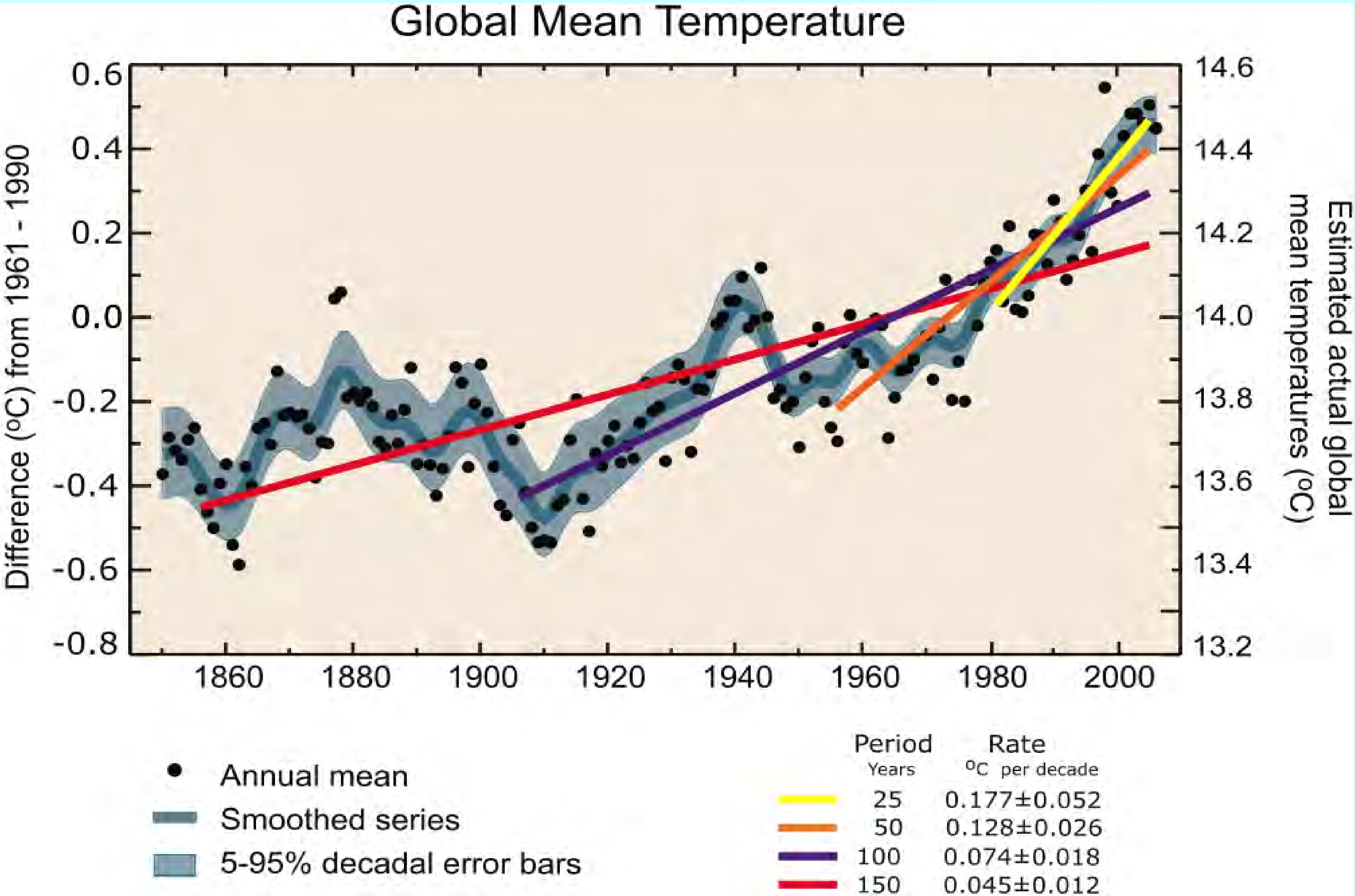
On Noise, Cycle and Trend in Climate Data

Norden E. Huang

Research Center for Adaptive Data Analysis
Center for Dynamical Biomarkers and translational Medicine
National Central University

IPCC Global Mean Temperature Trend



Problems

- The data is derived from simple annual mean of noisy monthly data.
- The trends are determined using simple straight lines.

- What are the cycles and the trend?

- How to reduce the noise (not just to eliminate specific features) ?

After more than 15 years of searching, I found the key to nonlinear and nonstationary data analysis is the proper definition for frequency.

Then:

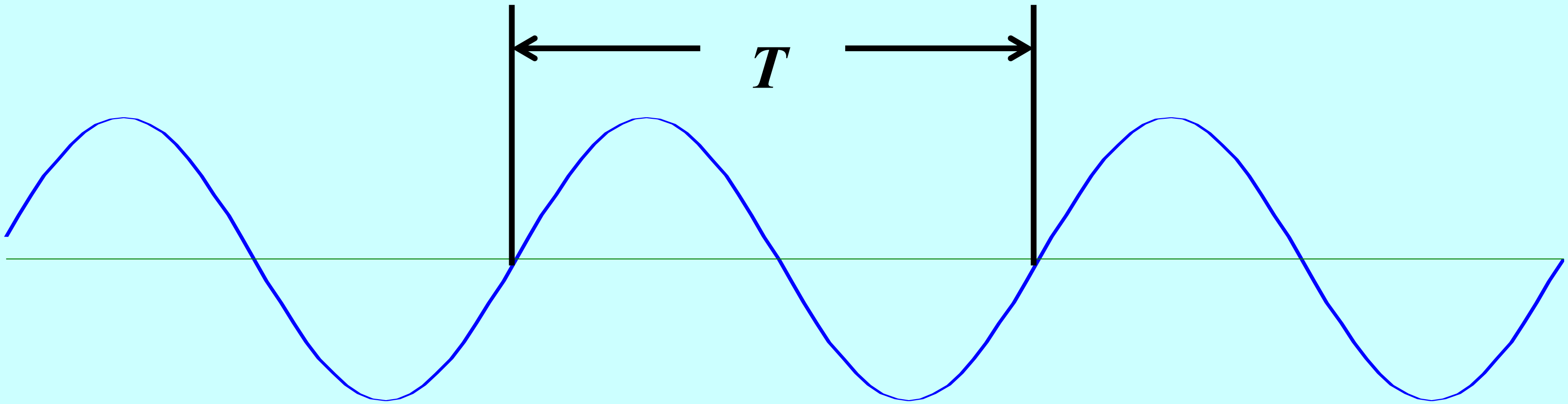
In search of frequency, I found the methods to cleanse the data, quantify nonlinearity and determine the trend.

What is Frequency?

Definition of Frequency

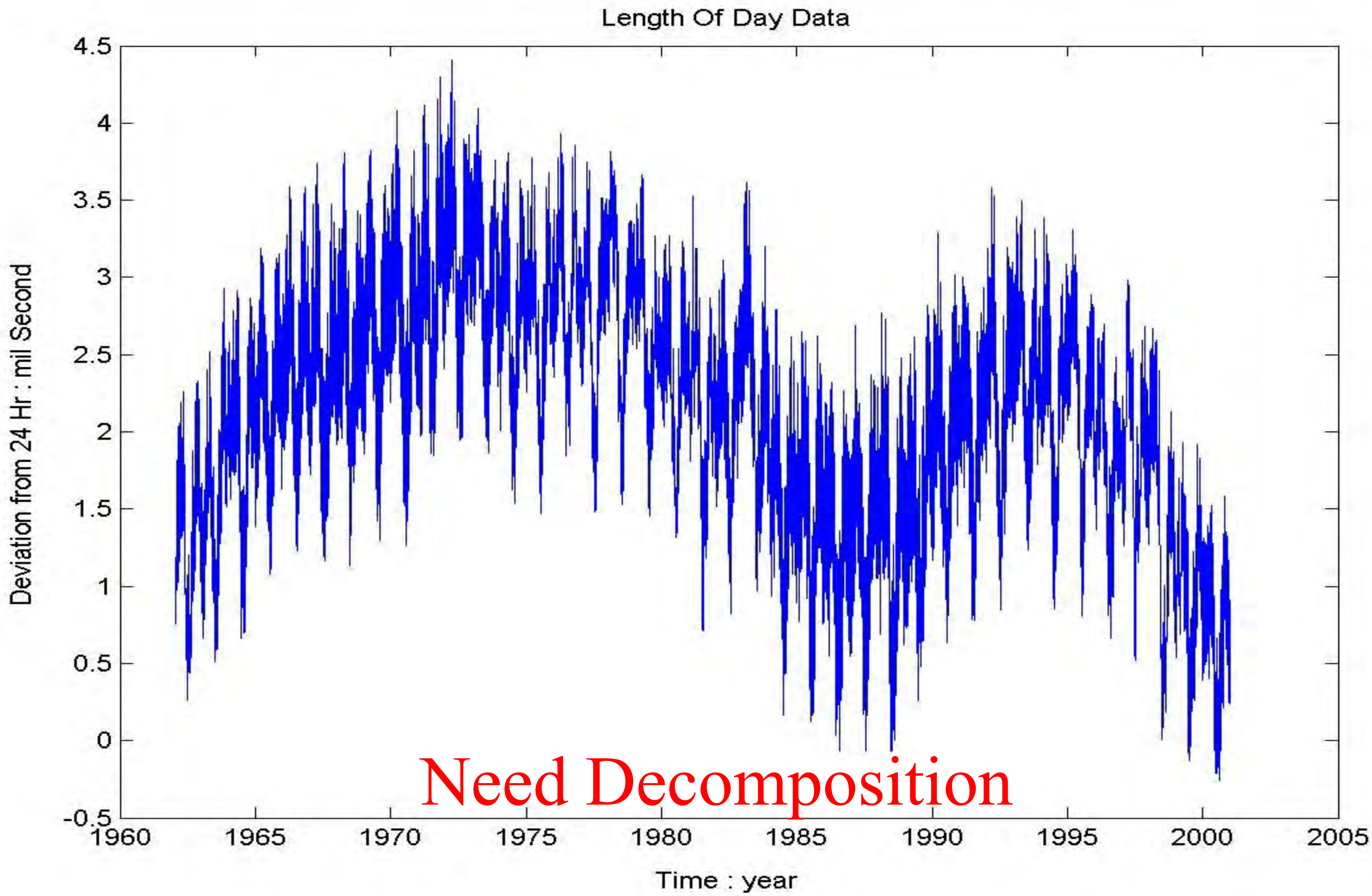
Given the period of a wave as T ; the frequency is defined as

$$\omega = \frac{1}{T}\,.$$



This definition is easy for regular sine wave, but not very practical for complicated oscillations.

Impossible for complicated waves



Jean-Baptiste-Joseph Fourier

Fourier's work is a great mathematical poem.
Lord Kelvin

1807 ***"On the Propagation of Heat in Solid Bodies"***

1812 ***Grand Prize of Paris Institute***

1817 ***Elected to Académie des Sciences***

1822 ***Appointed as Secretary of Math Section***

paper published

Ever since Fourier's ground breaking work, people always think of any signal in terms of waves, therefore cycles and frequency.

Fourier Analysis

Fourier proved that the Fourier expansion in terms of trigonometric function

$$x(t) = \sum_j a_j \, e^{i\omega_j t}$$

is complete, convergent, orthogonal and unique. Therefore, every signal could be think as combination of sinusoidal waves each with constant amplitude and frequency.

Are those frequency physically meaningful?

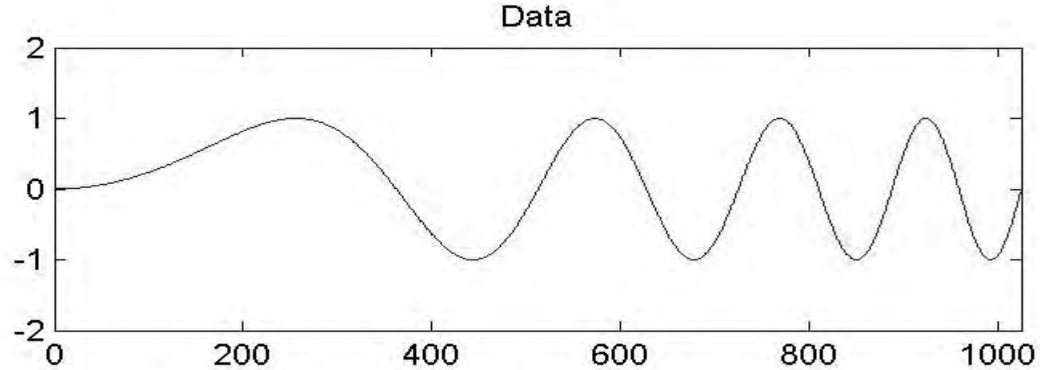
Definition of Power Spectral Density

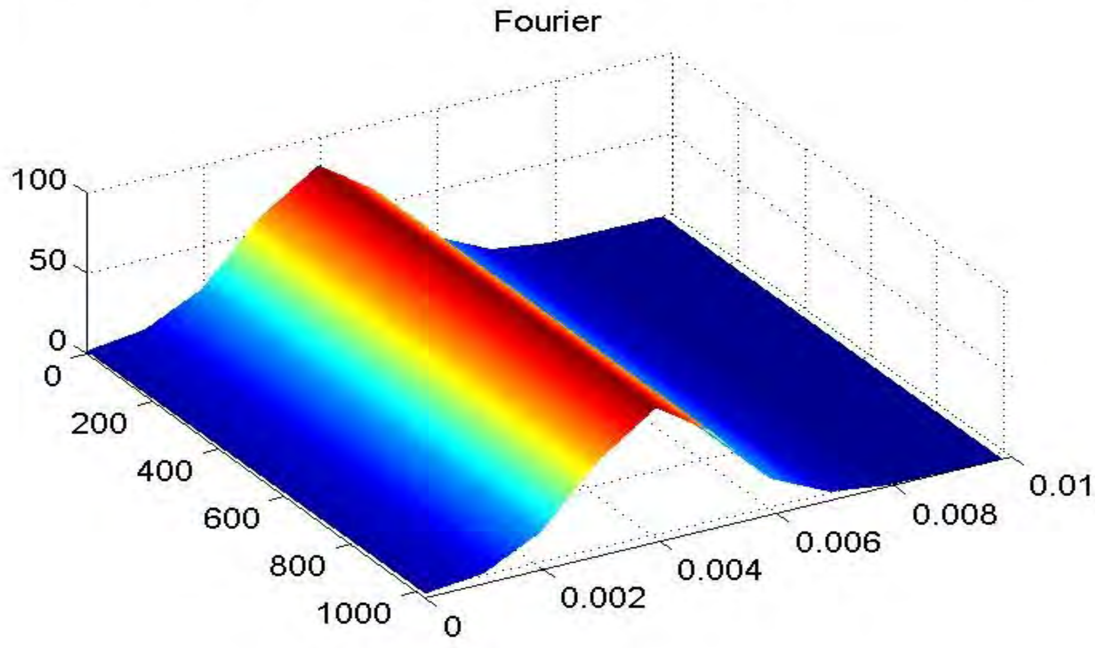
Since a signal with nonzero average power is not square integrable, the Fourier transforms do not exist in this case.

Fortunately, the Wiener-Khinchin Theroem provides a simple alternative. The PSD is the Fourier transform of the auto-correlation function, $R(\tau)$, of the signal if the signal is treated as a wide-sense stationary random process:

$$S(\omega) = \int_{-\infty}^{\infty} R(\tau) e^{-2i\omega\tau} d\tau \quad \Rightarrow \quad \int_{-\infty}^{\infty} S(\omega) d\omega = \overline{\varsigma^2(t)}$$

Fourier Spectrum

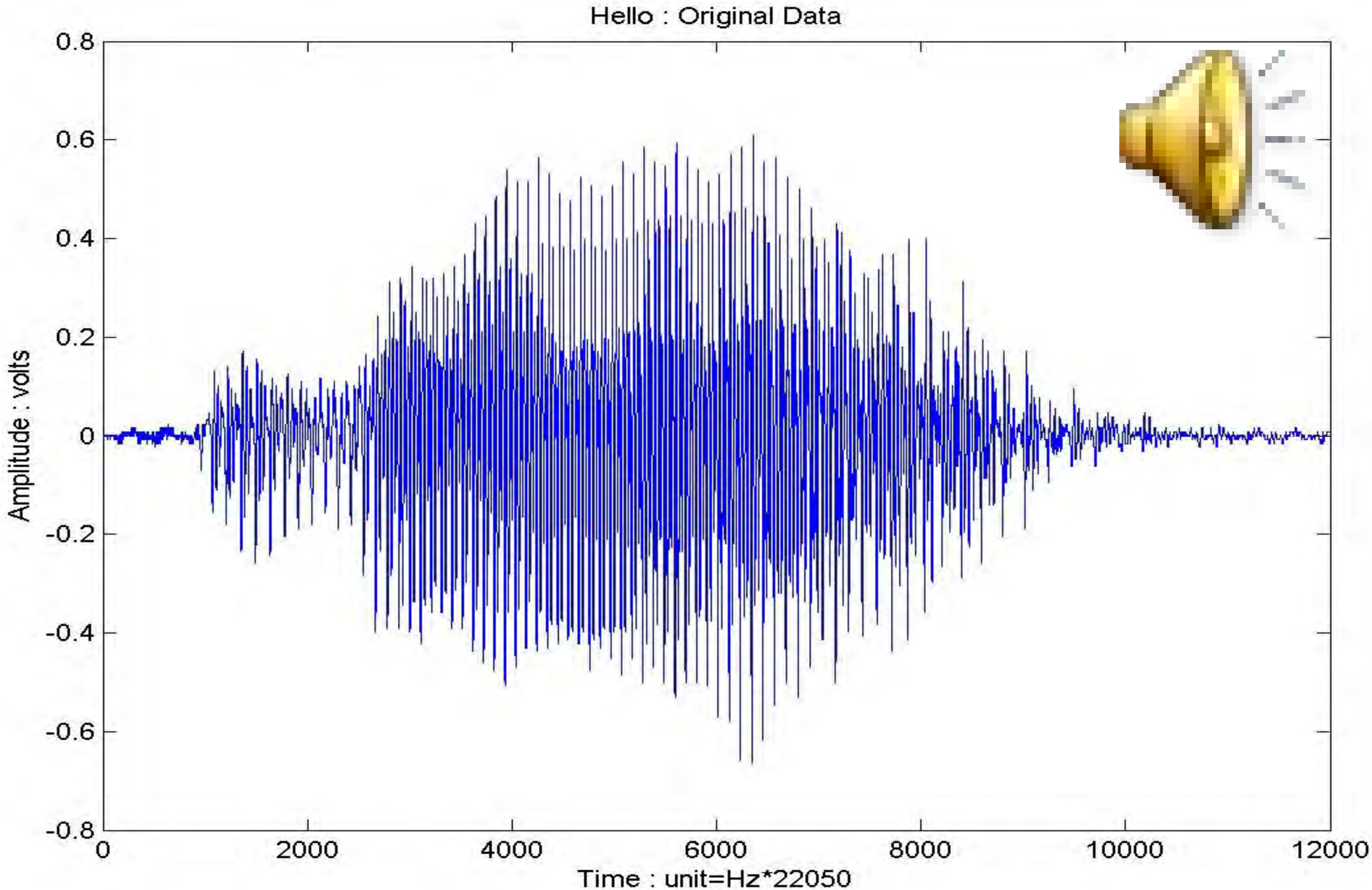




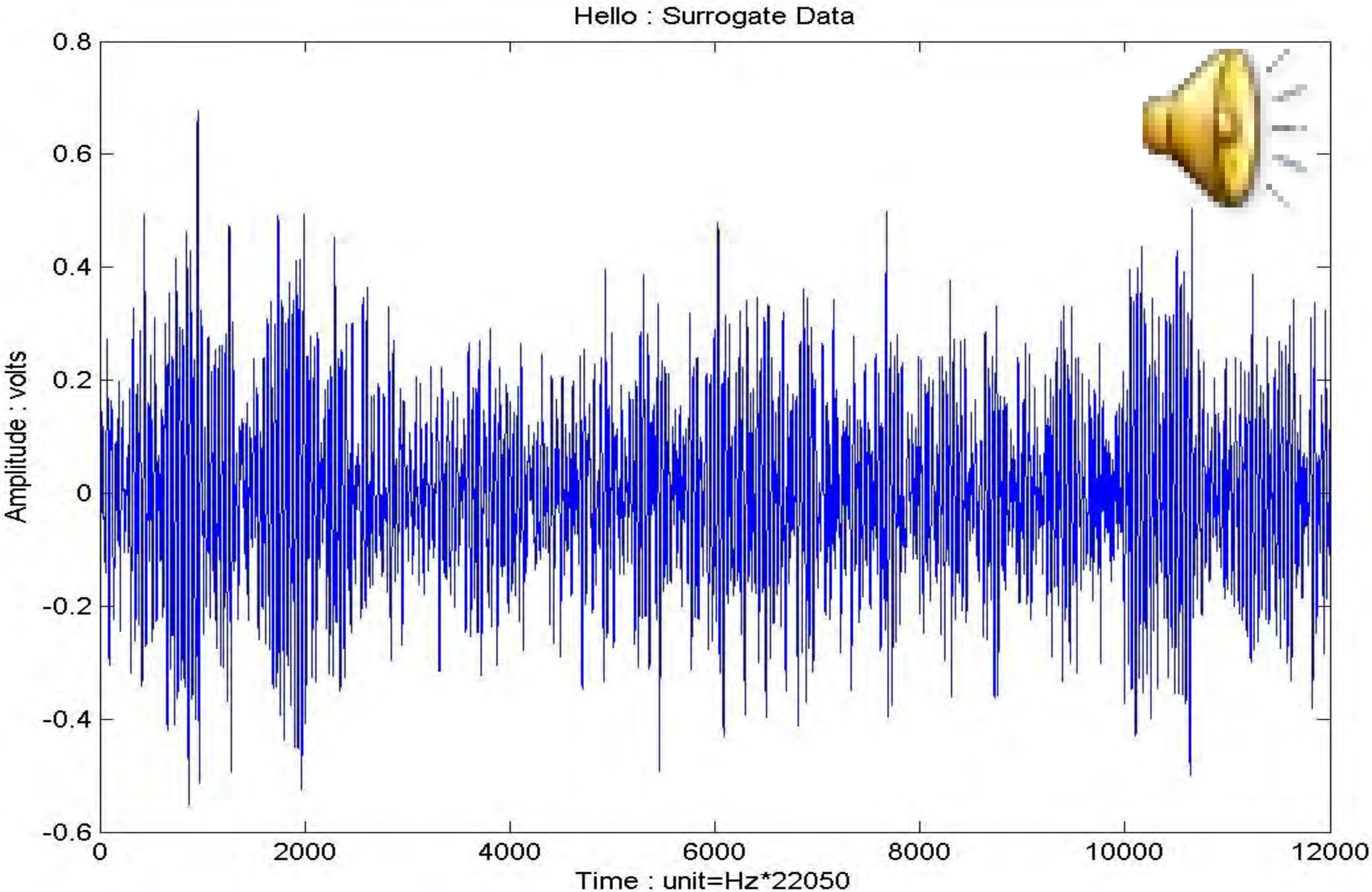
Surrogate Signal

I. Hello

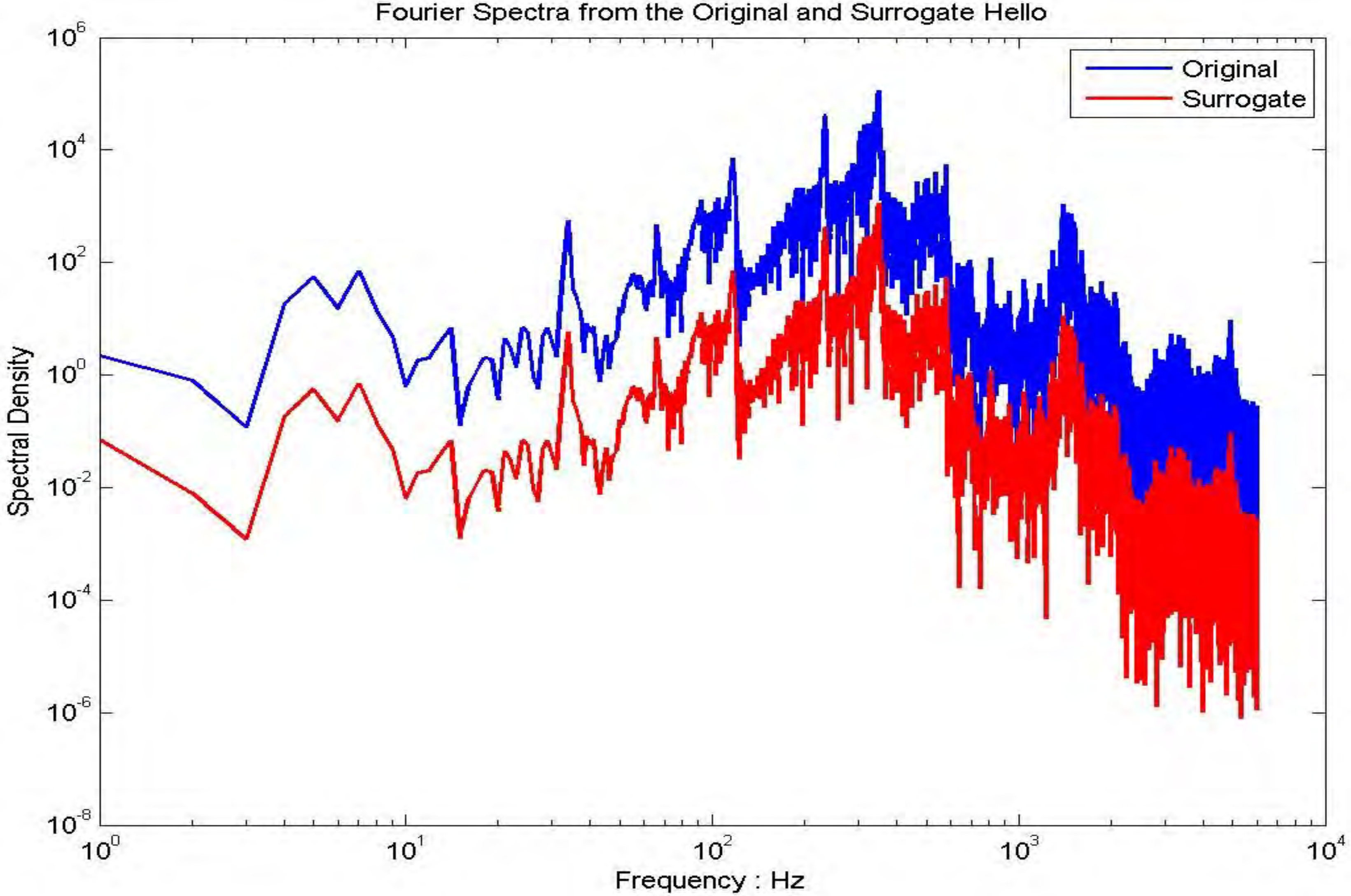
The original data : Hello



The surrogate data : Hello



The Fourier Spectra : Hello



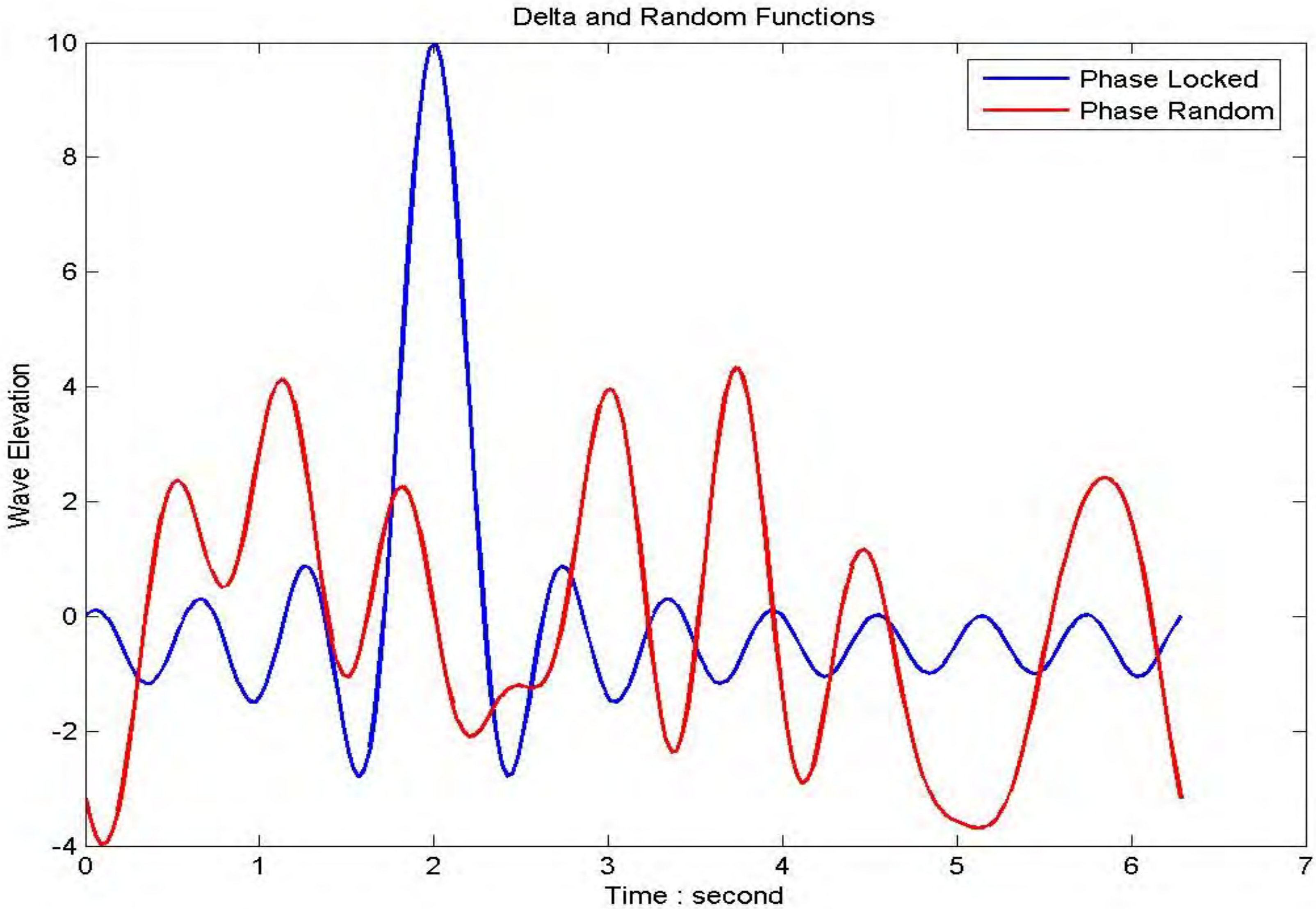
Surrogate Signal

II. delta function and white noise

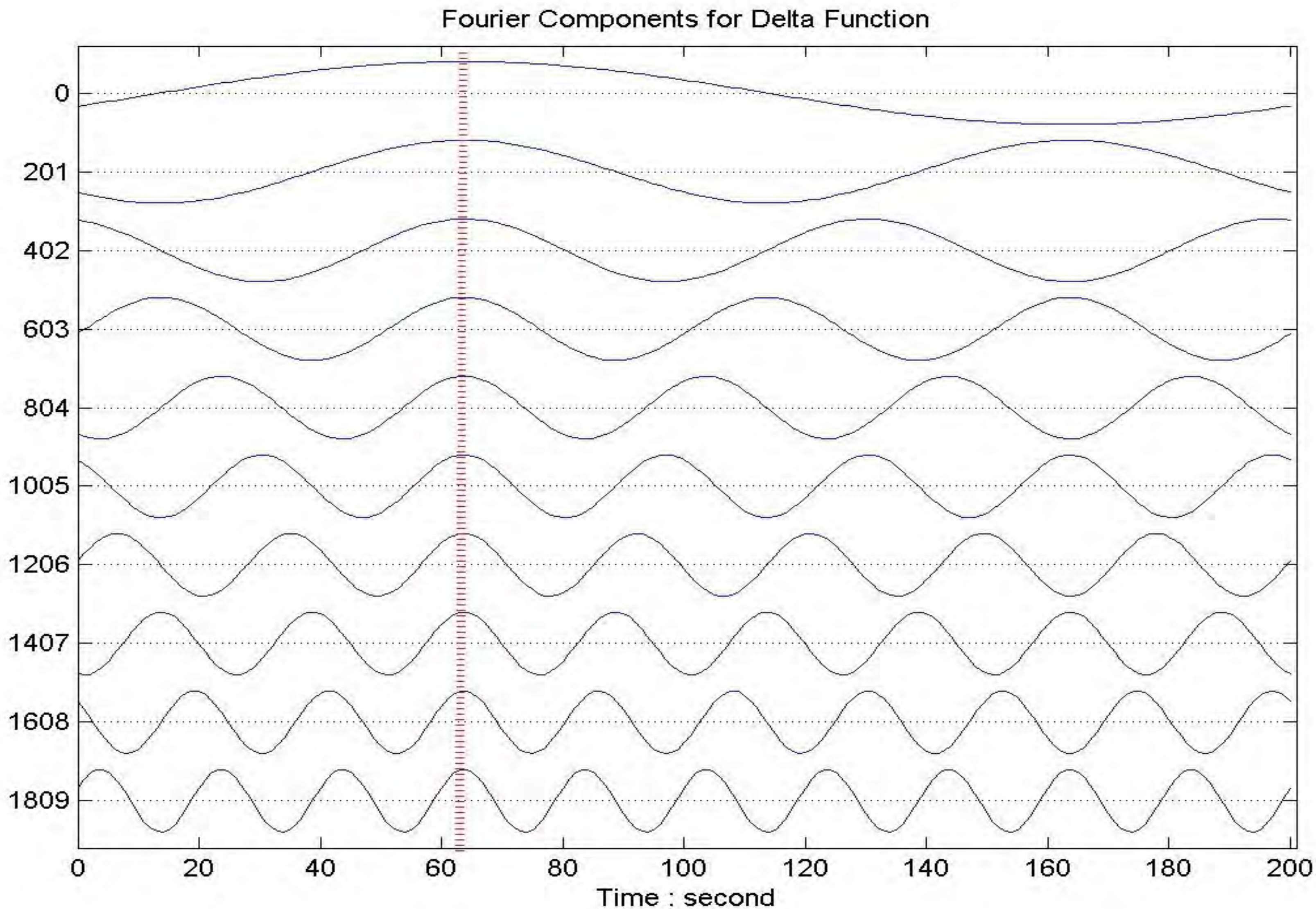
Non-causality:

Event involves both past and future

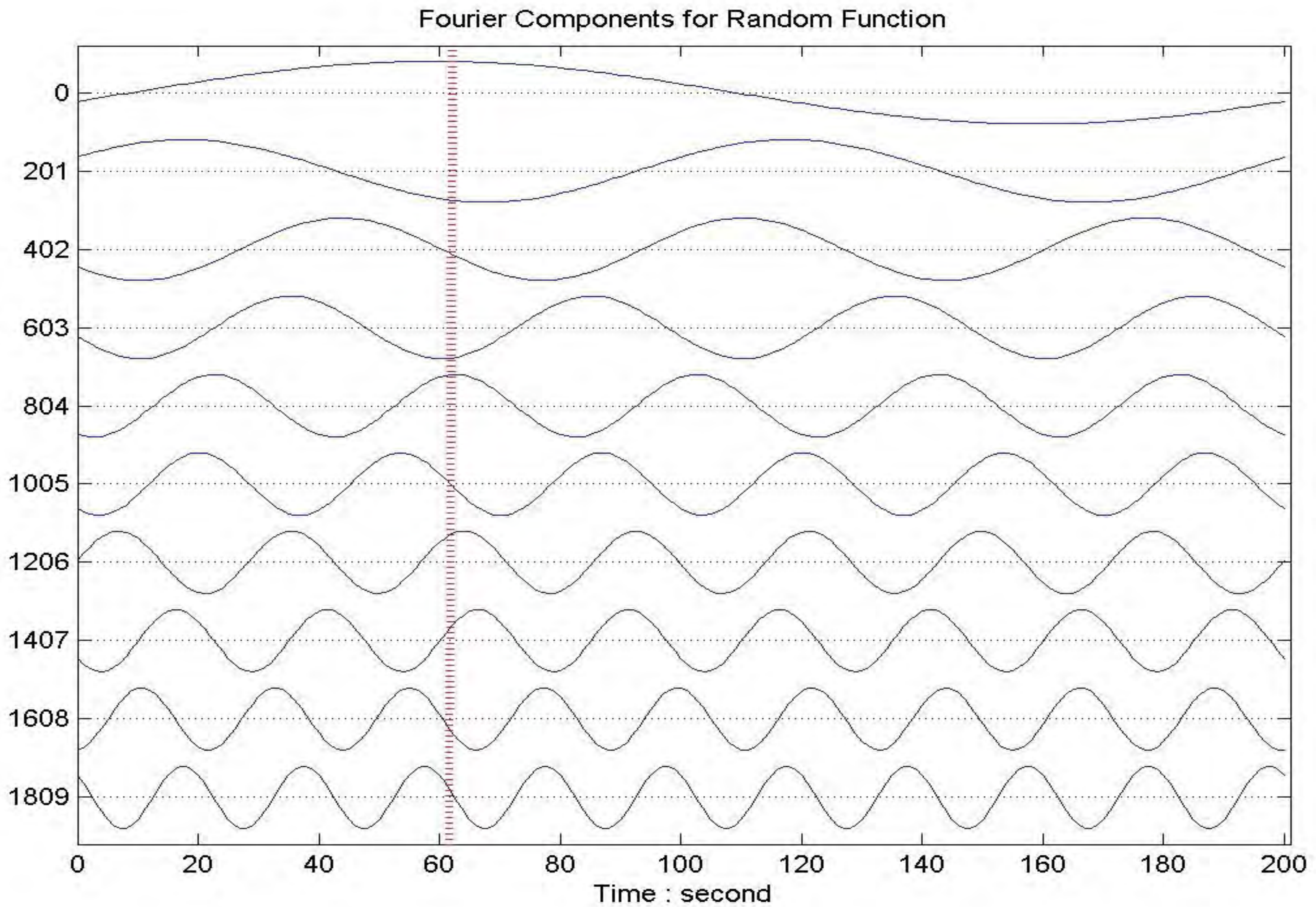
Random and Delta Functions



Fourier Components : Delta Function

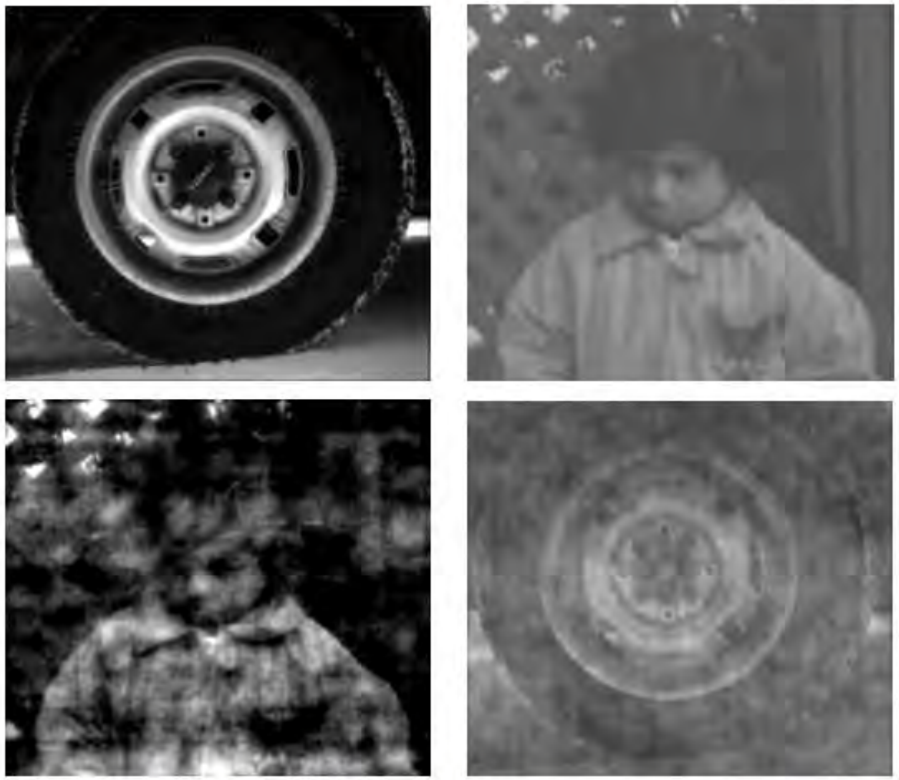


Fourier Components : Random Function



The Importance of Phase

Human Visual System – Importance of Phase Information



Surrogate images. *Top*: Original images I_1 and I_2; *Bottom*: Images $\hat{I}_1$ and $\hat{I}_2$ generated by exchanging the amplitude and phase spectra of the original images.

Can we just use phase to explore physical processes?

Yes, to some extent.

How to define frequency?

It seems to be trivial.

But frequency is an important parameter for us to understand many physical phenomena.

Instantaneous Frequency

$$Velocity = \frac{distance}{time}\ ;\ mean\ \ velocity$$

$$Newton \Rightarrow v = \frac{dx}{dt}$$

$$Frequency = \frac{1}{period}\ ;\ mean\ \ frequency$$

$$HHT\ defines\ the\ \ phase\ \ function\ \Rightarrow\ \ \omega = \frac{d\theta}{dt}$$

So that both v and ω can appear in differential equations.

The Idea and the need of Instantaneous Frequency

According to the classic wave theory, the wave conservation law is based on a gradually changing ***φ(x,t)*** such that

$$\vec{k} = \nabla \varphi \ , \qquad \omega = -\frac{\partial \varphi}{\partial t} \ ;$$

$$\Rightarrow \ \frac{\partial \vec{k}}{\partial t} + \nabla \ \omega = 0 \ .$$

Therefore, both wave number and frequency must have instantaneous values and differentiable.

This is the true definition of frequency.

Instantaneous Frequencies and Trends for Nonstationary Nonlinear Data

Hot Topic Conference
University of Minnesota,
Institute for Mathematics and Its Applications, 2011

Prevailing Views on *Instantaneous Frequency*

The term, Instantaneous Frequency, should be banished forever from the dictionary of the communication engineer.

J. Shekel, 1953

The uncertainty principle makes the concept of an Instantaneous Frequency impossible.

K. Gröchennig, 2001

Hilbert Transform : Definition

For any $x(t) \in L^p$,

$$y(t) = \frac{1}{\pi} \wp \int_{\tau} \frac{x(\tau)}{t-\tau} \, d\tau \, ,$$

then, $x(t)$ *and* $y(t)$ *form the analytic pairs:*

$$z(t) = x(t) + i\, y(t) = a(t)\, e^{i\theta(t)} \, ,$$

where

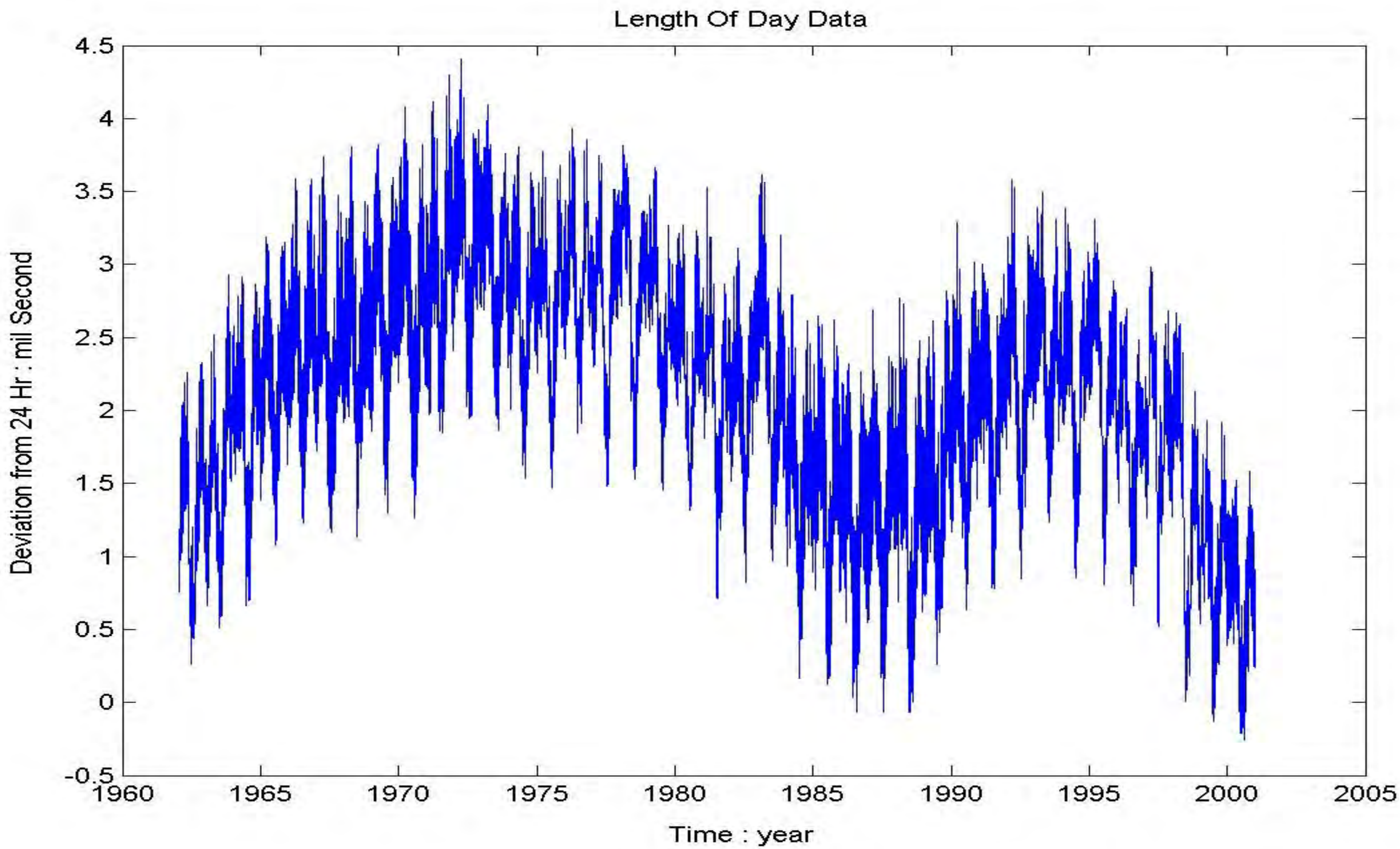
$$a(t) = \left(x^2 + y^2 \right)^{1/2} \text{ and } \theta(t) = \tan^{-1} \frac{y(t)}{x(t)} \, .$$

The Traditional use of the Hilbert Transform for Data Analysis

failed miserably
and gave IF a bad break.

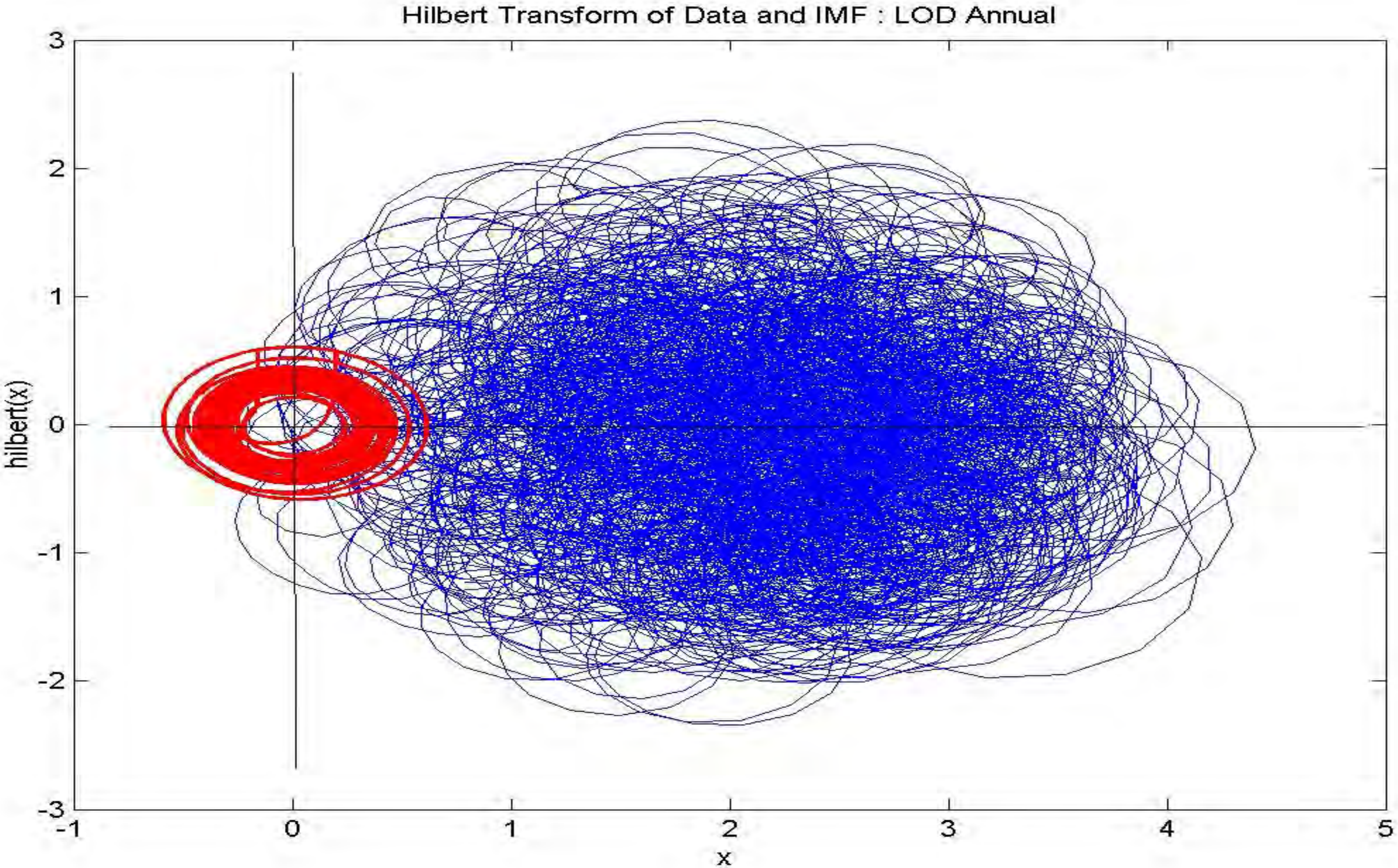
Traditional View

a la Hahn (1995) : Data LOD



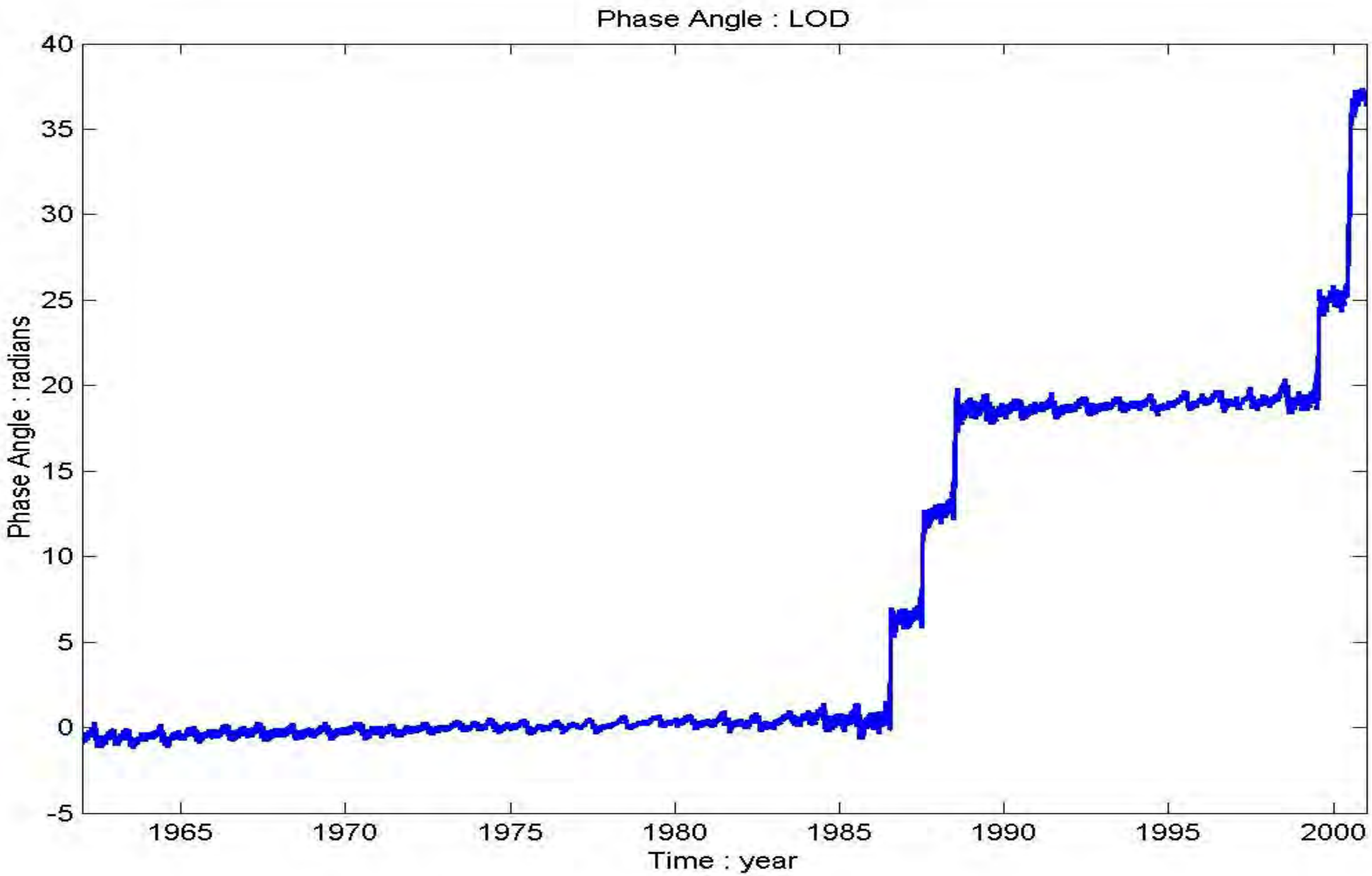
Traditional View

a la Hahn (1995) : Hilbert



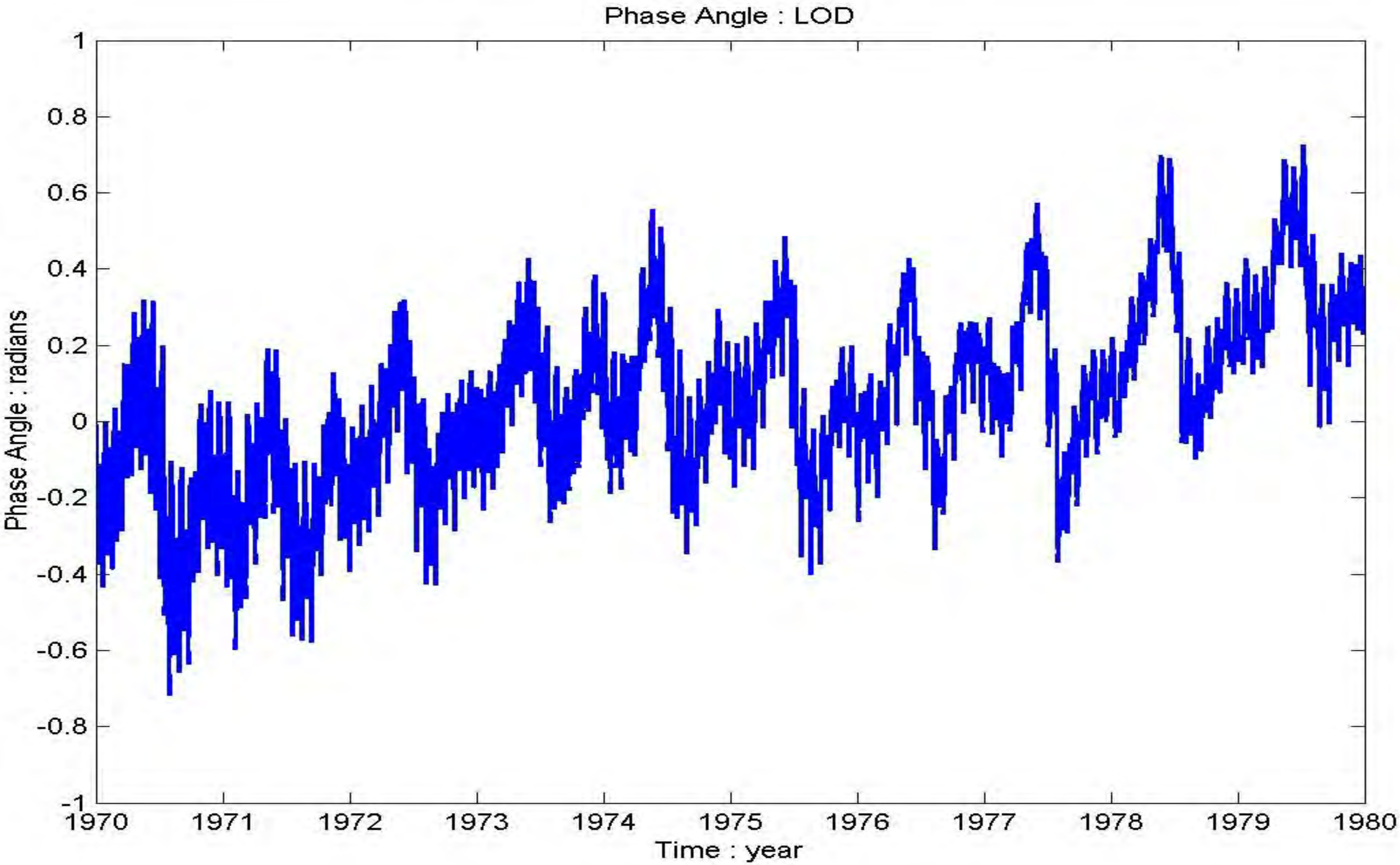
Traditional Approach

a la Hahn (1995) : Phase Angle



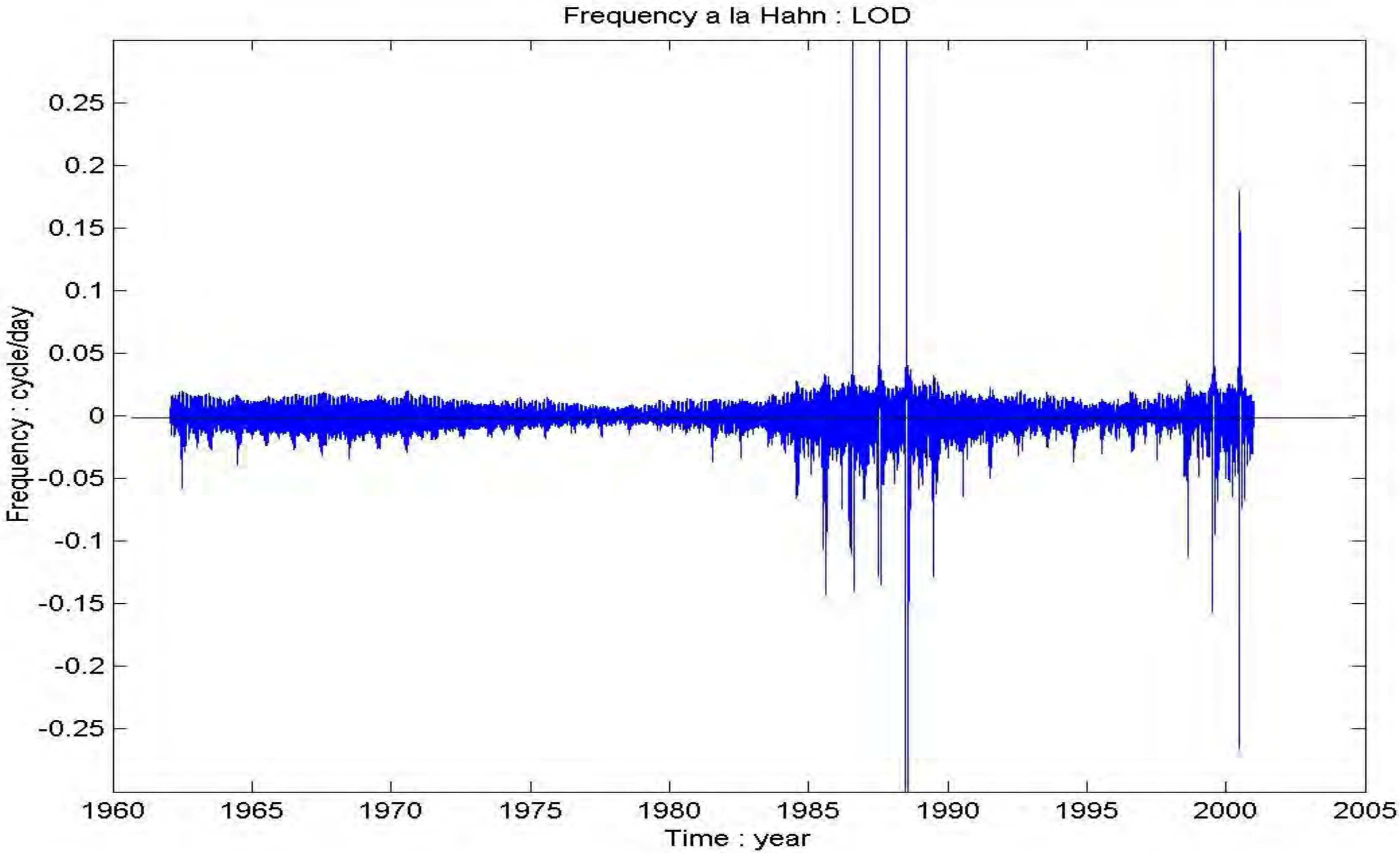
Traditional Approach

a la Hahn (1995) : Phase Angle Details



Traditional Approach

a la Hahn (1995) : Frequency



The combination of Hilbert Spectral Analysis and Empirical Mode Decomposition has been designated

by NASA as

HHT

(HHT vs. FFT)

Comparison between FFT and HHT

1. FFT :

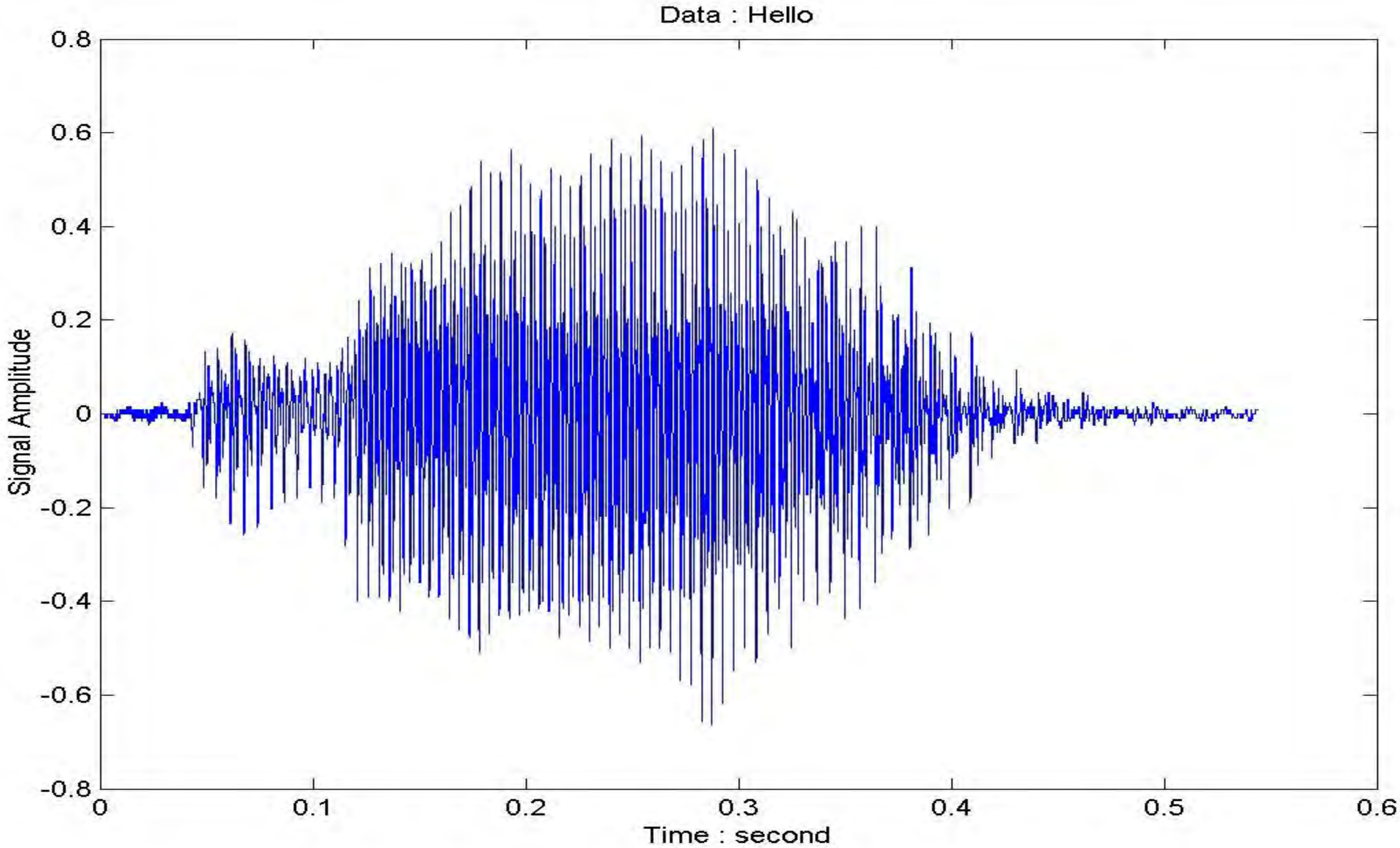
$$x(t) = \Re \sum_j a_j e^{i\omega_j t} \quad .$$

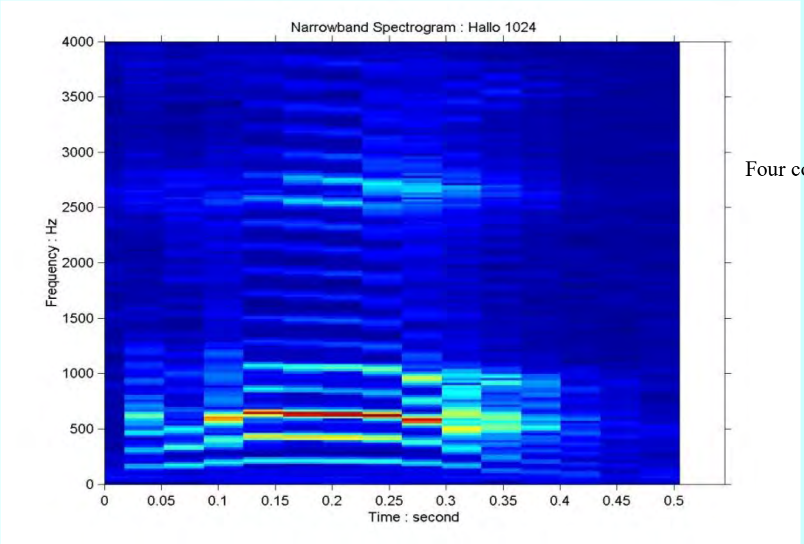
2. HHT :

$$x(t) = \Re \sum_j a_j(t)\, e^{i\int_t \omega_j(\tau)d\tau} \quad .$$

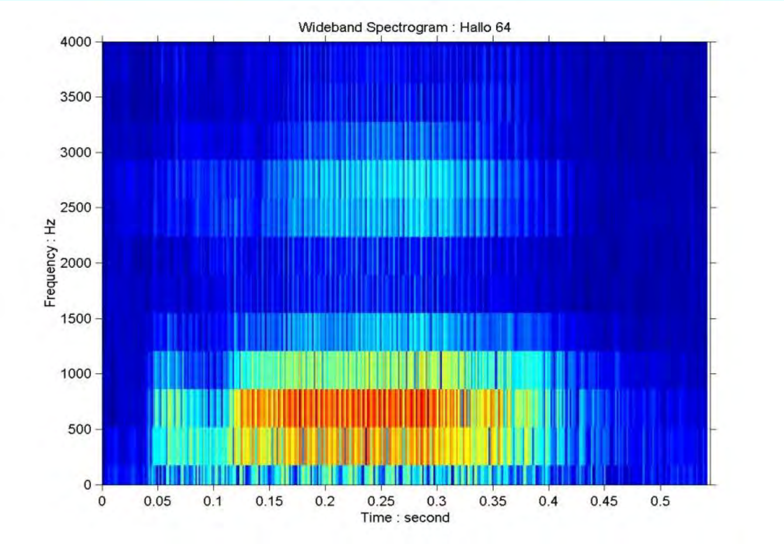
Speech Analysis

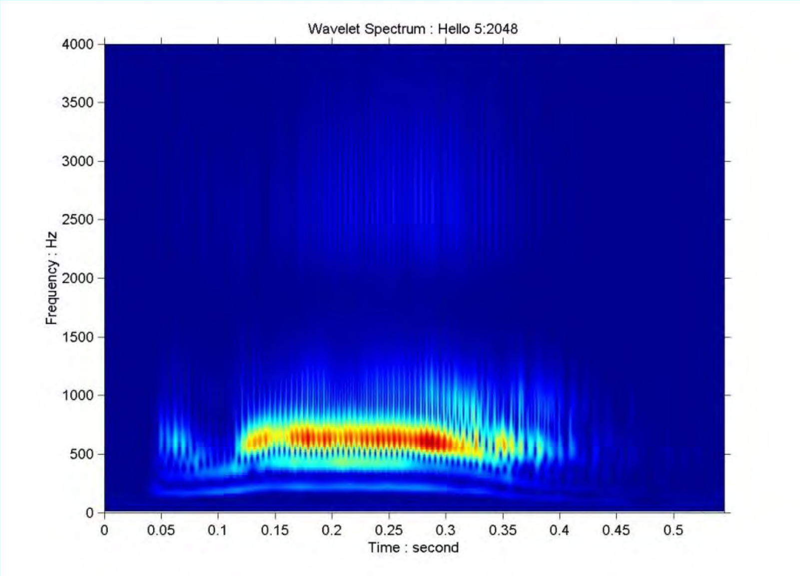
Hello : Data

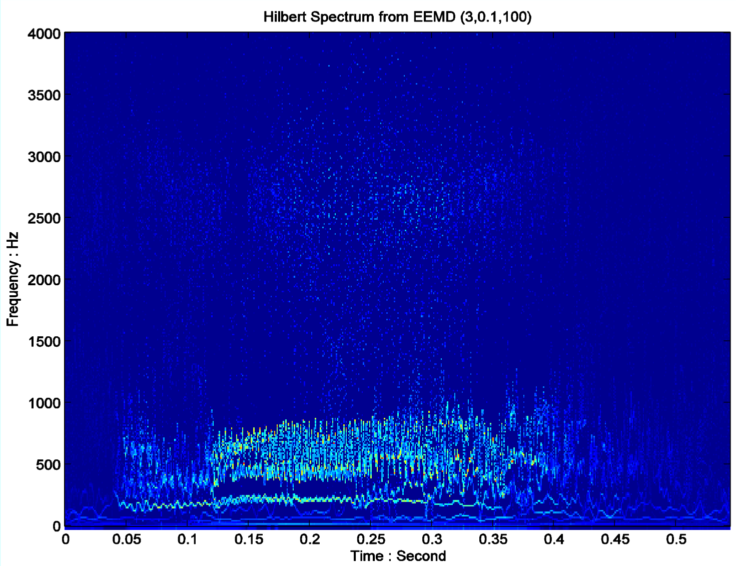




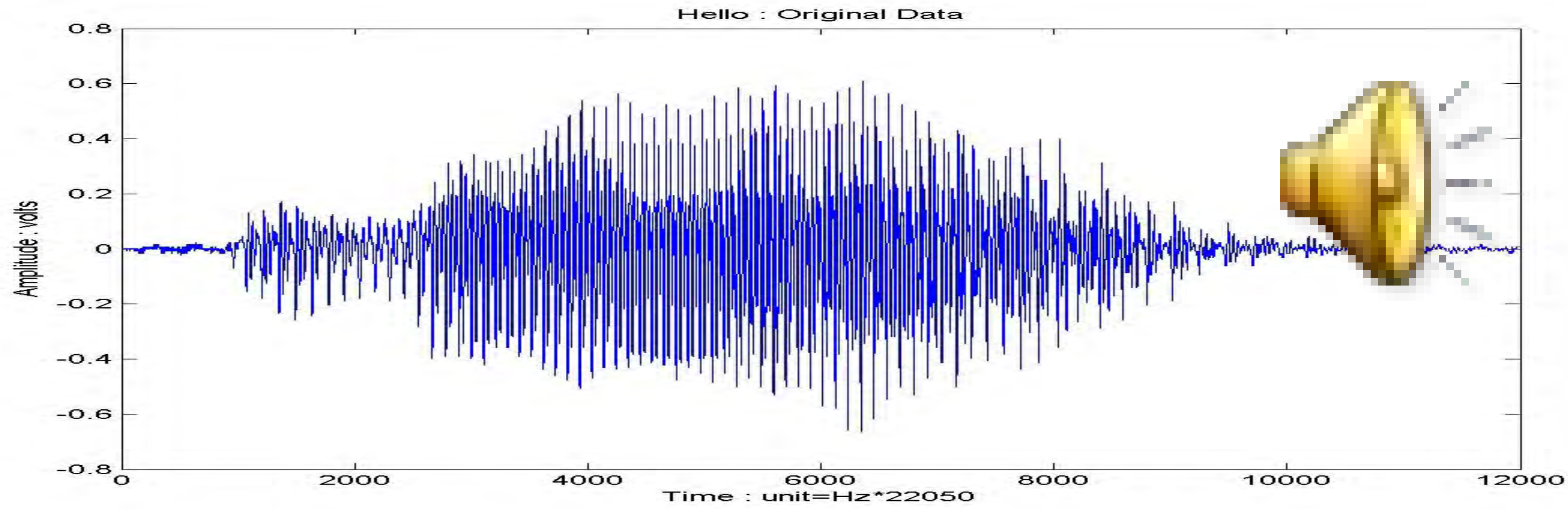
Four comparsions D

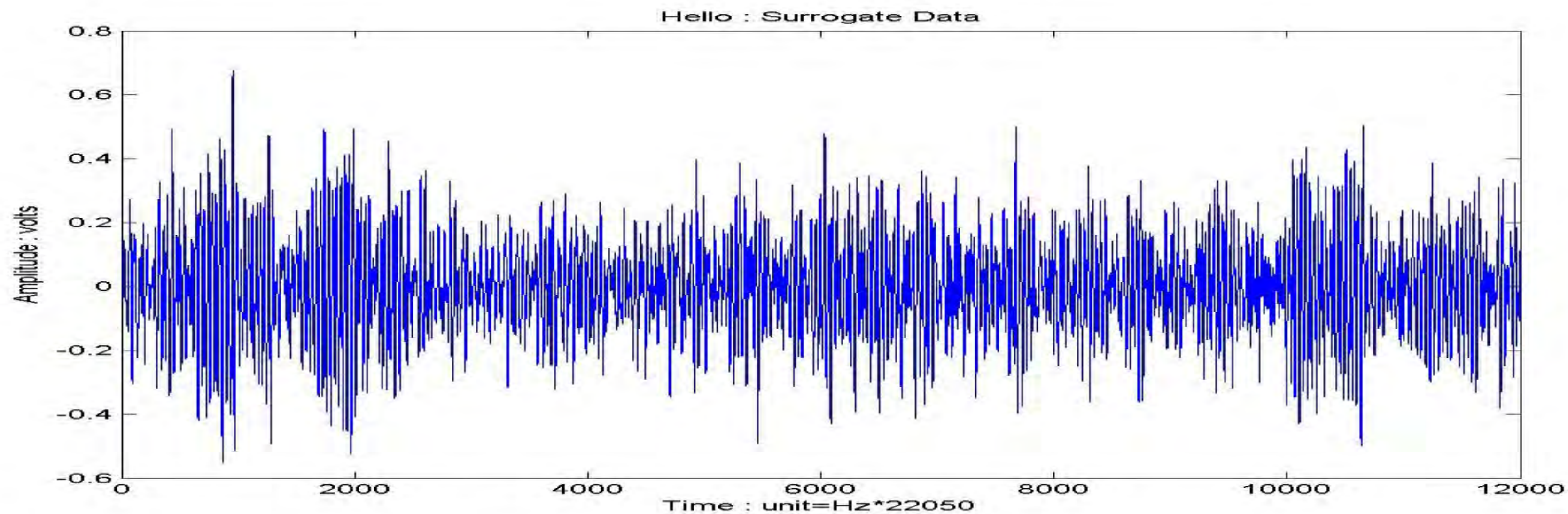


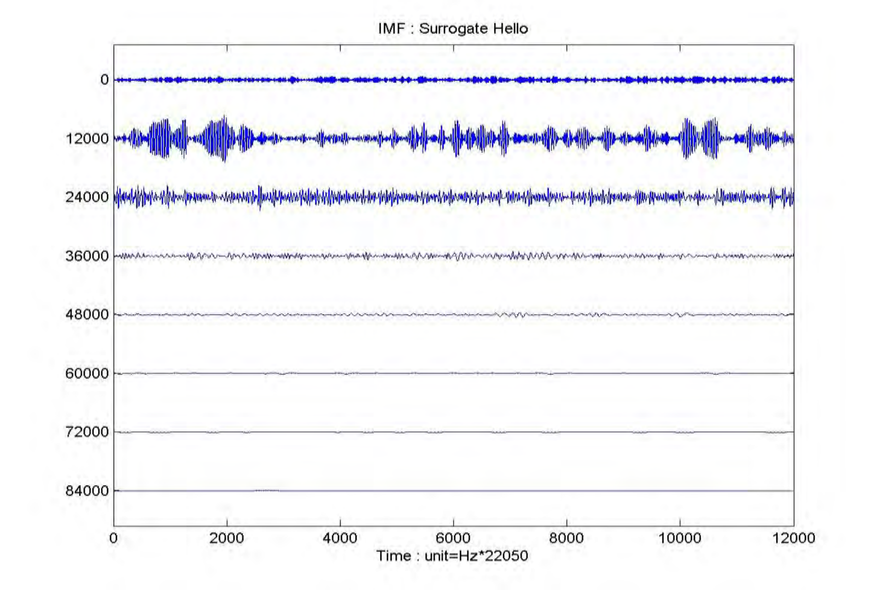




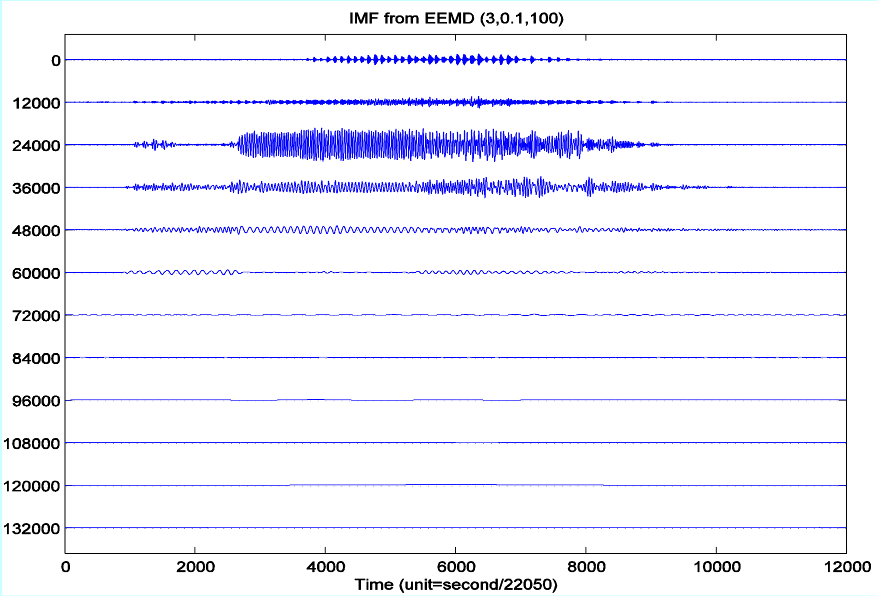
The original data : Hello



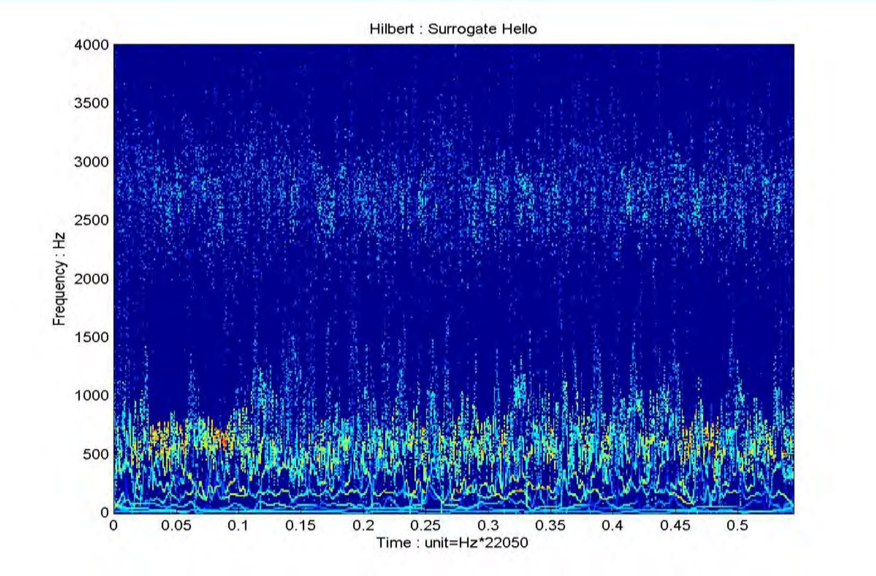




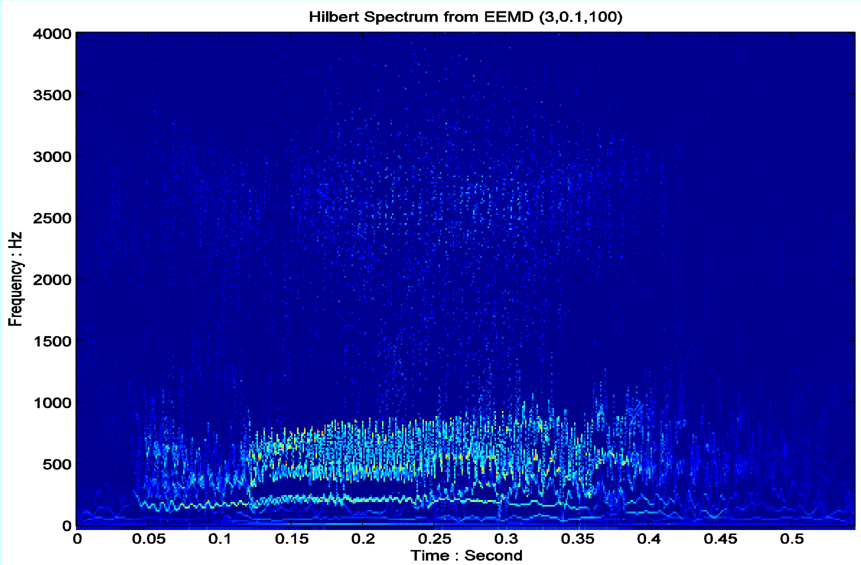
IMF : Surrogate Hello
0
12000
24000
36000
48000
60000
72000
84000
0 2000 4000 6000 8000 10000 12000
Time : unit=Hz*22050



IMF from EEMD (3,0.1,100)
0
12000
24000
36000
48000
60000
72000
84000
96000
108000
120000
132000
0 2000 4000 6000 8000 10000 12000
Time (unit=second/22050)



Hilbert : Surrogate Hello
Frequency : Hz
4000 3500 3000 2500 2000 1500 1000 500 0
0 0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5
Time : unit=Hz*22050



Hilbert Spectrum from EEMD (3,0.1,100)
Frequency : Hz
4000 3500 3000 2500 2000 1500 1000 500 0
0 0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5
Time : Second

Additionally

To quantify nonlinearity we also need instantaneous frequency.

The term, ‘Nonlinearity,’ has been loosely used, most of the time, simply as a fig leaf to cover our ignorance.

Can we be more precise?

How is nonlinearity defined?

Based on Linear Algebra: nonlinearity is defined based on input vs. output.

But in reality, such an approach is not practical: natural system are not clearly defined; inputs and out puts are hard to ascertain and quantify. Furthermore, without the governing equations, the order of nonlinearity is not known.

In the autonomous systems the results could depend on initial conditions rather than the magnitude of the 'inputs.'

The small parameter criteria could be misleading: sometimes, the smaller the parameter, the more nonlinear.

Linear Systems

Linear systems satisfy the properties of superposition and scaling. Given two valid inputs to a system ***H***,

$$x_1(t) \text{ and } x_2(t)$$

as well as their respective outputs

$$y_1(t) = H\{x_1(t)\} \text{ and}$$
$$y_2(t) = H\{x_2(t)\}$$

then a linear system, ***H***, must satisfy

$$\alpha y_1(t) + \beta y_1(t) = H\{\alpha x_1(t) + \beta x_2(t)\}$$

for any scalar values **α** and **β**.

How should nonlinearity be defined?

The alternative is to define nonlinearity based on data characteristics: Intra-wave frequency modulation.

Intra-wave frequency modulation is known as the harmonic distortion of the wave forms. But it could be better measured through the deviation of the instantaneous frequency from the mean frequency (based on the zero crossing period).

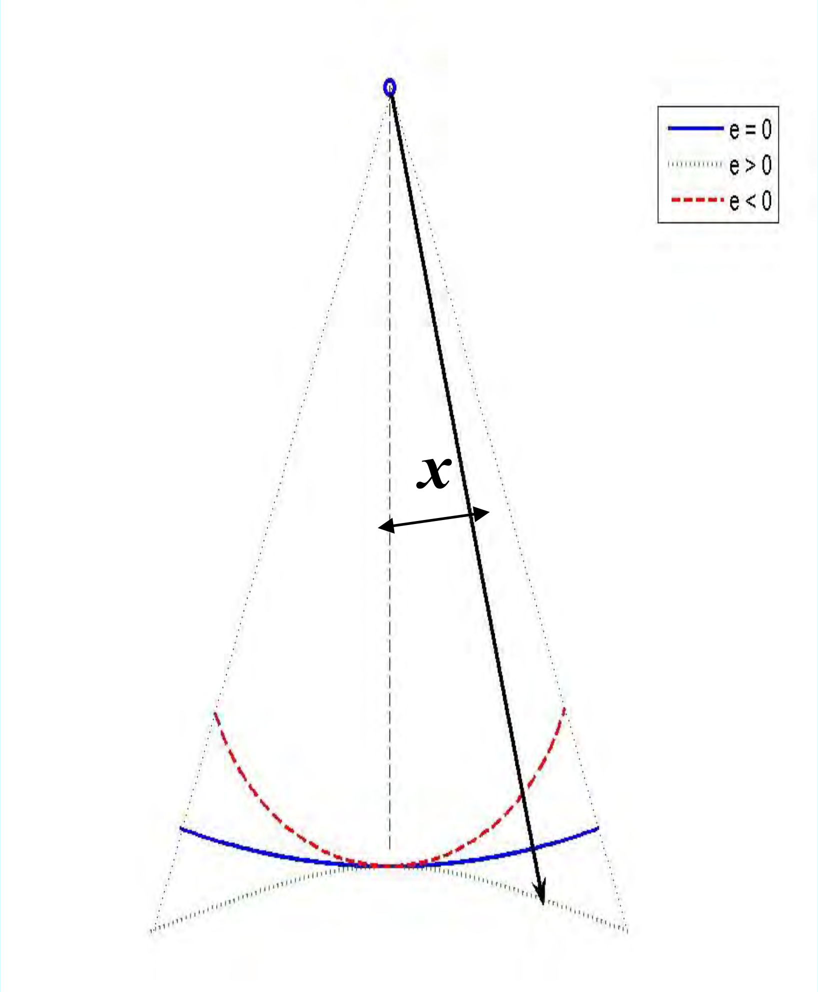
Characteristics of Data from Nonlinear Processes

$$\frac{d^2x}{dt^2} + x + \varepsilon x^3 = \gamma \cos \omega t$$

$$\Rightarrow \quad \frac{d^2x}{dt^2} + x\left(1 + \varepsilon x^2\right) = \gamma \cos \omega t$$

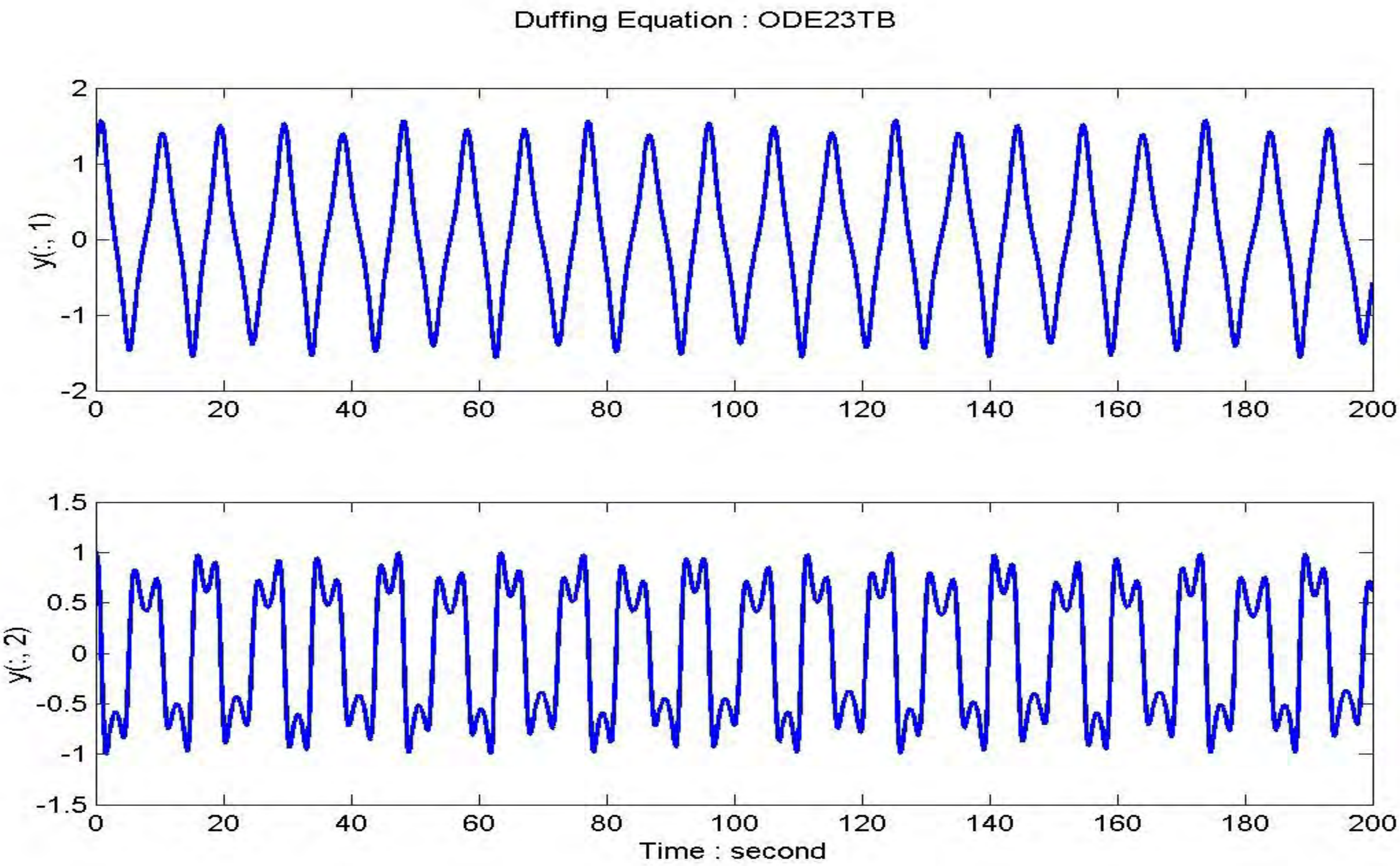
$\Rightarrow$ *Spring with position dependent cons tan t ,*
int ra – wave frequency mod ulation ;
therefore , we need ins tan tan eous frequency.

Duffing Pendulum



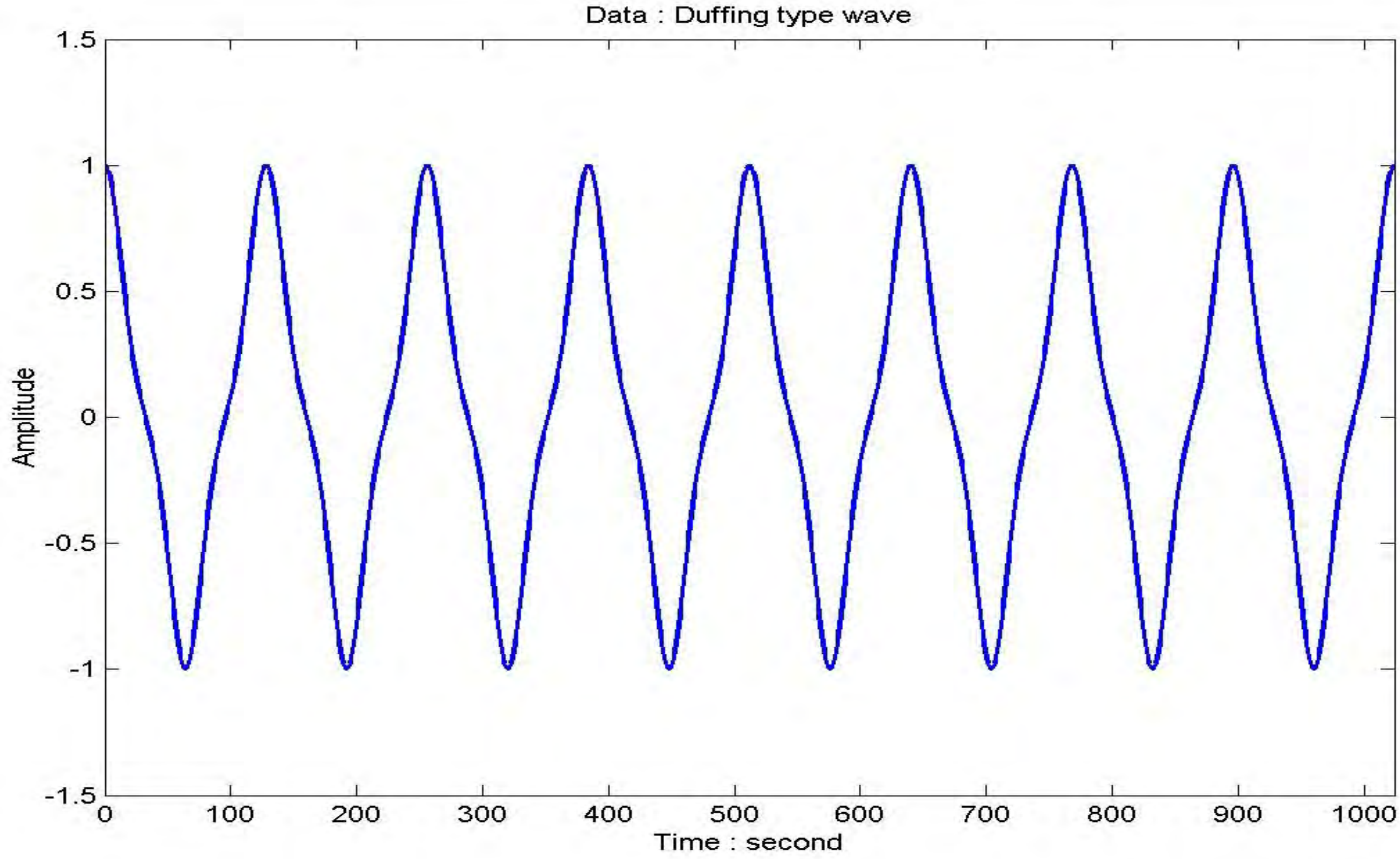
$$\frac{d^2x}{dt^2} + x(1 + \varepsilon x^2) = \gamma \cos \omega t \ .$$

Duffing Equation : Data



Duffing Type Wave

Data: $x = cos(wt+0.3\ sin2wt)$



Duffing Type Wave

Perturbation Expansion

For $\varepsilon \ll 1$, *we can have*

$$
\begin{aligned}
x(t) &= \cos\left(\omega t + \varepsilon \sin 2\omega t\right) \\
&= \cos\omega t \cos\left(\varepsilon \sin 2\omega t\right) - \sin\omega t \sin\left(\varepsilon \sin 2\omega t\right) \\
&= \cos\omega t - \varepsilon \sin\omega t \sin 2\omega t + ... \\
&= \left(1 - \frac{\varepsilon}{2}\right)\cos\omega t + \frac{\varepsilon}{2}\cos 3\omega t + ...
\end{aligned}
$$

This is very similar to the solution of Duffing equation .

Degree of nonlinearity

Let us consider a generalized intra-wave frequency modulation model as:

$$x(t) = \cos(\omega t + \delta \sin \eta \omega t) \quad \Rightarrow \quad IF = \frac{d\theta}{dt} = \omega\left(1 + \eta\delta \cos \eta \omega t\right)$$

Depending on the value of η, we can have either a up-down symmetric or a asymmetric wave form.

Degree of Nonlinearity

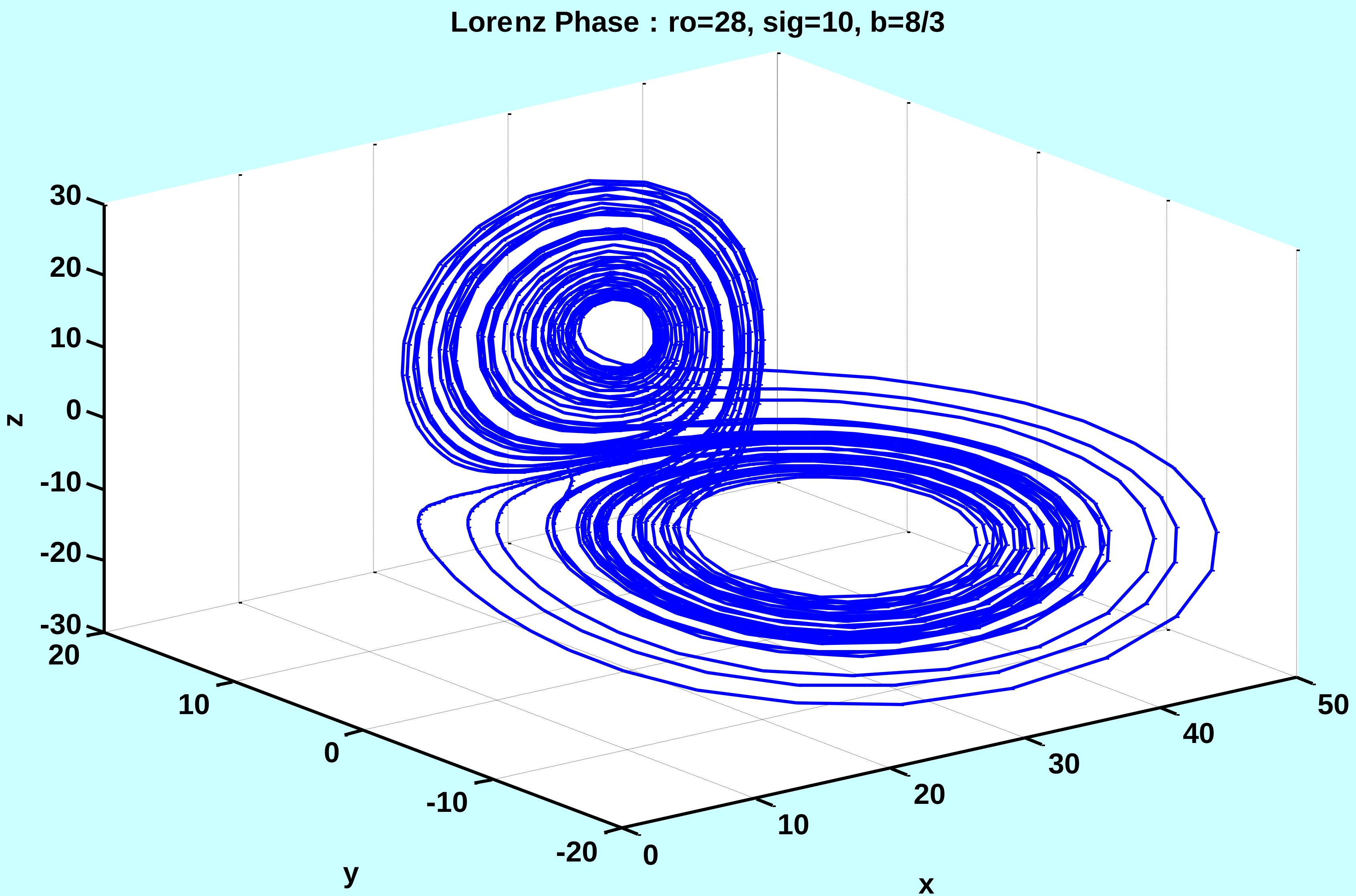
- DN is determined by the combination of $\delta\eta$ precisely with Hilbert Spectral Analysis. Either of them equals zero means linearity.
- We can determine δ *and* η separately:
  - η can be determined from the instantaneous frequency modulations relative to the mean frequency.
  - δ can be determined from DN with known η.

  NB: from any IMF, the value of $\delta\eta$ cannot be greater than 1.
- ***The combination of δ and η gives us not only the Degree of Nonlinearity, but also some indications of the basic properties of the controlling Differential Equation, the Order of Nonlinearity.***

Lorenz Model

- Lorenz is highly nonlinear; it is the model equation that initiated chaotic studies.
- Again it has three parameters. We decided to fix two and varying only one.
- There is no small perturbation parameter.
- We will present the results for $\rho=28$, the classic chaotic case.

Phase Diagram for ro=28

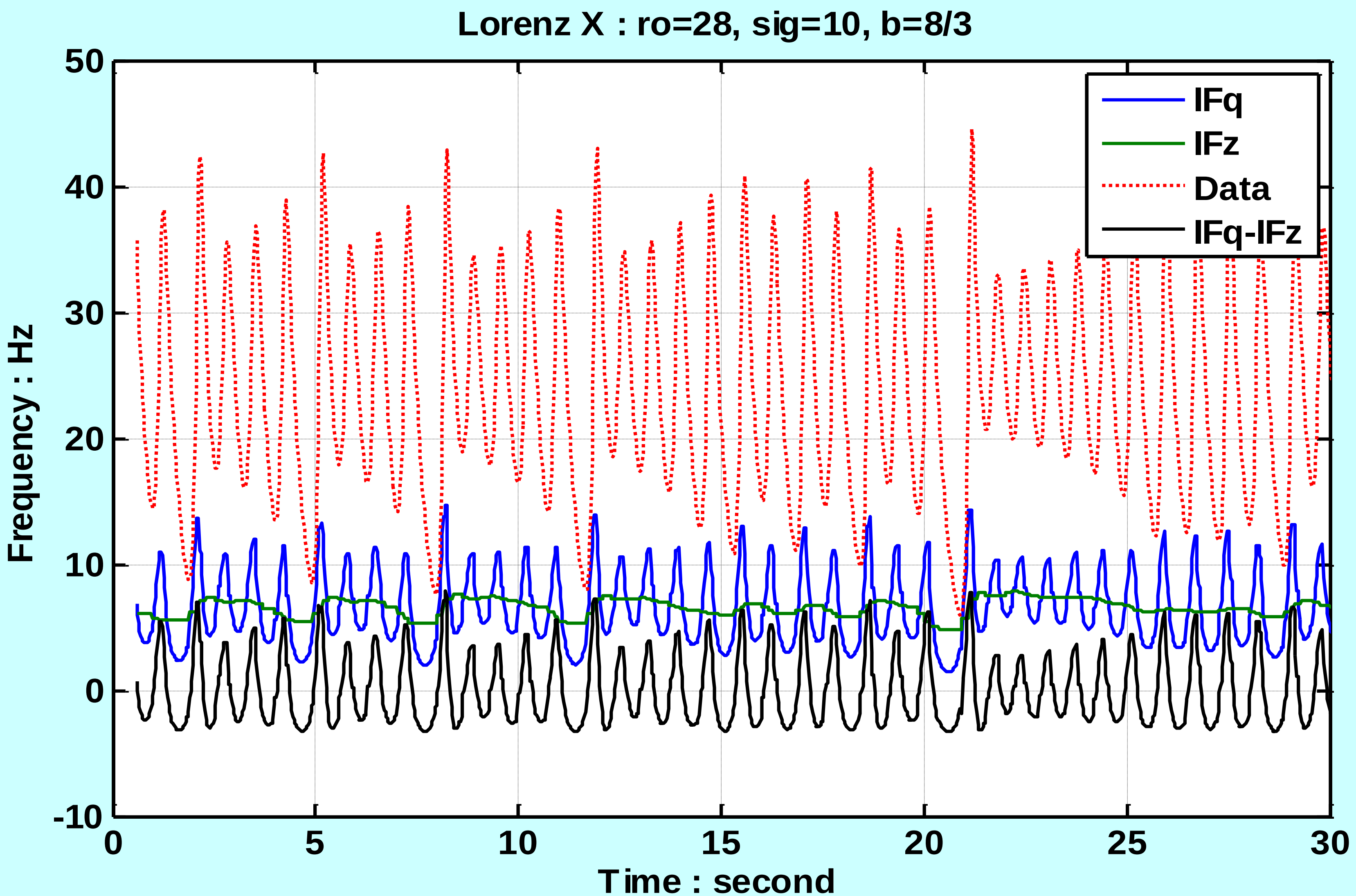


X-Component

DN1=0.5147

CDN=0.5027

Data and IF



Lorenz X : ro=28, sig=10, b=8/3
Frequency : Hz
Time : second
IFq
IFz
Data
IFq-IFz
50
40
30
20
10
0
-10
0
5
10
15
20
25
30

Lorenz Model

$$\frac{dx}{dt} = \sigma\left(y - x\right)$$

$$\frac{dy}{dt} = x\left(\rho - z\right) - y$$

$$\frac{dz}{dt} = xy - \beta z$$

with σ (Prandtl number)=10; β=8/3;

ρ (Rayleigh number) varying

How to define Trend ?

Parametric or Non-parametric?

Intrinsic vs. extrinsic approach?

The State-of-the arts: Trend

"One economist's trend is another economist's cycle"

Watson : Engle, R. F. and Granger, C. W. J. 1991 *Long-run Economic Relationships*. Cambridge University Press.

Philosophical Problem

名不正則言不順
言不順則事不成

——孔夫子

On Definition

Without a proper definition,
logic discourse would be impossible.
Without logic discourse,
nothing can be accomplished.

Confucius

Definition of the Trend

Proceeding Royal Society of London, 1998
Proceedings of National Academy of Science, 2007

Within the given data span, the trend is an intrinsically fitted monotonic function, or a function in which there can be at most one extremum.

The trend should be an intrinsic and local property of the data; it is determined by the same mechanisms that generate the data.

Being local, it has to associate with a local length scale, and be valid only within that length span, and be part of a full wave length.

The method determining the trend should be intrinsic. Being intrinsic, the method for defining the trend has to be adaptive.

All traditional trend determination methods are extrinsic.

Algorithm for Trend

- Trend should be defined neither parametrically nor non-parametrically.
- It should be the residual obtained by removing cycles of all time scales from the data intrinsically.
- Through EMD.

EMD as Filters

EMD is a dyadic filter bank.

The filtering is operating in time domain.

Need a Filter to Remove Alias

- Traditional Fourier filter is inadequate:
  - Removal of Harmonics will distort the fundaments
  - Noise spikes are local in time; signals local in time have broad spectral band
- HHT is an adaptive filter working in time space rather than frequency space.

EMD as filters

Once we have the EMD expansion : $\quad x(t) = \sum_{j=1}^{N} c_j \ ,$

we can define the filters as fellows :

Low Pass Filter : $\quad x_L(t) = \sum_{j=L}^{N} c_j \ ;$

High Pass Filter : $\quad x_H(t) = \sum_{j=1}^{H} c_j \ ;$

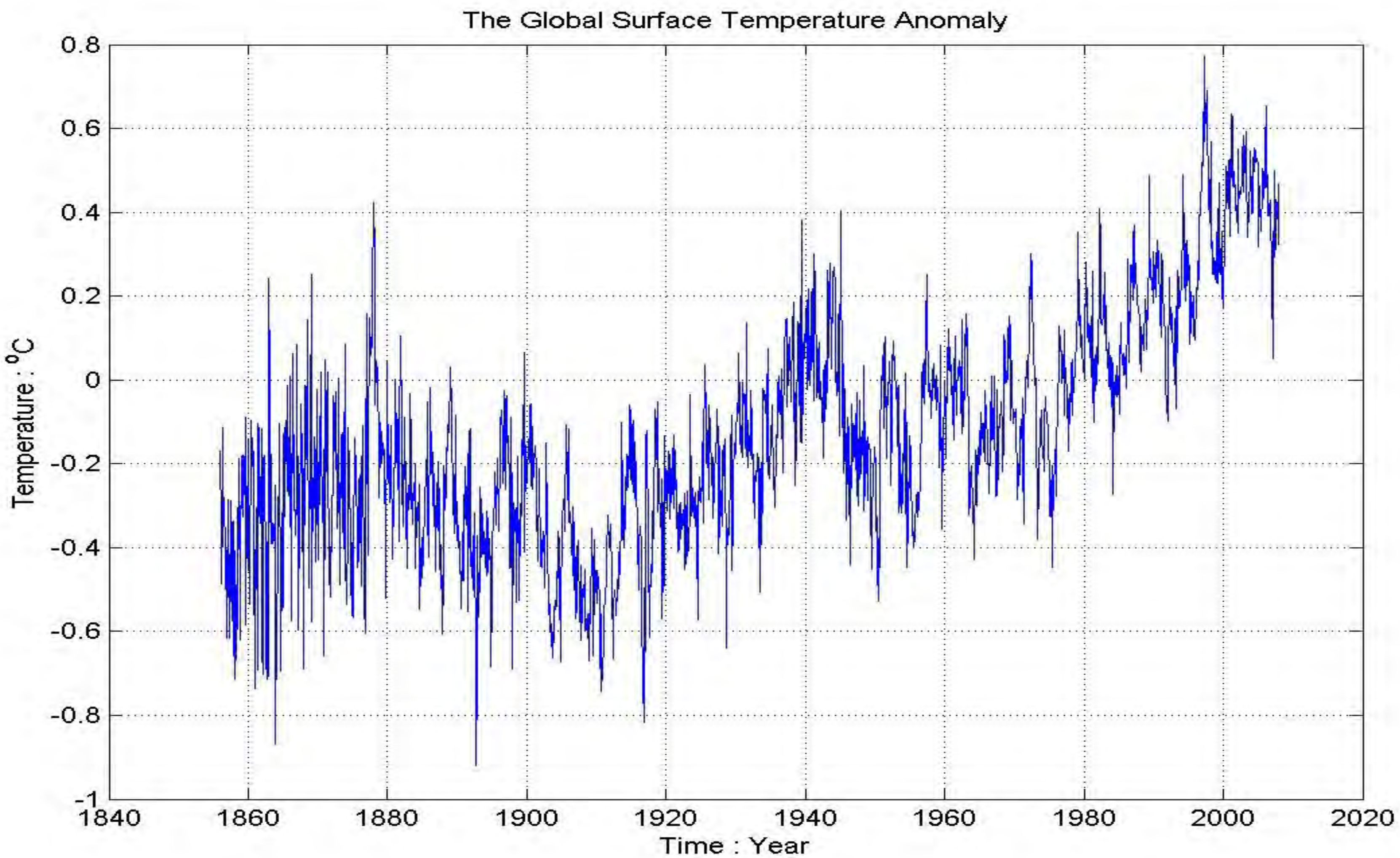
Band Pass Filter : $\quad x_B(t) = \sum_{j=B}^{M} c_j \ .$

How are GSTA annual data derived? Noise Reduction

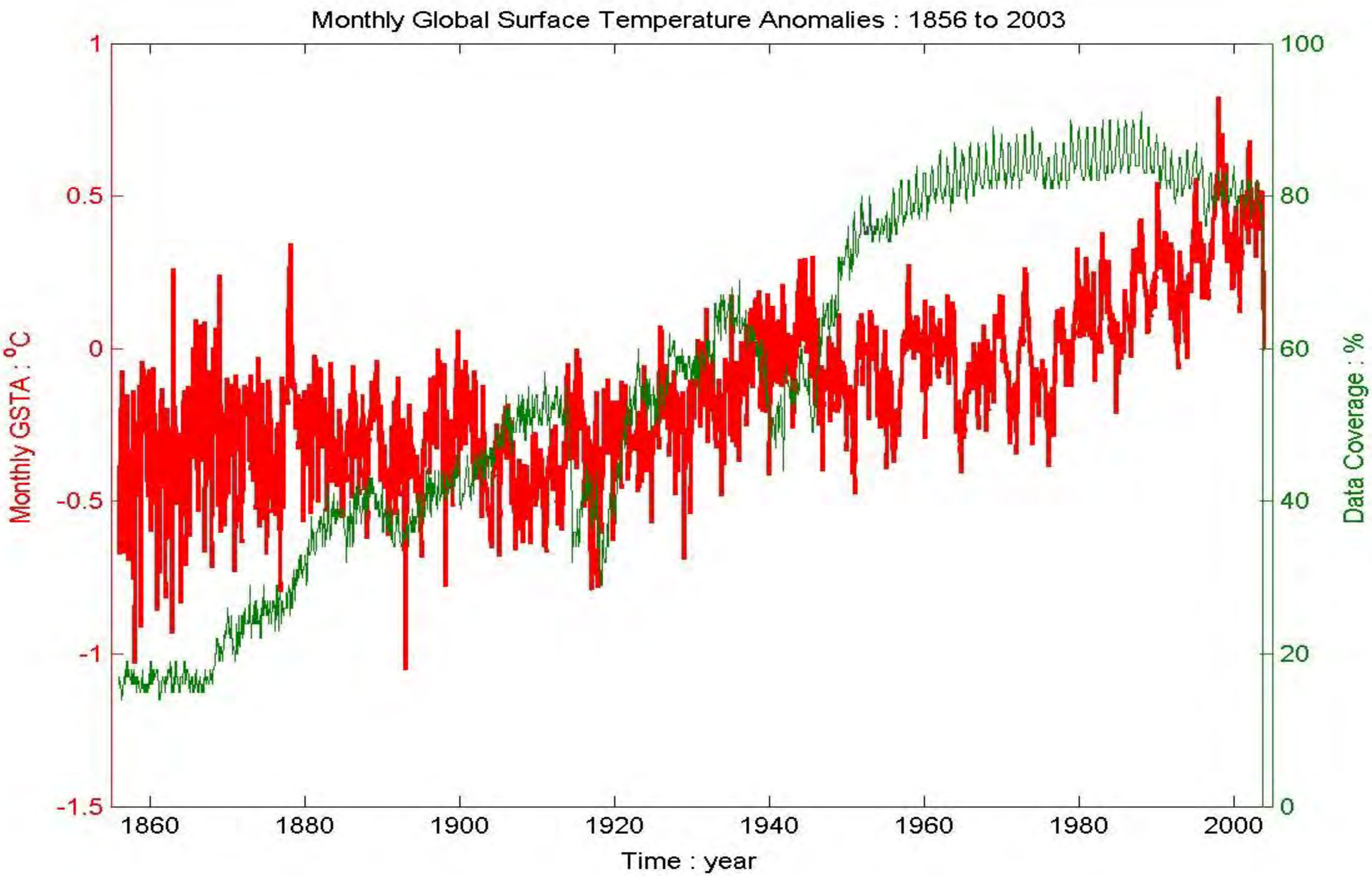
Using Global Surface Temperature Anomaly data

1856 to 2003

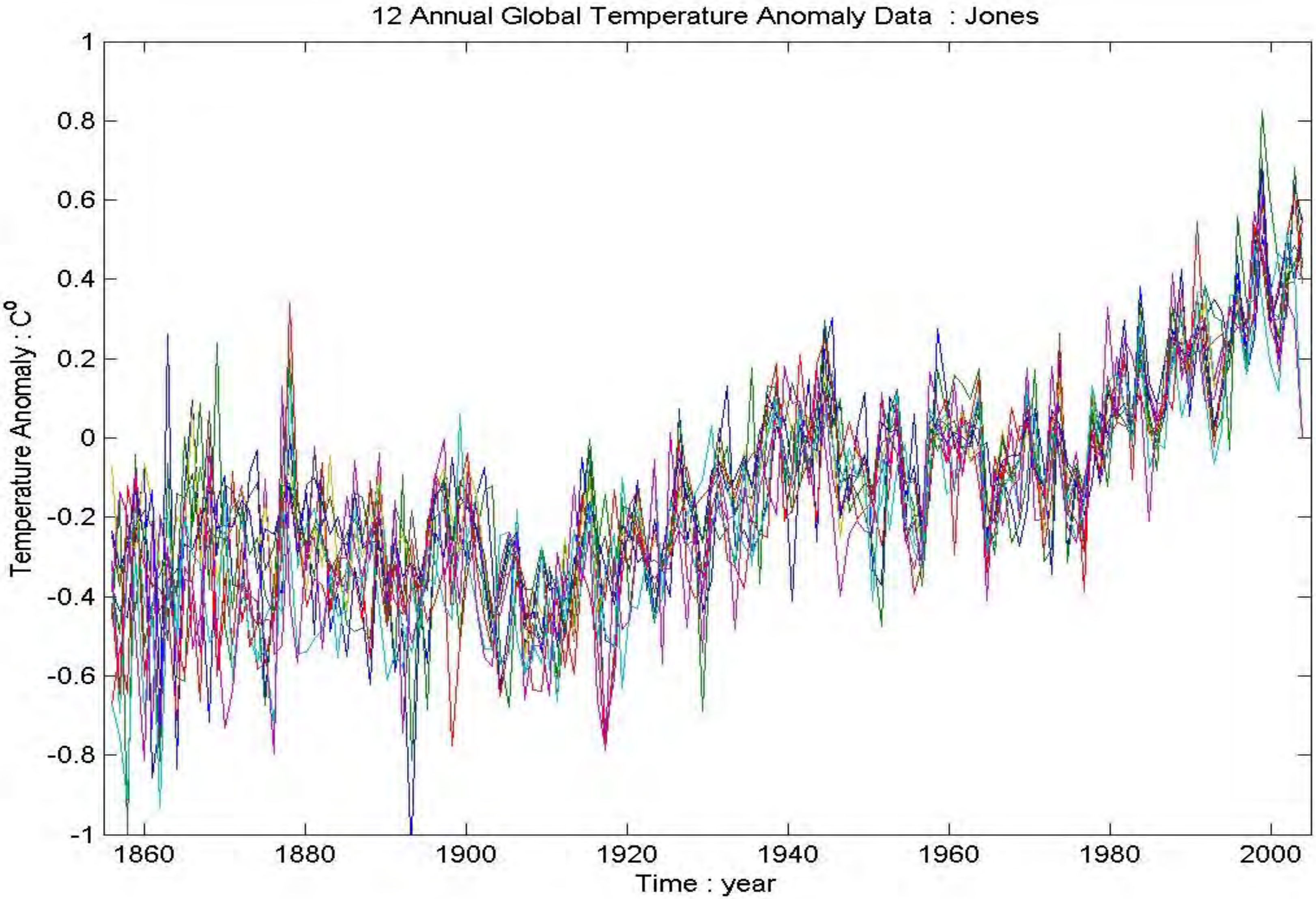
GSTA


The Global Surface Temperature Anomaly
Temperature : °C
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
1840
1860
1880
1900
1920
1940
1960
1980
2000
2020
Time : Year

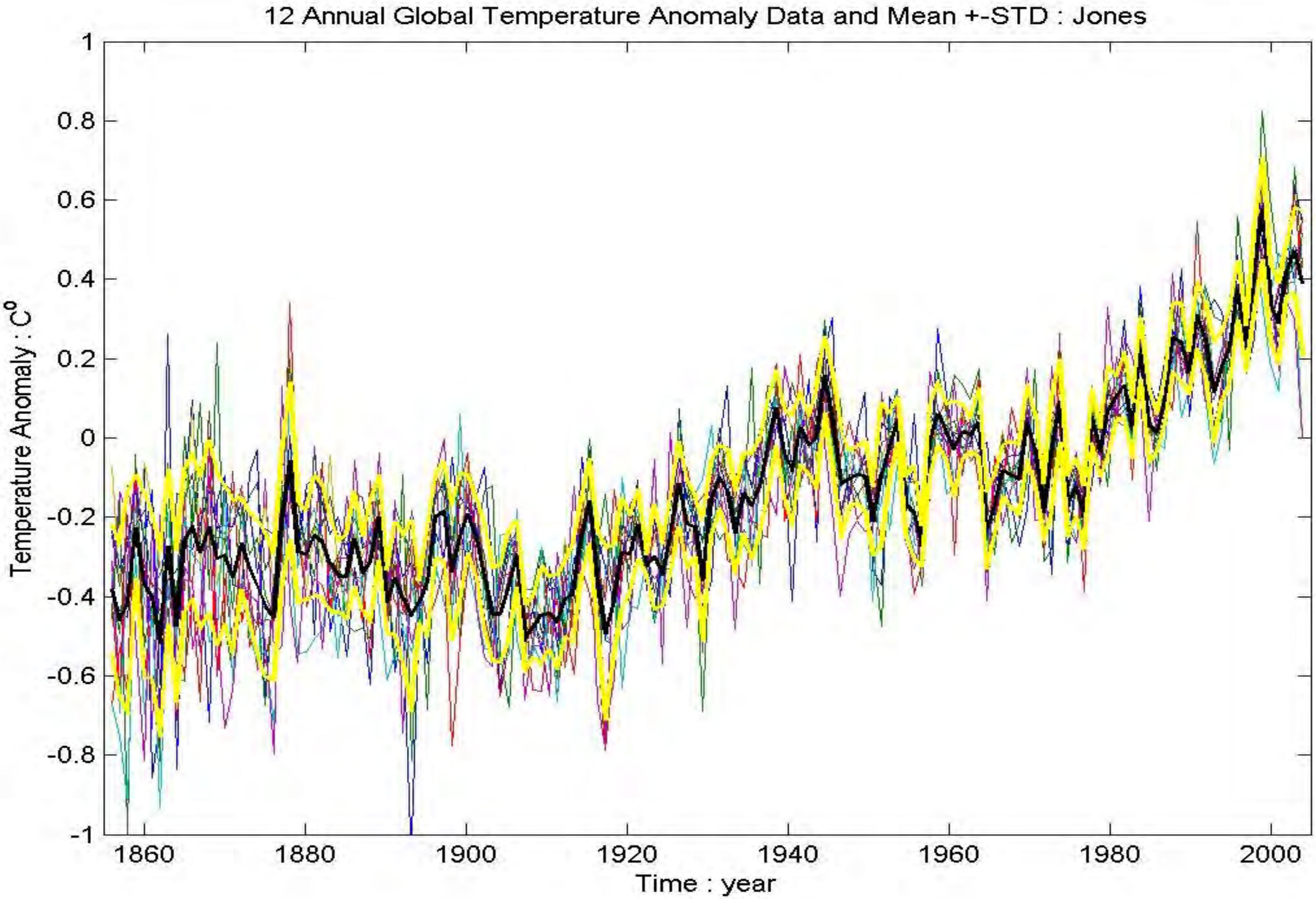
Jones (2003) Monthly GSTA Data



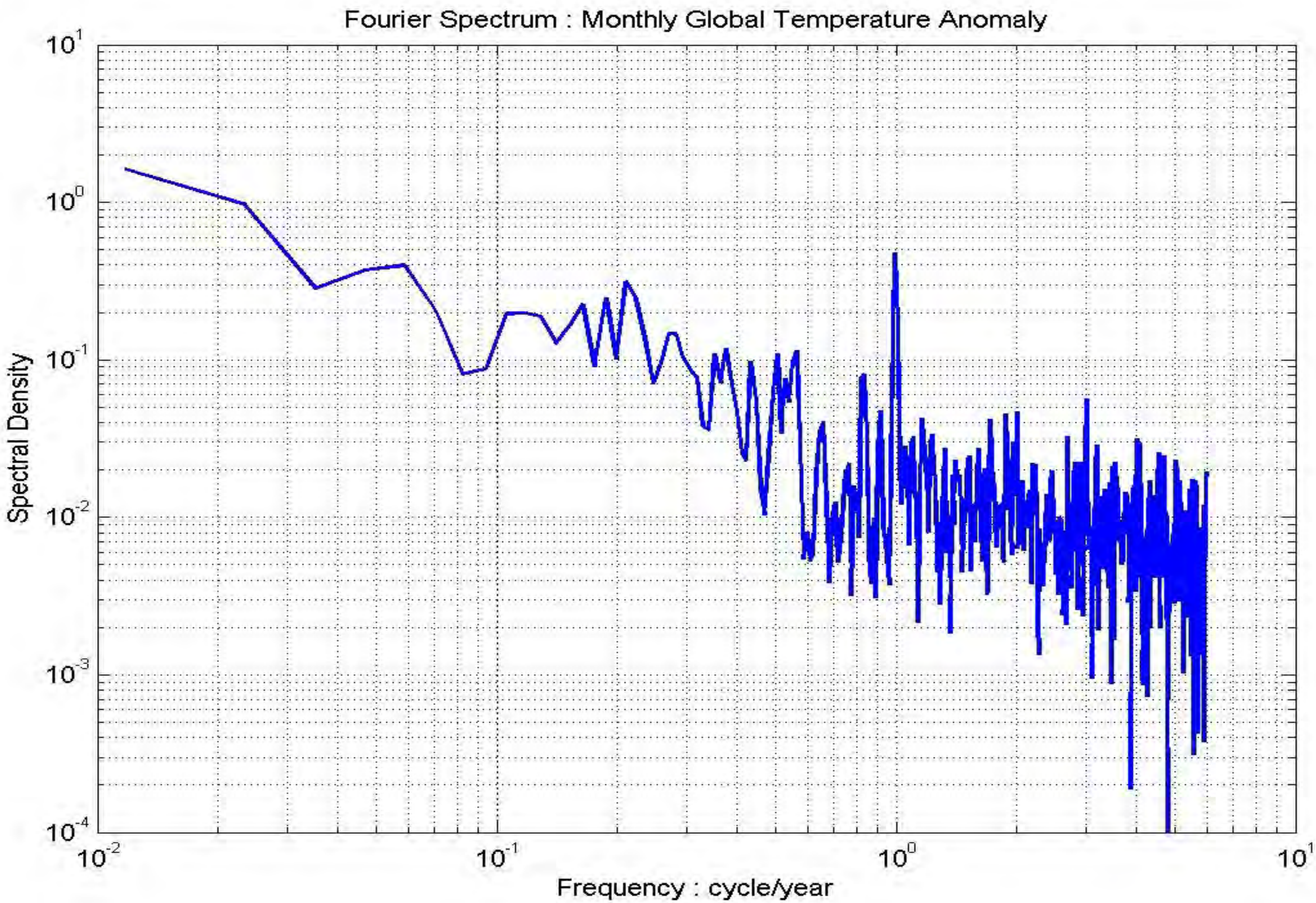
Jones (2003) 12 Monthly GSTA Data



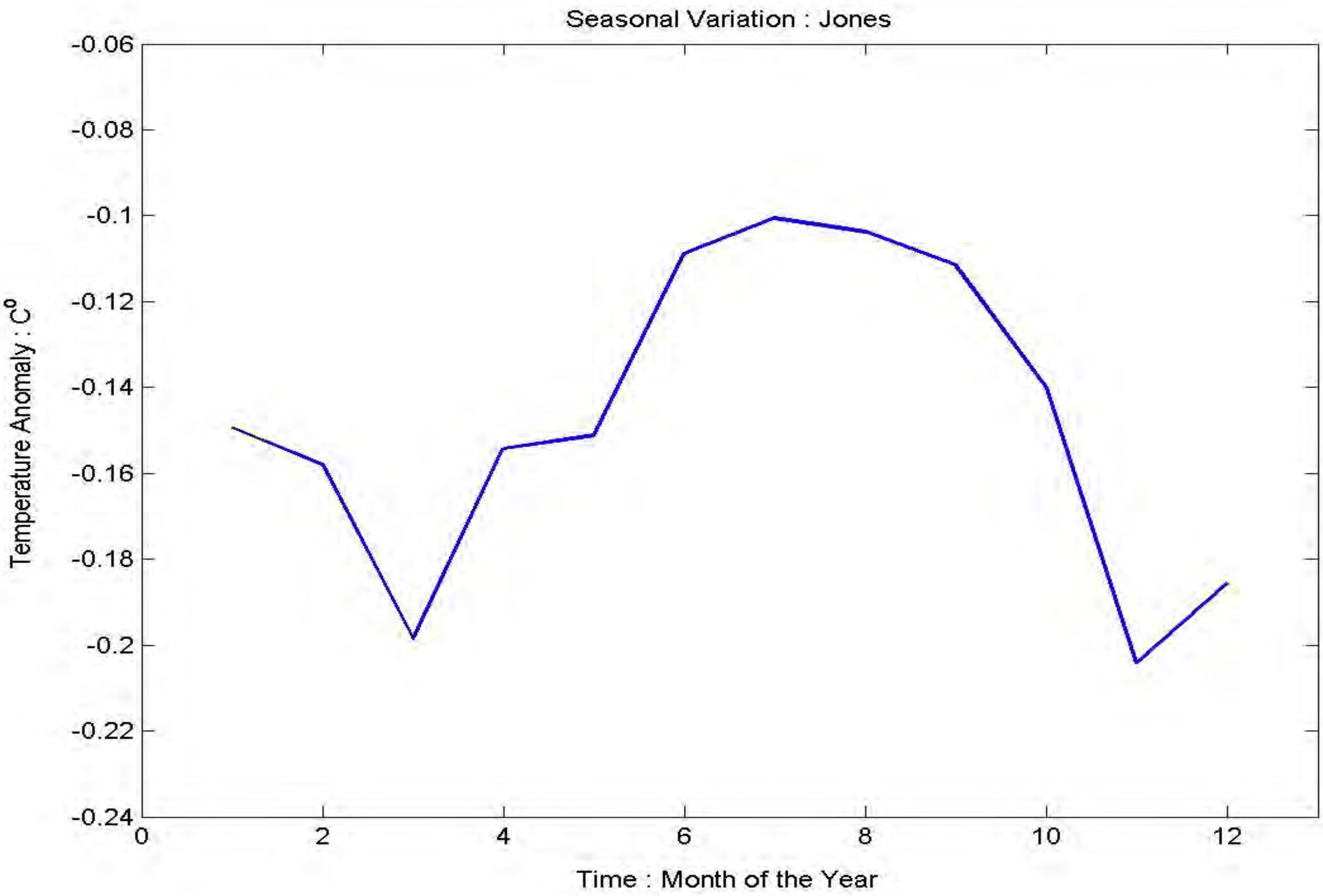
Jones (2003) 12 Monthly GSTA Data



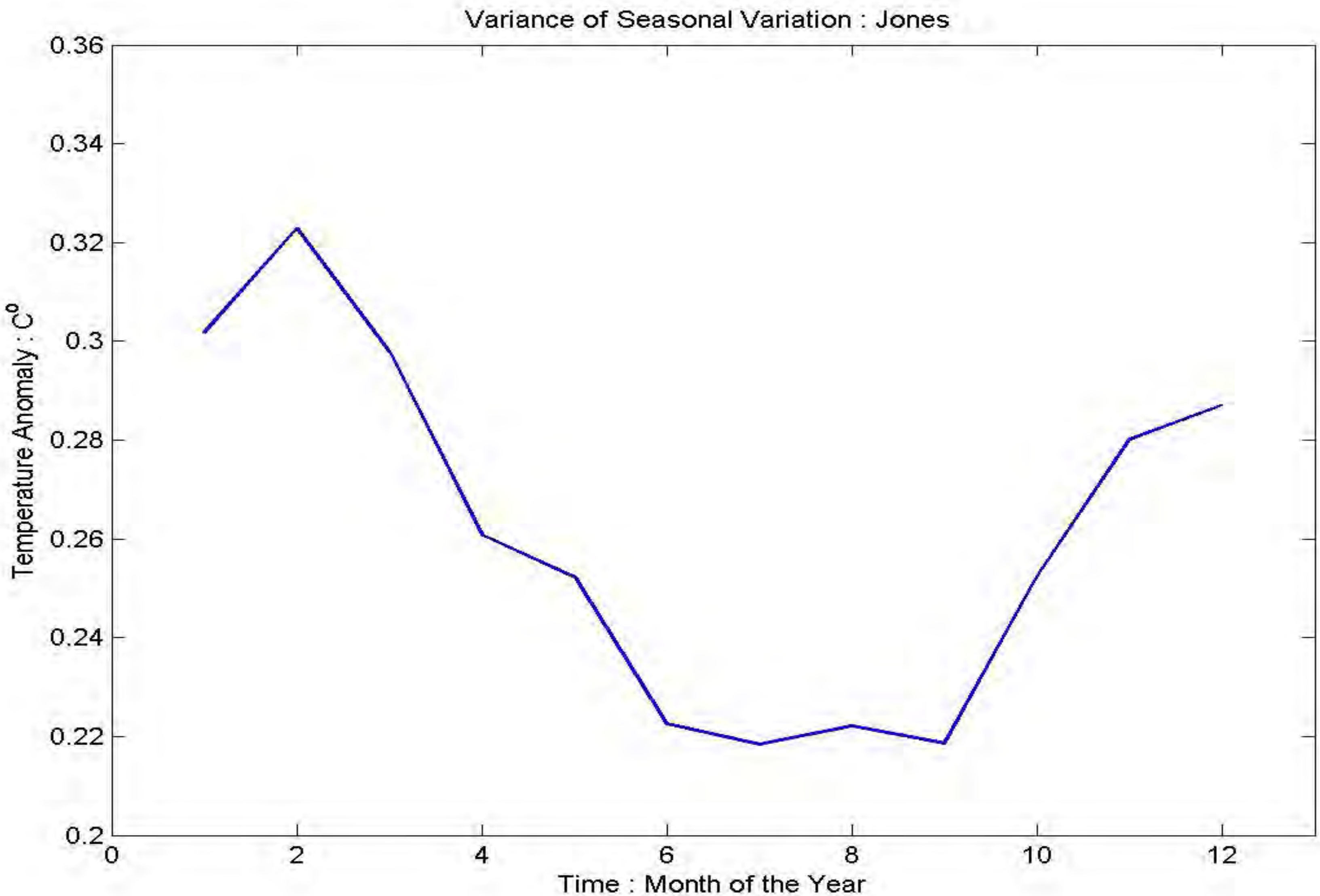
Jones Monthly GSTA Data : Fourier Spectrum



Jones (2003) GSTA Data Seasonal Variation



Jones (2003) GSTA Data Seasonal Variance



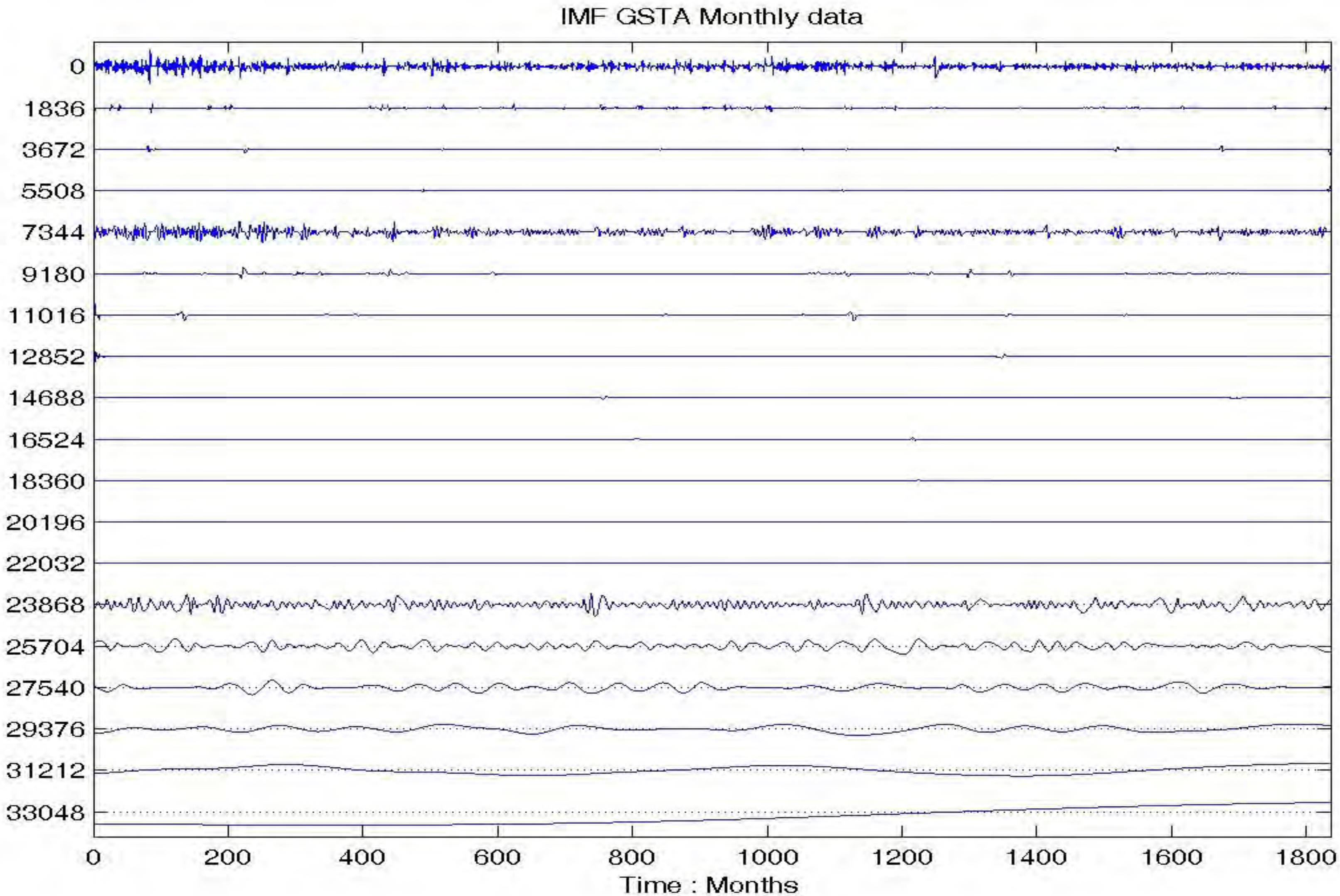
Observations

- Annual data is actually the mean of 12:1 down sample set of the original monthly data.
- In spite of the removal of climatologic mean, there still is a seasonal peak (1 cycle / year).
- Seasonal Variation and Variance are somewhat irregular.
- Data contain no information beyond yearly frequency, for higher frequency part of the Fourier spectrum is essentially flat.
- Decide to filtered the Data with HHT before down sample.

GSTA annual mean from Original and HHT Filtered

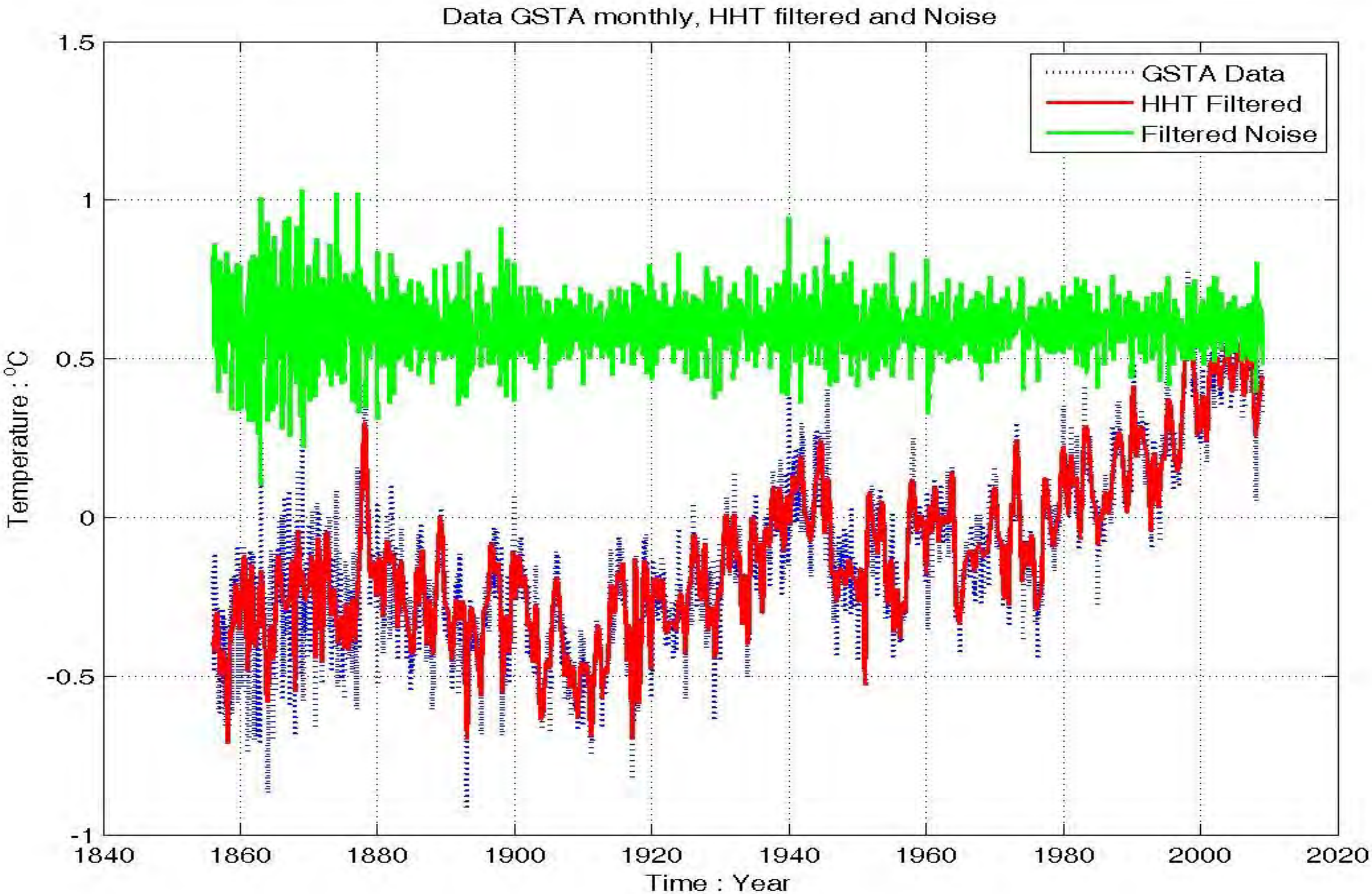
HHT with intermittence test

GSTA IMF

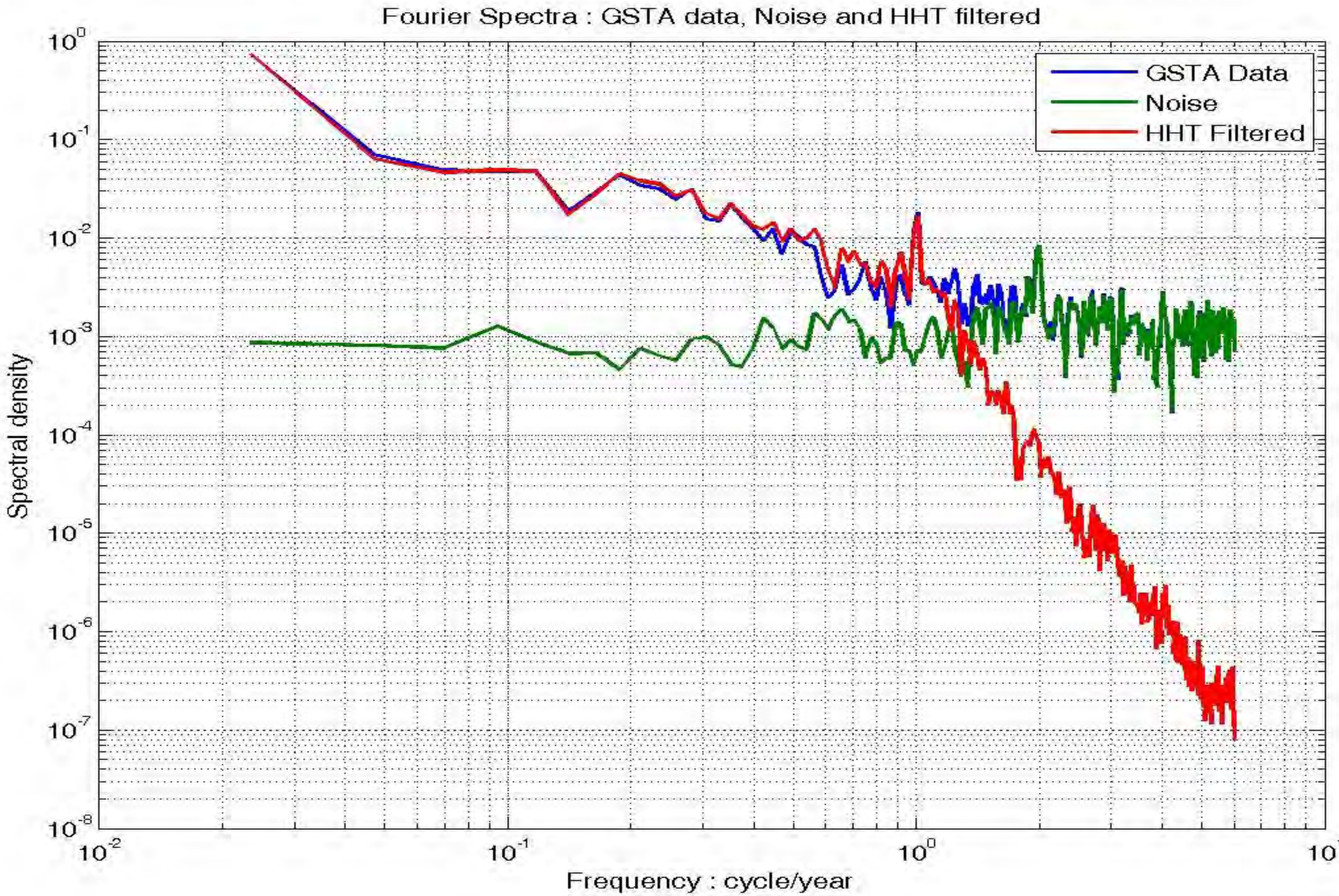


IMF GSTA Monthly data
0
1836
3672
5508
7344
9180
11016
12852
14688
16524
18360
20196
22032
23868
25704
27540
29376
31212
33048
0
200
400
600
800
1000
1200
1400
1600
1800
Time : Months

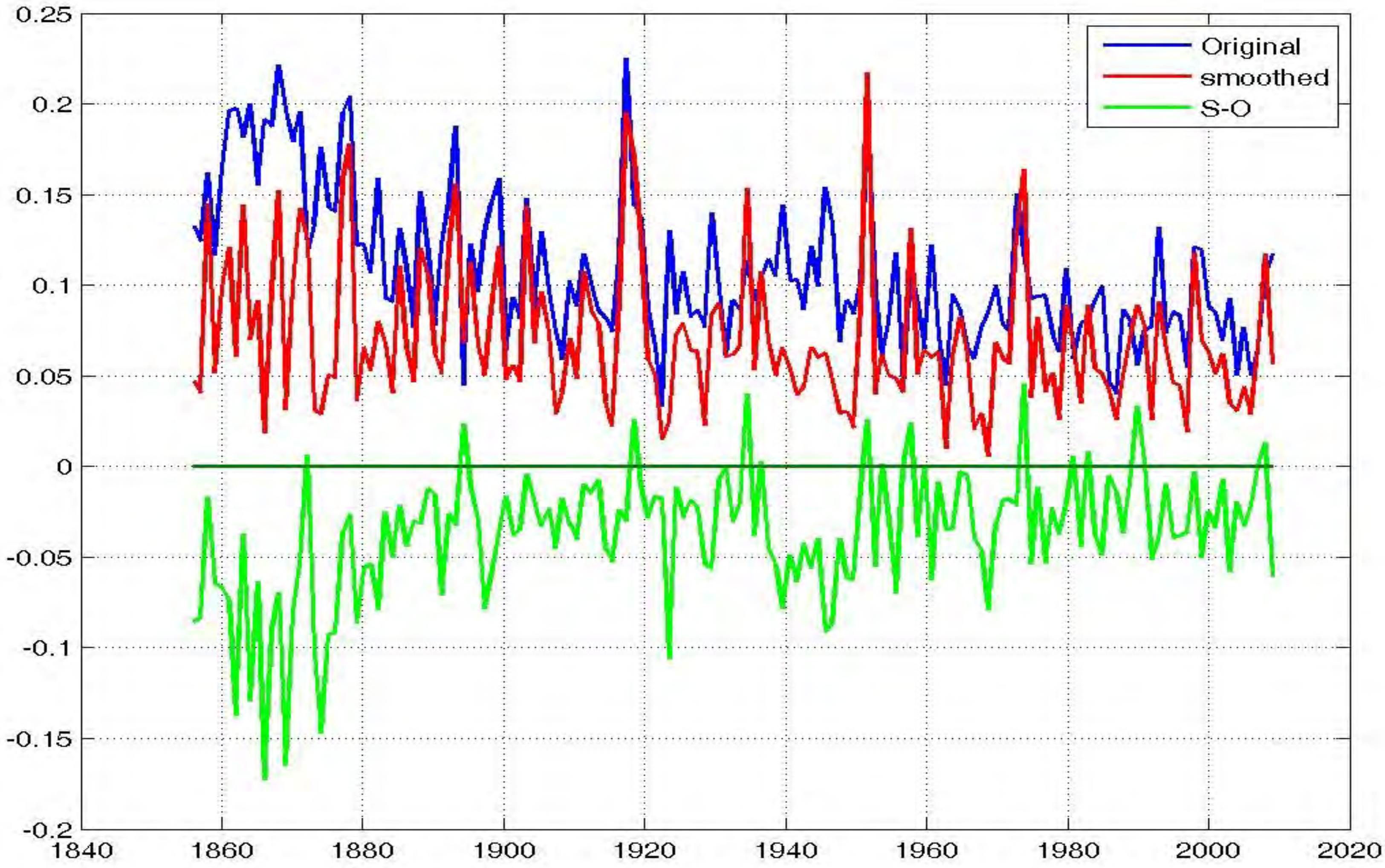
GSTA : Filtered and original data



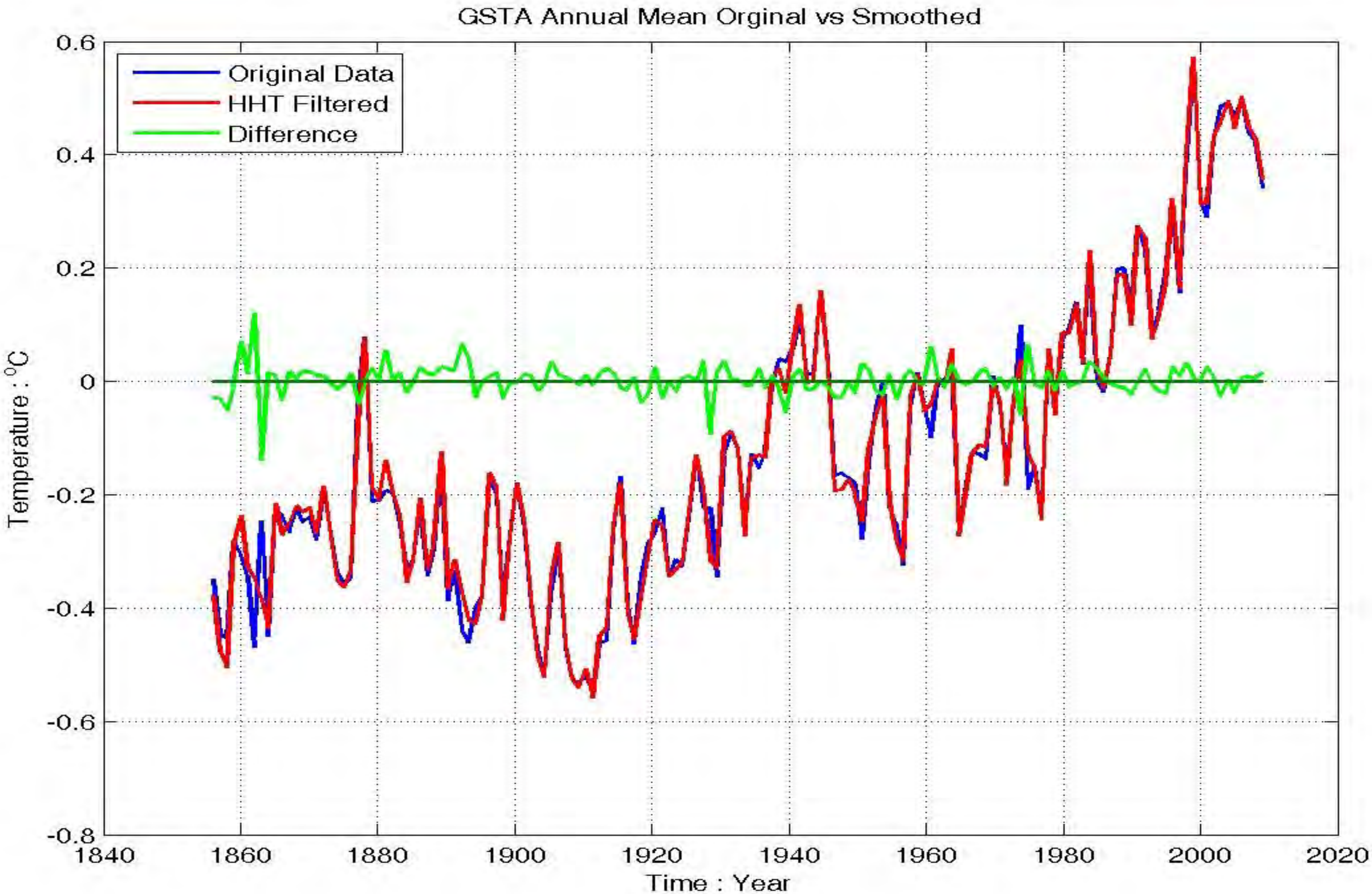
GSTA : Spectra Filtered and original data



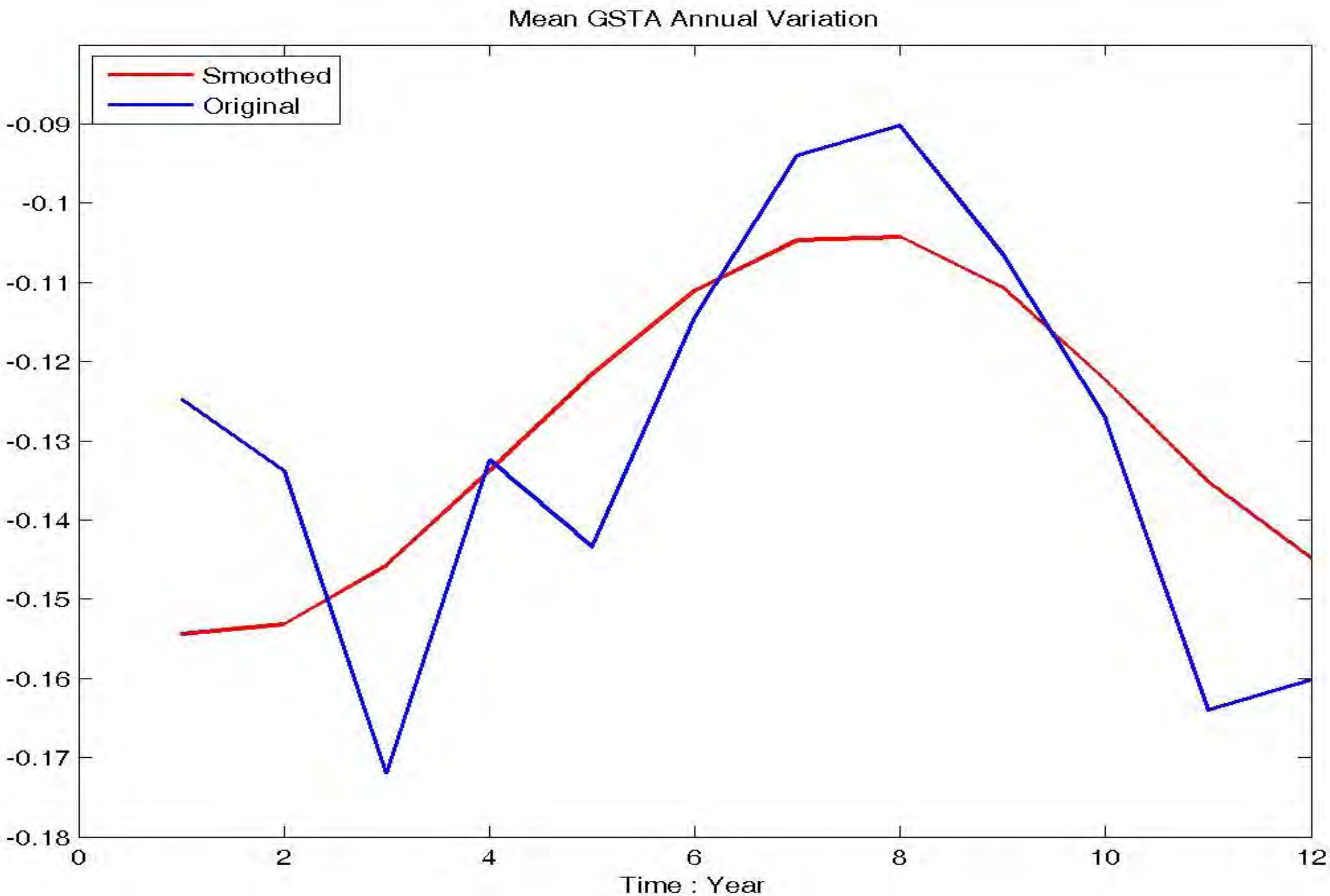
GSTA : STD Filtered and original data



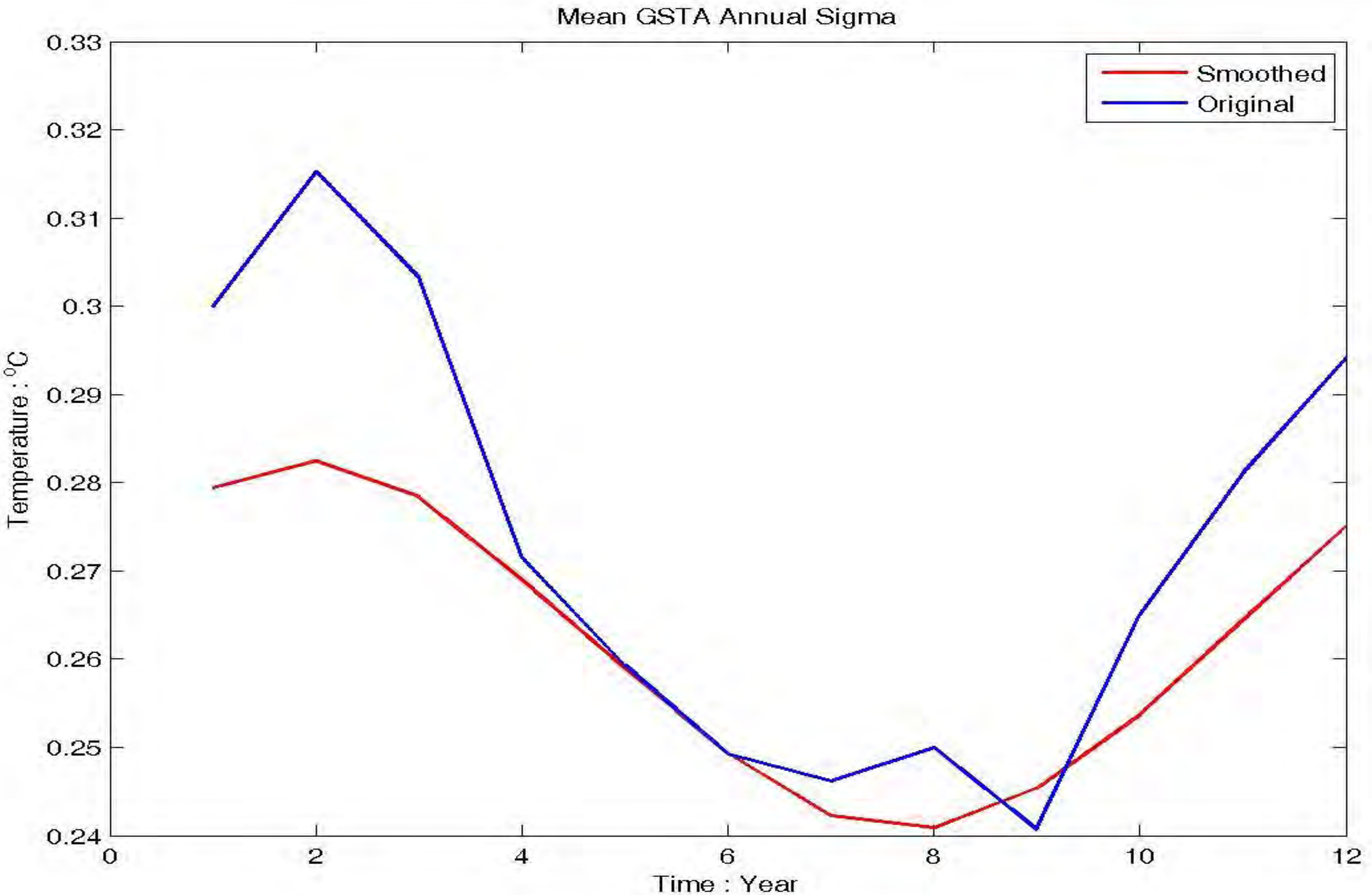
GSTA: Annual mean filtered and original data



GSTA: Variations for Annual mean filtered and original data



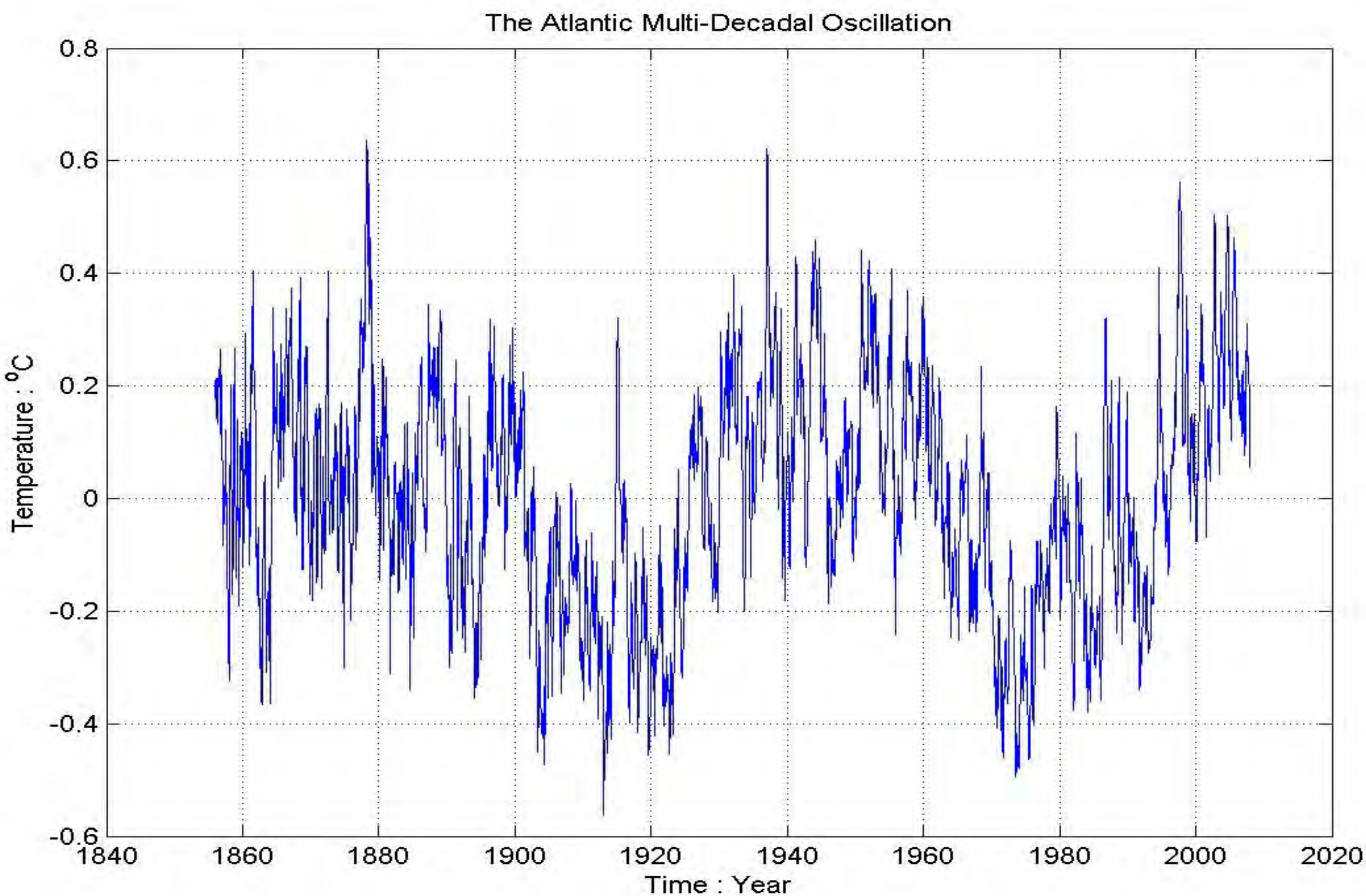
GSTA: Sigma for Annual mean filtered and original data



AMO annual mean from Original and HHT Filtered

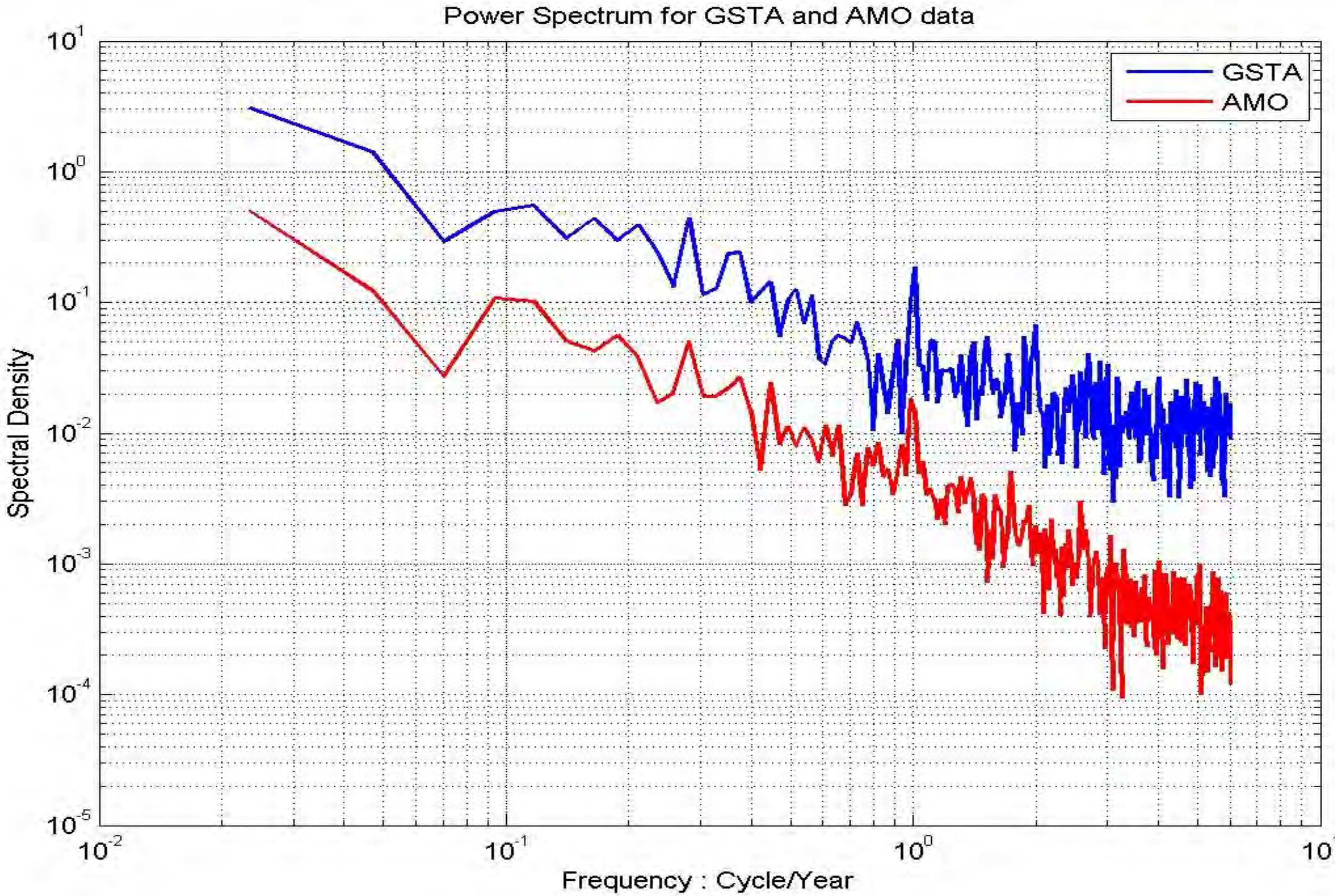
HHT with intermittence test

AMO

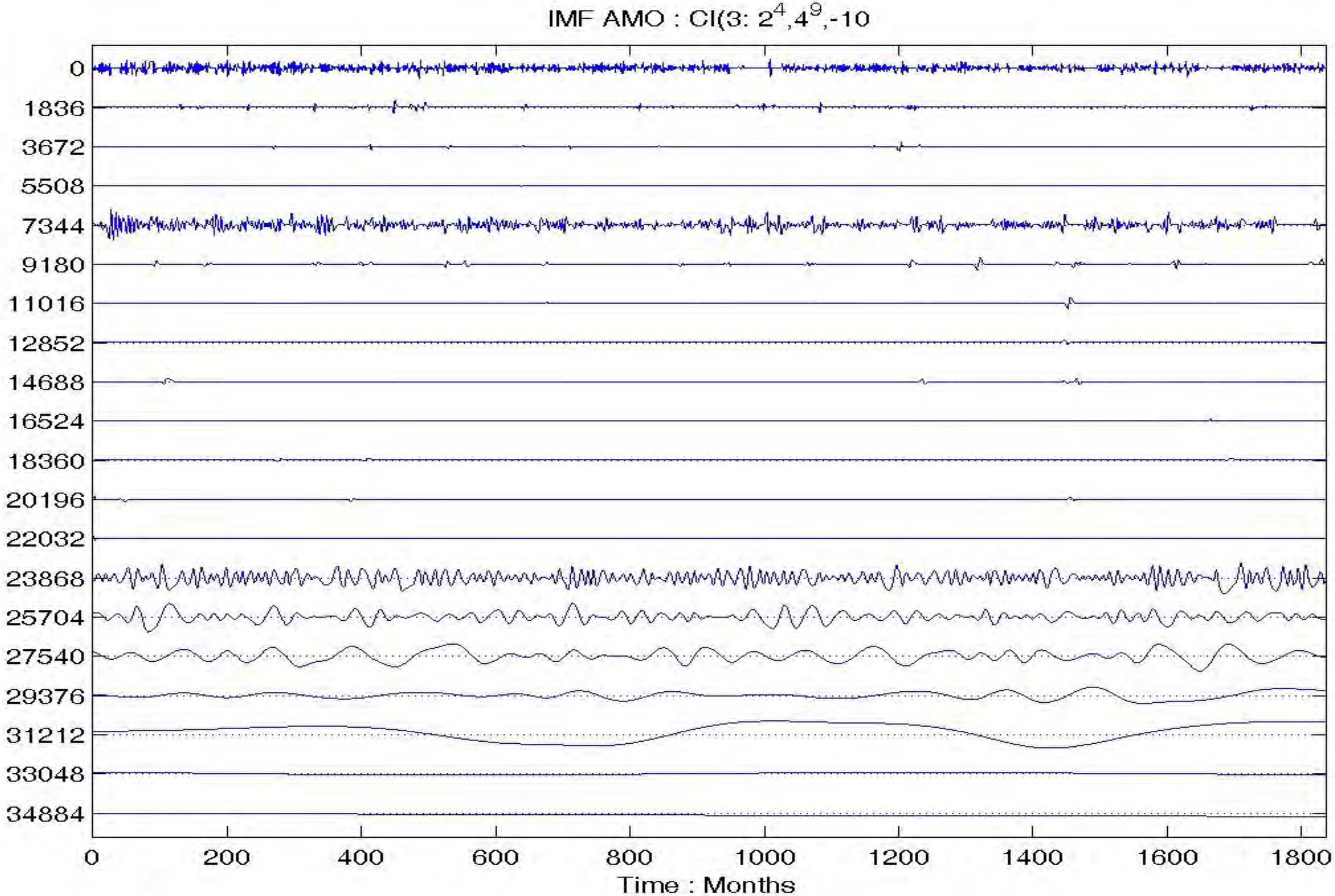


The Atlantic Multi-Decadal Oscillation
Temperature : °C
Time : Year
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
1840
1860
1880
1900
1920
1940
1960
1980
2000
2020

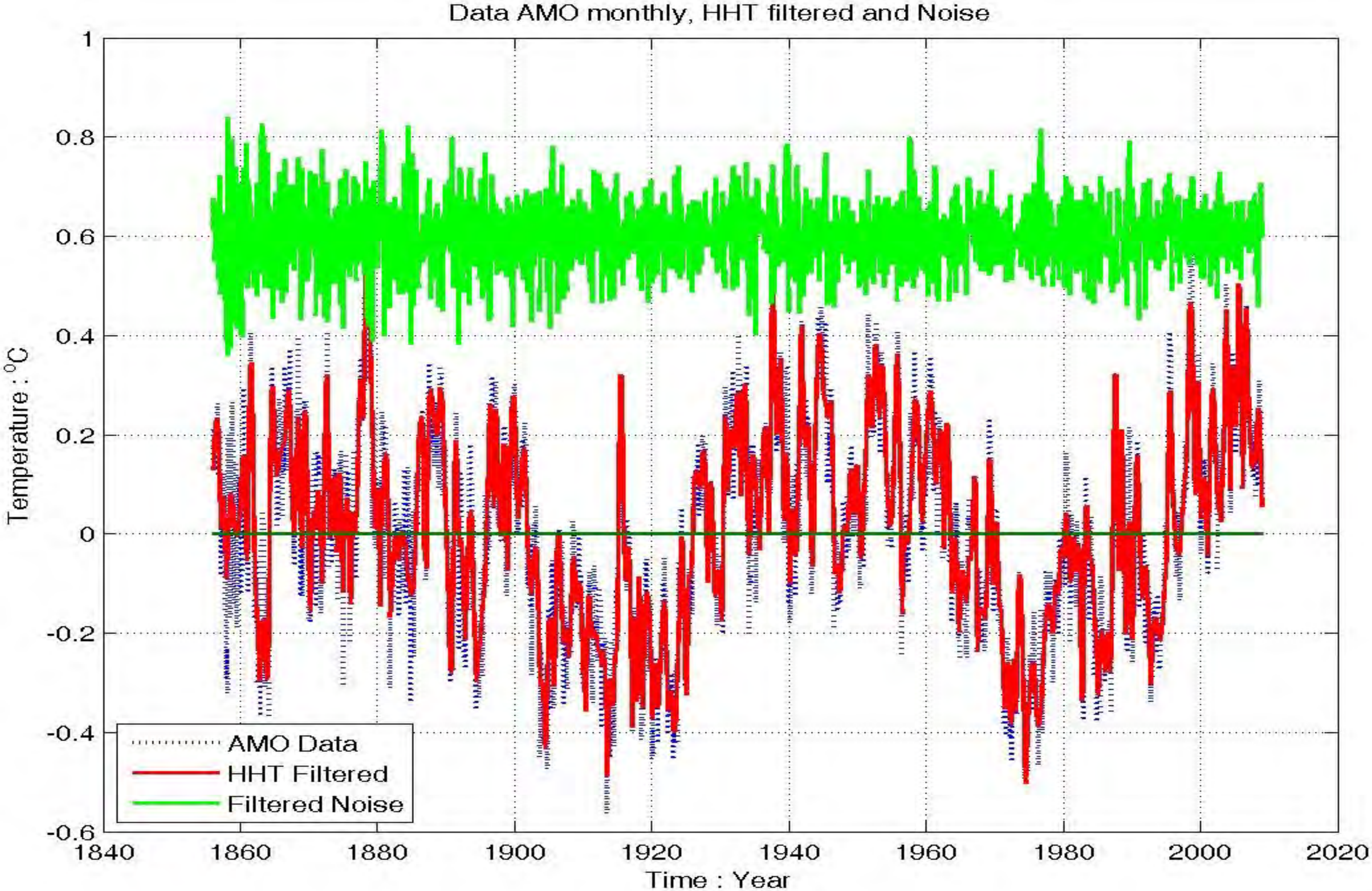
Fourier Spectra for GSTA and AMO



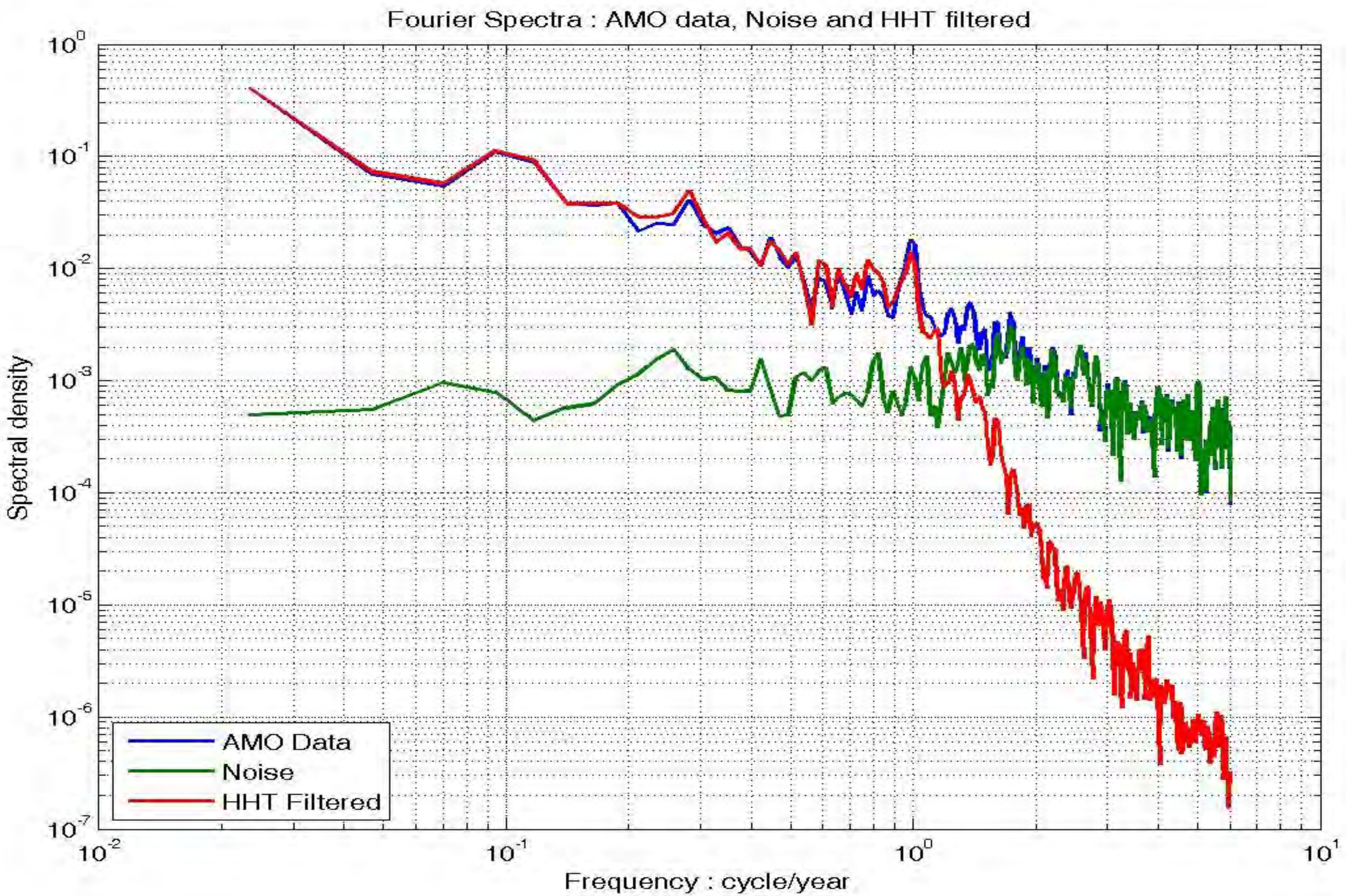
AMO IMF


IMF AMO : CI(3: 2^4,4^9,-10
0
1836
3672
5508
7344
9180
11016
12852
14688
16524
18360
20196
22032
23868
25704
27540
29376
31212
33048
34884
0
200
400
600
800
1000
1200
1400
1600
1800
Time : Months

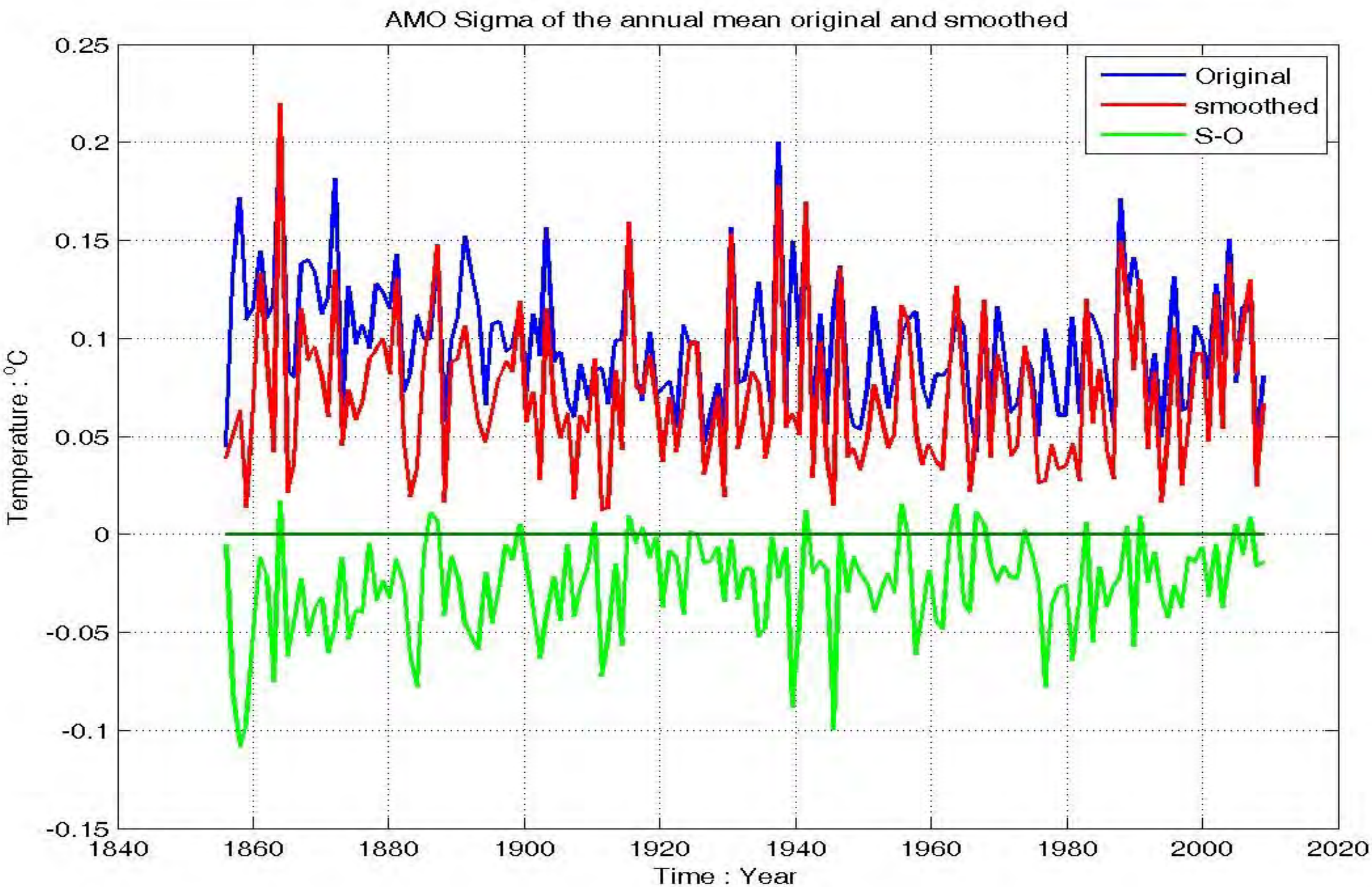
AMO : Filtered and original data



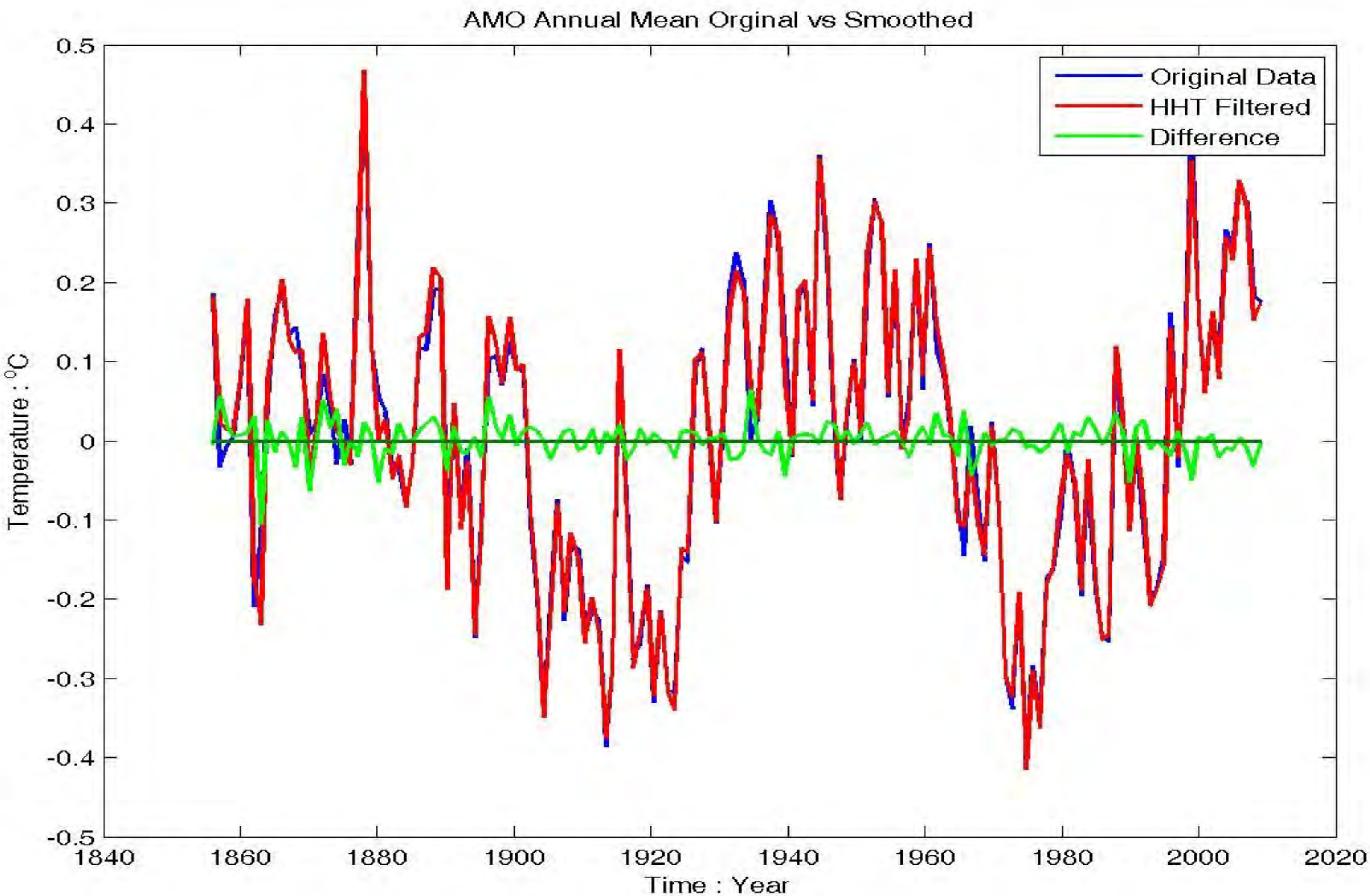
AMO : Spectra Filtered and original data



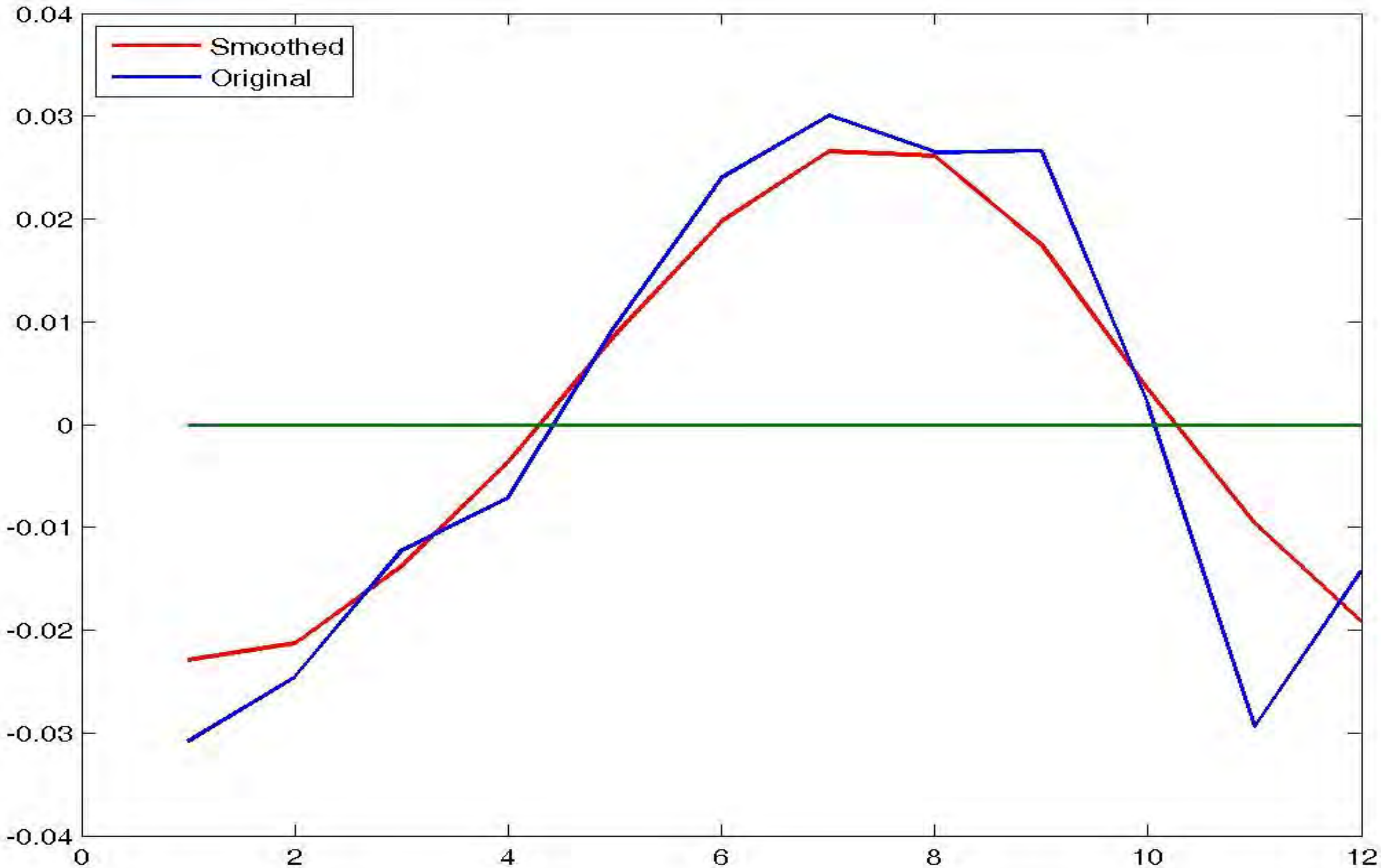
AMO : STD Filtered and original data



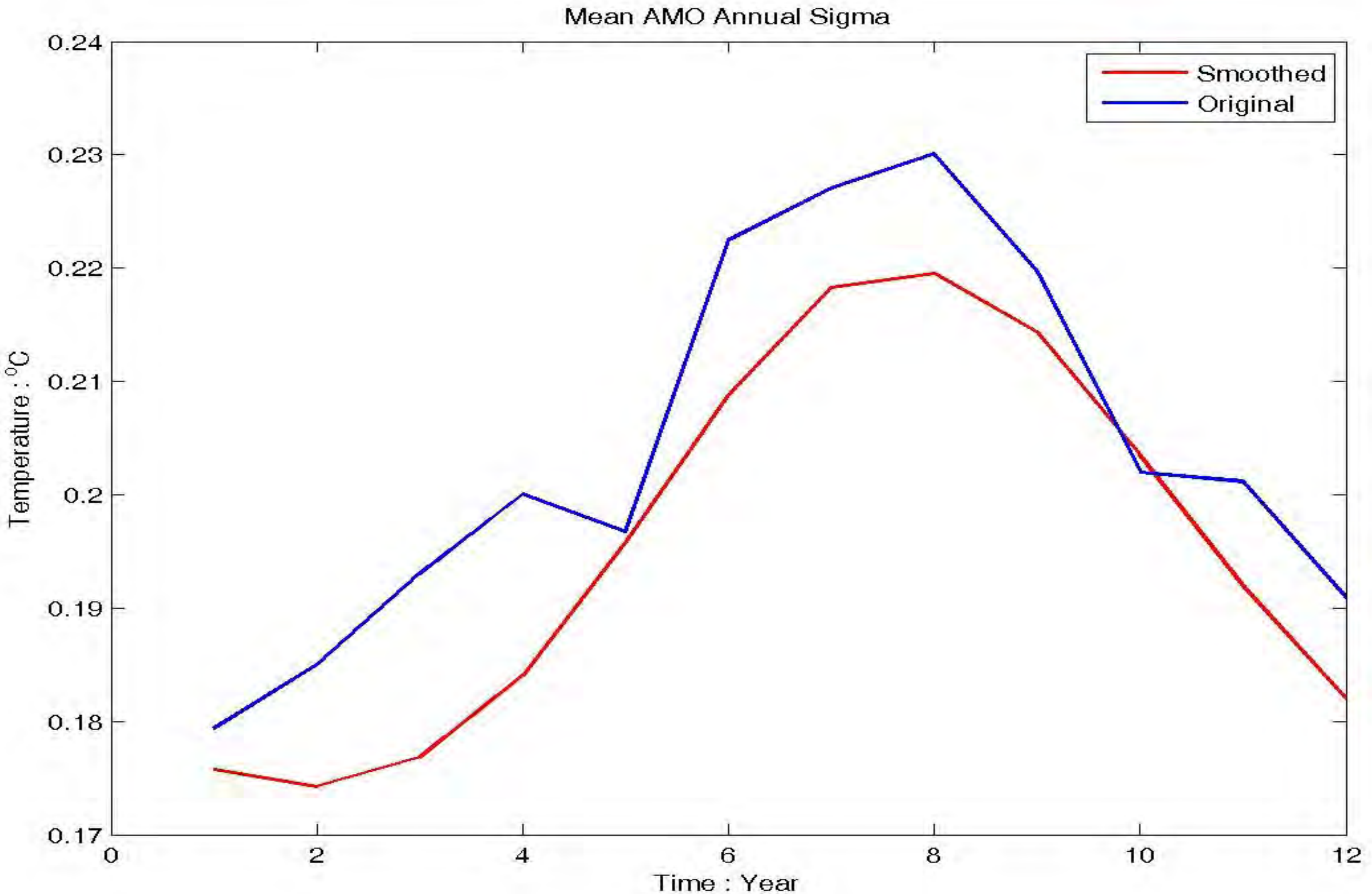
AMO: Annual mean filtered and original data



AMO: Variations for Annual mean filtered and original data



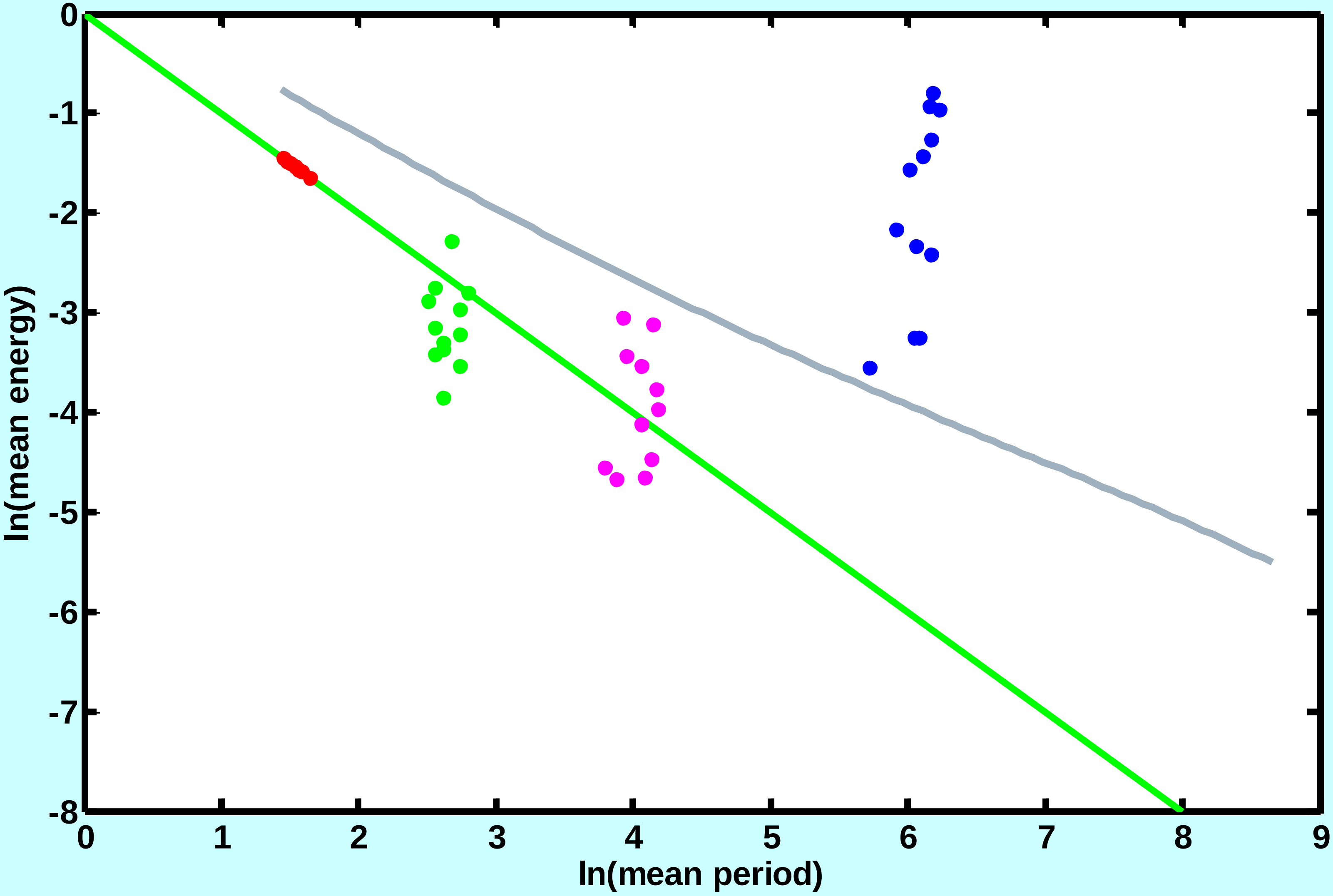
AMO: Sigma for Annual mean filtered and original data



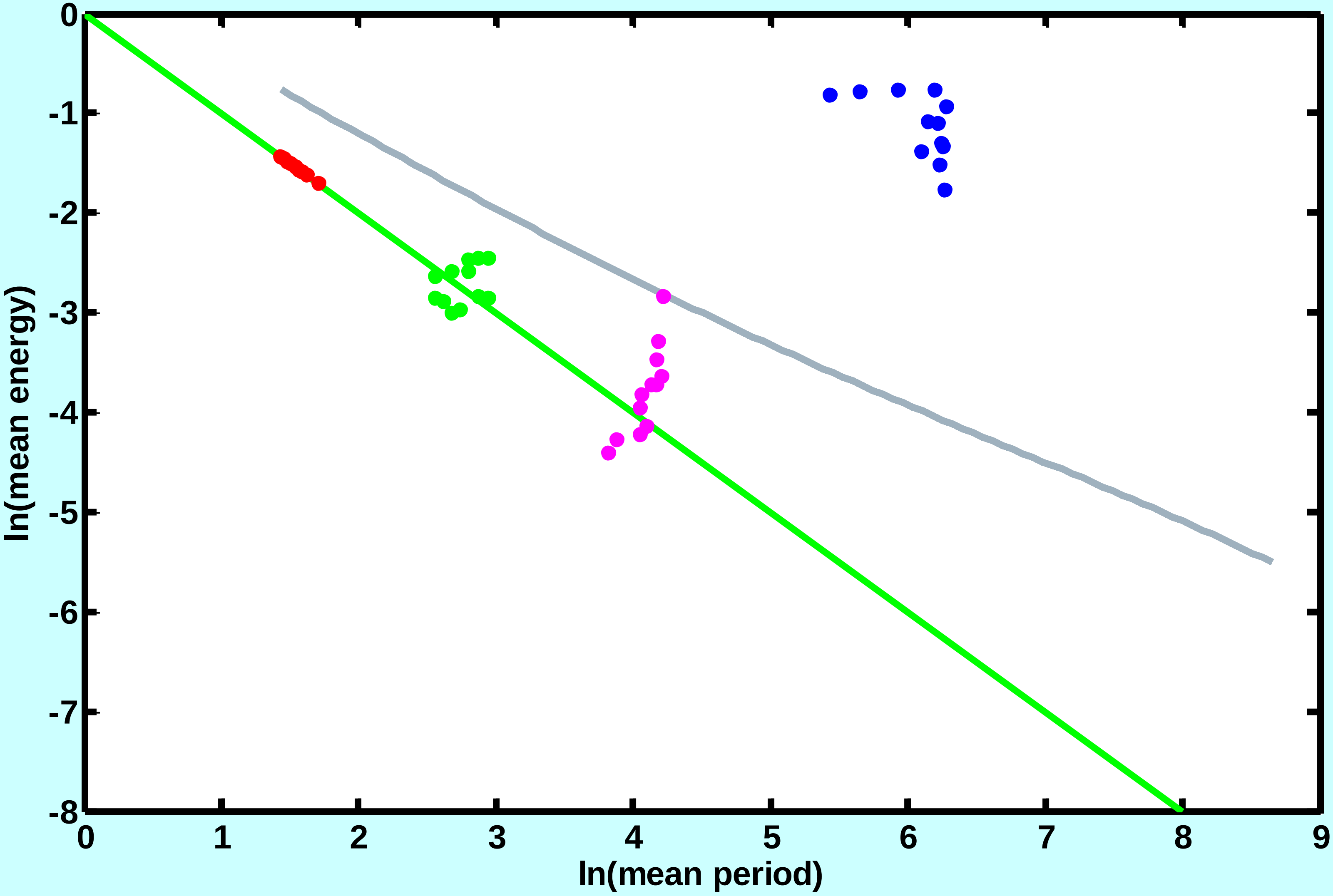
Similarity and Cross-correlations

Between GSTA and AMO

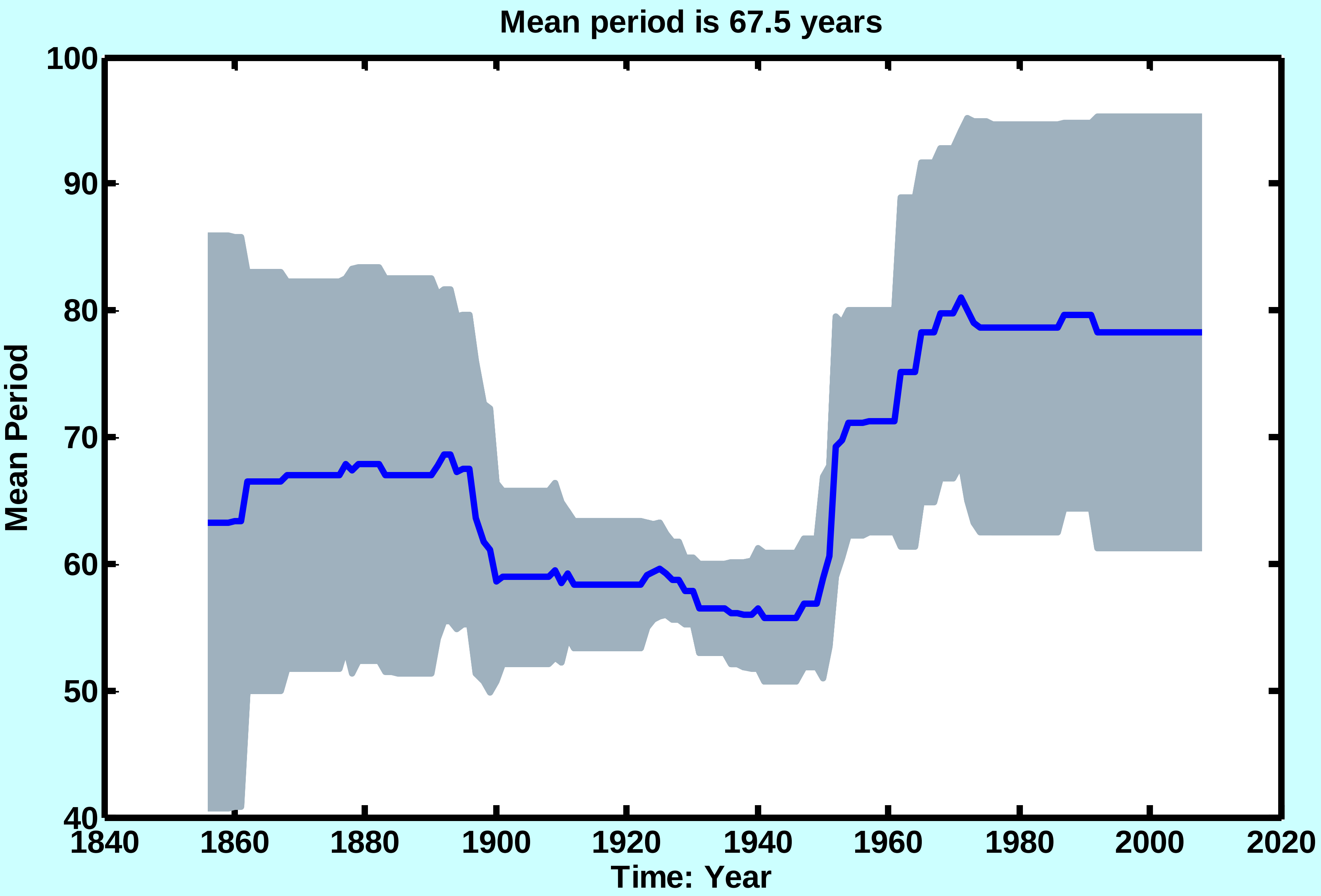
Significance Test of GSTA



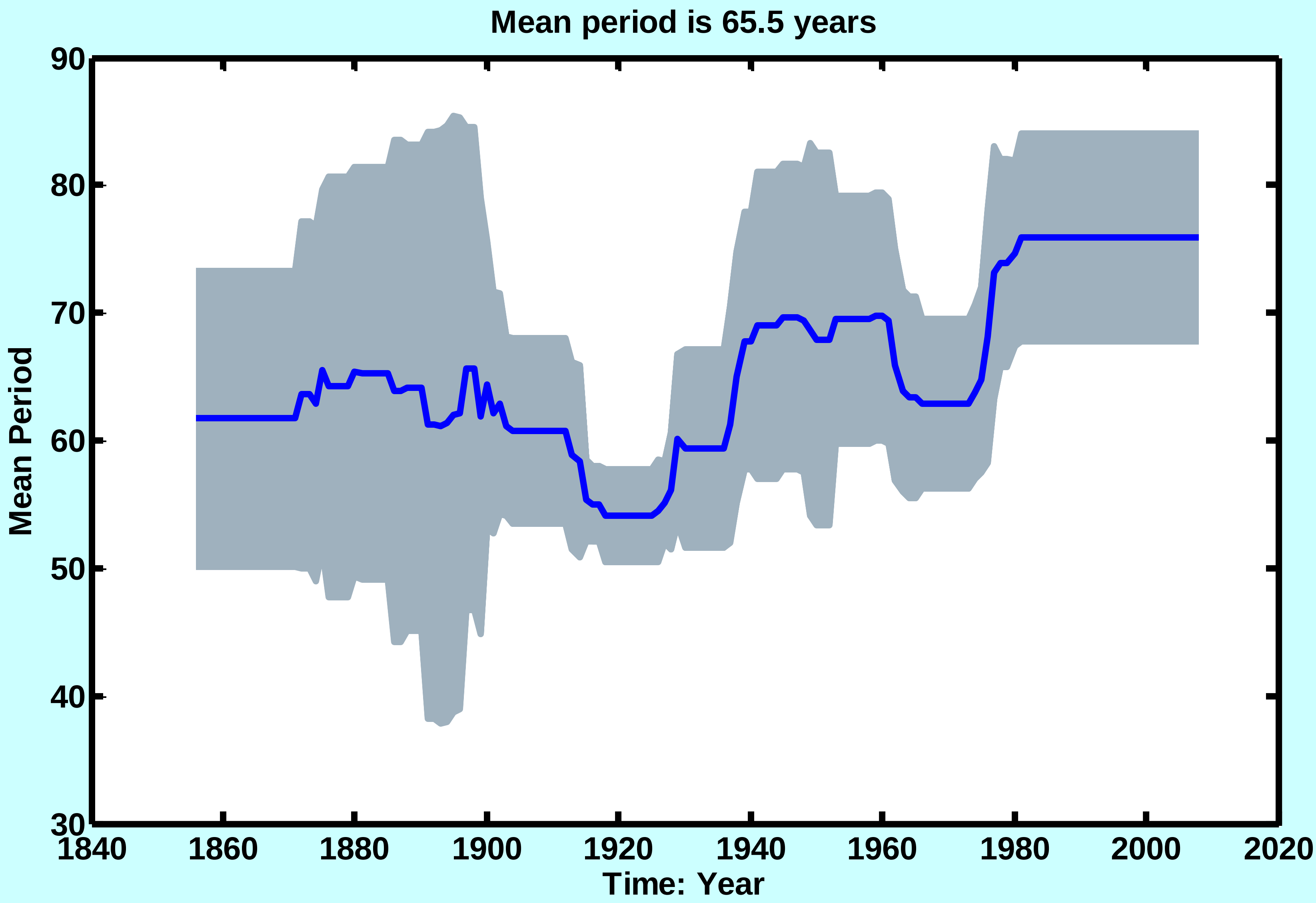
Significance Test of AMO



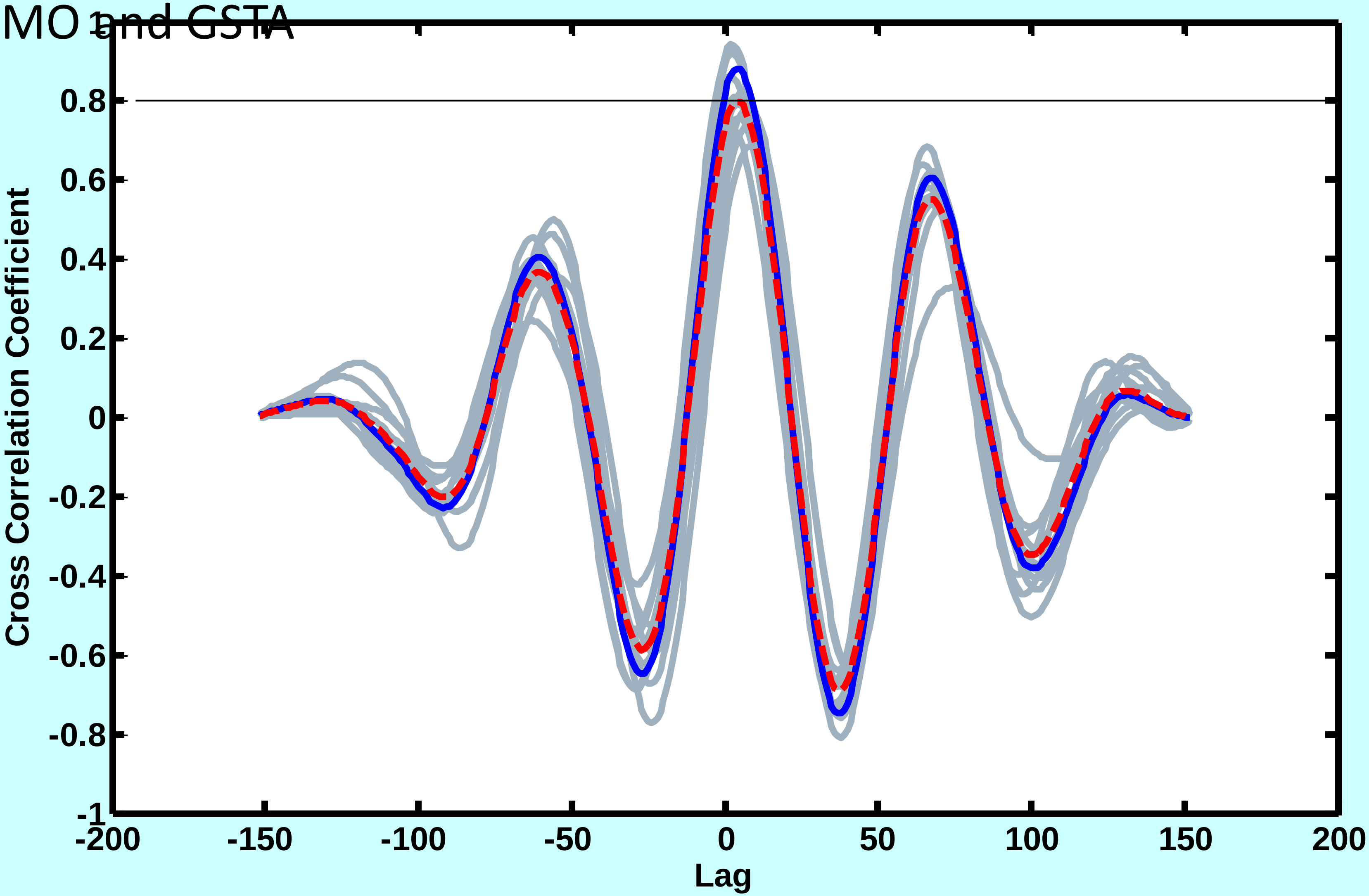
Mean Instantaneous Periods of IMF4 of GSTA



Mean Instantaneous Periods of IMF4 of AMO



Cross-Correlation between IMFs 4 of AMO and GSTA



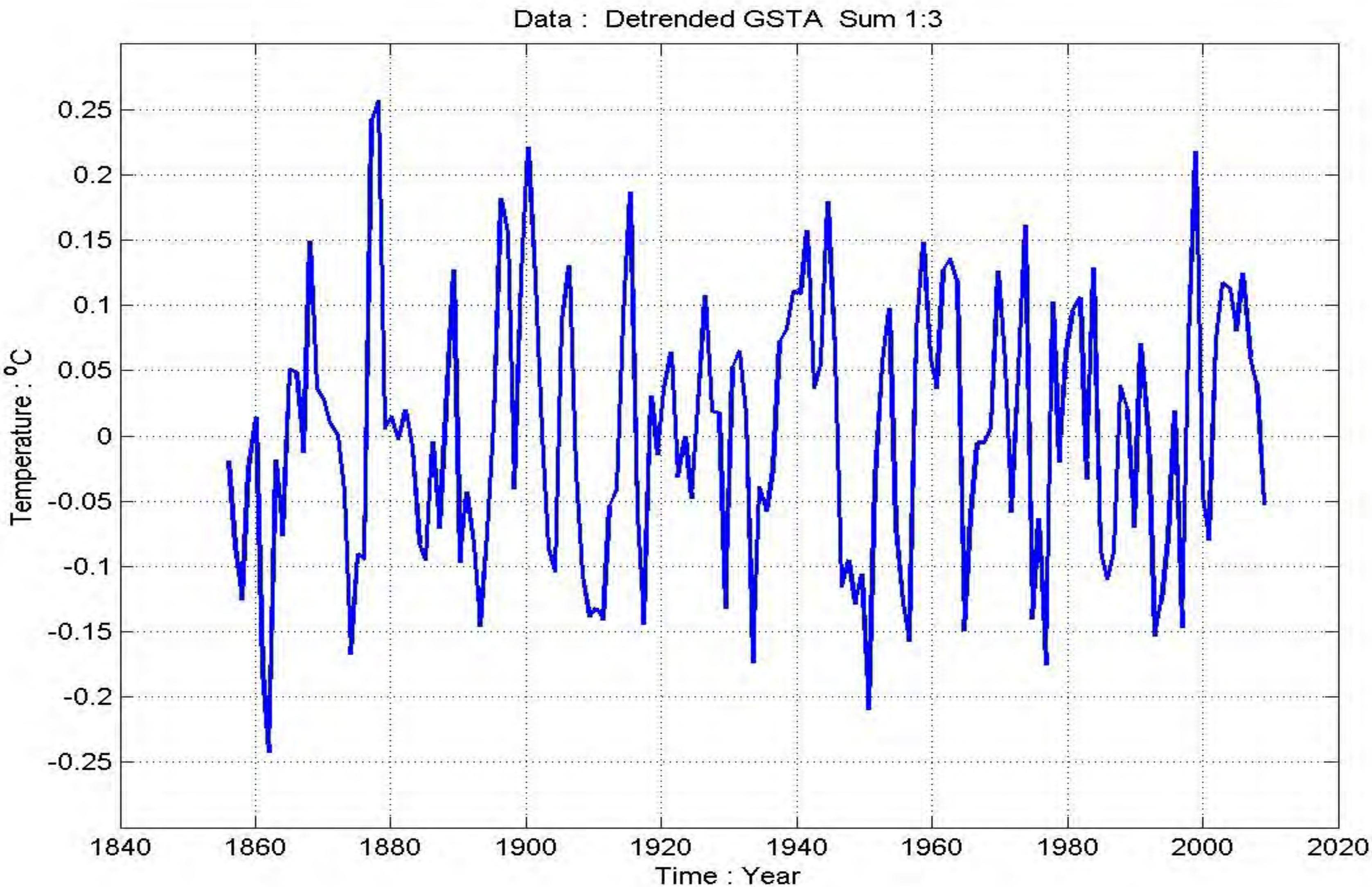
Blue line: correlation of annual mean of GSTA and AMO

Red line: mean of correlation of each downsample of GSTA and AMO

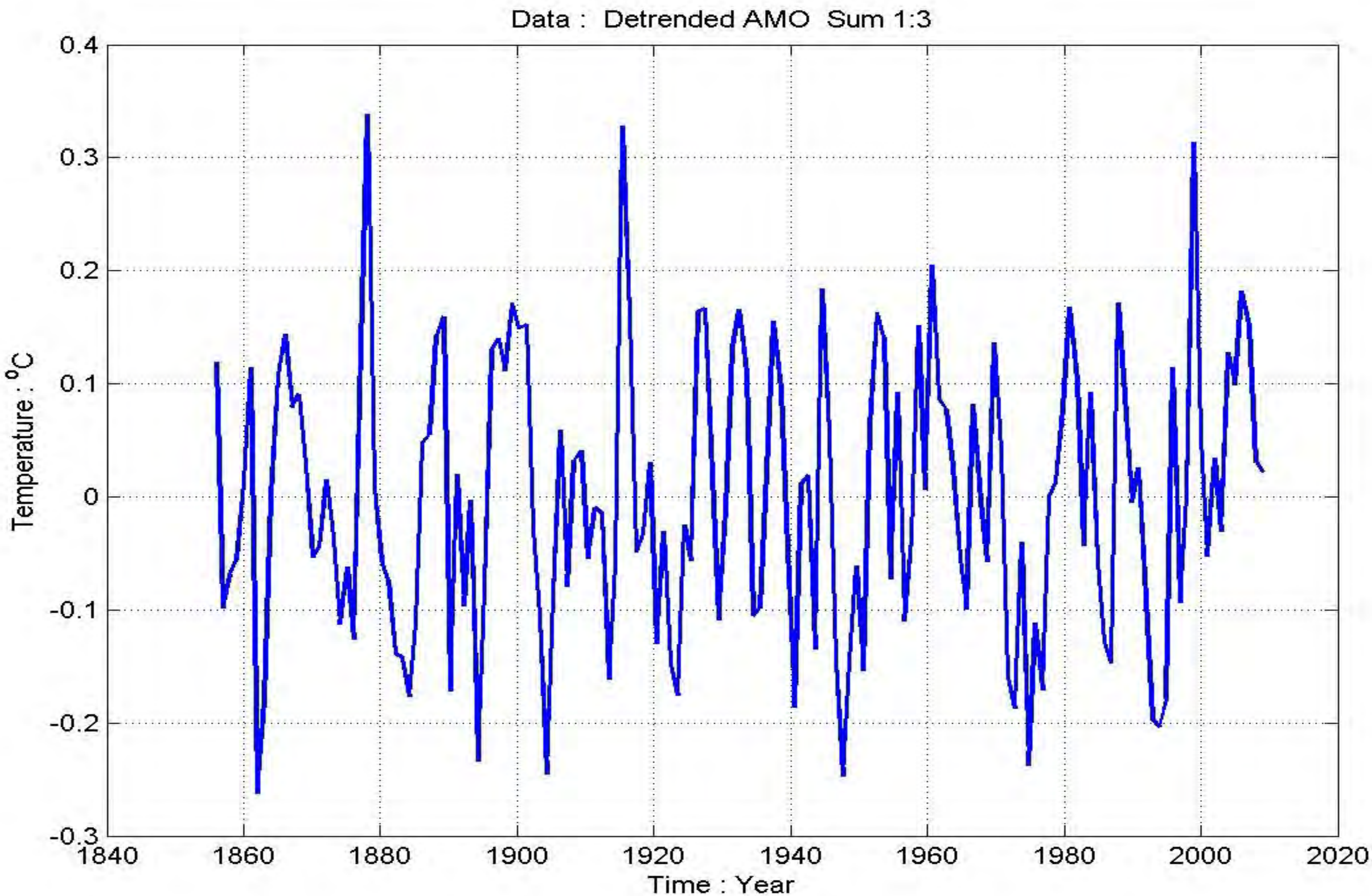
Comparison between

GSTA and AMO

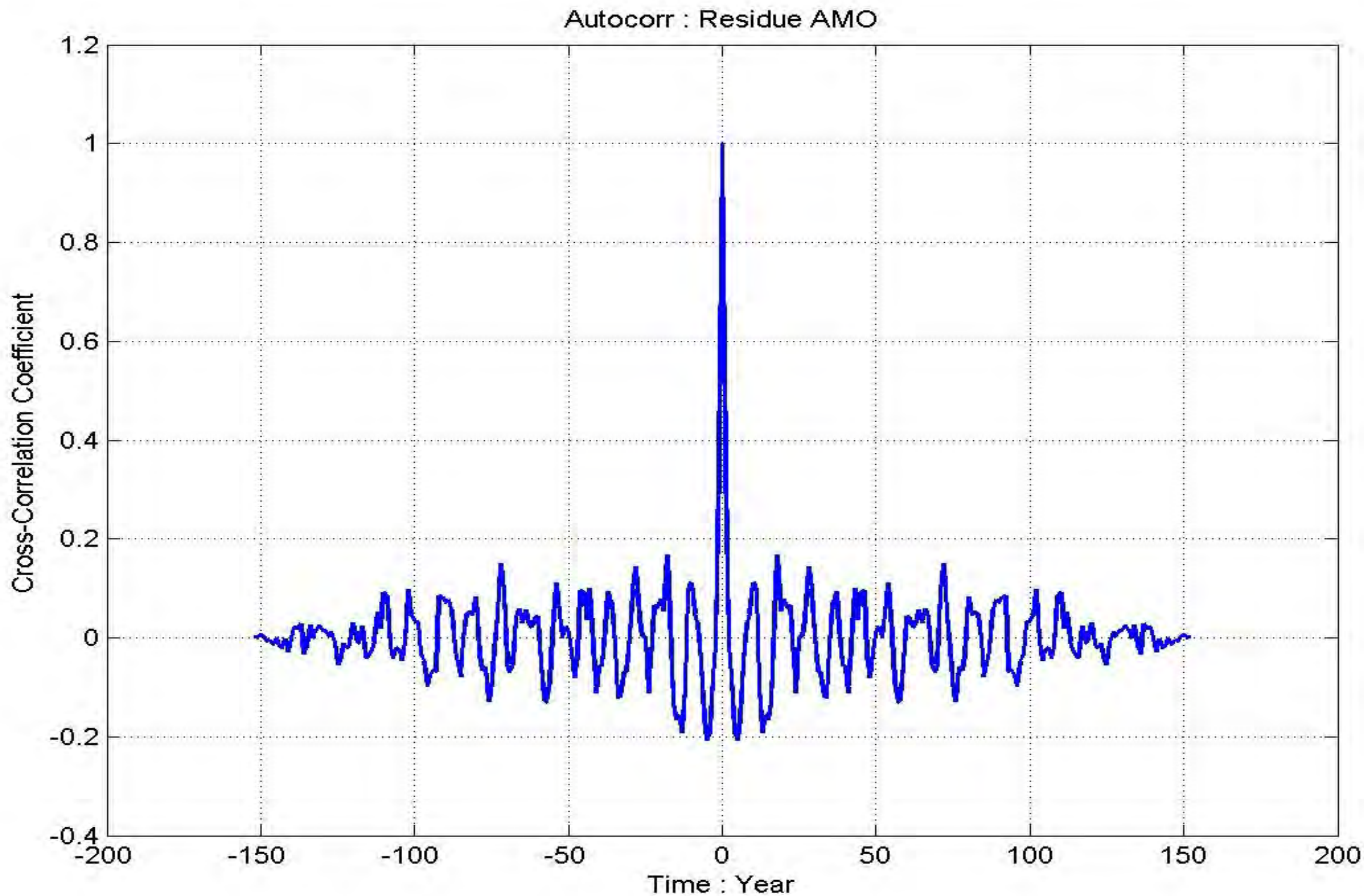
Detrended GSTA



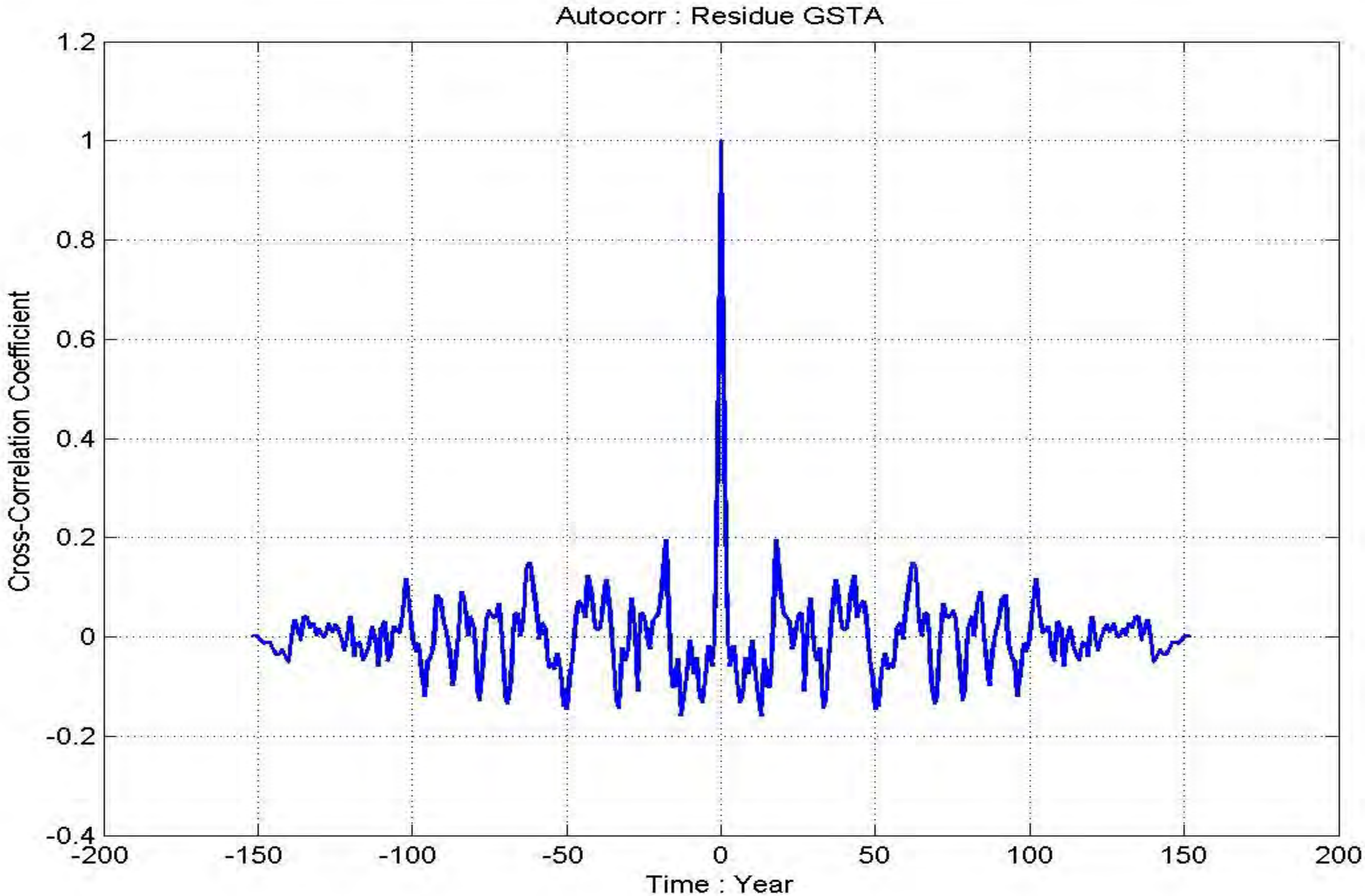
Detrended AMO



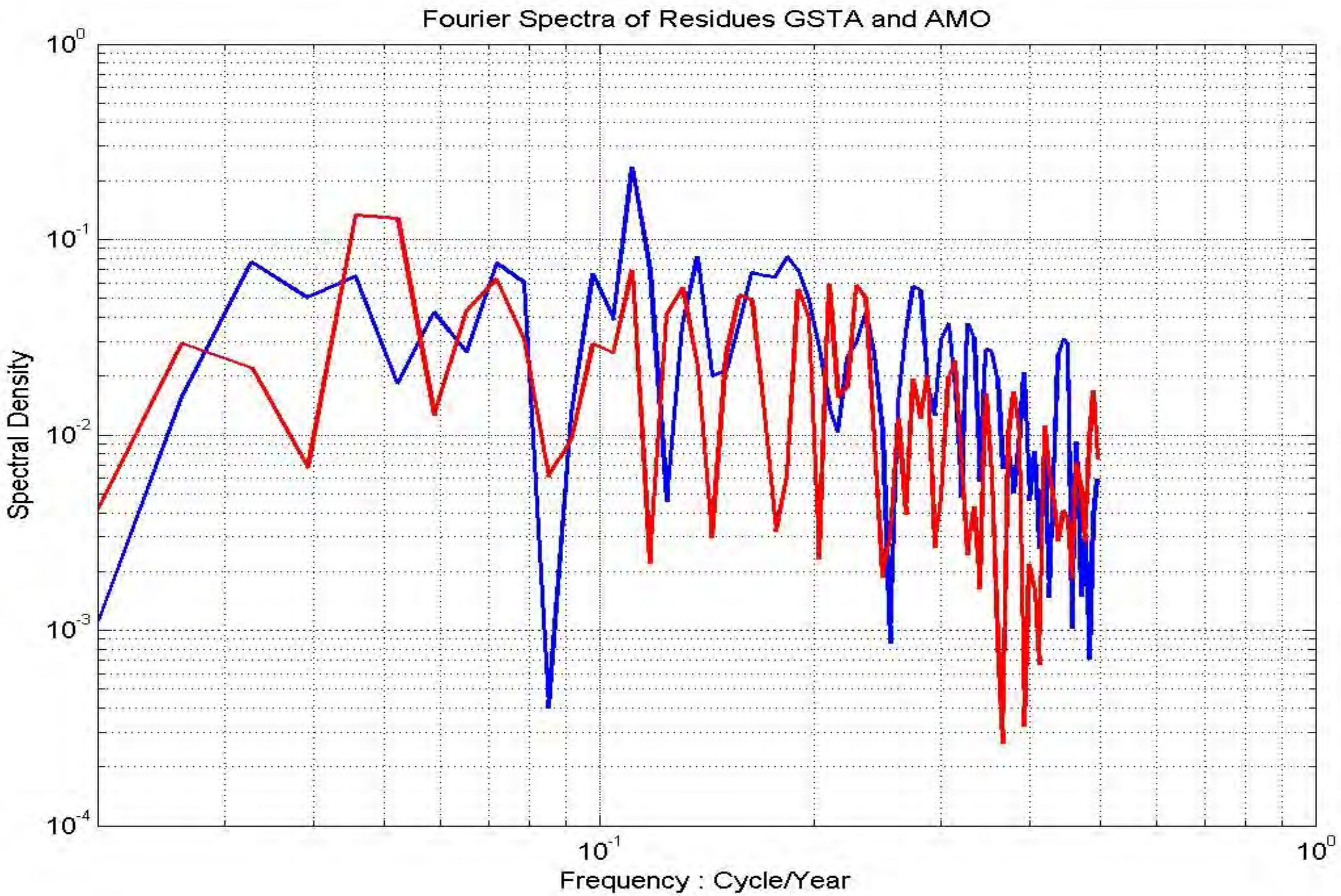
Autocorr : Residues AMO



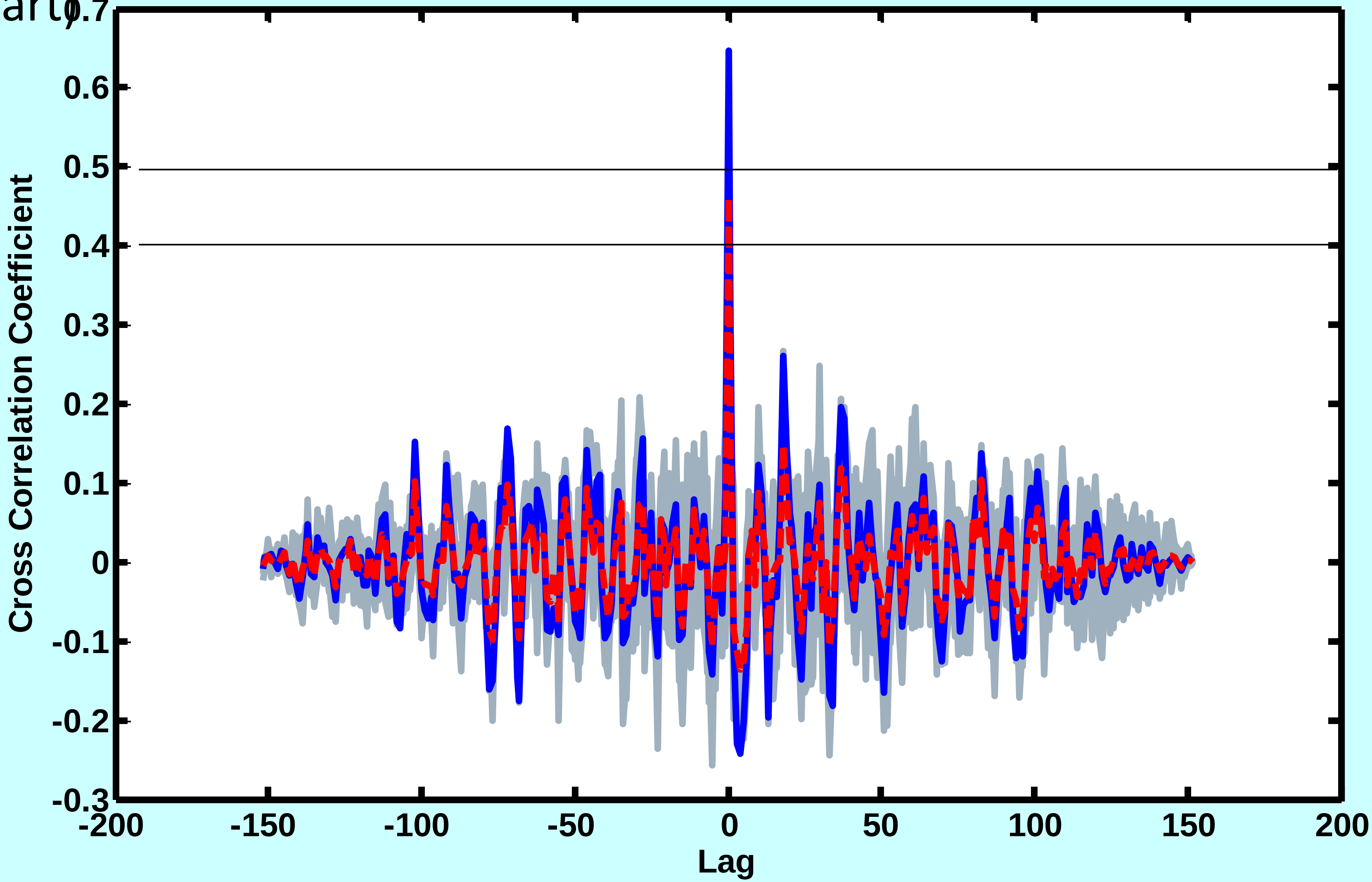
Autocorr : Residues GSTA



Fourier Spectra of Residues



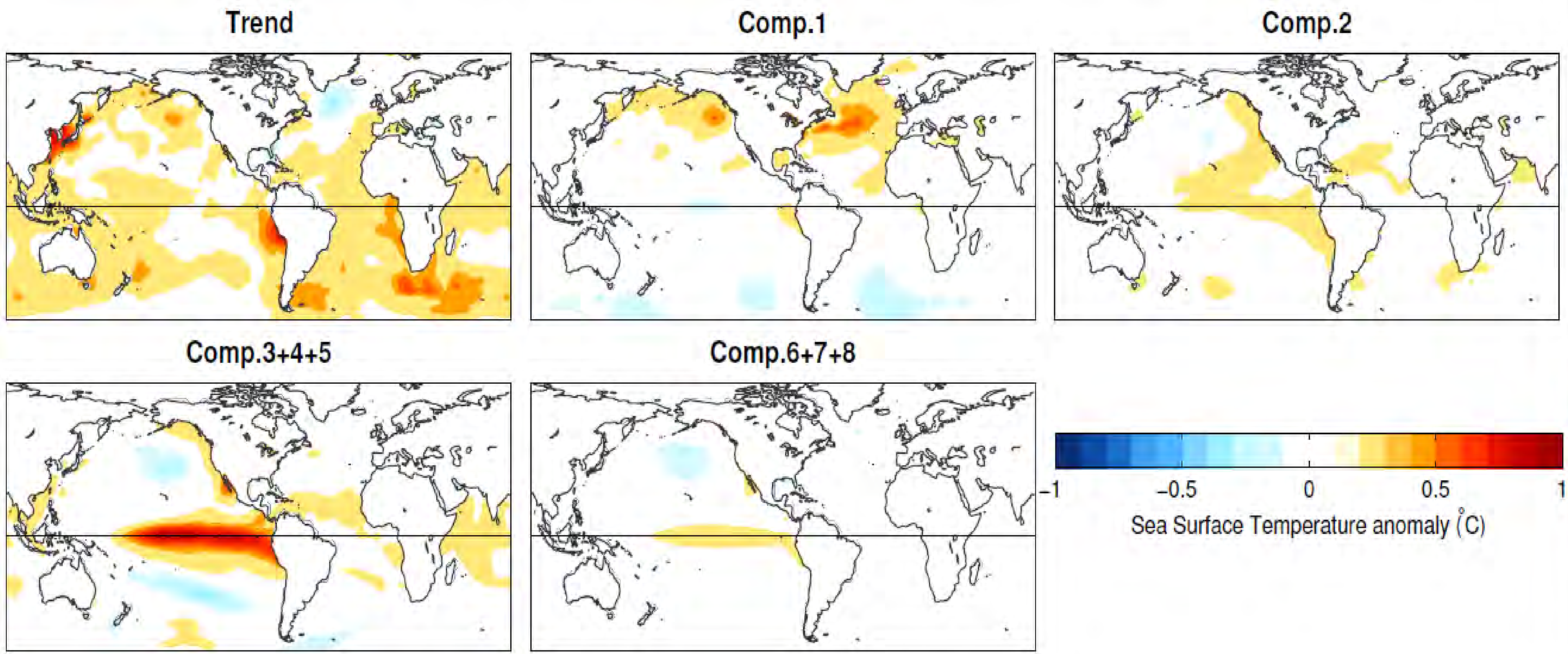
Cross-Correlation between IMFs 1-3 of AMO and GSTA (noise part)



Blue line: correlation of annual mean of GSTA and AMO

Red line: mean of correlation of each downsample of GSTA and AMO

Global Surface Temperature Data Regressions



What are statistically noisy components could still be physically meaningful and significant.

ENSO phenomenon is stochastic with statistical properties of noise, but it is a dominant factor in global climate change.

Question: Is long term ENSO predictable?

With HHT we can determine cycles, the trend and eliminate noise.

Thanks