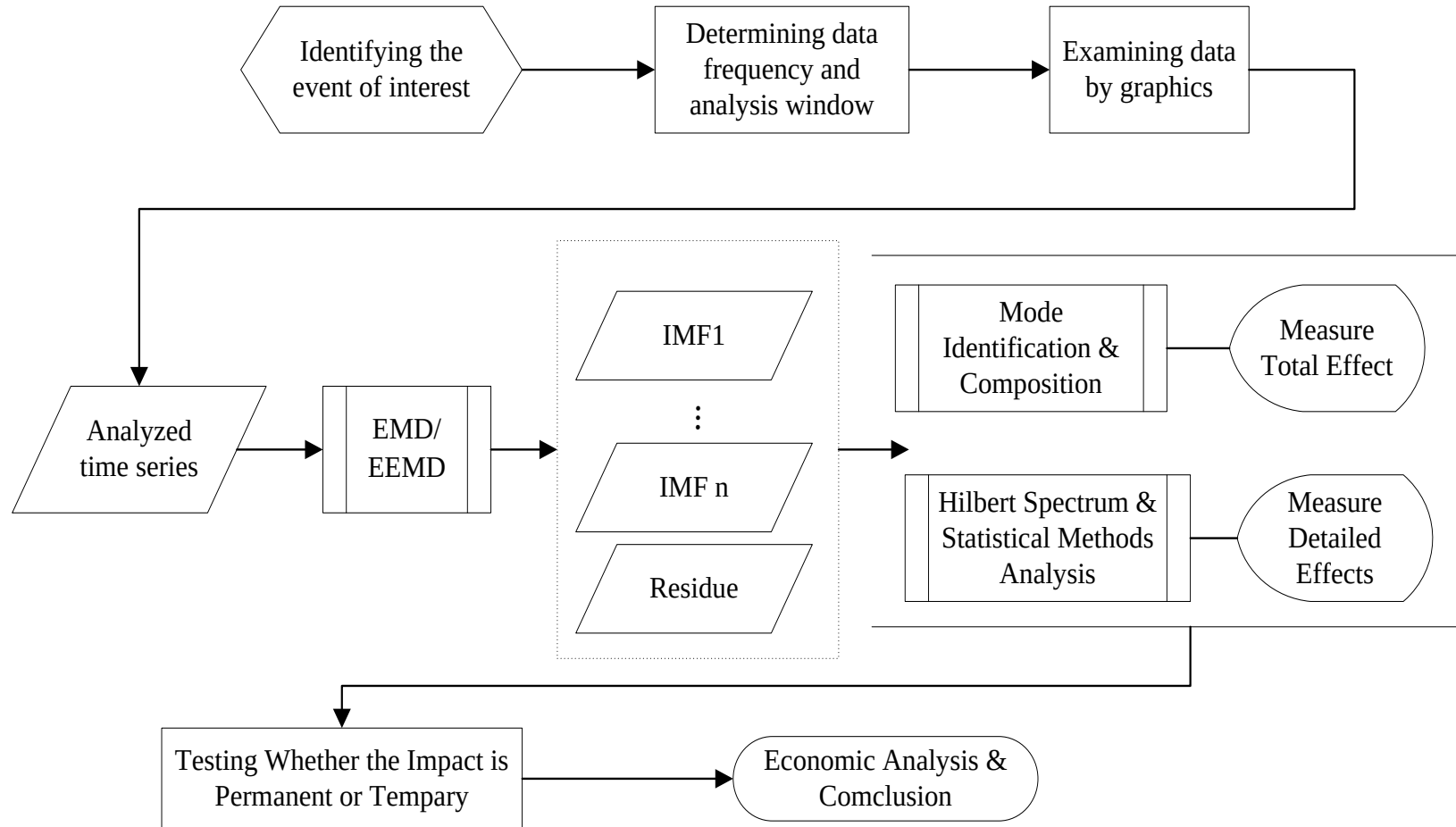


Fig. 2 Shifting Process of EMD

EMD-BASED EVENT ANALYSIS

- ◉ Decompose the original series into some independent modes, and then identify the fluctuations in the modes, which represent the effect of the event of interest.
- ◉ The mode, after getting rid of noises and trends, is used to approximate the pattern and degree of the impact caused by the event.
- ◉ Other modes provide useful information about changes triggered by the event at different time scales, such as volatilities differences between high and low frequency modes.
- ◉ **Advantages: precise, multi-scale, objective**

EMD-BASED EVENT ANALYSIS APPROACH



OUTLINE



Introduction

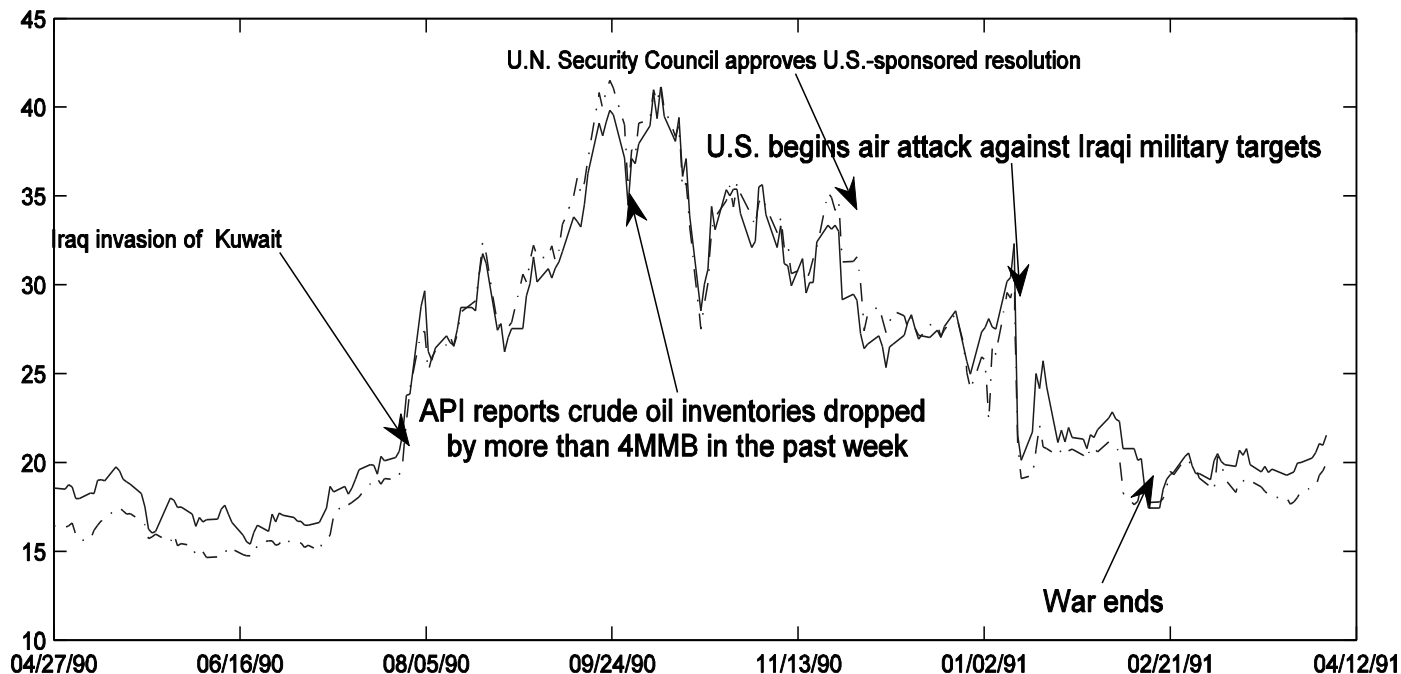
Methodology Formulation

Case Study

Discussions and conclusions

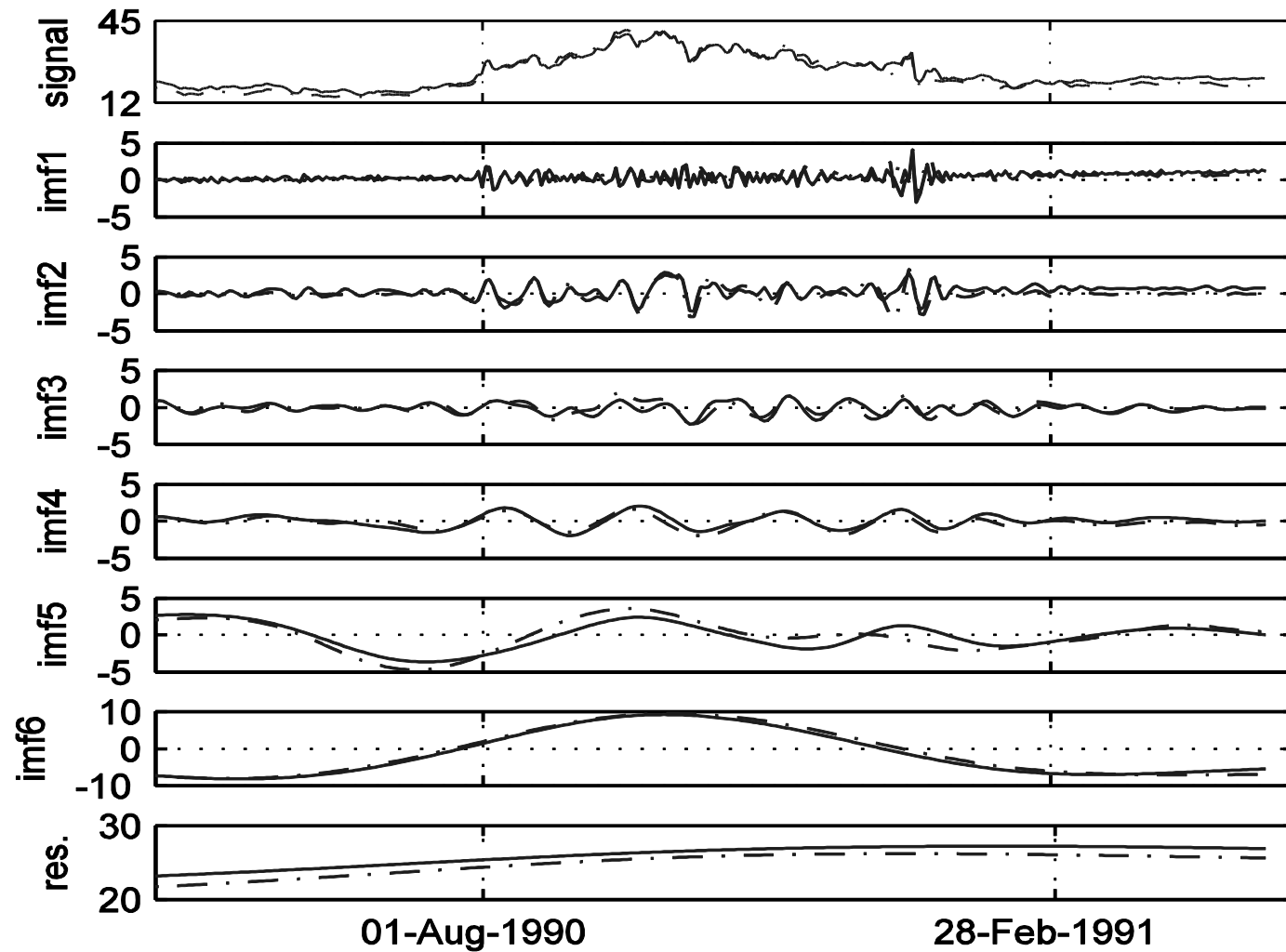
CASE I: THE PERSIAN GULF WAR IN 1991

- The event window: from August 2, 1990 to February 28, 1991.
- The estimation window: three months before and after the war: from April 30, 1990 to August 1, 1990 and from March 1, 1991 to May 31, 1991.
- Data: two benchmark crude oil prices - West Texas Intermediate spot



- The daily WTI (solid line) and Brent (dotted line) crude oil prices from March 30, 1990 to May 31, 1991

DECOMPOSING DATA BY ENSEMBLE EMD

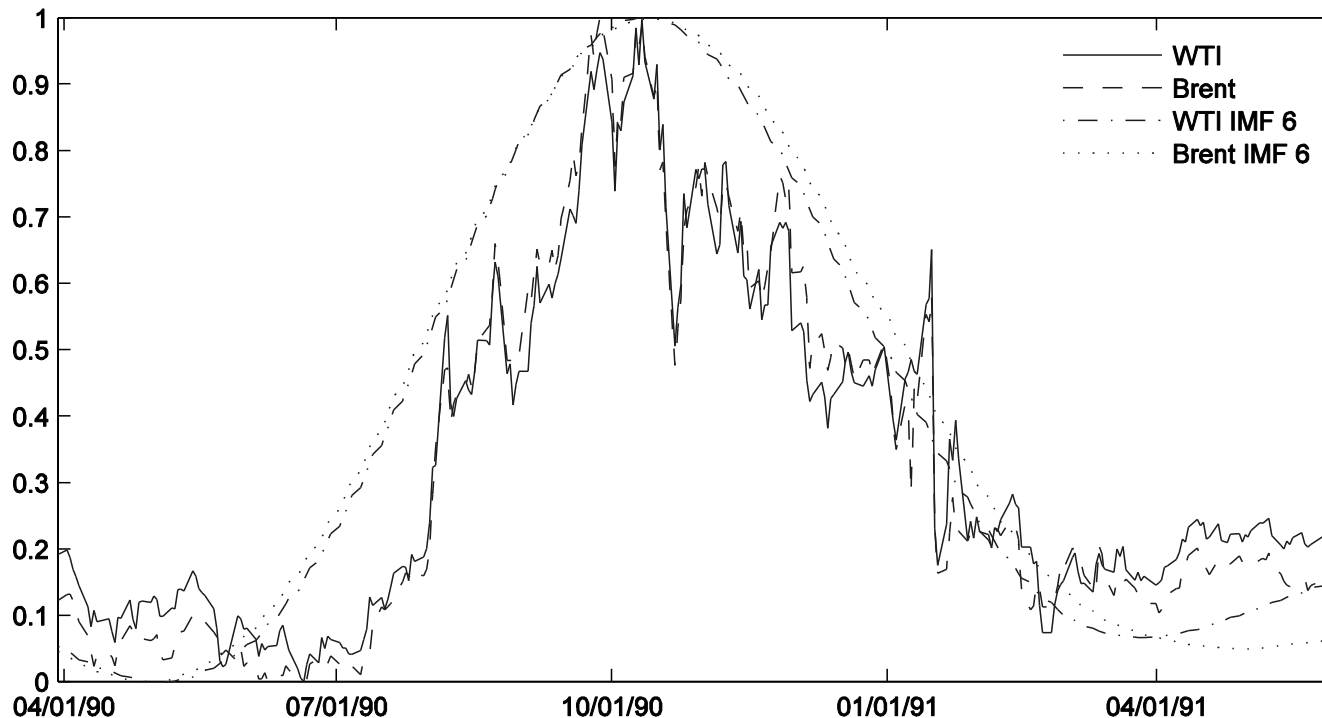


MEASURES OF IMFS AND THE RESIDUE FOR THE WTI AND BRENT DAILY CRUDE OIL PRICES IN 1991 PERSIAN GULF WAR

	WTI			Brent		
	Mean Period	Correlation Coefficient	Variance as % of Observed	Mean Period	Correlation Coefficient	Variance as % of Observed
IMF1	3.39	0.18	2.37	3.48	0.14	1.65
IMF2	9.66	0.39	3.58	11.45	0.34	4.44
IMF3	16.19	0.28	2.31	20.27	0.32	2.42
IMF4	31.15	0.19	3.36	33.08	0.18	2.37
IMF5	82.42	0.59	5.58	148.4	0.84	7.12
IMF6	120.83	0.89	64.72	130.16	0.92	47.37
Residue		-0.34	0.91		-0.24	0.69

THE DOMINANT MODE DURING THE EVENT WINDOW

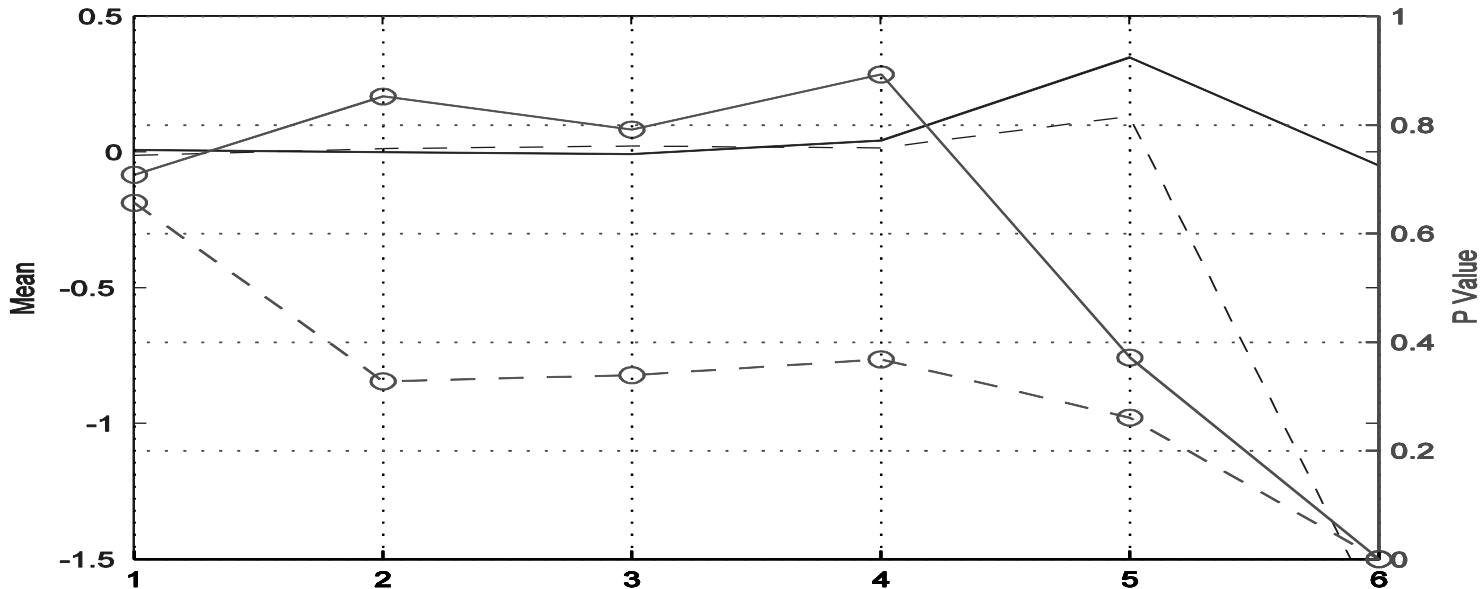
- ◉ IMF6s are identified as the dominant IMFs according to statistical measures.
- ◉ The distance between the first local minima and the following local maxima is used to measure the total impact of the invasion on crude oil prices : \$16



■ Fig. The normalized prices and IMF6s derived from EEMD.

KNOWLEDGE FROM THE HIGH FREQUENCY MODES

- The high frequency fluctuations in the price series are no larger than \$8.
- Most of the large and sudden jumps of the original prices are captured by IMF1s: the markets are with the mean-reversion trend in the short-term.
- The total effects of these high frequency IMFs are near zero.



- Fig. The mean of the fine-to-coarse reconstruction as a function of index K. At $K = 6$ the mean departs significantly from zero ($p < 0.01$). Solid line: mean of reconstruction (WTI); Dashed line: mean of reconstruction (Brent); Solid line with marker: P value (WTI); Dashed line with marker: P value (Brent).

DID THE EVENT AMPLIFY THE VOLATILITIES?

- At each time point, the energy is mainly distributed at low frequency.
- For the high frequency region, the energy clusters in the event windows are slightly denser than those in the estimation window.
- The energy in the event windows is higher than that in the estimation window.

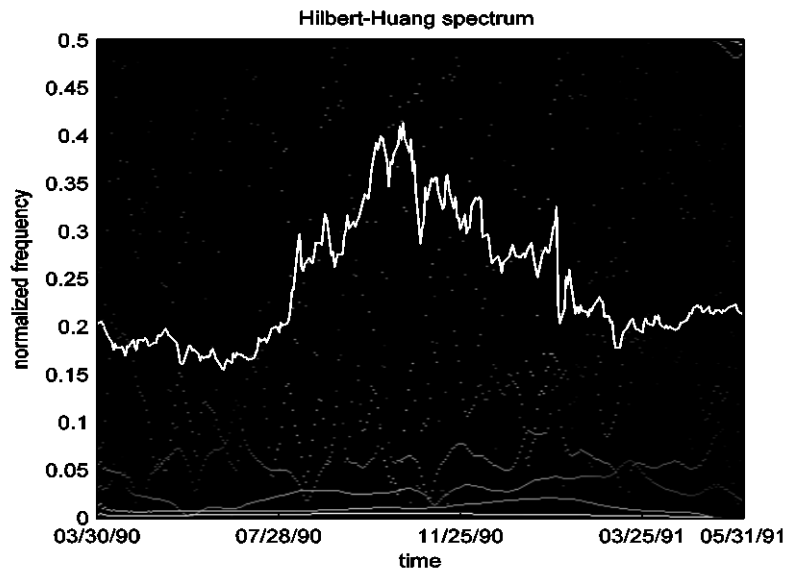


Fig. The Hilbert spectrum of IMFs

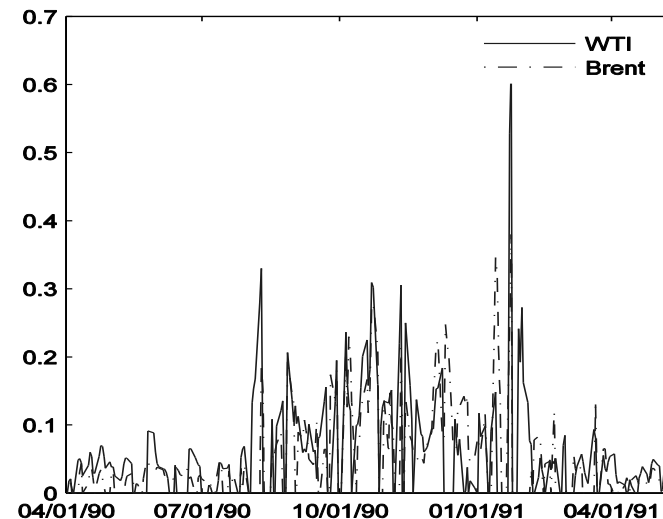


Fig. The marginal Hilbert spectrum for normalized frequency range [0.5,1].

TESTING WHETHER THE IMPACT IS TEMPORARY OR PERMANENT

- ◎ The BIC criterion detects eight structure breakpoints:
 - December 1973, July 1979, November 1982, December 1985, February 1989, **November 1992**, October 1999 and September 2004.
 - After the war, Iraq was prohibited to export any crude oil until 1995, the war did result in changes in the supply side of global crude oil markets.
 - Due to the **lag in effect**, the breakpoint didn't immediately appear after the end of the war.
 - It is reasonable to conclude that the Iraqi invasion of Kuwait and **the Persian Gulf War in early 1990s did have serious impacts on crude oil markets in the long run.**

ECONOMIC ANALYSIS

- ◎ The war had several types of impacts on crude oil prices:
 - (i) high frequency fluctuations with small amplitudes;
 - (ii) a large shock as a delta impulse, the time span of which is consistent with the event;
 - (iii) a level shift effect in the long run.
- ◎ Small fluctuations during the war do not have a long-term effect on crude oil prices.
- ◎ The total impact of the war on crude oil prices is like a delta impulse.
- ◎ Crude oil prices went upward prior to the war, and fell down after the uncertainties about future situations are cleared away. Here the change of prices satisfies the **crisis model** (Looney, 2003): prior to a crisis, oil prices generally go upward. With the advent of a crisis, future availability of crude oil is in doubt and traders add an uncertainty or "war premium" to the prices. The prices continued to increase. However, the price decreased dramatically after the situation was clear.

OUTLINE



Introduction

Methodology Formulation

Case Study

Conclusions

CONCLUSIONS

- ⦿ A new theory of crude oil price is introduced.
- ⦿ EMD is a powerful tool for complex data analysis.
- ⦿ In the proposed approach, the EMD method decomposes fluctuations of crude oil price series into different time scales, based on the local characteristics of the series itself.
- ⦿ Each mode (or the composition of a few modes) has clearly concrete implications.
- ⦿ High frequency IMFs with shorter periods correspond to the impacts of minor events on crude oil prices, while low frequency IMFs with longer periods correspond to the impacts caused by extreme events, and the final residue is the long-term trend.

CONCLUSION

- ◉ For the case of extreme events, there is a dominant mode (one single IMF or a composition of several IMFs) which outlines the pattern and degree of the impact, without considering sudden changes caused by minor events during the event, and factors, the scales of which are larger than the event window.
- ◉ Many traditional event analysis methods fail to deal with the overlapping events. The EMD's scale separation ability solves it efficiently. The effects of overlapping events with similar outbreak time but different durations are decomposed into different modes.

结论与启发

1. 中国的青年学生完全可以做出世界一流水平的研究工作；
2. 复杂数据分解方法为我们从事经济学研究提供了有力工具；
3. 好的研究选题+好的研究思路+好的研究方法
==》好的研究工作；
4. 其它复杂数据处理方法可为我们研究复杂经济系统提供可能（SSA等；数据分解+计量建模+深度学习==?）；
5. 许多全新的经济学研究问题有待我们去研究去攻克（新冠肺炎疫情给经济学研究带来的新问题、新挑战与新机会）；
6. 勇于去挑战“常识”或“不可能”，敢为人先（Interval Econometrics）

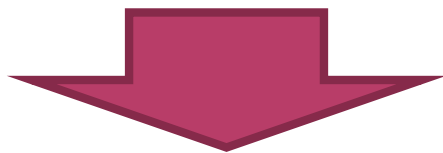
希望这个交流对同学们
有所启迪与帮助！

谢谢大家！

直接滤波方法DFA

信号提取(SIGNAL EXTRACTION)

- ◎ 对称滤子：过去、未来采用相等权重，输出无延迟
- ◎ 时间序列最右端无法使用



- ◎ 非对称滤子，在边界上与实际观测值有较大差距，数值的差别（幅值），转折点的延迟（相位）

模型基础的滤波方法的基本思想

- ◎ TRAMO/SEATS
- ◎ CENSUS X-12 ARIMA
- ◎ 利用模型对未来的预测，使用对称滤子
- ◎ 规则：单步的方差误差最小
- ◎ 实时估计的延迟－拐点的滞后
- ◎ 实时估计的噪声－结果不可靠

DFA的基本思想

- 预先定义的symmetric filter对原序列进行滤波，得到分析人员感兴趣的滤波 (Ideal trend)
- 对预先定义的非对称滤子 (ZPC滤子) 进行参数优化，优化准则是通过非对称滤子优化得到的滤波与与ideal trend之间的误差最小

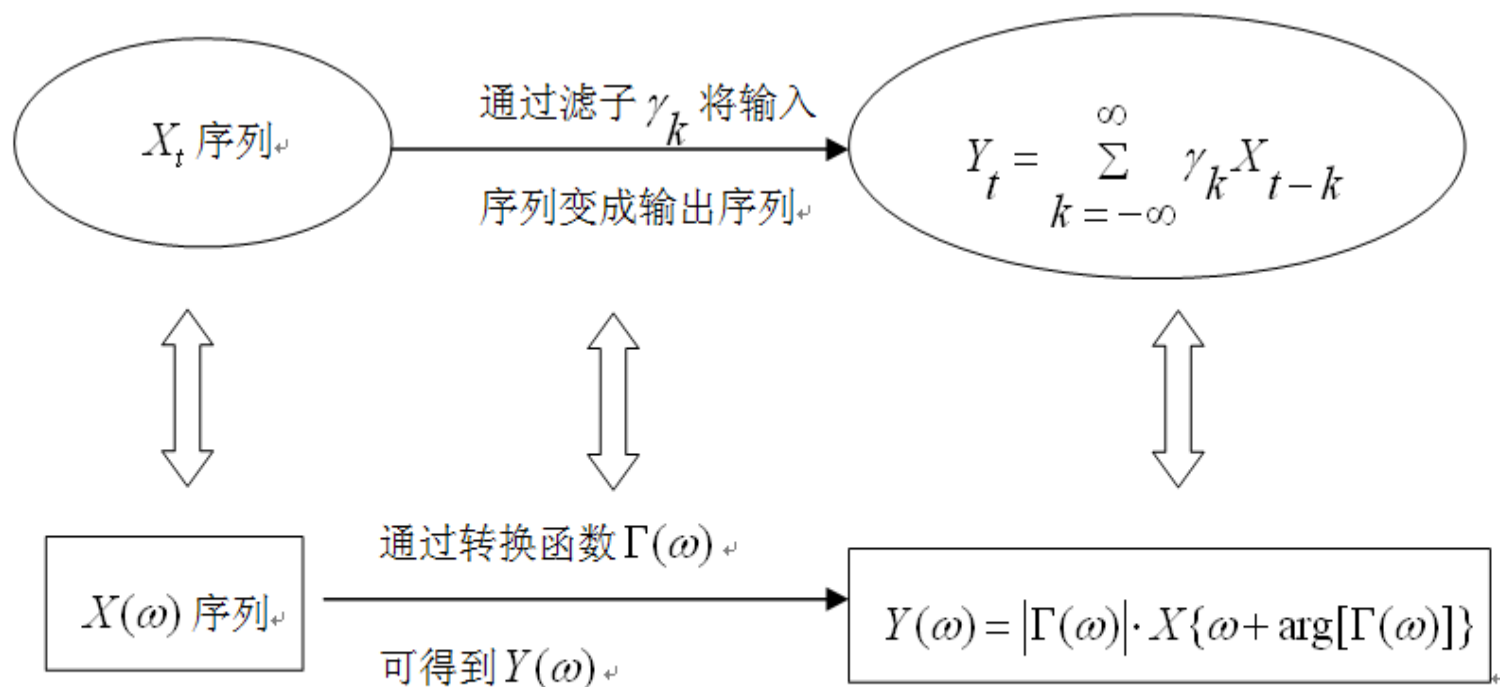


图 1 DFA 算法的对应关系

• 理论基础的变革：过滤误差而非预测误差

优化规则

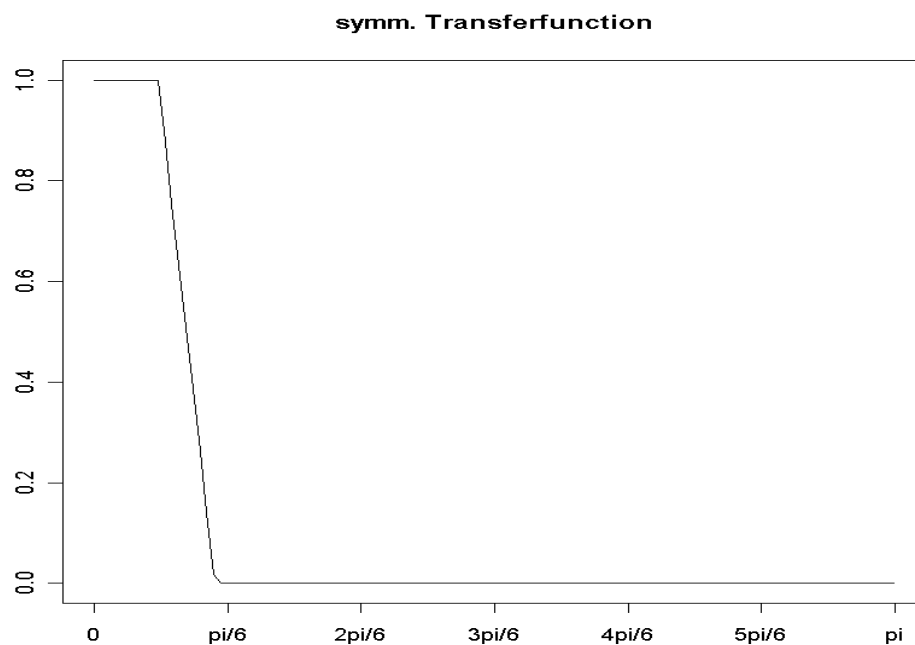
◎ 优化规则：

$$\min_{\hat{\Gamma}} \sum_{k=1}^{N/2} |\Gamma(\omega_k) - \hat{\Gamma}(\omega_k)|^2 I_{NX}(\omega_k)$$

- $\omega_k = 2\pi k/N$ ：离散的频率坐标集
- $\Gamma(\omega_k)$ ：对称算子的传递函数
- $\hat{\Gamma}(\omega_k)$ ：非对称算子的传递函数
- $I_{NX}(\omega_k)$ ：要过滤的时间序列周期

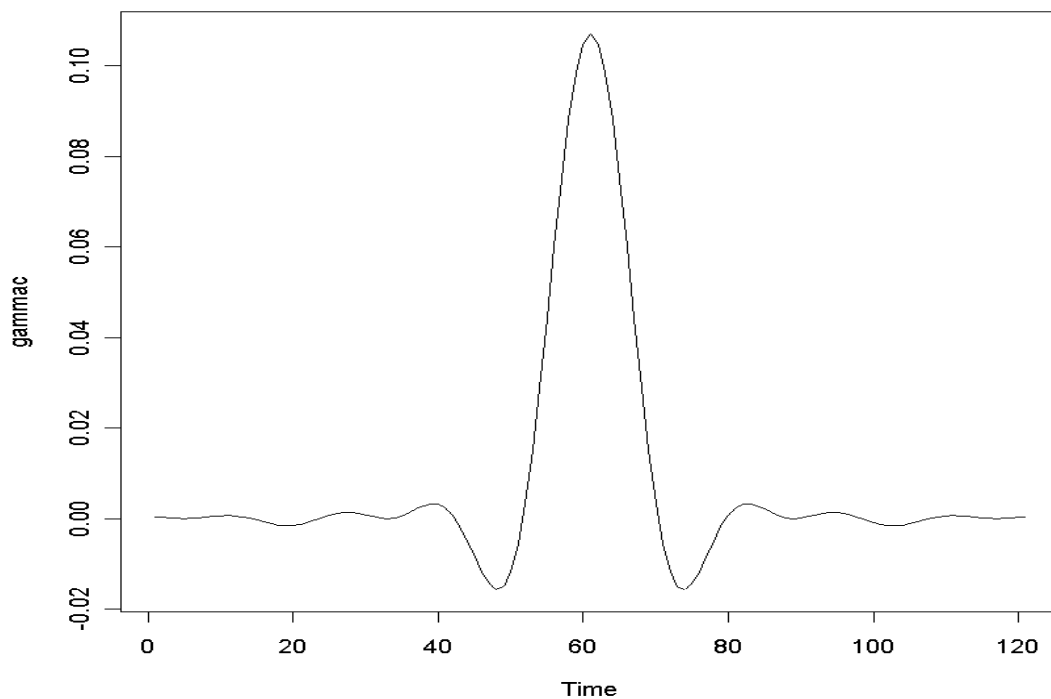
理想趋势 (IDEAL TREND) 的转换函数

$$\Gamma(\omega) := \begin{cases} 1 & 0 \leq |\omega| \leq \pi \cdot pb \\ \frac{\pi \cdot sb - |\omega|}{\pi \cdot sb - \pi \cdot pb} & \pi \cdot pb \leq |\omega| \leq \pi \cdot sb \\ 0 & \pi \cdot sb \leq |\omega| \leq \pi \end{cases}$$



理想趋势的滤子

$$\tilde{\gamma}_k = \begin{cases} -\frac{1}{\pi(\pi/7 - \pi/14)} \left[\frac{\cos(k\pi/7) - \cos(k\pi/14)}{k^2} \right] & k \neq 0 \\ \frac{1}{2} \left(\frac{1}{7} + \frac{1}{14} \right) & k=0 \end{cases}$$



ZPC滤子

$$Y_t = \sum_{k=1}^Q a_k Y_{t-k} + \sum_{k=-r}^q b_k X_{t-k}$$

$$\begin{aligned}\hat{F}(\omega) &= \frac{\sum_{k=-r}^q b_k \exp(-ik\omega)}{1 - \sum_{k=1}^Q a_k \exp(-ik\omega)} \\ &= C \exp(ir\omega) \frac{\prod_{j=1}^n (Z_{2j-1} - \exp(-i\omega))(Z_{2j} - \exp(-i\omega))}{\prod_{k=1}^{n'} (P_{2k-1} - \exp(-i\omega))(P_{2k} - \exp(-i\omega))} \\ &\quad \cdot \frac{\prod_{j=2n+1}^{q+r} (Z_j - \exp(-i\omega))}{\prod_{k=2n'+1}^Q (P_k - \exp(-i\omega))}\end{aligned}$$

$$y_t = ay_{t-1} + b_0 x_t + b_1 x_{t-1}$$

$$\hat{F}(\omega) = \frac{b_0 + b_1 \exp(-i\omega)}{1 - a \exp(-i\omega)}$$

$$= C \frac{Z - \exp(-i\omega)}{P - \exp(-i\omega)}$$

$$C = -b_1/a, P = 1/a, Z = -b_0/b_1$$

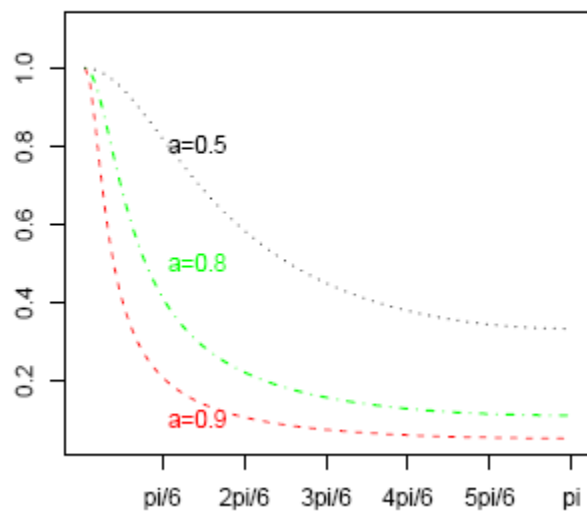
若Z不存在，则变成AR滤子

若P不存在，则变成MA滤子

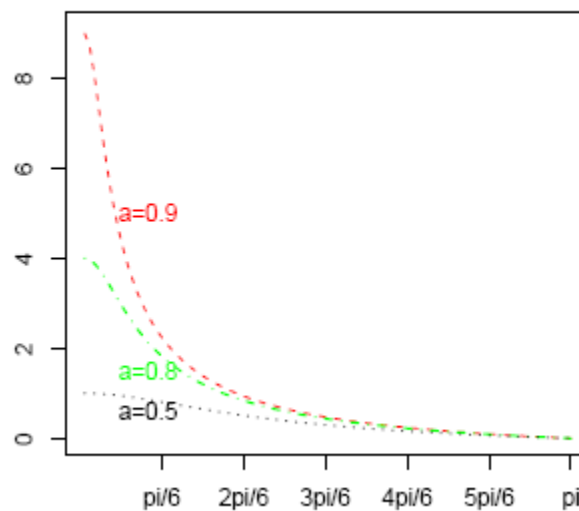
滤子: AR滤子

$$\Gamma_{AR} = \frac{C}{(P - \exp(-i\omega))(\overline{P} - \exp(-i\omega))}$$

Amplitude Functions AR(1)



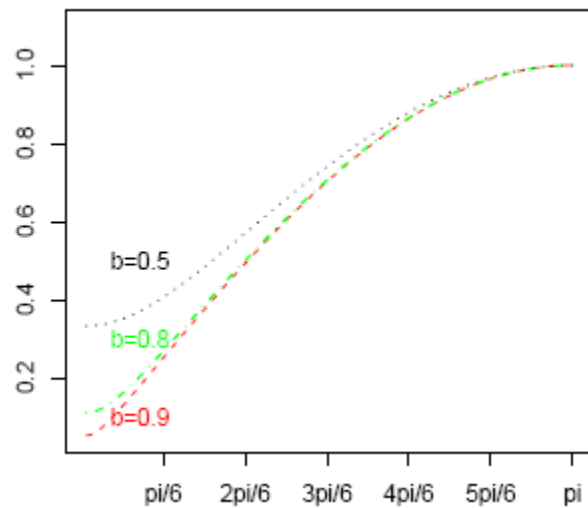
Time Delay Functions AR(1)



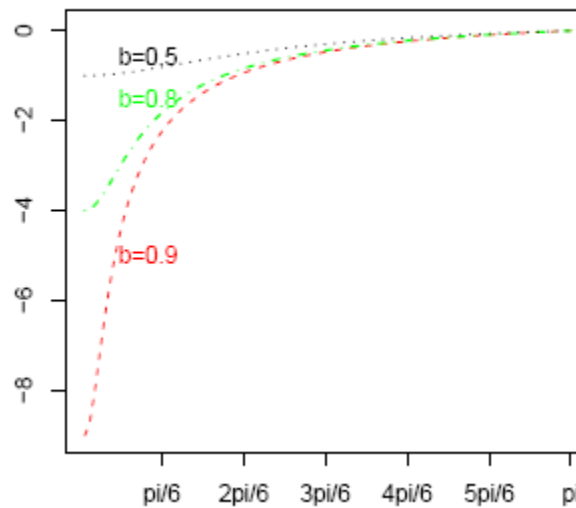
滤子: MA 滤子

$$\Gamma_{MA}(\omega) = C(Z - \exp(-i\omega))$$

Amplitude Functions MA(1)



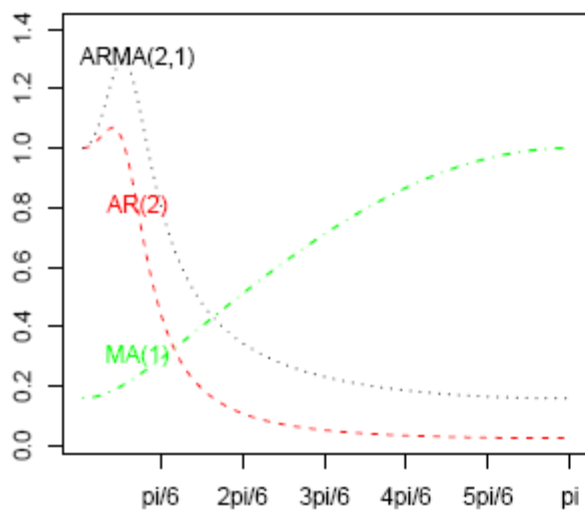
Time Delay Functions MA(1)



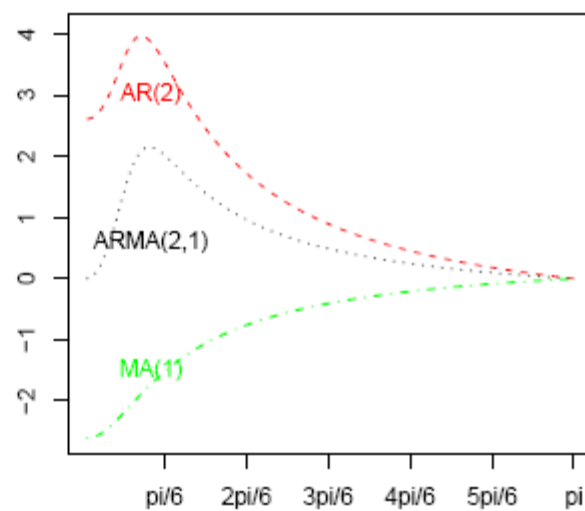
滤子: ARMA滤子

$$\Gamma_{ARMA} = C \frac{Z - \exp(-i\omega)}{(P - \exp(-i\omega))(\bar{P} - \exp(-i\omega))}$$

Amplitude Functions



Time Delay Functions



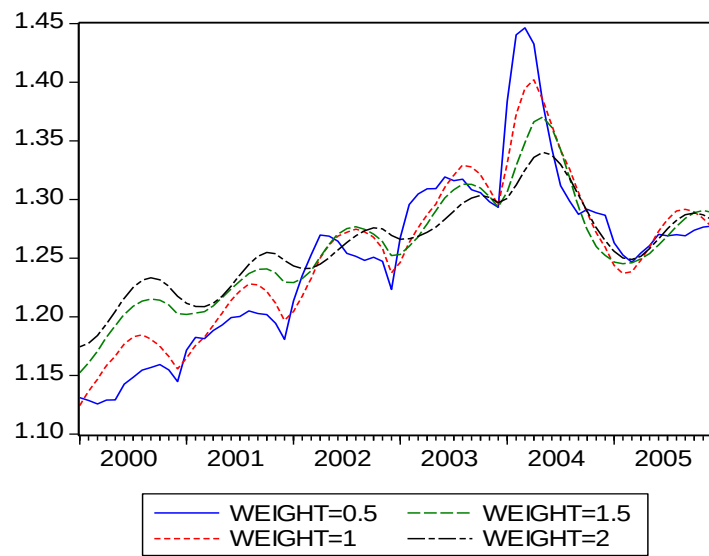
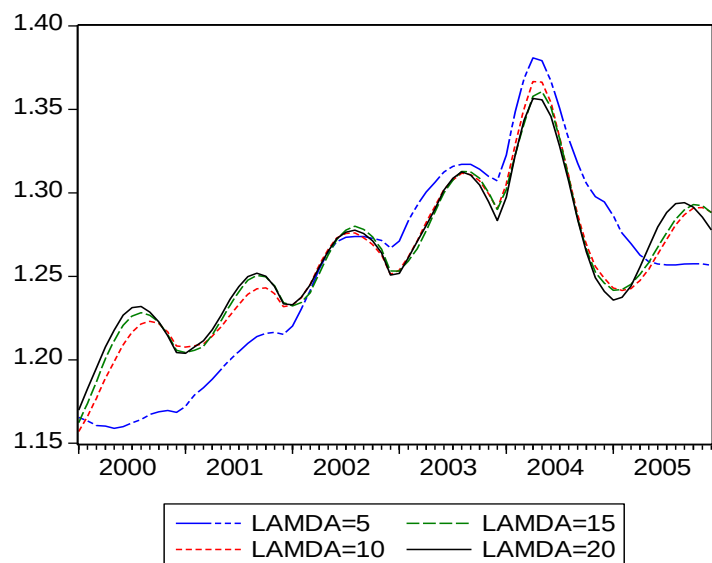
转折点的探测：一般的误差原则

$$\min_{\hat{\Gamma}} \sum_{k=1}^{N/2} |\Gamma(\omega_k) - \hat{\Gamma}(\omega_k)|^2 I_{NX}(\omega_k)$$

$$\min_{\hat{\Gamma}} \begin{cases} \frac{2\pi}{T} \sum_{k=0}^{[T/2]} W(\omega_k)^2 |\Gamma(\omega_k) - \hat{A}(\omega_k)|^2 I_{NX}(\omega_k) \\ + \lambda \frac{2\pi}{T} \sum_{k=0}^{[T/2]} 2W(\omega_k)^2 A(\omega_k) \hat{A}(\omega_k) (1 - \cos(\hat{\Phi}(\omega_k))) I_{NX}(\omega_k) \end{cases}$$

- ⊙ $W(\omega)$ ：更强的高频噪点的剔除力度，可靠性
- ⊙ λ ：减小时间延迟，速度
- ⊙ $W(\omega)=1$ and $\lambda=1$ ：均方差的近似（水平滤子）

结果比较



lamda越大，则时间延迟越小，探测拐点越及时；
越大；

weight越大，则高频噪点的剔除力度越强，输出信号越平滑，从而增加了探测拐点的可靠性。

模型应用：传统季调方法

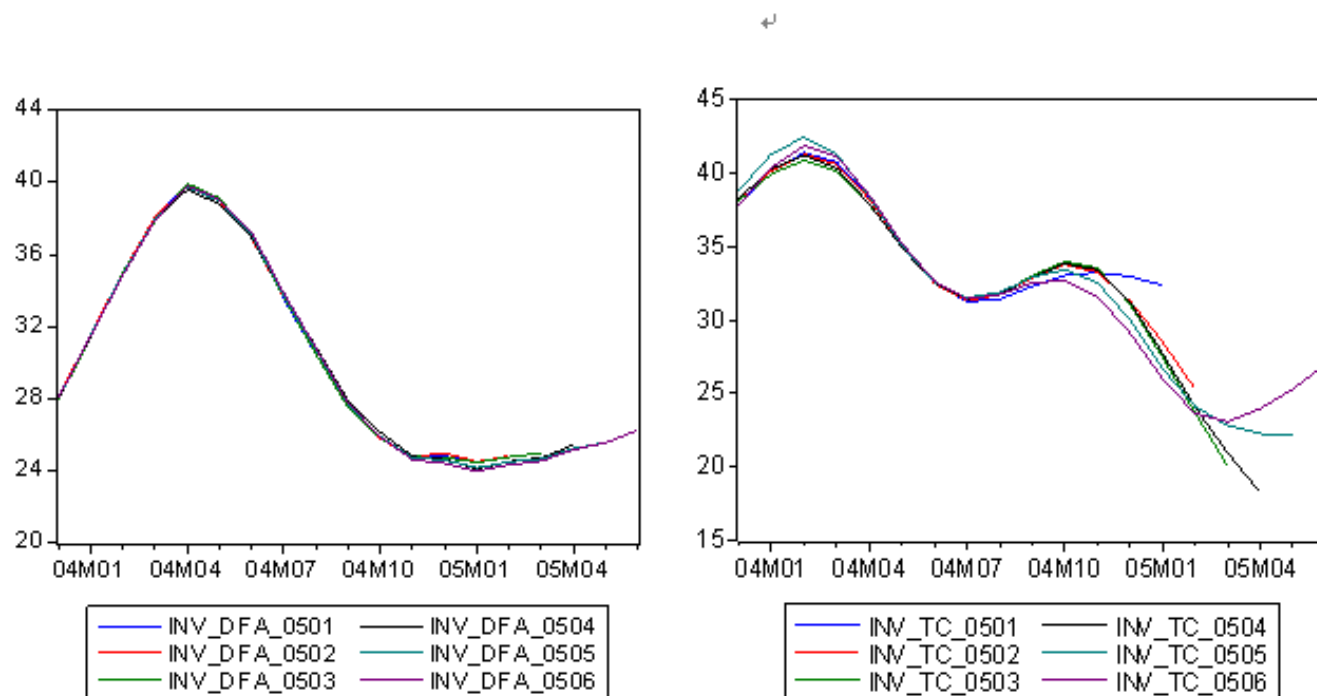


图 6 固定资产投资完成额累计同比的信息提取结果¹⁾

模型应用：传统季调方法

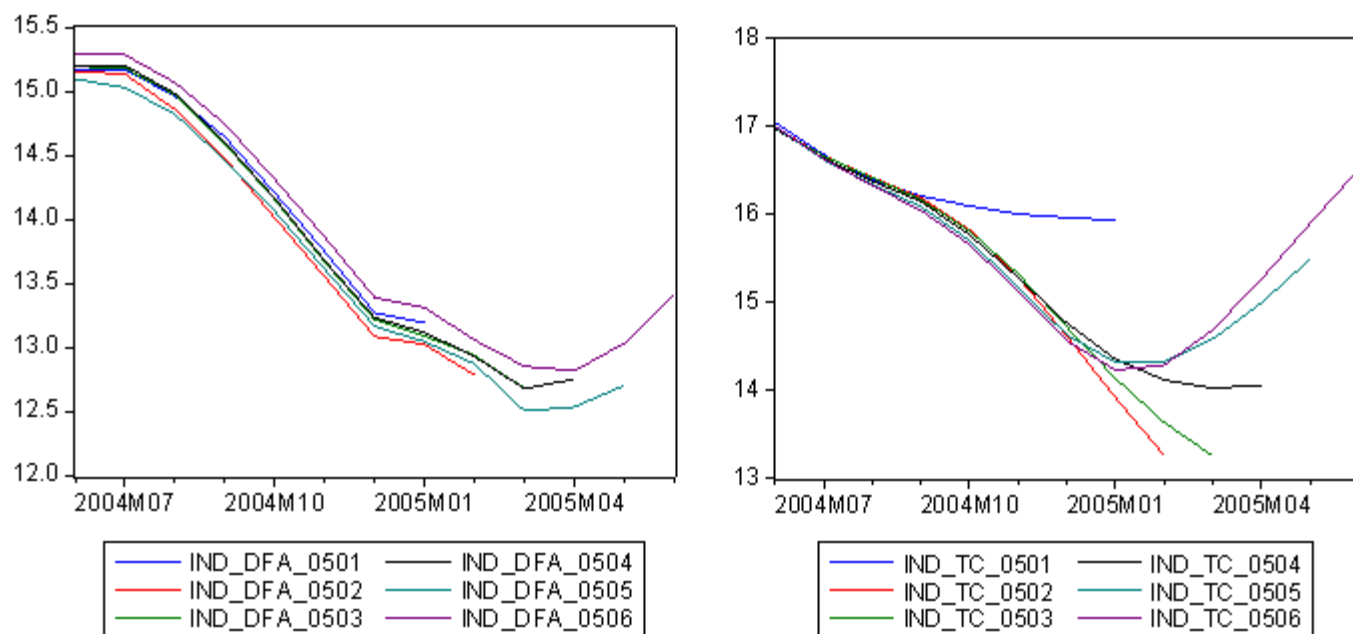
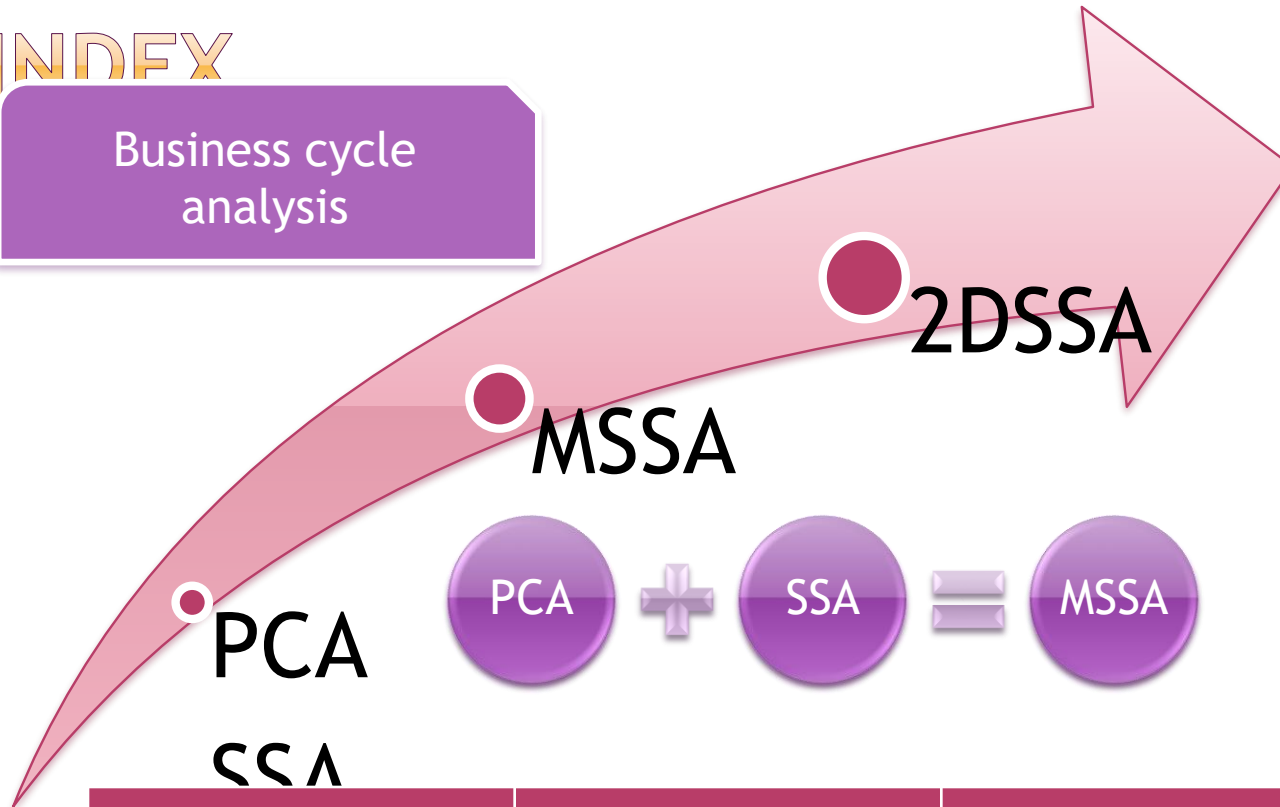


图 7 工业增加值当期同比的信息提取结果

SSA方法

THE USE OF SSA FOR PROSPERITY INDEX

Business cycle analysis



PCA	MSSA	2DSSA
Extract the business cycle.	Construction for composite index	Construction for multi-country composite index

A SHORT DESCRIPTION OF THE SSA

We consider a time series $Y_T = (y_1, \dots, y_T)$. Fix L ($L \leq T/2$), the window length, and let $K = T - L + 1$.

Step 1. (*Computing the trajectory matrix*)

transfers a one-dimensional time series Y_t into the multi-dimensional series X_k with $X_i = (y_i, \dots, y_{i+L-1})$, $K = T - L + 1$. $2 < L < T$

$$\mathbf{X} = (x_{ij})_{i,j=1}^{L,K} = \begin{pmatrix} y_1 & y_2 & y_3 & \dots & y_K \\ y_2 & y_3 & y_4 & \dots & y_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & y_{L+2} & \dots & y_T \end{pmatrix}.$$

A SHORT DESCRIPTION OF THE SSA

Step 2. (*Constructing a matrix for applying SVD*): compute the matrix $\mathbf{X}\mathbf{X}^T$.

Step 3. (*SVD of the matrix $\mathbf{X}\mathbf{X}^T$*): compute the eigenvalues and eigenvectors of the matrix $\mathbf{X}\mathbf{X}^T$ and represent it in the form $\mathbf{X}\mathbf{X}^T = P\Lambda P^T$. Here $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_L)$ is the diagonal matrix of eigenvalues of $\mathbf{X}\mathbf{X}^T$ ordered so that $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L \geq 0$ and $P = (P_1, P_2, \dots, P_L)$ is the corresponding orthogonal matrix of eigen-vectors of $\mathbf{X}\mathbf{X}^T$.

Step 4. (*Selection of eigen-vectors*): select a group of l ($1 \leq l \leq L$) eigen-vectors $P_{i_1}, P_{i_2}, \dots, P_{i_l}$.

The grouping step corresponds to splitting the elementary matrices $\mathbf{X}_i$ into several groups and summing the matrices within each group. Let $I = \{i_1, \dots, i_l\}$ be a group of indices $i_1, \dots, i_l$. Then the matrix $\mathbf{X}_I$ corresponding to the group I is defined as $\mathbf{X}_I = \mathbf{X}_{i_1} + \dots + \mathbf{X}_{i_l}$.

A SHORT DESCRIPTION OF THE SSA

Step 5. (*Reconstruction of the one-dimensional series*): compute the matrix $\tilde{\mathbf{X}} = ||\tilde{x}_{i,j}|| = \sum_{k=1}^l P_{i_k} P_{i_k}^T \mathbf{X}$ as an approximation to $\mathbf{X}$. Transition to the one-dimensional series can now be achieved by averaging over the diagonals of the matrix $\tilde{\mathbf{X}}$.

EXAMPLES-TIME SERIES

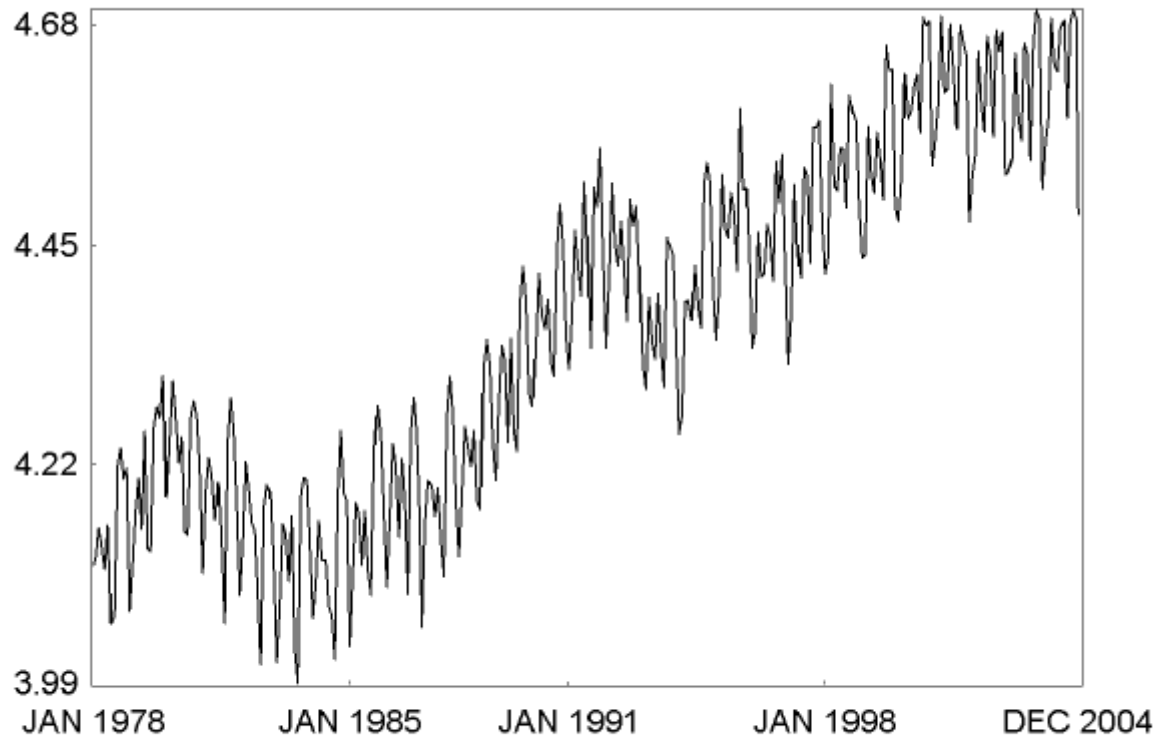


Figure 1: Fabricated metal series in Germany

SELECTION OF THE WINDOW LENGTH L

- ◉ the assumption that there is an annual periodicity in the series suggests that we must pay attention to the frequencies $k=12$ ($k = 1, \dots, 12$). As it is advisable to choose L *reasonably large (but smaller than $T/2$ which is 162 in this case)*, we choose $L = 120$.

SELECTION OF R

Choosing $L = 120$ and performing SVD of the trajectory matrix $\mathbf{X}$, we obtain 120 eigen-triples, ordered by their contribution (share) in the decomposition. Fig. 2 depicts the plot of the logarithms of the 120 singular values.

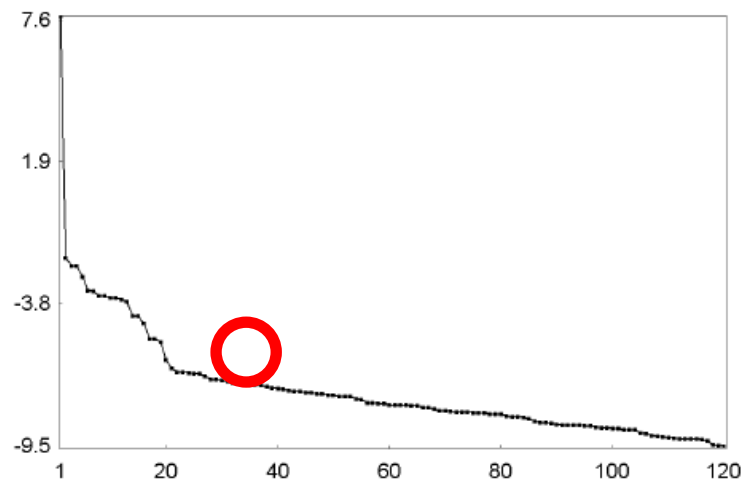


Figure 2: Logarithms of the 120 eigenvalues.

SELECTION OF R

Another way of grouping is to examine the matrix of the absolute values of the w -correlations. Fig. 3 shows the w -correlations for the 120 reconstructed components in a 20-grade grey scale from white to black corresponding to the absolute values of correlations from 0 to 1. Based on this information, we select the first 18 eigentriples for the reconstruction of the original series and consider the rest as noise.

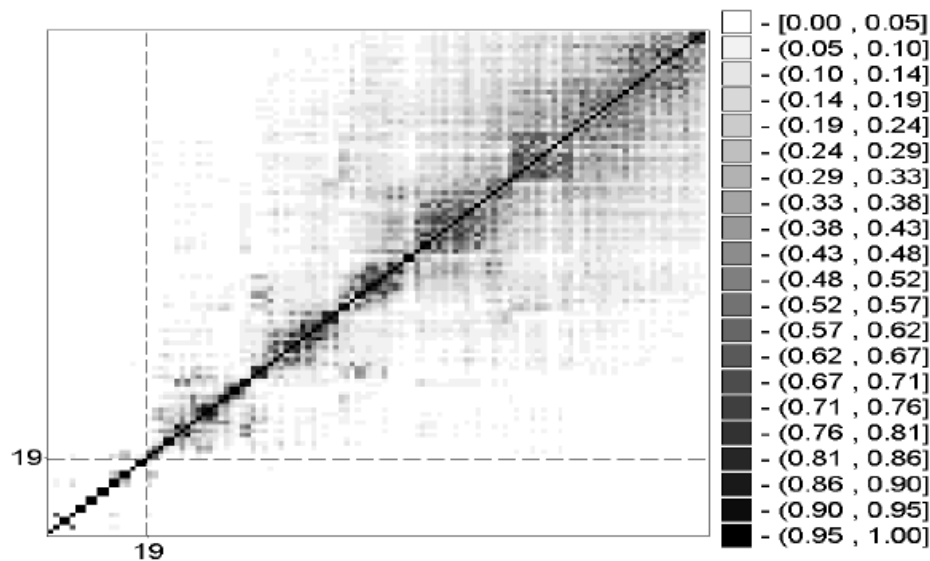
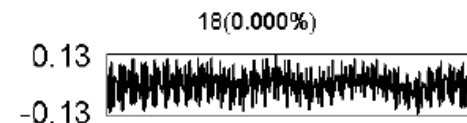
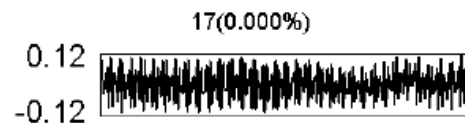
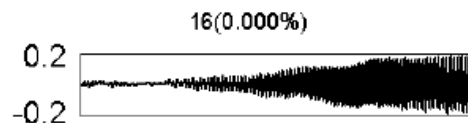
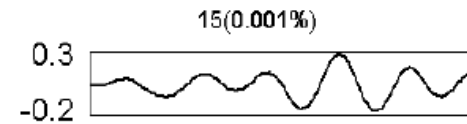
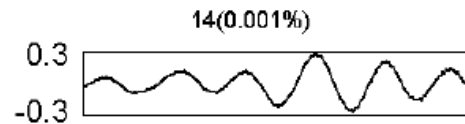
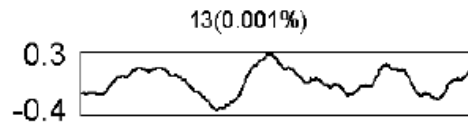
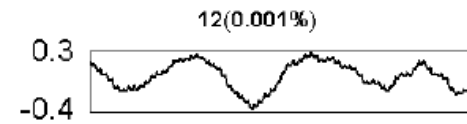
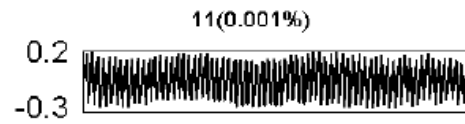
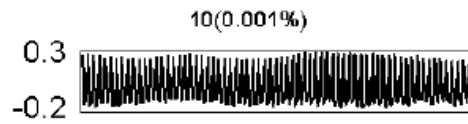
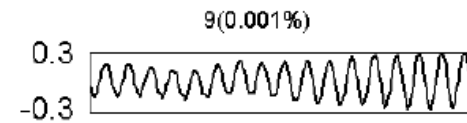
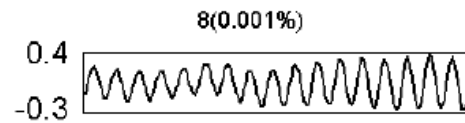
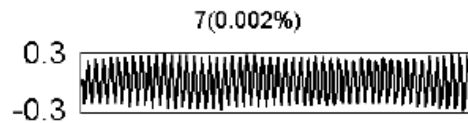
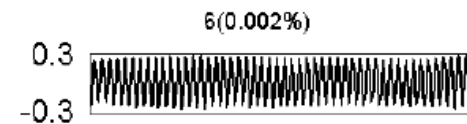
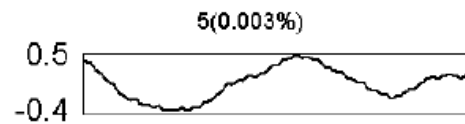
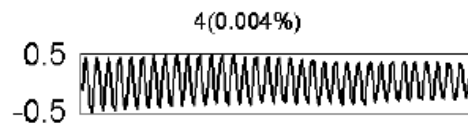
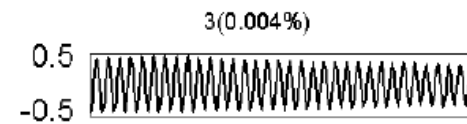
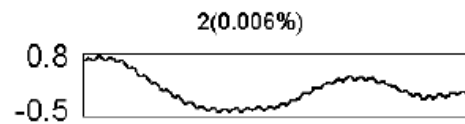
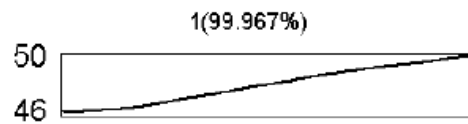


Figure 3: Matrix of w -correlations for the 120 reconstructed components.

GROUPING



GROUPING

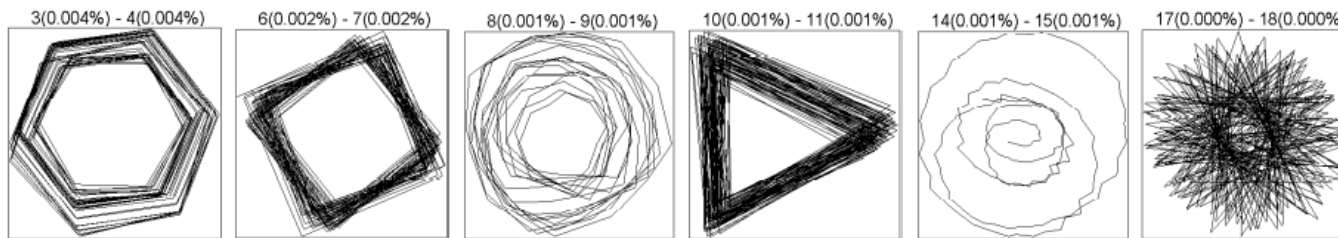
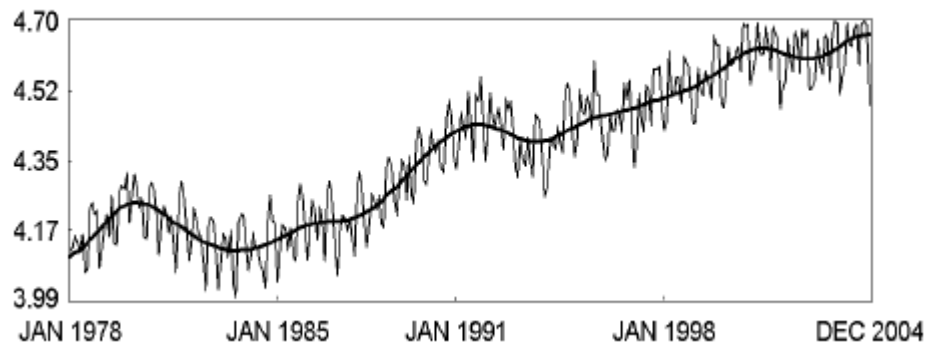


Figure 5: Scatterplots (with lines connecting consecutive points) corresponding to the paired harmonic principal components.

The periodograms of the paired eigentriples (3-4 , 6-7, 8-9, 10-11 and 17-18) also confirm that the eigentriples correspond to the periods of 6, 4, 12, 3, 36 and 2.4 months.

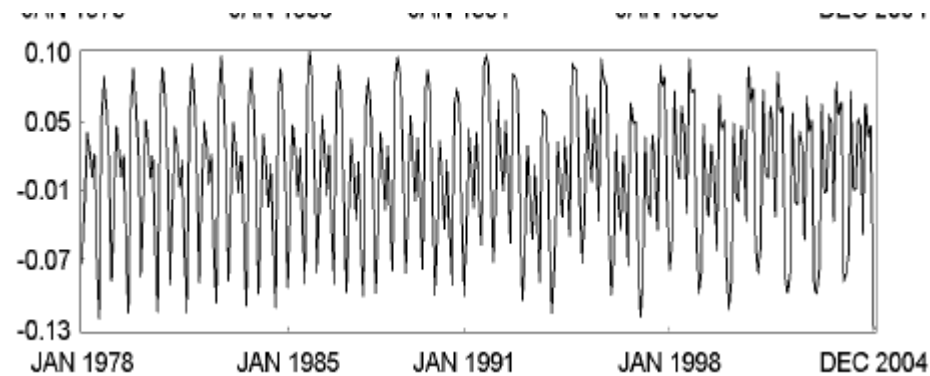
RECONSTRUCTION

- ◉ extracted trend which is obtained from the
- ◉ eigentriples 1, 2, 5, and 12-13. It clearly follows the main tendency in the series.

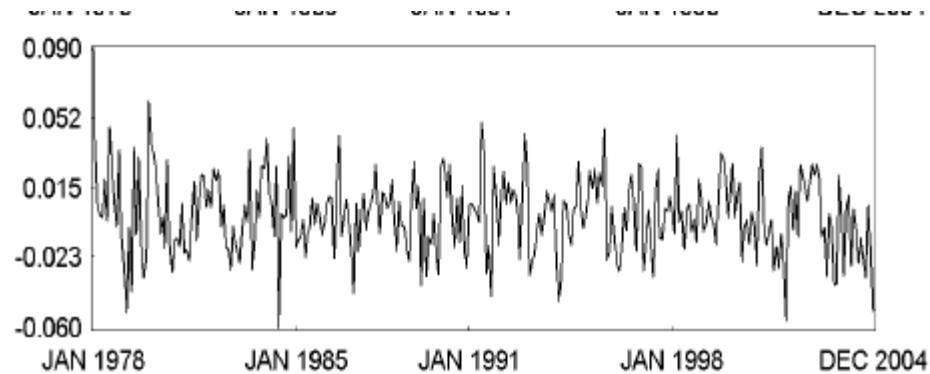


RECONSTRUCTION

- selected harmonic components (3,4, 6-11, 14-18)



- rest of the eigentriples components (19-120)
as noise



THE USE OF SSA FOR PROSPERITY INDEX

- Model based prosperity index: SSA approach
 - PCA's Empirical Orthogonal Function (EOF)

$$X_i = \sum_{k=1}^M a_i^k E_i^k$$

- SSA's EOFs

$$X_{t+j} = \sum_{j=1}^L a_t^k E_j^k$$

- MSSA's EOFs

$$X_{i,t+j} = \sum_{j=1}^M \sum_{i=1}^L a_t^k E_{ij}^k$$

THE USE OF SSA FOR PROSPERITY INDEX

◎ PCA

- Extracts the dominant "spatial" patterns of the variability.
- The first EOFs represents the dominant, pure "spatial" patterns of variability, and their associated PCs are the coefficients that regulate these patterns.

◎ SSA

- Removes incoherent noise (white noise).
- Extracts regular oscillations (with a narrow spectral peak).
- Coherent nonlinear trends are identified as the lower frequency modes.
- Efficient on short signal.

THE USE OF SSA FOR PROSPERITY INDEX

- ◎ MSSA has the advantage of
 - PCA for extracting the dominant "spatial" patterns of the variability.
 - The spectral filtering capabilities of SSA.
 - All identified modes have spatio-temporal properties.
 - *This advanced spatial and spectral filtering is helpful to identify the most coherent (and more especially oscillatory) spatio-temporal modes in a short noisy signal.*

BASIC KNOWLEDGE ON 2D-SSA

- Given a 2D data set

$$F = \begin{pmatrix} f(1,1) & f(1,2) & \dots & f(1,N_c) \\ f(2,1) & f(2,2) & \dots & f(2,N_c) \\ \vdots & \vdots & \ddots & \vdots \\ f(N_r,1) & f(N_r,2) & \dots & f(N_r,N_c) \end{pmatrix}.$$

- Choose two window length $L1$ and $L2$
- Construction of the trajectory matrix of the field F by moving $L1 \times L2$ -windows
- In the 1D case, one transforms a 1D object into a 2D matrix. In the 2D case, a 2D object will be embed into a four-dimensional space
- SVD & reconstruction.

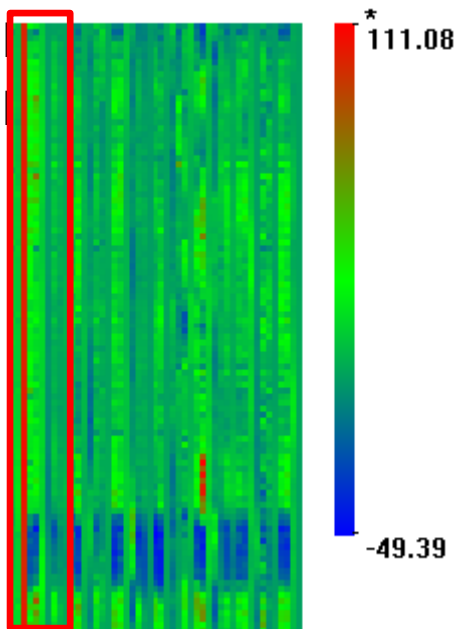
THE USE OF SSA FOR PROSPERITY INDEX

- Data transformation

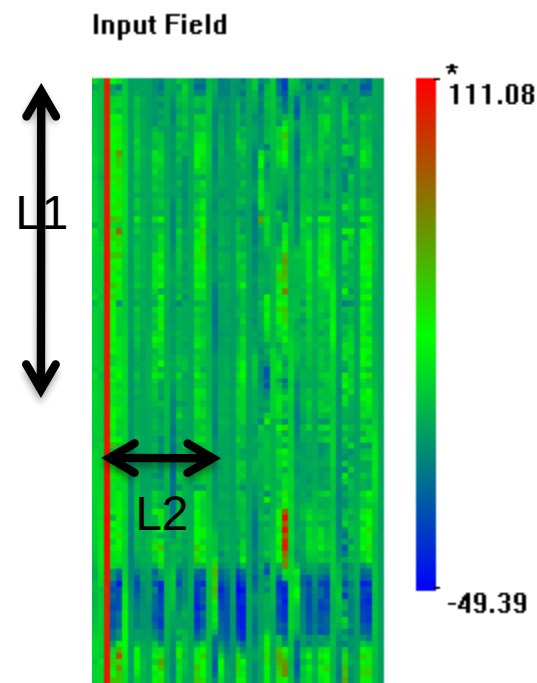
- Input data:

- (Input Field

- -
 -



- Choosing the window lengths.

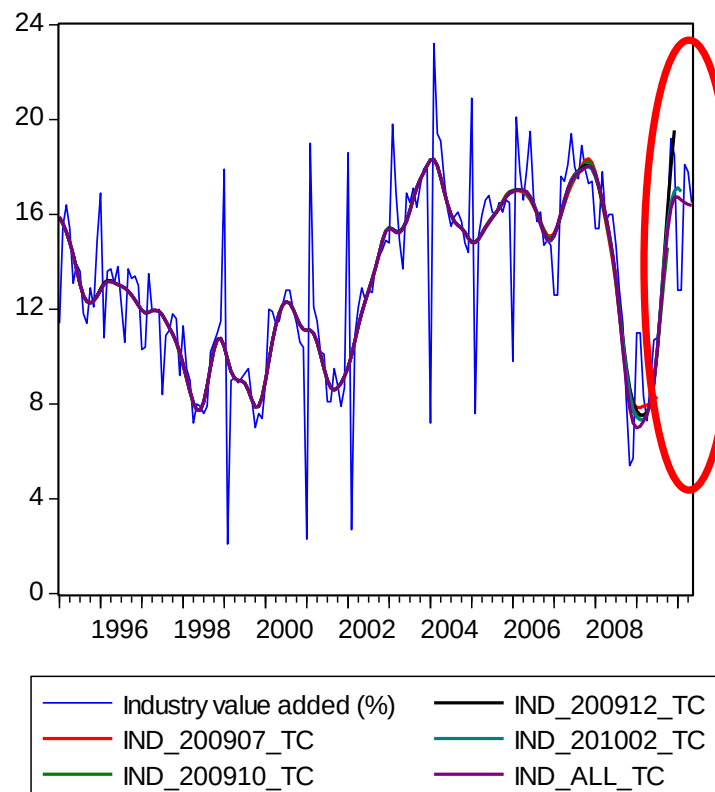


THE USE OF SSA FOR PROSPERITY INDEX

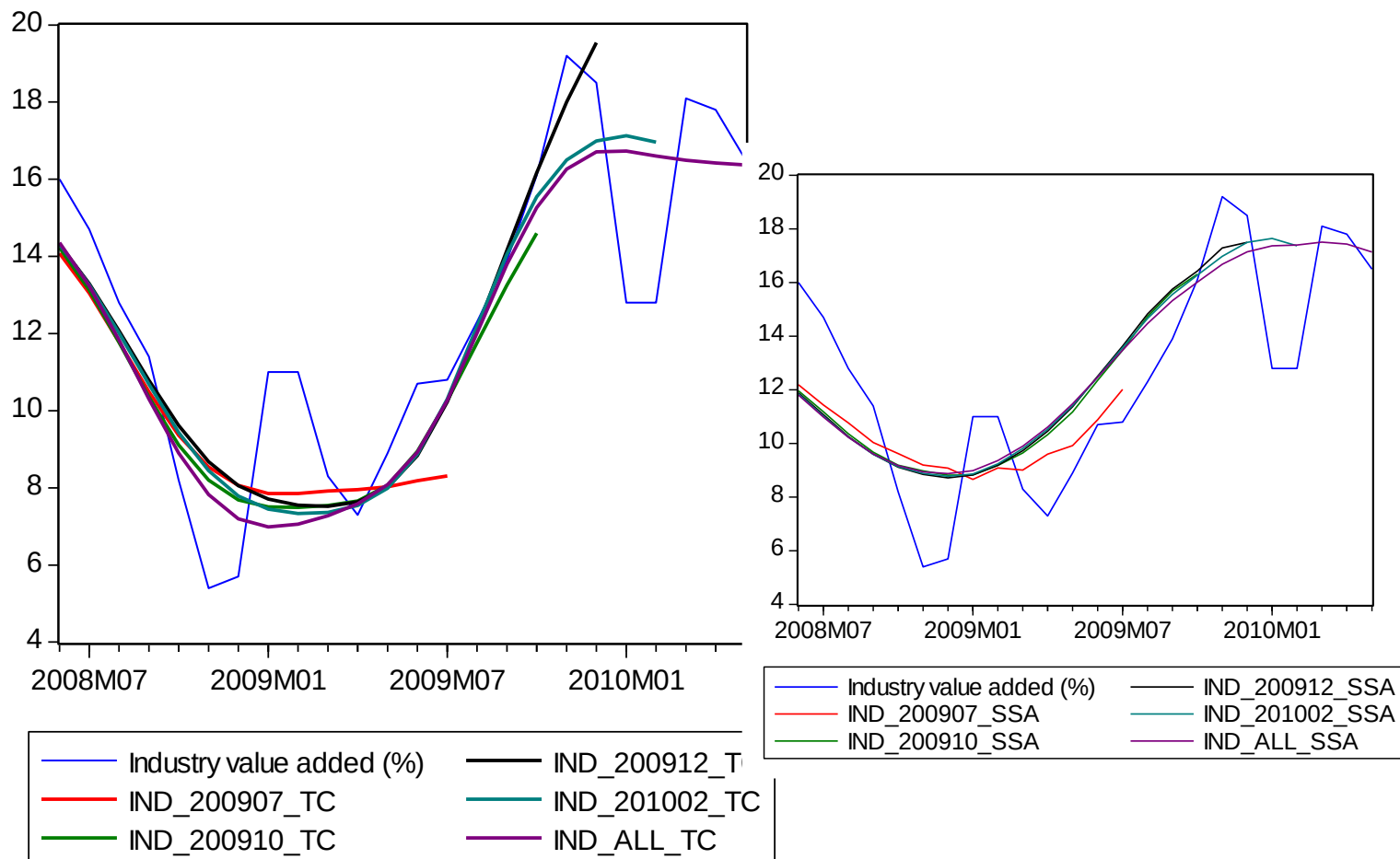
- ◎ 2DSSA's advantage:
 - Take the advantage of MSSA, which combined SSA and PCA.
 - Extract common factors
 - Filter noise
 - Identify and make good use of the interaction among different countries

EMPIRICAL STUDIES: COMPARE WITH X12

- Compare SSA's performance with X12-ARIMA for data filter.
- The sample is widen piece by piece.



EMPIRICAL STUDIES: COMPARE WITH X12



EMPIRICAL STUDIES: 2DSSA-BASED INDEX

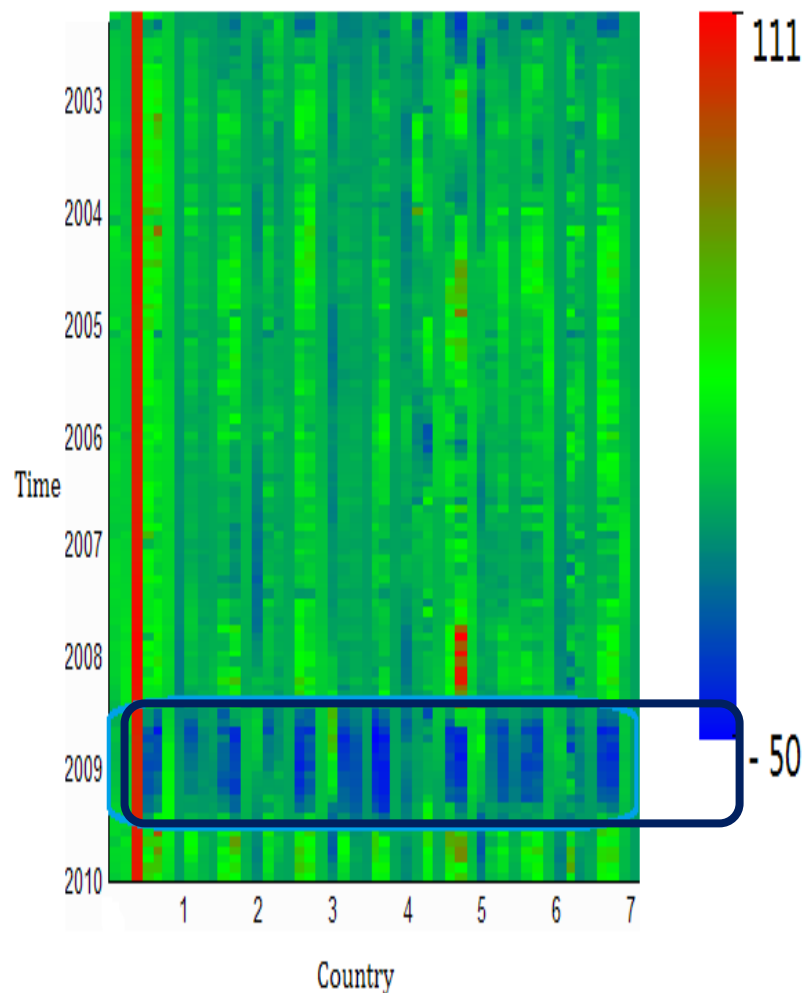
Fig. 2 Input data of Asian countries

⊙ Countries from Asia:

China, Indonesia, Japan, Korea, Malaysia, Singapore and Thailand.

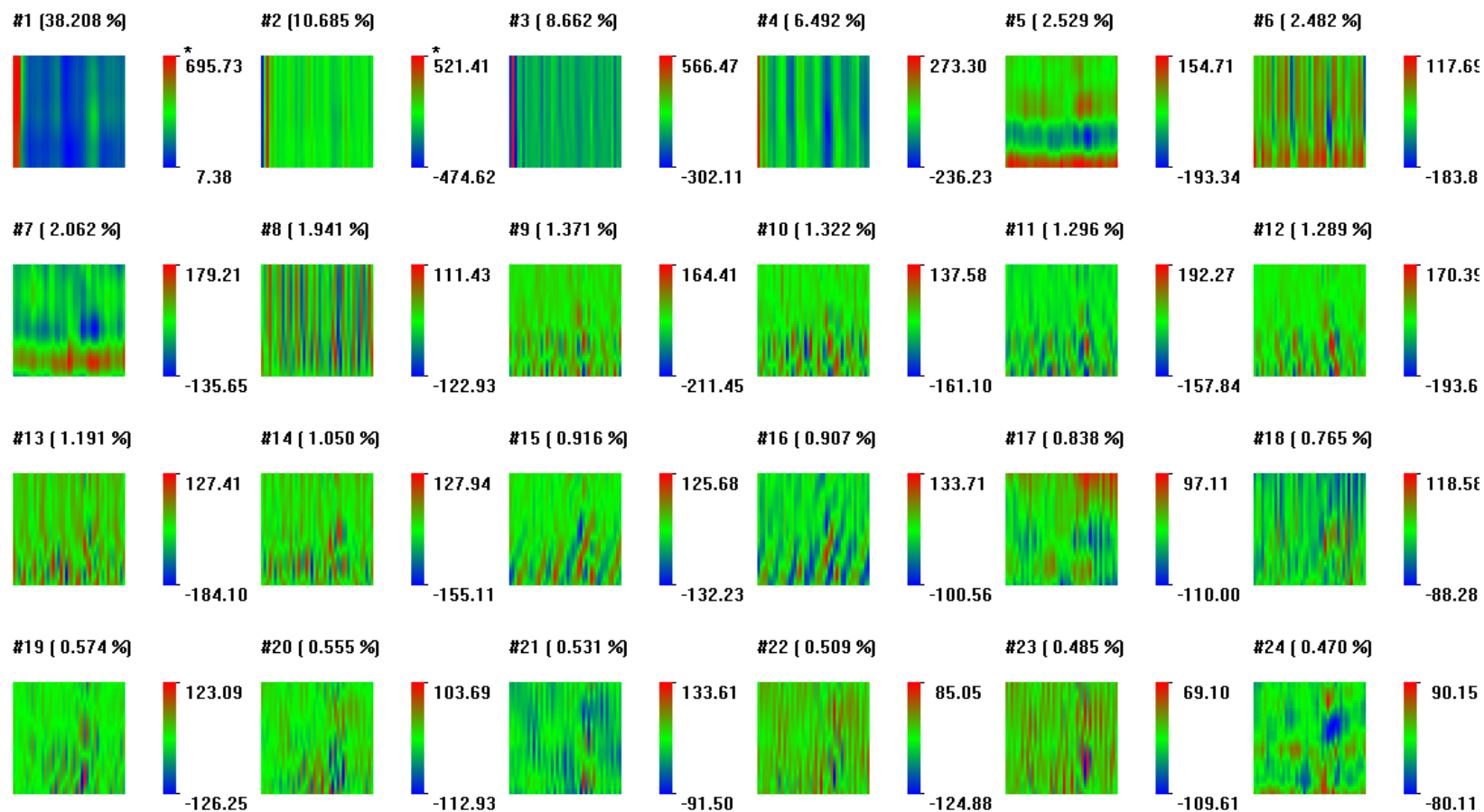
⊙ Input data:

economic growth, financial data, price, trade, industry and so on.



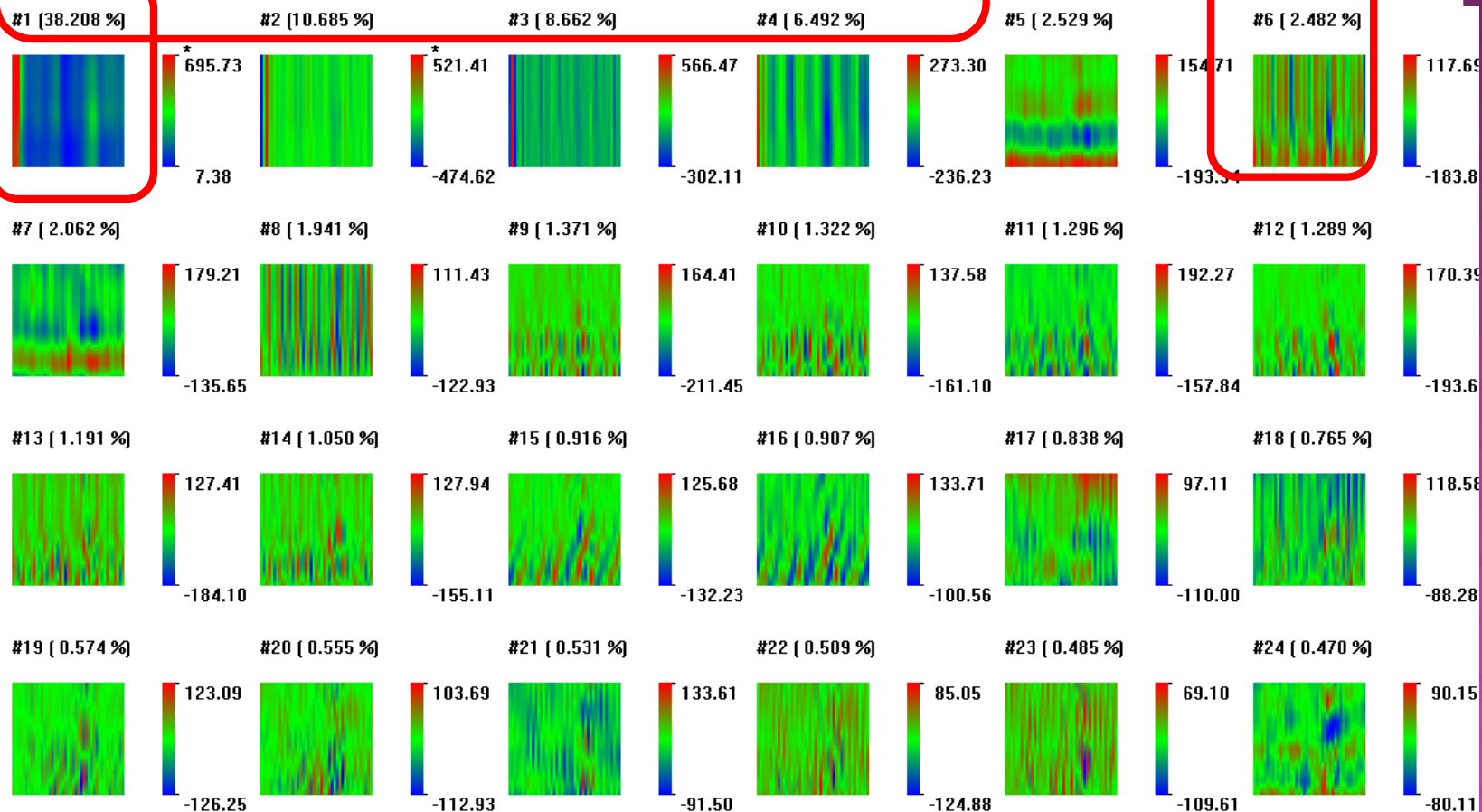
EMPIRICAL STUDIES: 2DSSA-BASED INDEX

◎ 2D Principal Components

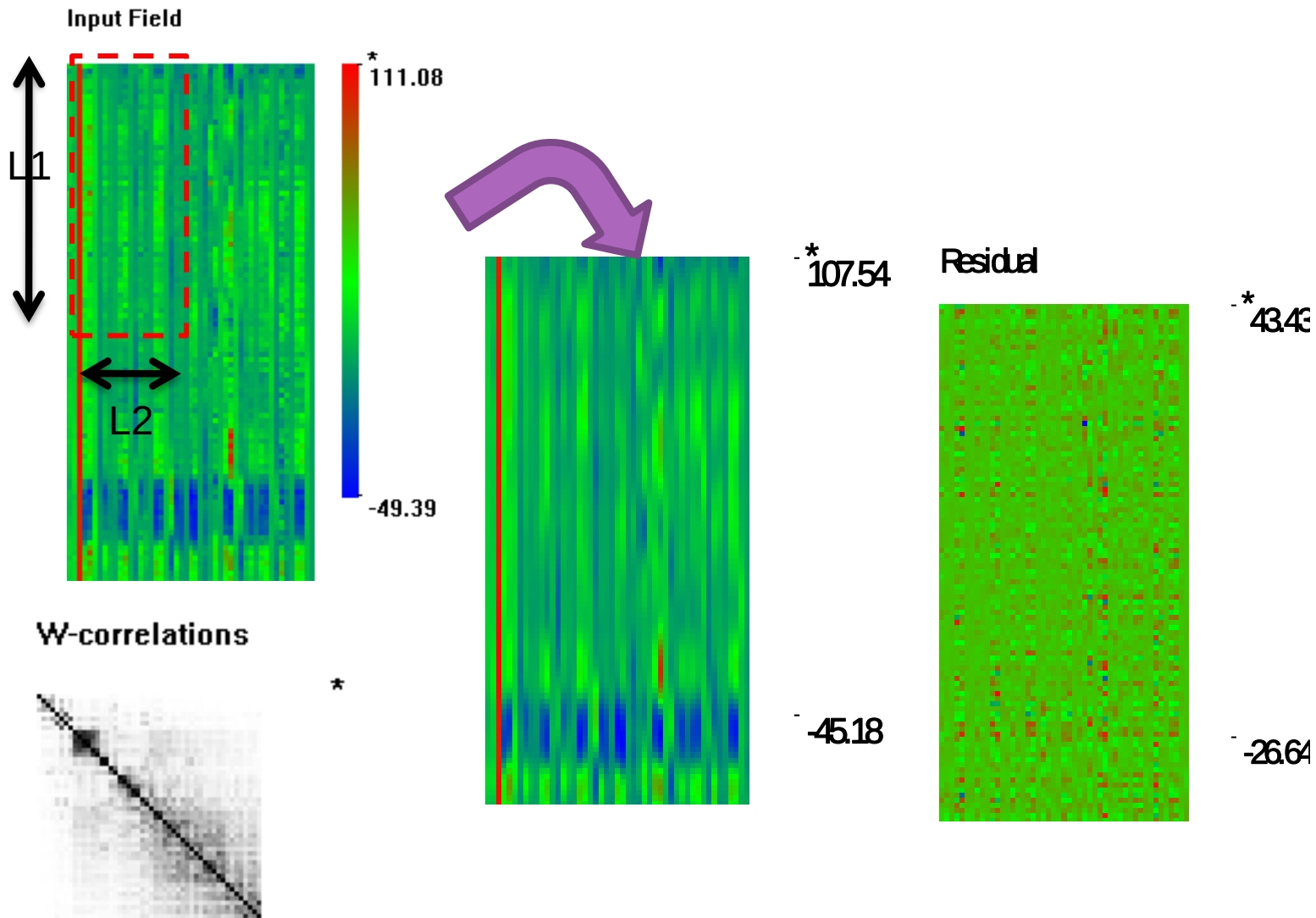


EMPIRICAL STUDIES: 2DSSA-BASED INDEX

2D Principal Components



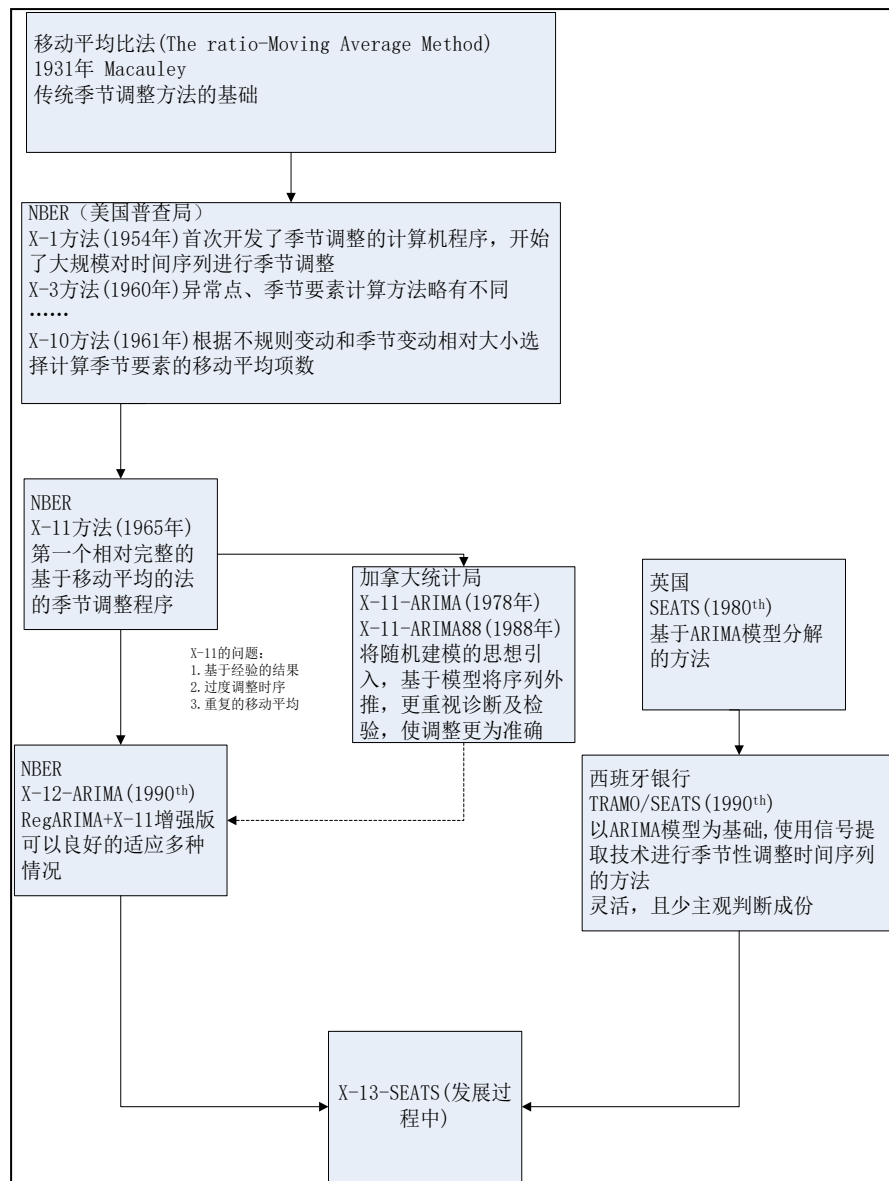
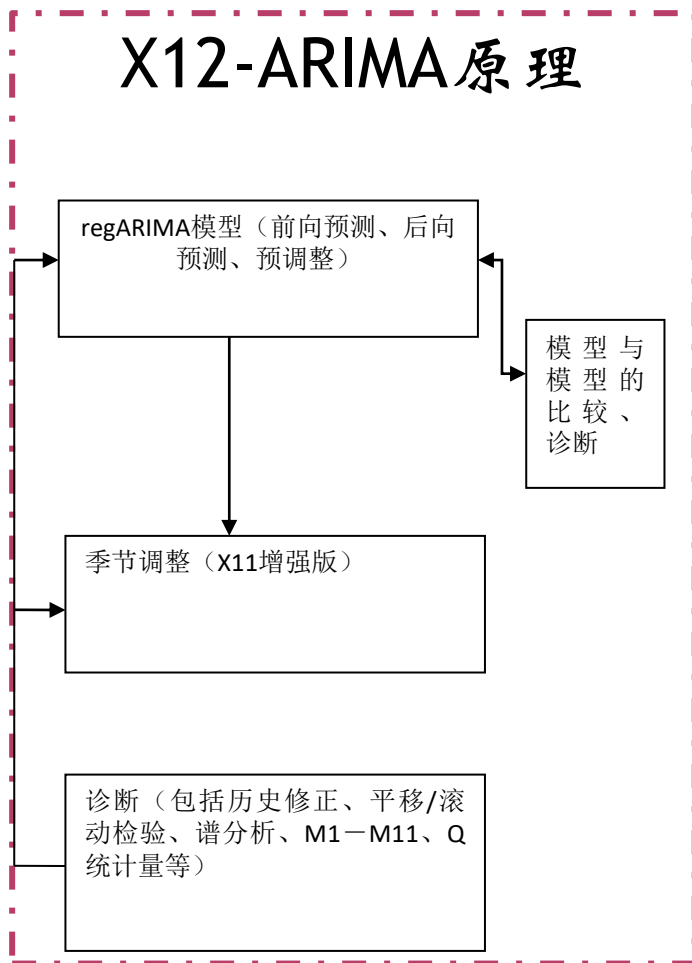
EMPIRICAL STUDIES: 2DSSA-BASED INDEX



其它滤波方法：季节调整方法

季节调整方法发展历程

X12-ARIMA原理



REGARMIA建模原理

通过识别、估计和诊断建立 ARIMA 模型并用于预测，从而实现时间序列的延拓。↵

设原序列为 y_t ，则对 y_t 建立模型如下：↵

$$(1-L)^d(1-L^s)^D y_t = \sum_i \beta_i (1-L)^d(1-L^s)^D x_{it} + w_t \quad \leftarrow$$

其中 w_t 服从平稳 ARMA 模型，即 $\varphi(L)\Phi(L^s)w_t = \theta(L)\Theta(L^s)a_t$ ， s 为季节周期， $\varphi(L)$ 为非季节自回归 (AR) 算子， $\Phi(L^s)$ 是季节移动平均 (AR) 算子， $\theta(L)$ 是非季节移动平均 (MA) 算子， $\Theta(L^s)$ 是季节移动平均 (MA) 算子， a_t 为白噪声。 $(1-L)^d(1-L^s)^D$ 意味着非季节差分次数为 d ，季节差分次数为 D 。↵

REGARMA建模原理

x_{it} 是与 y_t 同时预测的回归变量，主要包括各种离群值以及与日历相关的影响因素等。在传统的 X-11 方法中，这些成分的估计是在季节调整过程中的完成的，而在 X-12-ARIMA 中将这些功能集中到了新增的 regARIMA 模块中，同时在 X-11 中仍保留了这些功能，但仍推荐使用 regARIMA 进行预调整。而在 EVIEWS 中，只能在其中某一个模块进行识别和处理，以避免重复。↵

离群值：离群值点 AO；水平漂移 LS；暂时变化 TC；斜线上升 Ramp Effect。↵

日历效应：固定季节效应；闰年效应；月份长度效应；季度长度效应；交易日效应；工作日效应；移动假日效应。↵

X-11季节调整模块原理与过程

- X-11季节调整模块基于多次迭代的移动平均方法进行成分分解，以表格形式输出计算结果，一般包括A-G 7个基本部分，如下图所示：

数据预处理	X11计算原型	简明统计
• A: 先验调整计算	• B: 初步估计极端值和日历 • C: 初步估计极端值和日历 • D: 计算最终的S、I、SI和TC项等	• E: 各类序列之间的比较 • F: 季节调整质量的衡量 • G: 各序列的谱分析图

- 其中B、C、D部分是主要的迭代计算部分，采用的是相同的计算原型，将在后文中作详细介绍。
- F部分是传统的诊断检验部分，包括11个M统计量和2个Q统计量，M统计量的值都应当小于1才是可以接受的；Q统计量服从分布。
- G部分包括原序列，季节调整后序列及不规则因子序列的谱分析结果，可以直观的看出季节调整的结果是否理想。

X-11季节调整模块原理与过程

第 一 阶 段	1. 使用2×12移动平均估计 TC 项: $C_t^{(1)} = M_{2 \times 12}(Y_t)$, 可消除 12 阶不变季节性, 并减小不规则成分的方差 2. 估计 SI: $(S+I)_t^{(1)} = Y_t - C_t^{(1)}$ 3. 对每个月份应用3×3移动平均估计预备季节成分: $S_t^{(1)} = M_{3 \times 3}[(S+I)_t^{(1)}]$, 可保留对应 12 个月的季节因子 标准化得 $S_t^{(1)} = \bar{S}_t^{(1)} - M_{2 \times 12}(\bar{S}_t^{(1)})$, 使得因子之和在每一个连续的 12 个月内都近似为 0 4. 估计季节调整后序列: $A_t^{(1)} = (C+I)_t^{(1)} = Y_t - S_t^{(1)}$
------------------	---

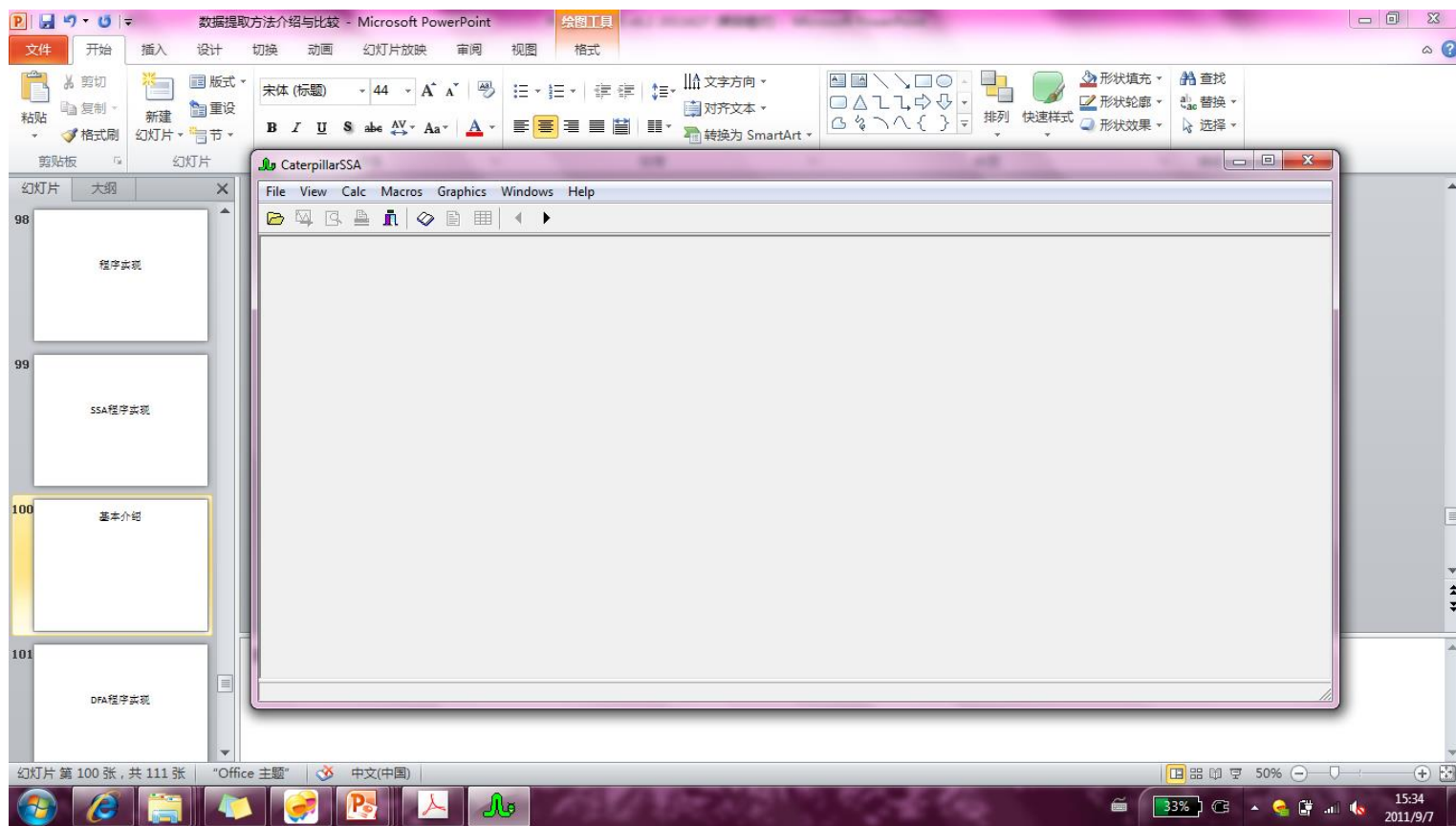
第 2 阶段	5. 用 13 项 Henderson 移动平均估计 TC 项: $C_t^{(2)} = H_{13}(A_t^{(1)})$, 估计季节—不规则成分: $(S+I)_t^{(2)} = Y_t - C_t^{(2)}$, Henderson 移动平均更接近于理想的“低通滤波器”, 具有比普通移动平均更好的平滑能力 6. 对每个月份应用3×5移动平均估计最终季节成分: $\bar{S}_t^{(2)} = M_{3 \times 5}[(S+I)_t^{(2)}]$, 标准化得 $S_t^{(2)} = \bar{S}_t^{(2)} - M_{2 \times 12}(\bar{S}_t^{(2)})$ 7. 估计季节调整后的序列: $A_t^{(2)} = (C+I)_t^{(2)} = Y_t - S_t^{(2)}$
第 3 阶段	8. 估计最终趋势: $C_t^{(3)} = H_{2H+1}(A_t^{(2)})$, 这里的 H 根据数据特点和 Henderson 准则共同决定, 可取 5、7、9、13、23 等, H 越大, 则序列越平滑, 但有可能损失的信息也越多。 9. 估计最终不规则成分: $I_t^{(3)} = A_t^{(2)} - C_t^{(3)}$

有关具体算子的更详细的介绍与说明可参考中国人民银行调查统计司 (2006) [1]

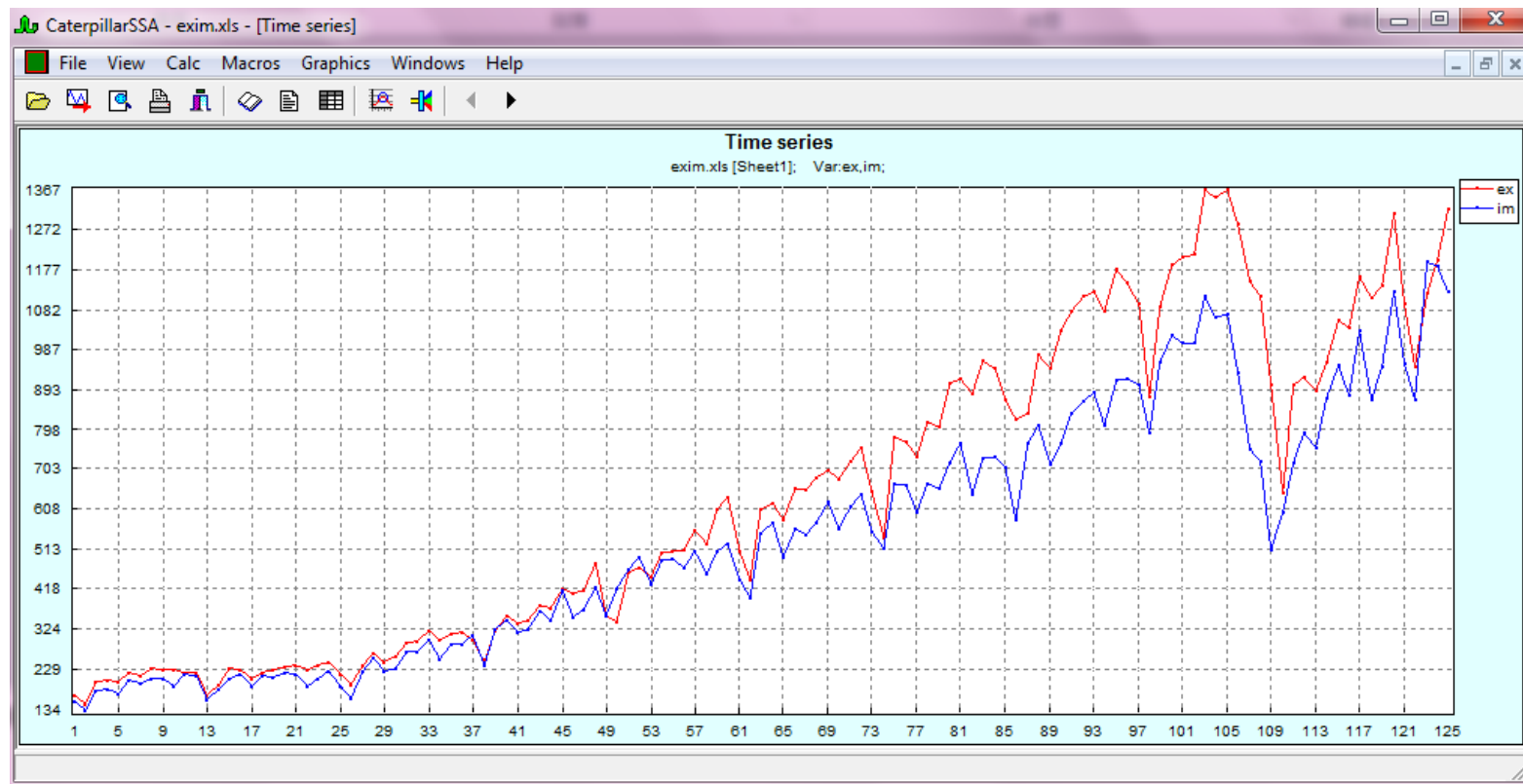
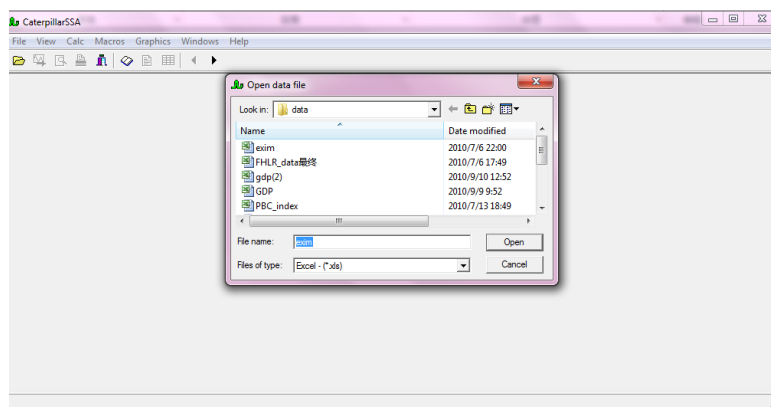
程序实现

SSA程序实现

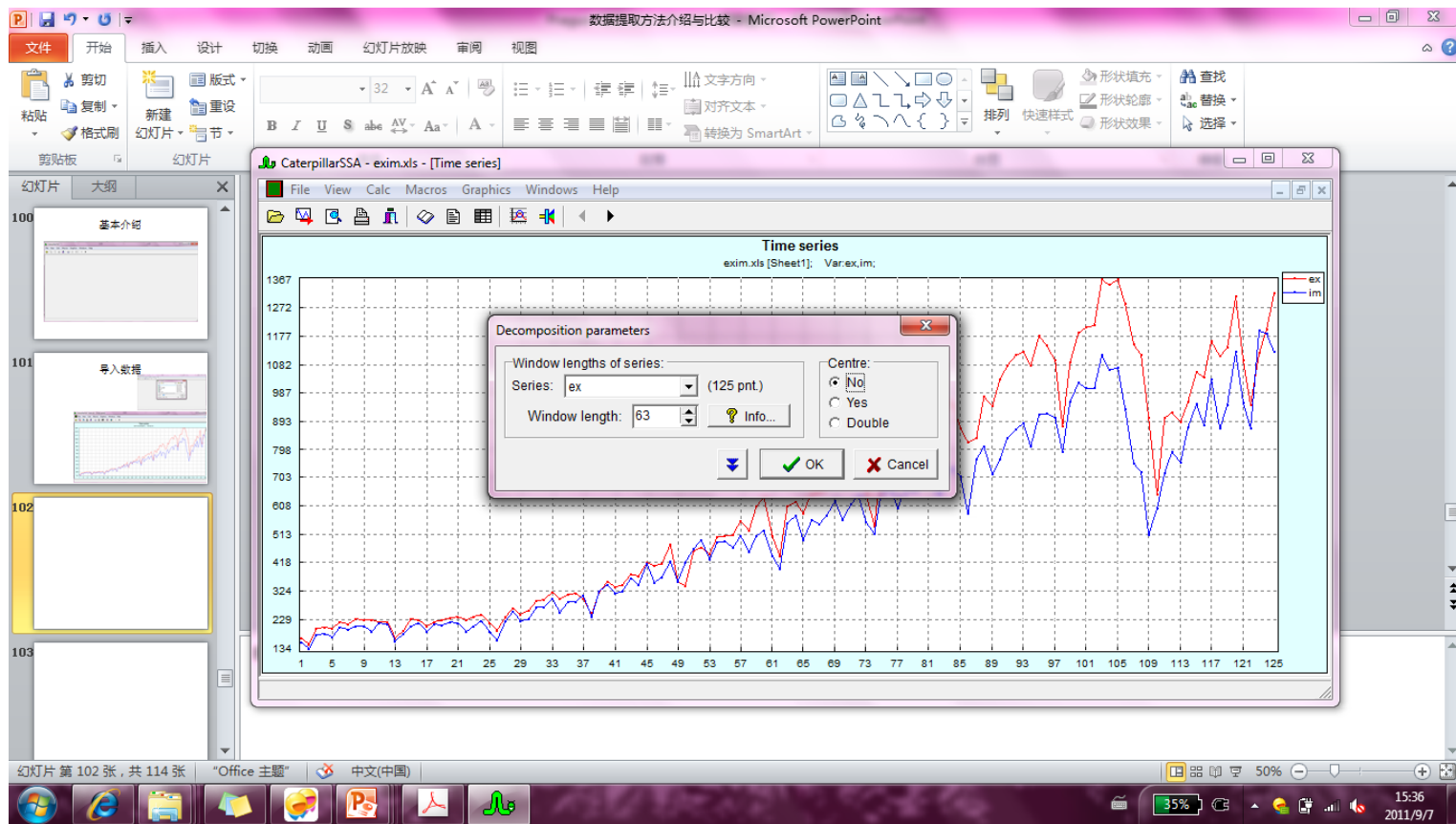
基本介绍



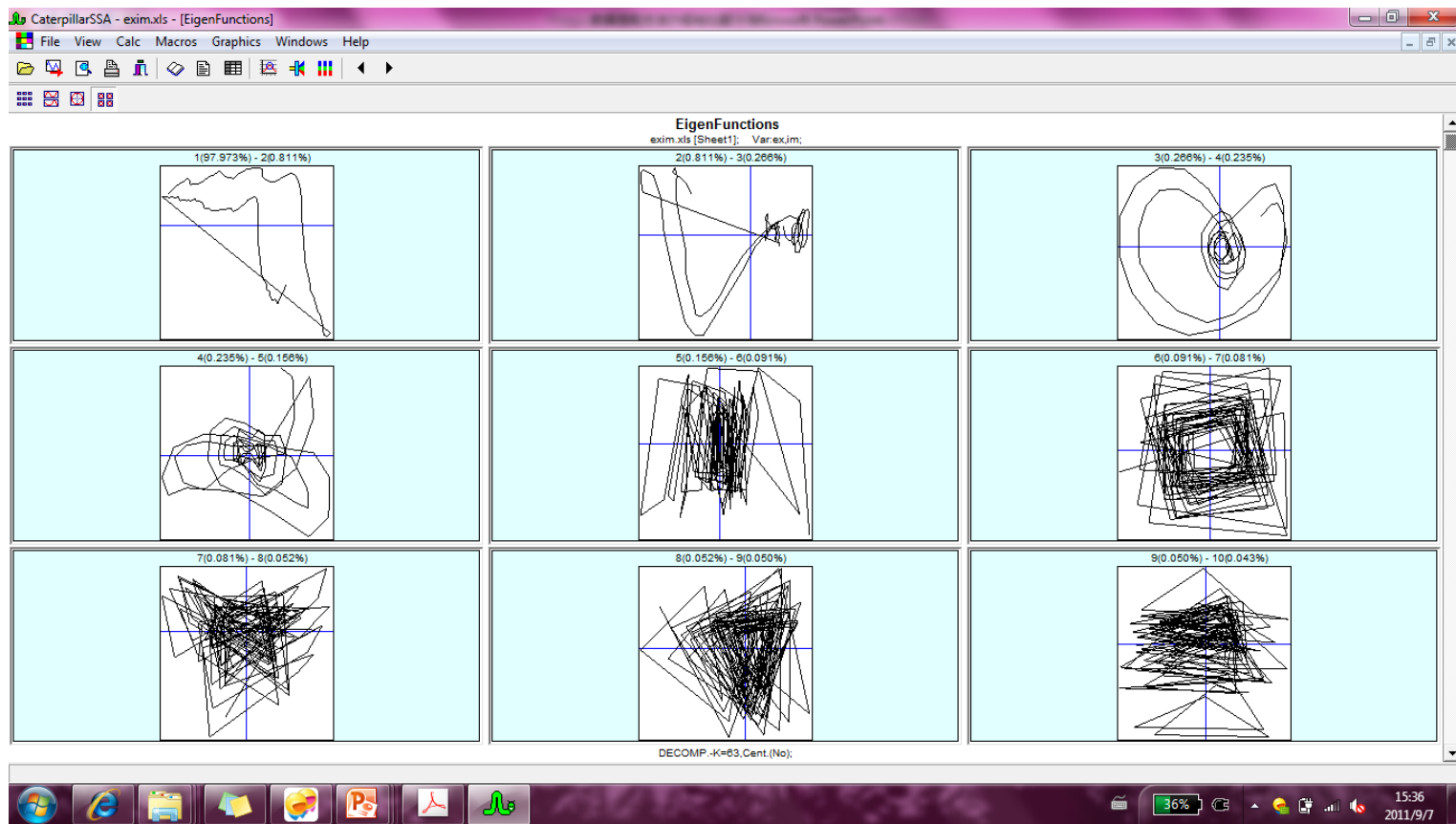
导入数据



数据分解与参数选择



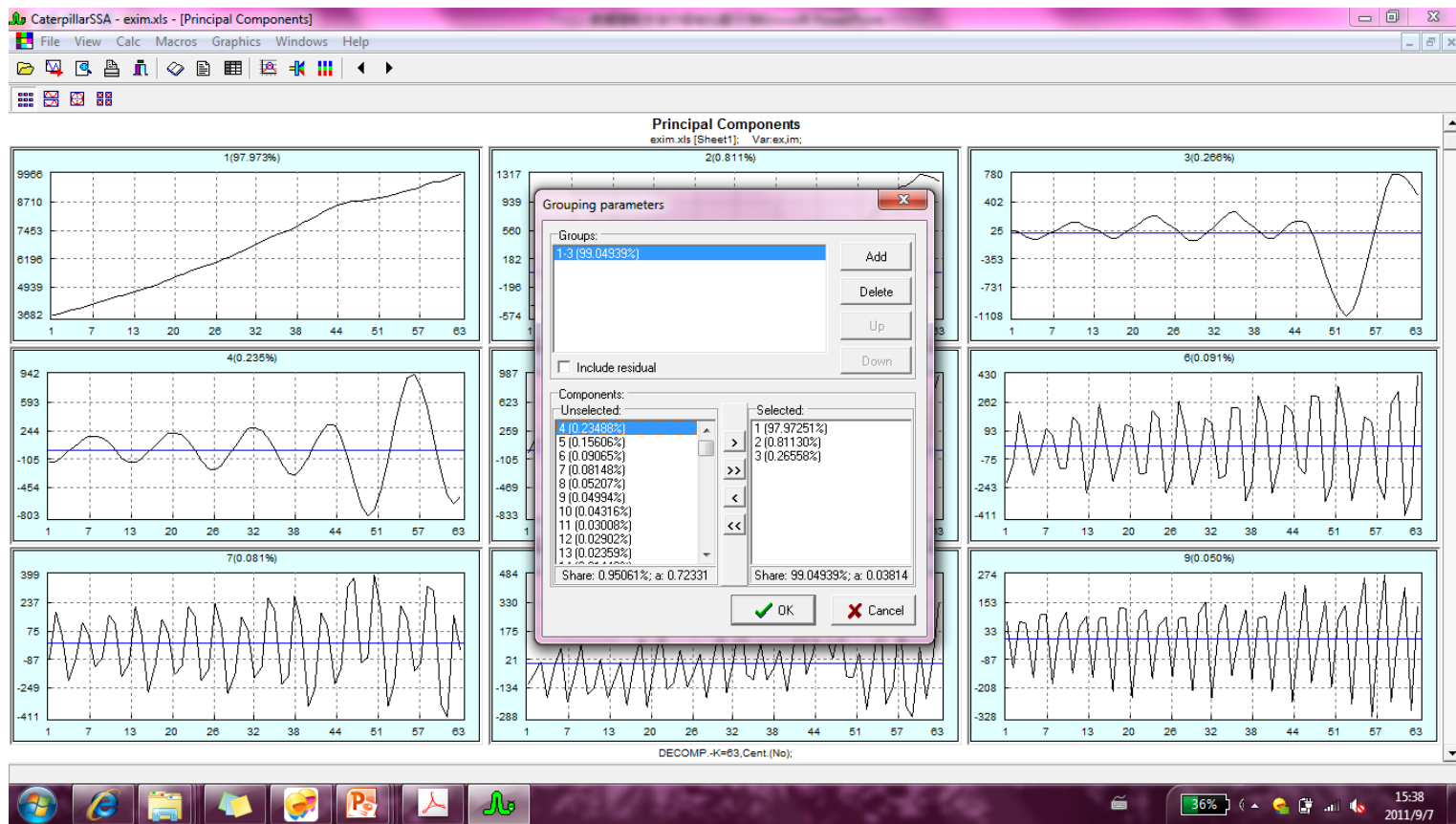
数据分解结果



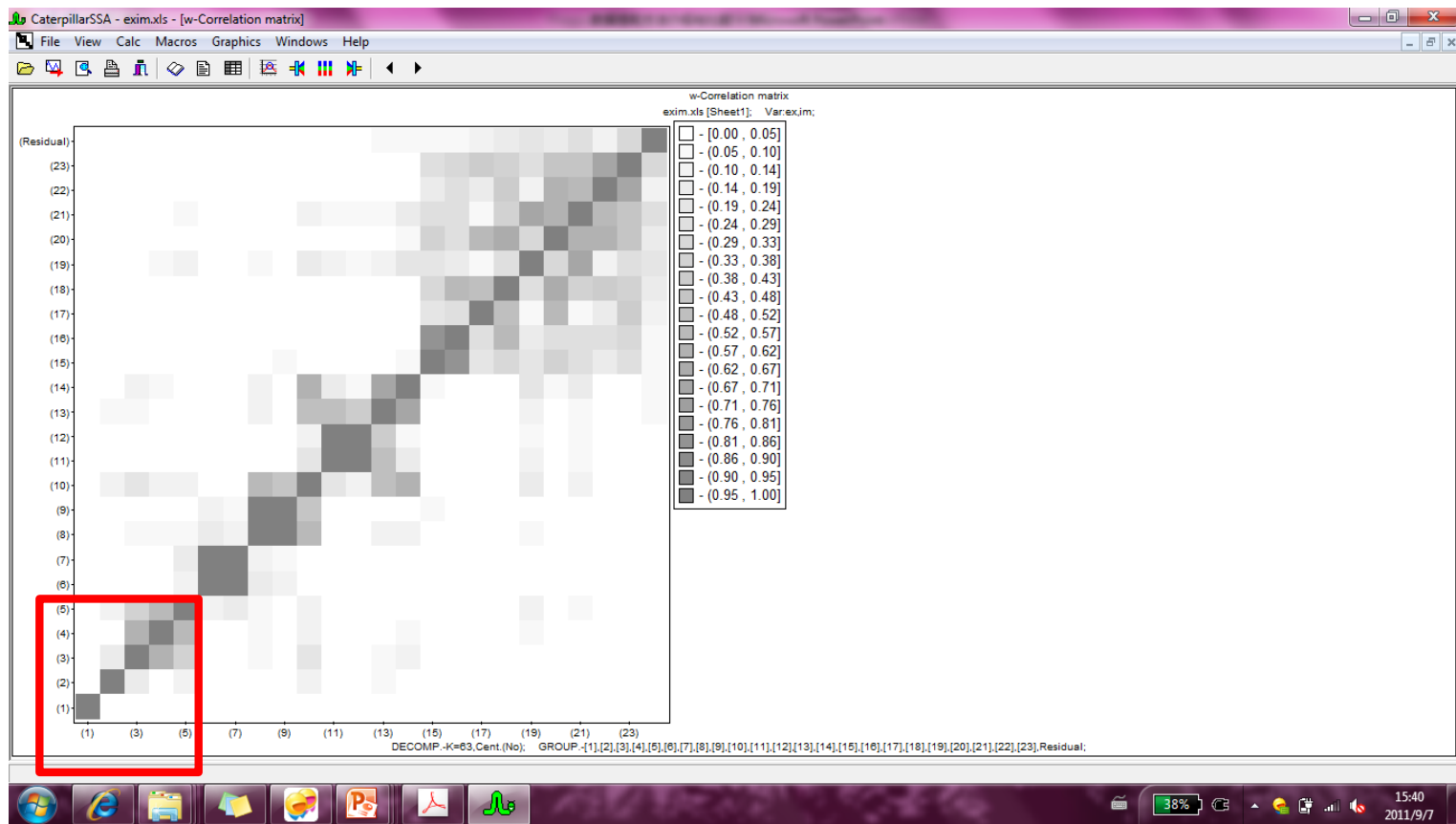
选择其它模式查看



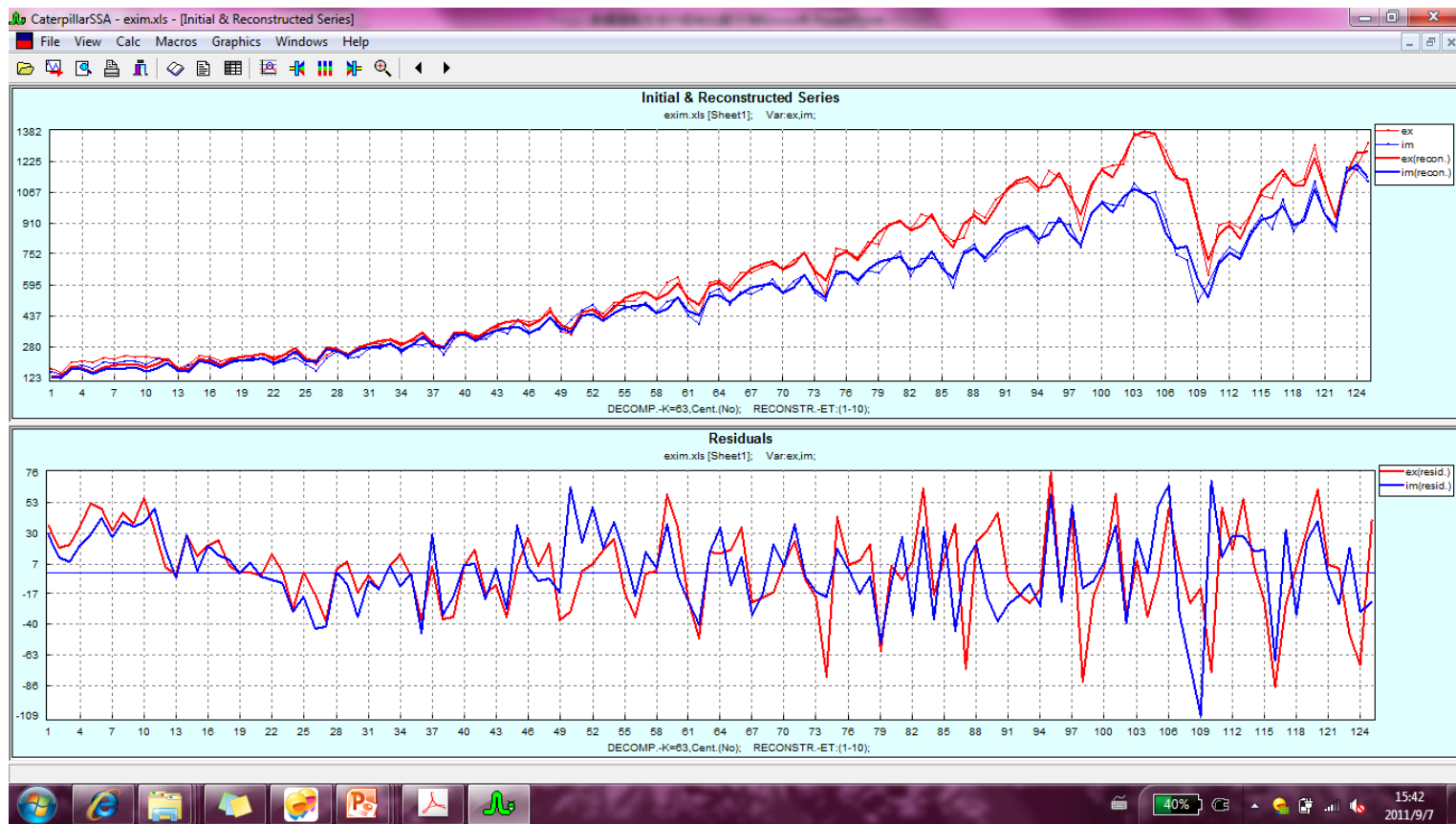
数据分组



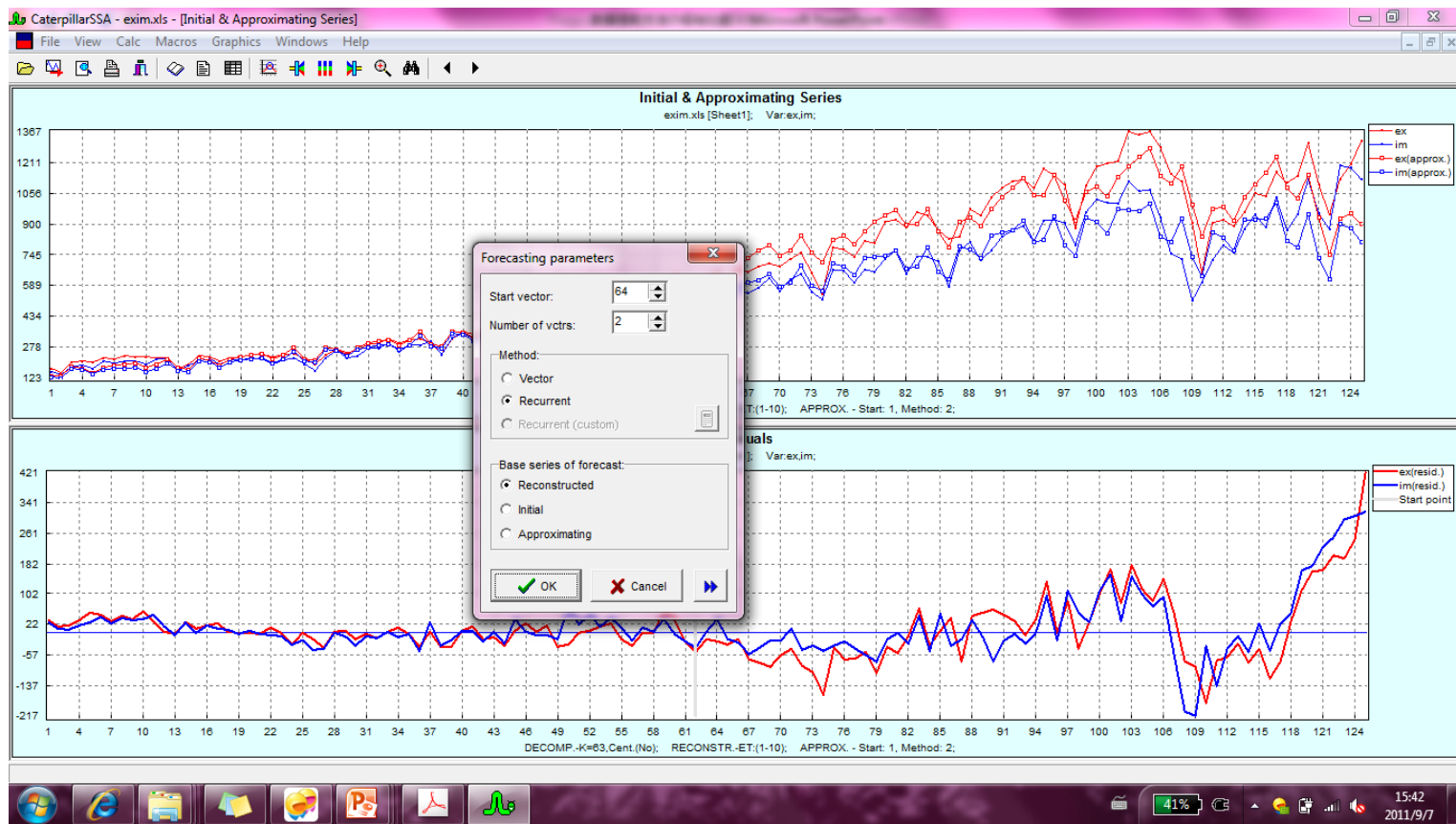
数据分组参数选择



数据重组



预测



DFA程序实现

DOWNLOADS

- ◉ R-package on CRAN
- ◉ A suitable text-editor
 - Winedit for R
 - Tinn for R: **single user setting (SDI)**
 - ...
- ◉ Install the package signalextraction from CRAN
- ◉ Load the package signalextraction into R
- ◉ Ready to go!

EXAMPLES

- ◉ Based on the example-file contained in
`vignettes' (contained in package
signalextraction)
- ◉ Level-criteria (KOF Economic Barometer)
 - Various settings emphasizing time series dynamics (d)
and user-preferences (i2)
- ◉ Turning-point criterion
 - Various settings emphasizing
 - Speed
 - Reliability/smoothness
 - Risk-aversion

RUNNING THE PACKAGE

◉ Fit<-dfa(x,...)

- '`<-`' means 'assign' (can be replaced by '`=`')
- The optimization is launched
- 'Fit' is an object of class 'dfa' which contains
 - Filtered series
 - Filter characteristics: amplitude and time delay
 - Filter coefficients
 - ...
- This information can be accessed by addressing the components through `Fit$name`

PLOTS

◉ Plot(Fit):

- Original and filtered data
- Periodograms of original and filtered data
- Filter characteristics
 - Amplitude
 - Time delay

FUNCTION-CALL AND PARAMETERS

- ⊙ `Fit<-dfa(x, quart = FALSE, d = 0, tpfiler = FALSE, lambda = 3, expweight = 1.5, pbd = 1.02, pb = 1/14, sb = 1/7, n.loops = 10, verbose = 1, limamp = 3, i2 = TRUE)`
- ⊙ `x`:
 - data: univariate time series
- ⊙ `quart`
 - FALSE: monthly data, TRUE: quarterly data

FUNCTION-CALL AND PARAMETERS

◉ d:

- 0: optimization criterion 3.7 (best for bounded data)
- 1: optimization criterion 6.7 (best for asymptotically unbounded data)

◉ tpfilter

- FALSE: level criteria 3.7 or 6.7 (depending on d)
- TRUE: TP-criterion 5.4 (irrespective of the value of d)

FUNCTION-CALL AND PARAMETERS

⊙ lambda:

- activated for TP=TRUE only
- **controls the speed** of the real-time filter
 - the weight attributed to the **time delay in the pass-band** in criterion 5.4

⊙ expweight:

- activated for TP=TRUE only
- controls the reliability (smoothness) of the real-time filter
 - the weight attributed to the **amplitude function in the stop-band** in criterion 5.4

FUNCTION-CALL AND PARAMETERS

⊙ Pbd (>1)

- a regularity-constraint ensuring stability of the filter, see 3.20
- default values are based on experience

⊙ pb and sb

- determine the signal (ideal trend)
- width of pass- and stop-bands

⊙ n.loops

- controls numerical optimization
- number of loops (genetic algorithm)
- the smaller n.loops the faster the optimization

FUNCTION-CALL AND PARAMETERS

- ⦿ **verbose**

- Controls output generated by numerical optimization routine

- ⦿ **limamp (>1)**

- regularity constraint
- controls the amplitude function in the pass-band

- ⦿ **i2**

- filter constraint
- i2=TRUE: the time delay of the real time filter vanishes in frequency zero

X12-ARIMA

The screenshot displays the EViews software interface. The main window shows a list of series on the left, including 'cpi', 'cpi_f_m1', 'eq_cpi_s', 'm1', 'm2', 'r_gdp', 'resid', 'rpi', 'series01', 'tabforeca', 'wage_in', and 'annual'. The 'CPI_F_HB' series is selected, and its data is displayed in a table. The table has columns for time periods and values. The 'X12 Options' dialog box is open, showing the 'ARIMA Options' tab. The 'Data Transformation' section has 'None' selected. The 'ARIMA Spec' section has 'Select from file' selected. The 'Specification from File' section has 'File: c:\progra~1\views5\x12a.mdl' and 'Select by out-of-sample fit (default in-sample)' checked. The 'Path = d:\cpi-prg DB = none WF = data_prg' is shown at the bottom.

Program: CPI_FORECAST - (d:\cpi-prg\cpi_forecast.prg)

Workfile: DATA_PRG - (d:\cpi-prg\data_prg.wf1)

Range: 1952 2012 -- 61 obs
Sample: 2000 2012 -- 13 obs

Series: CPI_F_HB Workfile: DATA_PRG/month_nf

View/Proc/Object/Properties/Print/Name/Freeze/Default/Sort/Edit+/-/Smpl+/-/Label

CPI_F_HB

Last updated: 09/07/11 - 15:44
Imported from 'd:\cpi-prg\data_prg.xls'

Time Period	Value
2001M01	0.9
2001M02	0.1
2001M03	-0.6
2001M04	0.2
2001M05	-0.5
2001M06	-1.2
2001M07	-0.5
2001M08	0.1
2001M09	0.9
2001M10	0.3
2001M11	-0.2
2001M12	0.1
2002M01	0.3
2002M02	

X12 Options

Trading Day/Holiday Seasonal Adjustment Outliers Diagnostics

ARIMA Options

Data Transformation

- ☒ None
- ☐ Auto (log or none)
- ☐ Logistic
- Box-Cox power: [0] (0 for log)

Regressors

- ☐ Constant
- ☐ Seasonal dummies

ARIMA Spec

- ☐ No ARIMA
- ☐ Specify in-line
- ☒ Select from file

ARIMA Estimation Sample

(subsample of workfile sample)

Specification from File

File: c:\progra~1\views5\x12a.mdl Browse...

- ☐ Select best (default is first acceptable)
- ☒ Select by out-of-sample fit (default in-sample)

OK Cancel

Path = d:\cpi-prg DB = none WF = data_prg

42% 15:45 2011/9/7