On the Analysis of Linear and Quadratic Chirp Processes Using Time Deformation

Liangang Liu, Henry L. Gray and Wayne A. Woodward*

Abstract

A new class of non-stationary processes called linear chirp stationary (L-C) process is developed, whose frequencies change approximately linearly in time. The Wigner-Ville time-frequency distribution is used to estimate the parameters in the L-C stationary process. The spectral analysis and forecasting performance of the L-C modeling is shown to be much better than the traditional AR modeling. The framework of filtering L-C stationary process is proposed based on the work of filtering M-stationary process (Kohlmia, 2004). Both simulated and real seismic data are analyzed using the L-C model. Another class of non-stationary processes called quadratic-chirp stationary (Q-C) process is discussed and some properties are given as well.

*Department of Statistical Science, Southern Methodist University, 3225 Daniel Avenue, P.O.Box 750332, Dallas, TX 75275-0332.

1 Introduction

The Fourier Transformation (FT) provides the basis for the classical tools of time series analysis under a weak stationarity assumption. However, stationarity is a mathematical idealization that in some cases may be invalid. Examples can be found in bat echolocation, whale clicks, sonar and radar data in which frequencies are changing in time and thus classical tools based on a stationarity assumption are invalid. FT is not suitable for these data since it does not explicitly show the time localization of frequency component, i.e., the FT spectrum loses all the time information. The short-time Fourier transform (STFT) and the wavelet transform (WT) are two of the most popular tools to analyze this type non-stationary data. The STFT and WT are both window based methods and they suffer from the same time-frequency resolution limitations, i.e., time resolution and frequency resolution can not be made arbitrarily good simultaneously, which is known as uncertainty principle (Boggess and Narcowich, 2001). Another way to look at non-stationary time series data of this type is to transfer it to stationarity by time deformation. Gray and Zhang (1988) first defined the continuous parameter M-stationary process and Euler process for applications ini which frequency is continuous decreasing with time. Gray and Zhang (1988) showed that many such series can be transferred to stationarity By use of the logarithmic time deformation. Gray, Vijverberg and Woodward (2004) extended their results to discrete M-stationary and Euler processes. Jiang, Gray and Woodward (2004) generalized the "Logarithmic" transform to a Box-Cox type transform and defined G(λ)-stationary process.

In this work, two new classes of non-stationary processes called linear chirp stationary processes and quadratic chirp stationary process are discussed, which have approximately linearly and quadratically changing frequency, respectively. Both of

the processes can be transformed to stationarity by certain time deformation. Some simulated and real data examples are given as well.

2 Linear Chirp Stationary Processes

2.1 Definitions and Properties

In the following, we define the concept of Linear Chirp Stationary Processes.

Definition 2.1 *Let $X(t)$ be a stochastic process defined for $t \in (0, \infty)$ such that for any $(\frac{-b+\sqrt{(2at+b)^2+4a\tau}}{2a}) \in (0, \infty)$ and constatns $a > 0$ and $b \geq 0$ the following hold:*

(i) $E[X(t)] = \mu$

(ii) $Var[X(t)] = \sigma^2 < \infty$

(iii) $E[(X(t) - \mu)(X((\frac{-b+\sqrt{(2at+b)^2+4a\tau}}{2a}) - \mu)] = R_X(\tau),$

Then $X(t)$ is called a Linear Chirp (L-C) stationary process, $R_X(\tau)$ is the L-C auto-covariance function, and τ is the L-C time lag.

The L-C autocorrelation function is defined by $\rho_X(\tau) = \frac{R_X(\tau)}{var(X(t))}$. Obviously, the L-C stationary process is non-stationary in the usual sense.

Definition 2.2 *Let $X(t)$ be an L-C stationary process, and let $Y(u) = X(t)$, where $u = g(t) = at^2 + bt$. Then $Y(u)$ will be called the stationary dual of $X(t)$.*

Note that $Y(u)$ is weakly stationary in the traditional sense.

Theorem 2.1 *A stochastic process $\{X(t), t > 0\}$ is L-C stationary if and only if it has a stationary dual process $Y(u)$. Also, $R_X(\tau) = R_Y(\tau)$.*

Proof: The constant and finite mean and variance hold obviously. W.l.o.g., let $\mu = 0$. For the autocovariance,

$\Leftarrow$: Suppose $Y(u)$ is stationary. Then it follows that

$$
\begin{aligned}
R_X(\tau) &= E[X(t)X(\frac{-b + \sqrt{(2at + b)^2 + 4a\tau}}{2a})] \\
&= E[X(g^{-1}(u))X(g^{-1}(u + \tau))] \\
&= E[Y(u)Y(u + \tau)] \\
&= R_Y(\tau).
\end{aligned}
$$

Hence $R_X(\tau)$ depends only on τ. Therefore $X(t)$ is an L-C stationary process.

$\Rightarrow$: Suppose $X(t)$ is L-C stationary. Then

$$
\begin{aligned}
R_Y(\tau) &= E[Y(u)Y(u + \tau)] \\
&= E[Y(g(t))Y(g(\frac{-b + \sqrt{(2at + b)^2 + 4a\tau}}{2a}))] \\
&= E[X(t)X(\frac{-b + \sqrt{(2at + b)^2 + 4a\tau}}{2a})] \\
&= R_X(\tau).
\end{aligned}
$$

Hence $R_Y(\tau)$ depends only on τ, and $Y(u)$ is a weakly stationary process.

Definition 2.3 *Let $X(t)$ be an L-C stationary process. Then the L-C spectrum is defined as the Fourier Transform of the L-C autocovariance function, i.e.,*

$$
G_X(f) = \int_{-\infty}^{\infty} e^{-2\pi f\tau} R_X(\tau) d\tau. \tag{1}
$$

The L-C spectral density is defined as the Fourier Transform of L-C autocorrelation function, i.e., $s_X(f) = \frac{G_X(f)}{var(X(t))}$. Furthermore, the L-C spectrum is equal to the power spectrum of the dual process $Y(u)$, i.e., $G_X(f) = G_Y(f)$ since $R_X(\tau) = R_Y(\tau)$.

4

Definition 2.4 *The instantaneous ACF of an L-C stationary process is defined as the usual ACF of the process at a time t. That is,*

$$
\begin{aligned}
\rho_X^*(\tau, t) &= E[X(t)X(t+\tau)]/var(X(t)) \\
&= \rho_X\left(\frac{[2a(t+\tau)+b]^2 - (2at+b)^2}{4a}\right), \quad for \quad \tau \geq 0
\end{aligned}
\tag{2}
$$

and

$$
\rho_X^*(\tau, t) = \rho_X\left(\frac{(2at+b)^2 - [2a(t+\tau)+b]^2}{4a}\right), \quad for \quad -t < \tau < 0
\tag{3}
$$

Note that the instantaneous ACF of L-C stationary at time t and lag τ is just the L-C ACF at different L-C time lag.

Definition 2.5 *The instantaneous period of the ACF of an L-C stationary process at time t is defined as*

$$
\begin{aligned}
IP(t) &= \frac{-b + \sqrt{(2at+b)^2 + 4a\tau_l}}{2a} - t \tag{4} \\
&= \frac{(2at+b)}{2a}\left\{-1 + \sqrt{1 + \frac{4a\tau_l}{(2at+b)^2}}\right\}, \tag{5}
\end{aligned}
$$

where τ_l is the L-C period of the ACF of the L-C stationary process. Actually, τ_l is also the period of the ACF of the dual process.

Definition 2.6 *The instantaneous frequency of the ACF of an L-C stationary process is defined as the reciprocal the IP, i.e.,*

$$
\begin{aligned}
IF &= \frac{1}{\frac{-b+\sqrt{(2at+)b^2+4a\tau_l}}{2a} - t} \tag{6} \\
&= \frac{2a}{(2at+b)\{-1 + \sqrt{1 + \frac{4a\tau_l}{(2at+b)^2}}\}} \tag{7} \\
&= \frac{(2at+b)}{2\tau_l}\left\{1 + \sqrt{1 + \frac{4a\tau_l}{(2at+b)^2}}\right\}. \tag{8}
\end{aligned}
$$

Note that the IF is asymptotically linear with respect to t. Instantaneous frequency is commonly defined in the engineer community as the derivative of the phase of the signal (Boashash, 1992). In the L-C process, this is $f(t) = 2at + b$. Note that $f(t)$ is asymptotically equivalent to IF since

$$\lim_{t \to \infty} \frac{f(t)}{IF} = 1.$$

Furthermore, Jiang(2003) showed that $f(t)$ is the first term of Taylor series expansion of IF.

Definition 2.7 *Let $\varepsilon(u)$ be continuous white noise, i.e., $E[\varepsilon(u)] = 0$ and*

$$E[\varepsilon(u)\varepsilon(u + \tau)] = C\delta(\tau),$$

where C is a positive constant and δ is the dirac delta function. Also let $a(t) = \varepsilon(at^2 + bt)$ for $t > 0$. Then $a(t)$ will be referred to as continuous L-C white noise.

Theorem 2.2 *Let $a(t)$ be continuous L-C white noise, then*

$$E\left[a(t)a\left(\frac{-b + \sqrt{(2at + b)^2 + 4a\tau}}{2a}\right)\right] = C\delta(\tau).$$

Definition 2.8 *Let $X(t)$ be a continuous parameter L-C stationary process, and let $G_X(f^*)$ be the L-C spectrum. Then the instantaneous spectrum of $X(t)$ is*

$$S_X(f;t) = G_X\left(\frac{f^2}{(2at + b)f + a}\right) = G_Y\left(\frac{f^2}{(2at + b)f + a}\right), \tag{9}$$

where f is the instantaneous frequency and G_Y is the spectrum of the dual.

Theorem 2.3 *If $X(t)$ is a continuous parameter L-C stationary process, then the instantaneous ACF and spectrum are related by the incomplete Fourier Transform:*

$$S_X(f;t) = \int_{-t}^{\infty} \exp\{-i2\pi f^* l\}\rho_X^*(\tau, t)[2a(t + \tau) + b]d\tau, \tag{10}$$

where $f^* = \frac{f^2}{(2at+b)f+a}$ and $l = \frac{[2a(t+\tau)+b]^2 - (2at+b)^2}{4a}$.

Proof:

$$
\begin{aligned}
S_X(f;t) &= G_X\left(\frac{f^2}{(2at+b)f+a}\right) \\
&= G_Y\left(\frac{f^2}{(2at+b)f+a}\right) \\
&= G_Y(f^*) \\
&= \int_{-\infty}^{\infty} \exp\{-i2\pi f^* l\}\rho_Y(l)dl
\end{aligned}
$$

Letting $l = \frac{[2a(t+\tau)+b]^2-(2at+b)^2}{4a}$, then it follows that

$$
S_X(f;t) = \int_{-t}^{\infty} \exp\{-i2\pi f^* l\}\rho_X^*(\tau,t)[2a(t+\tau)+b]d\tau.
$$

For discrete observations, the instantaneous spectrum will be calculated by

$$
\begin{aligned}
S_X(f;t) &= G_X\left(\frac{f^2}{(2at+b)f+a}\right) \\
&= G_Y\left(\frac{f^2 h}{(2at+b)f+a}\right) \\
&= \sum_{k=-\infty}^{\infty} \exp\left\{-i2\pi\left(\frac{f^2 h}{(2at+b)f+a}\right)k\right\}\rho_Y(k). \qquad (11)
\end{aligned}
$$

Therefore, the relationship between the instantaneous spectrum and the instantaneous ACF based on an discrete realization of a L-C process is given by

$$
\begin{aligned}
\rho_X^*(\tau,t) &= \sum_{f} \exp\left\{i2\pi\left(\frac{f^2 h}{(2at+b)f+a}\right)\frac{[2a(t+\tau)+b]^2-(2at+b)^2}{4ah}\right\}S_X(f;t) \\
&= \sum_{f} \exp\left\{i2\pi\left(\frac{f^2\{[2a(t+\tau)+b]^2-(2at+b)^2\}}{4a[(2at+b)f+a]}\right)\right\}S_X(f;t). \qquad (12)
\end{aligned}
$$

Definition 2.9 *Let $X(t)$ be a discrete L-C stationary process, and let $Y(u)$ denote the dual based on sample rate h. Then the Instantaneous Nyquist Frequency of $X(t)$ at time t is defined as*

$$
f_{Ny,X} = \frac{(2at+b)+\sqrt{(2at+b)^2+8ah}}{4h}. \qquad (13)
$$

7

One class of L-C processes are the L-C Autoregressive processes.

Definition 2.10 *A process $X(t)$, $t > 0$ defined by*

$$\prod_{i=1}^{p}(\frac{1}{2at+b}D - \alpha_i)X(t) = a(t), \tag{14}$$

where D is the differential operator, $a > 0$, $b \geq 0$, α_i are constants, $p > 0$ and $a(t)$ is L-C white noise, is called a continuous parameter p^{th} order L-C Autoregressive process. In this case the L-C autocorrelation function, $R_X(\tau)$, satisfies the differential equation

$$\prod_{i=1}^{p}(\frac{1}{2at+b}D - \alpha_i)R_X(\tau) = 0.$$

The dual $Y(u) = X(t)$, where $u = at^2 + bt$, satisfies

$$\prod_{i=1}^{p}(D - \alpha_i)Y(u) = \varepsilon(u),$$

where $\varepsilon(u) = a(\frac{-b+\sqrt{b^2+4au}}{2a})$. The condition for $X(t)$ to be L-C stationary is that the real part of the $\alpha_i's$ are negative.

The L-C Autoregressive process can be easily extended to the L-C Autoregressive Moving-average process. This process is defined by

$$\prod_{i=1}^{p}(\frac{1}{2at+b}D - \alpha_i)X(t) = \prod_{i=1}^{q}(\frac{1}{2at+b}D - \beta_i)a(t).$$

The condition for the Linear-Chirp Autoregressive Moving-average process to be stationary is that $p > q \geq 0$ and the real part of the $\alpha_i's$ are negative.

Theorem 2.4 *If the data $X(t_k)$ are sampled from an L-C stationary process at the time points t_k, where $t_k = \frac{-b+\sqrt{b^2+4akh}}{2a}$ with k be integers and $h > 0$, then $Z_k = X(t_k)$ is a discrete stationary process. The process Z_k will be referred to as the discrete dual of $X(t)$ with Linear-Chirp sampling interval h.*

Proof: Suppose μ and σ^2 are the mean and variance of the L-C stationary process $X(t)$, and w.l.o.g, we let $\mu = 0$. The three conditions for stationarity of $Z(k)$ can be verified as follows:

(i) $E[Z_k] = E[X(t_k)]] = \mu = 0$

(ii) $Var[Z_k] = Var[X(t_k)] = \sigma^2$

(iii) $R_Z(\tau) = E[Z_k Z_{t+\tau}] = E[X(t_k)X(t_{k+\tau})]$
  Since $t_k = \frac{-b+\sqrt{b^2+4akh}}{2a}$, then k can be expressed as $\frac{(2at_k+b)^2-b^2}{4ah}$. Then it follows that

$$
\begin{aligned}
t_{k+\tau} &= \frac{-b + \sqrt{b^2 + 4a(k+\tau)h}}{2a} \\
&= \frac{-b + \sqrt{b^2 + 4a(\frac{(2at_k+b)^2-b^2}{4ah} + \tau)h}}{2a} \\
&= \frac{-b + \sqrt{(2at_k + b)^2 + 4ah\tau}}{2a}.
\end{aligned}
$$

From the Definition 2.1, $R_Z(\tau) = E[X(t_k)X(\frac{-b+\sqrt{(2at_k+b)^2+4ah\tau}}{2a})] = R_X(h\tau)$, which only depends on τ given h. Therefore, Z_K is a stationary process.

Remark: $X(t_k)$ will be referred to as the discrete L-C stationary process, and Z_k will be called the dual of $X(t_k)$. If $Z(u)$ is the dual of $X(t)$, then Z_k is a sampled series from $Z(u)$ at the equally spaced interval h. This theorem says that if the L-C stationary process is sampled properly, then we obtain a discrete stationary realization. Moreover, if we discretize a pth order continuous L-C autoregressive process with sample rate h such that the highest angular frequency corresponding to the complex roots ($\alpha's$) is less than π/h, then $Z(k) = X(t_k)$ is a uniquely defined $ARMA(p, p-1)$ process, and the parameters ϕ and θ depend on both α and h. This

9

result is based on the result of discretizing a continuous AR(p) process by Phake and Wu (1974).

2.2   Wigner-Ville Time-Frequency Distribution

In this section, a brief introduction of Wigner-Ville Time-Frequency Distribution (WVD) is given. Originally, the Wigner distribution (WD) was derived by Wigner (1932) for the calculation of the quantum correction terms to the Boltzmann formula. Claasen and Mecklenbräuker (1980) developed a comprehensive approach for the application of the WD to joint time-frequency analysis. The WVD is used to analyze an observed discrete realization from an L-C process.

2.2.1   Deterministic Signals

Given a signal $x(t)$, the WD can be expressed as

$$W_x(t, f) = \int_{-\infty}^{\infty} x(t + \frac{\tau}{2}) x^*(t - \frac{\tau}{2}) e^{-i2\pi f\tau} d\tau, \tag{15}$$

where the superscripted asterisk (*) denotes the complex conjugate. The WD maps the one-dimensional signal of time to the two-dimensional function of time and frequency. It can also be thought of as a short-time Fourier transform (STFT) where the windowing function is a time-scaled, time-reversed copy of the original signal. A good discussion of the properties of the WD is given by Hlawatsch and Boudreaux-Bartels (1992). The WVD simply substitutes the analytic signal $z(t)$ for the real signal $x(t)$, i.e., $z(t) = x(t) + iH\{x(t)\}$, where

$$
\begin{aligned}
H\{\ldots\} &= \text{ the Hilbert Transform operator} \\
&= \mathcal{F}^{-1}\{(-i\,\text{sgn}f)\mathcal{F}\{\ldots\}\},
\end{aligned}
$$

where $\mathcal{F}\{\ldots\}$ denotes the FT. In other words, the Hilbert transform of $x(t)$ is calculated as follows:

1 Take the FT $X(f)$ of $x(t)$.

2 Multiply $X(f)$ by $-i$ for positive f, by i for negative f, and by zero for $f = 0$.

3 Take the inverse FT.

2.2.2 Random Processes

The WVD can be applied to realizations from random processes. The magnitude of the WVD at any time-frequency location is then a random variable, and provides information solely about the particular realization of the process. For more meaningful time-frequency analysis of random process, it is necessary to develop representations that characterize the statistical distribution of time-frequency energy or power, and thus act as time-frequency power spectra. The Wigner-Ville spectrum is defined as the expectation of the WVD, i.e.,

$$EW_z(t, f) = \int_{-\infty}^{\infty} R_z(t + \frac{\tau}{2}, t - \frac{\tau}{2})e^{-i2\pi f\tau}d\tau. \tag{16}$$

and it can be seen that it is the FT of the time dependent autocovariance function.

2.3 Model Fitting Procedure

Similar to the M- and $G(\lambda)$-stationary processes, in a real world application, the observed time series is most likely a subset of an L-C stationary process. In other words, the observed data begin at time $\Lambda + 1$, where the time shift, Λ, will be referred to as the origin offset. Unlike the M- and $G(\lambda)$-stationary processes, it is not necessary to estimate the Λ in L-C modeling. By adjusting b in the model, same

11

information will be obtained. Note that the instantaneous frequency at time t of an L-C process with origin offset Λ is $\frac{(2a(t+\Lambda)+b)}{2\tau}\{1 + \sqrt{1 + \frac{4a\tau}{(2a(t+\Lambda)+b)^2}}\}$, which can be written as $\frac{(2at+(2a\Lambda+b))}{2\tau}\{1 + \sqrt{1 + \frac{4a\tau}{(2at+(2a\Lambda+b))^2}}\}$. Letting $b1 = 2a\Lambda + b$, then the L-C process without origin offset will have the same instantaneous frequency as the one with origin offset. Given a discrete equally spaced data which is from a unknown L-C process, the procedure to analyze the data is as follows:

1. Calculate the WVD of the data.

2. Estimate $\mathrm{IF}(t)$ by peak detection of WVD.

3. Since $\mathrm{IF}(t) = 2at + b$, we then calculate $\hat{a}$ and $\hat{b}$ by least-squares.

4. Use $\hat{a}$ and $\hat{b}$ as the initial values to find the optimal a_o and b_o in terms of minimum cross-entropy criterion.

5. Interpolate the data onto the points $t_k = \frac{-b_o + \sqrt{b_o^2 + 4a_o kh}}{2a_o}$, where k and h are calculated from the following set of equations with $\Lambda = 0$.

$$
\begin{cases}
\Lambda + 1 = \frac{-b + \sqrt{b^2 + 4a(k_1+1)h}}{2a} \\
\\
\Lambda + n = \frac{-b + \sqrt{b^2 + 4a(k_1+n)h}}{2a}.
\end{cases}
\tag{17}
$$

6. Fit AR model to the dual.

7. Do the spectral analysis, forecasting and filtering ...

To obtain the dual in this procedure, linear interpolation is used. That is, the dual Y_k is given by

$$
Y_k = X(t_k) = \begin{cases}
X_{t_k}, & \text{if } t_k \text{ is integer} \\
([t_k] + 1 - t_k)X_{[t_k]} + (t_k - [t_k])X_{[t+k]+1}, & \text{otherwise}
\end{cases}
\tag{18}
$$

where $[t_k]$ represents the integer part of t_k.

2.4 Some Simulated L-C Examples

Three simulated data sets are discussed in this section to illustrate the L-C stationary process.

Example 2.1

Consider the L-C process generated from the model

$$X(t) = A\cos(2\pi\beta(a(t+\Lambda)^2 + b(t+\Lambda) + \phi)) + a(t), \qquad (19)$$

where $t = 1, 2, ...200$, A and β are constants, Λ is the origin offset, $\phi \sim \text{Uniform}(0, 2\pi)$ is the phase shift, and $a(t)$ is white noise. Figure 1 shows the realization of this model, where $A = 1$, $\beta = 0.0002$, $a = 1$, $b = 1$, $\phi = 0$. Figure 2 is the WVD of the data, which clearly shows the increasing frequency with time. Figure 3 is the dual $Y(k)$, which is computed by linear interpolation of the realization to the time points $\frac{-b+\sqrt{b^2+4akh}}{2a}$ for k from $k_1 + 1$ to $k_1 + n$, where n is the length of data and k_1 and h are calculated from (17), where $\Lambda + 1$ is the starting point of the realization. It can be seen that the dual has a frequency structure that does not change in time.

Example 2.2

This example is a realization of length 200 from a LC(2) process $X(t)$ with $a = b = 0.00015$ and $\Lambda = 100$. The discrete dual of this process with sample rate $h = 0.0603$ is the $AR(2)$ process

$$(1 - 1.732B + 0.9B^2)Y(k) = \varepsilon(k)$$

13

where $\sigma_\varepsilon^2 = 1$. Figure 4 shows the realization of $X(t)$. The instantaneous ACF is shown in figure 5. Figure 6 is the dual, which is calculated from the true parameters. As noted previous, if the dual is calculated using $a = 0.00015$, $b = 2a\Lambda + 0.00015 = 0.03015$ and $\Lambda = 0$, it will be exactly the same as the dual using the true parameters. Figure 7 shows the WVD and Figure 8 displays the L-C instantaneous spectrum. It can be seen that the L-C instantaneous spectrum has much better power concentration than the WVD, especially at the two edges of the data.

Example 2.3

This example is a realization of length 200 from a LC(4) process $X(t)$ with $a = b = 0.00015$ and $\Lambda = 100$. The discrete dual of this process with sample rate $h = 0.0603$ is an $AR(4)$ process

$$(1 - 1.732B + 0.95B^2)(1 - 0.6B + 0.9B^2)Y(k) = \varepsilon(k)$$

where $\sigma_\varepsilon^2 = 1$. Figure 9 shows the data and Figure 10 displays the dual calculated using the true parameters. The WVD is shown in Figure 11, and the L-C instantaneous spectrum is shown in Figure 12. From the L-C instantaneous spectrum, we can see that there are two dominant frequencies linearly increasing with time clearly. On the other hand, the WVD only displays one frequency changing linearly with time. The frequency associated with lower power is missing in WVD. Furthermore, the WVD has poorer power concentration.

2.5 Forecasting L-C Stationary Process

One of the most important problems in the study of time series is that of "predicting" a future value of a series, given a record of its past values. This problem is clearly

14

of interest in the context of economic systems as well as in the physical systems. When a set of equally spaced data, which is assumed to be from an L-C process, is observed, the model identification procedure discussed previously is used to estimate the parameters a and b. The forecasting is the done on the discrete dual scale. The forecasts on the dual scale are then transferred back to the original time scale by interpolation. The procedure for obtaining the l-step ahead forecasting follows.

1 Fit an L-C stationary process to the discrete observations.

2 Forecast using the discrete dual based on an AR(p) model fit to the dual.

3 Convert the forecasts to the original time scale by interpolation.

2.5.1 Simulation Study

The purpose of this simulation study is to show that the L-C modeling and forecasting procedure given in this thesis produce much better forecast on L-C data than using standard AR modeling. In section 4.8, we will make a similar comparison with $G(\lambda)$ processes. This is not surprising, since the L-C analysis procedure adapts to the time-varying frequency behavior in the data while standard AR methods do not. In the simulations, several realizations are generated from known L-C stationary processes. Then both AR and L-C processes are used to fit the data and produce forecasts. Forecasting performance is compared in terms of the mean of squared errors of the forecasts (MSE). One hundred realizations are generated from each of the following four L-C processes:

(a) LC(2) process with the discrete dual $(1-1.732B+0.98B^2)X_k = \varepsilon_k$, with sample rate $h = 0.0633$ and $var(\varepsilon_k) = 1$. The realization is from $t = 61$ to $t = 360$.

15

(b) LC(2) process with the discrete dual $(1-1.732B+0.98B^2)X_k = \varepsilon_k$, with sample rate $h = 0.0753$ and $var(\varepsilon_k) = 1$. The realization is from $t = 101$ to $t = 400$.

(c) LC(4) process with the discrete dual $(1-2.332B+2.999B^2-2.285B^3+0.960B^4)X_k = \varepsilon_k$, with sample rate $h = 0.0633$ and $var(\varepsilon_k) = 1$. The realization is from $t = 61$ to $t = 360$.

(d) LC(4) process with the discrete dual $(1-2.332B+2.999B^2-2.285B^3+0.960B^4)X_k = \varepsilon_k$, with sample rate $h = 0.0753$ and $var(\varepsilon_k) = 1$. The realization is from $t = 101$ to $t = 400$.

For each realization, the first 80% or 90% of the data will be used for model identification, and the remainder of the data is used to compare forecast performance. Figure 13 and Figure 14 show the results of the simulation. It is clear that the L-C model is much better than the AR model in terms of forecasting.

2.6 Filtering L-C Stationary Processes

Cohlmia et al.(2004) proposed a theoretic framework for filtering M-stationary and $G(\lambda)$-stationary processes. This technique is based on the strategy of applying the filter to the dual data instead of the data on the original time scale. This allows traditional filtering methods, such as the Butterworth filter, to be utilized since the dual data are stationary. After the filters are applied to the dual data, the filtered data is transformed back to the original time scale by interpolation. The same idea can be used to filter L-C stationary processes as well. The procedure is summarized below.

1 Estimate a, b and the best L-C stationary process.

2 Interpolate the data to the dual scale.

3 Filter the dual by an appropriate filter.

4 Interpolate the filtered data back to the original time scale.

Example 2.4

This example is a realization of a LC(4) process $X(t)$. In this case $a = b = 0.00015$, and the data are observed from $t = 101$ to $t = 400$. The discrete dual of this process with sample rate $h = 0.0753$ is an $AR(4)$ process

$$(1 - 1.732B + 0.98B^2)(1 + 0.6B + 0.95B^2)Y(k) = \varepsilon(k).$$

Figure 15 (a) shows the data, and Figure 16 shows the instantaneous spectrum based on the L-C modeling. There are clearly two dominant frequencies changing with time. Suppose the higher frequency is desired to be filtered out. By applying the procedures mentioned above, the filtered data is obtained. Figure 15 (b) shows the filtered data, and Figure 17 displays the instantaneous spectrum of the filtered data. It can be seen that the high frequency has been successfully removed, and only the low frequency component is left.

2.7 L-C Stationary Process and $G(\lambda)$-Stationary Process

In the real world applications, if there is some prior information about the data, and it is certain that the data is from a L-C stationary process, then the procedures described in previous sections can be used to analyze the data with the L-C model. However, sometimes, it is not certain whether the true underline process is L-C stationary or not. In this case the $G(\lambda)$-process can be used to model the data, since

the $G(\lambda)$-process provides a good approximation to the L-C process. The comparison between L-C processes and $G(\lambda)$-processes is made here in terms of both spectral analysis and forecasting performance.

2.7.1 Spectral Analysis

Two examples are studied to compare the spectrum for L-C-processes and $G(\lambda)$-processes. Both data sets are simulated from L-C stationary process. The data sets are modeled using L-C and $G(\lambda)$ processes separately, and the instantaneous spectrum is calculated based on both models. The first example is the data in Example 2.2, and the second example is the data in Example 2.3. The $G(\lambda)$ instantaneous spectrum of the first and the second data sets are shown in Figure 18 and Figure 19, respectively. Comparing these to the L-C modeling, which are in Figure 8 and 12, shows that the instantaneous spectrum based on the $G(\lambda)$ modeling is similar to those obtained from L-C modeling.

2.7.2 Forecasting

Forecast performance is compared here for different model settings. Two simulation studies are conducted, and forecast MSE is calculated based on the L-C model and the $G(\lambda)$ model separately. The data in the first simulation study are generated from the following models:

(a) LC(2) process with the discrete dual $(1 - 1.732B + 0.98B^2)X_k = \varepsilon_k$, where the sample rate $h = 0.0513$ and $var(\varepsilon_k) = 1$. The realization is from $t = 21$ to $t = 320$. The first 90% of the data are used to fit a model, and the last 10% of the data are used to calculate the forecast MSE.

(b) LC(2) process with the discrete dual $(1 - 1.732B + 0.98B^2)X_k = \varepsilon_k$, where the sample rate $h = 0.0633$ and $var(\varepsilon_k) = 1$. The realization is from $t = 61$ to $t = 360$. the first 90% of the data are used to fit a model, and the last 10% of the data are used to calculate the forecast MSE.

(c) LC(2) process with the discrete dual $(1 - 1.732B + 0.98B^2)X_k = \varepsilon_k$, where the sample rate $h = 0.0753$ and $var(\varepsilon_k) = 1$. The realization is from $t = 101$ to $t = 400$. The first 90% of the data are used to fit a model, and the last 10% of the data are used to calculate the forecast MSE.

(d) Same as model (a) except the first 80% of the data are used to fit a model, and the last 20% of the data are used to calculate the forecasting MSE.

(e) Same as model (b) except the first 80% of the data are used to fit a model, and the last 20% of the data are used to calculate the forecast MSE.

(f) Same as model (c) except the first 80% of the data are used to fit a model, and the last 20% of the data are used to calculate the forecast MSE.

The second simulation study is the same as the first study except $var(\varepsilon_k) = 2$. The results are shown in Figure 20 and Figure 21. From the simulations, it can be seen that the forecast MSE of the L-C models and the $G(\lambda)$ models are very close.

3   Quadratic Chirp Stationary Processes

In this section, we will discuss the quadratic chirp stationary process (Q-C), since the frequency is changing approximately quadratically with time. Definitions and some properties are given as well as some simulated examples.

3.1 Definitions and Properties

Definition 3.1 *Let $X(t)$ be a stochastic process defined for $t \in (0, \infty)$ and let $Q(\tau, t) \in (0, \infty)$. Then $X(t)$ is called Q-C stationary process if*

(i) $E[X(t)] = \mu$

(ii) $Var[X(t)] = \sigma^2 < \infty$

(iii) $E[(X(t) - \mu)(X(Q(\tau, t)) - \mu)] = R_X(\tau),$

where $Q(\tau, t) = \{ \frac{A^{\frac{1}{3}}}{6a} - \frac{6ac - 2b^2}{2aA^{\frac{1}{3}}} - \frac{b}{3a} \}$, where $A = 36abc + 108a^2u - 8b^3 + 12\sqrt{3}(4ac^3 - b^2c^2 + 18abcu + 27a^2u^2 - 4b^3u)^{\frac{1}{2}}a$, and $u = at^3 + bt^2 + ct + \tau$ with $a > 0$, and $b, c \geq 0$ are 3 constants. $R_X(\tau)$ is referred to as the Q-C autocovariance function.

Note: we can treat L-C stationary process as a special case of Q-C stationary process with $a = 0$.

Definition 3.2 *Let $X(t)$ be a Q-C stationary process, and let $Y(u) = X(t)$, where $u = g(t) = at^3 + bt^2 + ct$. Then $Y(u)$ will be called the stationary dual of $X(t)$.*

Theorem 3.1 *Stochastic process $X(t)$ is Q-C stationary if and only if it has a stationary dual process $Y(u)$ and $R_X(\tau) = R_Y(\tau)$.*

Proof: Follow the same logic as the corresponding result about the L-C stationary process.

Definition 3.3 *Let $X(t)$ be a Q-C stationary process. Then the Q-C spectrum is defined as the Fourier Transform of the Q-C autocovariance function, i.e.,*

$$G_X(f) = \int_{-\infty}^{\infty} e^{-2\pi f \tau} R_X(\tau) d\tau. \tag{20}$$

Definition 3.4 *Let $X(t)$ be a continuous parameter Q-C stationary process, and let $G_X(f^*)$ be the Q-C spectrum. Then the instantaneous spectrum of $X(t)$ at time t is*

$$S_X(f;t) = G_X(f^*) = G_Y(f^*), \tag{21}$$

where f is the instantaneous frequency, G_Y is the spectrum of the dual, and f^* is the solution of the equation $Q(\frac{1}{f^*}, t) = f$.

Definition 3.5 *Let $\varepsilon(u)$ be continuous white noise, and let $a(t) = \varepsilon(at^3 + bt^2 + ct)$ for $t > 0$ and $a > 0$, $b, c \geq 0$. Then $a(t)$ will be referred to as continuous Q-C white noise.*

A special case of a Q-C process is the pth order Quadratic-Chirp process.

Definition 3.6 *A process $X(t)$ for $t > 0$ defined by*

$$\prod_{i=1}^{p} (\frac{1}{3at^2 + 2bt + c} D - \alpha_i) X(t) = a(t), \tag{22}$$

where D is the differential operator, $a > 0$, $b, c \geq 0$, α_i are constants, $p > 0$ and $a(t)$ is Q-C white noise, is called a continuous parameter p^{th} order Q-C process (QC(p)).

The dual $Y(u) = X(t)$, where $u = at^3 + bt^2 + ct$, satisfies

$$\prod_{i=1}^{p} (D - \alpha_i) Y(u) = \varepsilon(u),$$

where $\varepsilon(u) = a\left(\frac{A^{\frac{1}{3}}}{6a} - \frac{6ac - 2b^2}{2aA^{\frac{1}{3}}} - \frac{b}{3a}\right)$, where $A = 36abc + 108a^2u - 8b^3 + 12\sqrt{3}(4ac^3 - b^2c^2 + 18abcu + 27a^2u^2 - 4b^3u)^{\frac{1}{2}}a$. The condition for $X(t)$ to be Q-C stationary is that the real part of the $\alpha_i's$ are negative.

It is easy to extend QC(p) processes to QC(p,q) processes. This process is defined by

$$\prod_{i=1}^{p} (\frac{1}{3at^2 + 2bt + c} D - \alpha_i) X(t) = \prod_{i=1}^{q} (\frac{1}{3at^2 + 2bt + c} D - \beta_i) a(t).$$

The condition for the QC(p,q) process to be stationary is that $p > q \geq 0$ and that the real part of the $\alpha_i's$ are negative.

Similar to the L-C stationary process, if the data are sampled properly from a Q-C stationary process, then a discrete weakly stationary dual process is obtained. The following theorem gives the sampling scheme.

Theorem 3.2 *If the data $X(t_k)$ is sampled from Quadratic-Chirp stationary process at the time points t_k, where $t_k = \frac{B^{\frac{1}{3}}}{6a} - \frac{6ac - 2b^2}{2aB^{\frac{1}{3}}} - \frac{b}{3a}$, where $B = 36abc + 108a^2kh - 8b^3 + 12\sqrt{3}(4ac^3 - b^2c^2 + 18abckh + 27a^2k^2h^2 - 4b^3kh)^{\frac{1}{2}}a$ with k an integer and $h > 0$, then $Z_k = X(t_k)$ is a discrete stationary process. Z_k will be referred to as the discrete dual of $X(t)$ with Quadratic-Chirp sampling interval h.*

Proof: Follow the same logic as the corresponding result about the L-C stationary process.

3.2 Examples

Two simulated data sets are given to illustrate the Q-C stationary process

Example 3.1

This example is a realization of length 400 from a QC(2) process $X(t)$, where $a = 0.0000015$, $b = 0.0001$ and $c = 0.1$. The realization is for $t = 101$ to $t = 500$. The discrete dual of this process with sample rate $h = 0.6261515$ is an $AR(2)$ process

$$(1 - 1.732B + 0.86B^2)Y(k) = \varepsilon(k),$$

where $var(\varepsilon(k)) = 1$. Figure 22 shows the data, and the WVD is given in Figure 23. It can be seen that the data is contracting, and the frequency is changing approximately

quadratically in time. Figure 24 shows the dual process. The dual is calculated using linear interpolation of the observations to the time point t_k in Theorem 3.2 for k from $k_1 + 1$ to $k_1 + 400$, where k_1 and h are calculated from

$$\begin{cases} t_{k_1+1} = 101 \\ t_{k_1+400} = 500 \end{cases} \tag{23}$$

Figure 25 shows the instantaneous spectrum based on Q-C modeling. It clearly shows that the dominant frequency is increasing approximately quadratically in time. It can be seen that the Q-C instantaneous spectrum has much better power concentration than the WVD.

Example 3.2

This example is a realization of length 400 from a QC(4) process $X(t)$, where $a = 0.00001$, $b = 0.001$ and $c = 0.5$. The realization is for $t = 101$ to $t = 500$. The discrete dual of this process with sample rate $h = 3.10701$ is an $AR(4)$ process:

$$(1 - 1.732B + 0.98B^2)(1 - 0.6B + 0.9B^2)Y(k) = \varepsilon(k),$$

where $var(\varepsilon(k)) = 1$. Figure 26 shows the data. The WVD is given in Figure 27. It can be seen that the data are contracting and the frequency is increasing approximately quadratically in time. Figure 28 shows the dual process. Figure 29 displays the instantaneous spectrum based on Q-C modeling. There are two dominant frequencies changing quadratically. It can be seen that Q-C instantaneous spectrum not only has better power concentration than the WVD, but also catch the second dominant frequency changing with time where the WVD missed.

4 Real Data Examples

Two real data sets are modelled using the L-C model in this chapter. Both of the data sets exhibit approximately linearly increasing frequency.

4.1 MNTA Data

The data to be analyzed is the MNTA data shown in Figure 3.11(a). We have analyzed this data using $G(\lambda)$ model in chapter 3. The WVD is given in Figure 30. The frequency changes approximately linearly, therefore the L-C process can be used to fit the data. By selecting $a = 0.00002722745$ and $b = 0.0149129$ and sampling rate is $h = 0.03399935$, the dual is calculated, which is shown in Figure 31. An AR(6) model is picked by the AIC criteria to fit the dual data, and the model is

$$(1 - 1.646B + 0.514B^2 + 0.088B^3 + 0.068B^4 + 0.071B^5 + 0.08B^6)(Y_k - 126.3827) = a_k,$$

where the white noise variance is 1705.54. The $G(\lambda)$ process with $\lambda = 2$ and origin offset $\Lambda = 166$ is also used to model the data. The comparison of the ACF plots of the MNTA data and the duals are given in Figure 32. L-C modeling seems to have made the data somewhate more stationary since the ACF of the first and second half of the dual are almost the same, which indicates that the frequency is not changing in time. Therefore, from the ACF point of view, the L-C process is somewhat better than the $G(\lambda)$ process. The instantaneous ACF based on L-C modeling is plotted in Figure 33. The instantaneous spectrum of L-C and $G(\lambda)$ models are in Figures 34 and 35, respectively, where it can be seen that they are quite similar. Both of them show the frequency is increasing linearly with time and have better power concentration than the WVD, especially for $t < 200$ and $t > 600$, where the WVD does not show the dominant frequency. The forecast performance of AR modeling and L-C modeling is

compared for defferent lags by the forecasting MSE, and results are given in Table 1. It can be seen that the L-C model has better forecasting for all the cases except when lags is 10. In Table 3.3, we have shown the forecasting performance of $G(\lambda)$ modeling. We can see that the $G(2)$ model is better that the L-C model in terms of forecasting. The improvement is calculated by

$$\frac{\text{MSE of AR Model} - \text{MSE of L-C Model}}{\text{MSE of AR Model}} \times 100\%.$$

Table 1: Forecast performance: AR Model and L-C Model for MNTA Data.

Lags	AR	LC	Improvement
10	2657.2	52490.7	-1875.4%
20	440049.3	235495.5	46.5%
30	985058.0	628160.8	36.2%
40	640151.6	392681.6	38.7%
50	1088099.4	852334.8	21.7%
60	1326102.4	978641.1	26.2%

4.2 Sperm Whale Click Data

This is an animal sonar signal called whale click data. The sound signal is produced by a sperm whale used for communication and echolocation.

The data in Figure 36 is part of the click. The WVD of the data is calculated in Figure 37. It is pretty clear that the data have decreasing period (i.e. increasing frequency), and the frequency is increasing approximately linearly. By using the L-C modeling procedure, the parameter estimates of the L-C model are found to be

Table 2: Forecast performance: AR Model, L-C Model and $G(\lambda)$ Model for Whale Click Data.

Lags	AR	LC(Improvement)	$G(\lambda)$(Improvement)
10	0.0001042	0.0000742(28.8%)	0.00008332(20.0%)
20	0.0001177	0.0000611(48.1%)	0.00003518(70.1%)
30	0.0001225	0.0000627(48.8%)	0.00003446(71.9%)
40	0.0001209	0.0001077(10.9%)	0.00005535(54.2%)

$a = 0.0006116748$, $b = 0$ and $h = 0.09847964$. The dual process is plotted in Figure 38. An AR(6) model is chosen to fit the dual using the AIC criteria, and the model is given by

$$(1 - 0.878B + 0.499B^2 - 0.132B^3 + 0.327B^4 - 0.028B^5 + 0.243B^6)(Y_k + 0.000054) = a_k,$$

where the variance of white noise is 0.00001098766. The instantaneous ACF and spectrum based on L-C modeling are shown in Figures 39 and 40, which characterize the data's frequency changing with time property very well. Comparing it with the WVD in Figure 37, we can see that the L-C instantaneous spectrum clearly shows the dominant frequency, while the WVD does not do good at the beginning and the end of the data. The forecast performance of AR modeling and L-C modeling is compared for defferent lags by the forecasting MSE, and the results are given in Table 2. The AIC criterion determines the order of the AR models. For all cases, the L-C model has better performance than the AR model. The forecast by $G(\lambda)$ modeling is given in Table 2 for comparison. It can be seen that the $G(\lambda)$ process is better than the L-C process in forecasting.

References

[1] Bartlett, M. S. (1946), "On the Theoretical Specification and Sampling Properties of Autocorrelated Time Series", *Journal of the Royal Statistical Society*, B8, 27.

[2] Boashash, B.(1992), "Estimating and Interpreting the Instantaneous Frequency of a Signal-Part; Part I: Fundamentals, II: Algorithms", *Proceedings of the IEEE*, 80, 4, 519-569.

[3] Choi, E., Gray, H. L. and Woodward, W. A. (2003), "Euler(p, q) Processes with Spectral Analysis of Non-stationary Time Series", *SMU Tech Report*, #313.

[4] Gray, H. L., and Zhang, N. (1988), "On a Class of Nonstationary Processes", *Journal of Time Series Analysis*, 9,2, 133-154.

[5] Gray, H. L., Woodward, W. A. and Vijverberg, C. P. (2003), "Time Deformation And the M-spectrum", *Submitted for publication*.

[6] Gray, H. L. and Vijverburg, C. P.(2002), "Time Deformation and The M-spectrum—A New Tool for Cyclical Analysis", *SMU TECH Report*, #314.

[7] Gray, H. L., Vijverburg, C. P., and Woodward, W. A.(2002), "Nonstationary Data Analysis by Time Deformation", *Submitted for publication*.

[8] Phadke, M. S. and Wu, S. M. (1974), "Modeling of Continuous Stochastic Processes from Discrete Observations with Application to Sunspots Data", *Journal of the American Statistical Association*, 69, 325-329.

[9] Priestley, M. B. (1981), Spectral Analysis and Time Series, Vols. 1 and 2, Academic Press, New York.

[10] Stock, J. (1988), "Estimating Continuous — Time Processes Subject to Time Deformation: An Application to Postwar U.S. GNP", *Journal of the American Statistical Association*, 83, 401, 77-85.

A   Figures
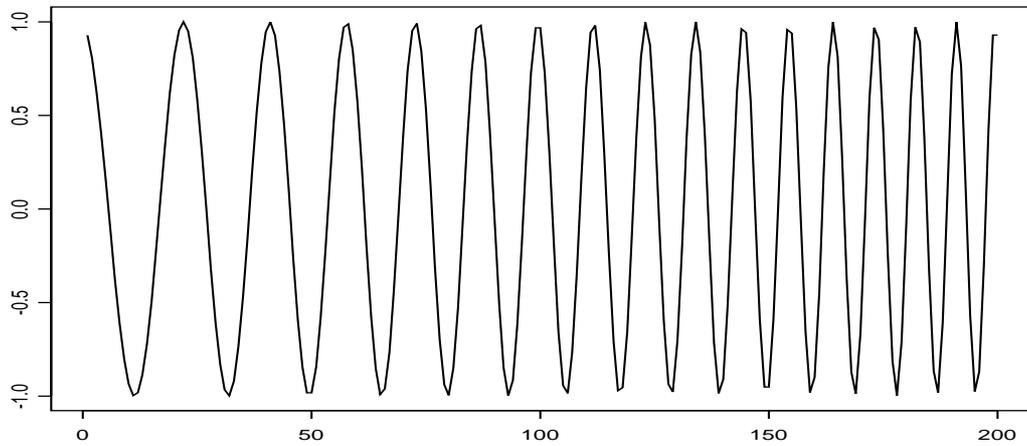
Figure 1: Realization of the Model in Example 2.1: $A = 1$, $\beta = 0.0002$, $a = 1$, $b = 1$, $\phi = 0$, $n = 200$ and $t_1 = 101$.
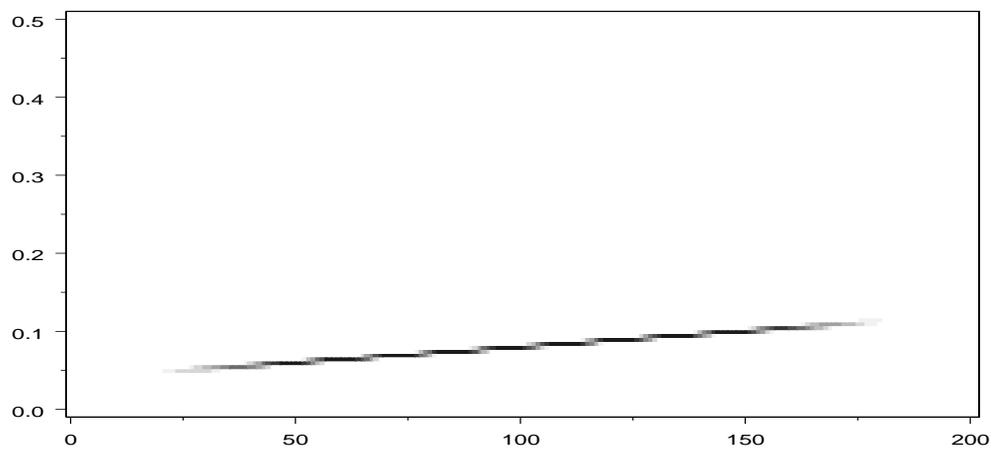
Figure 2: The WVD of the Data in Figure 1.
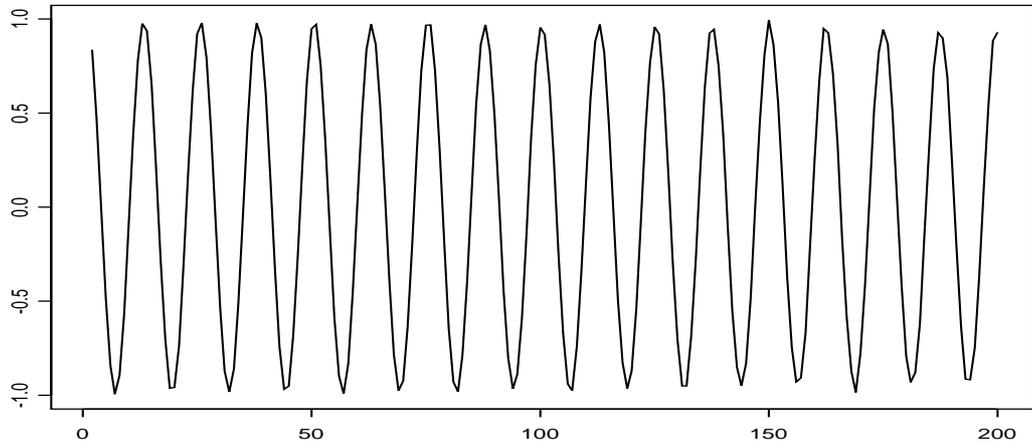
Figure 3: The Dual of Data in Figure 1.
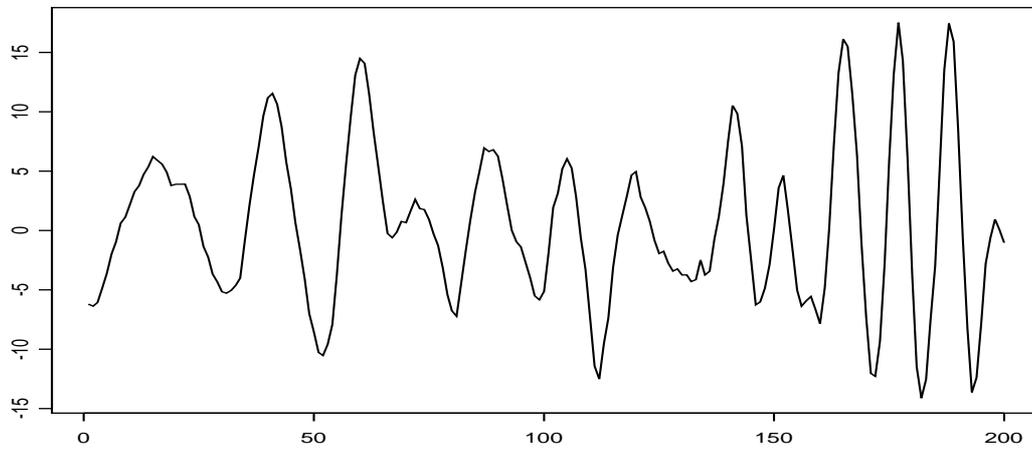
Figure 4: A Realization of length 200 in Example 2.2.

31
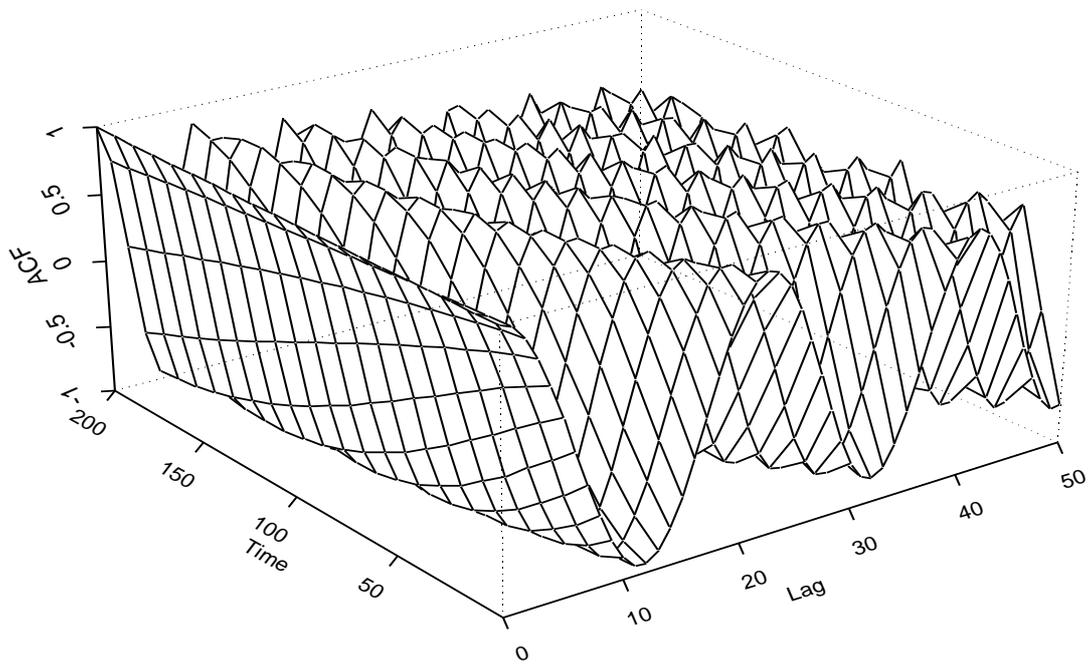
Figure 5: The Instantaneous ACF of Data in Example 2.2.
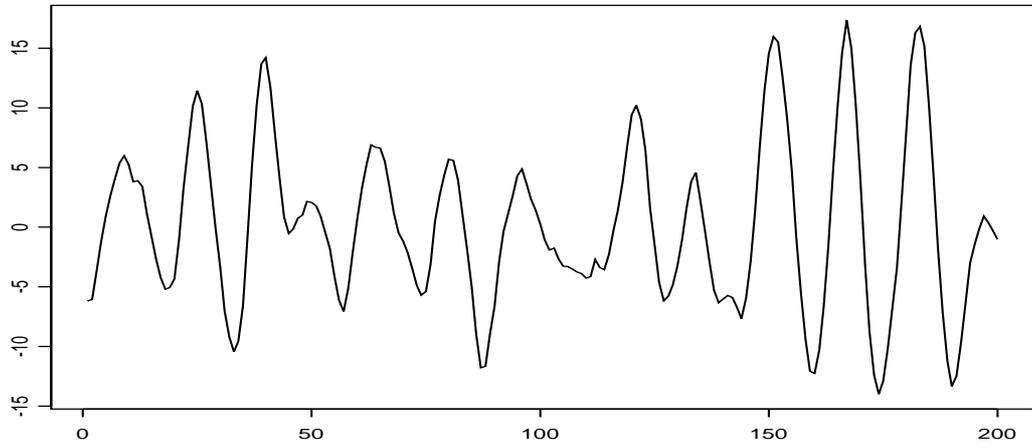
Figure 6: Dual of Data in Example 2.2.
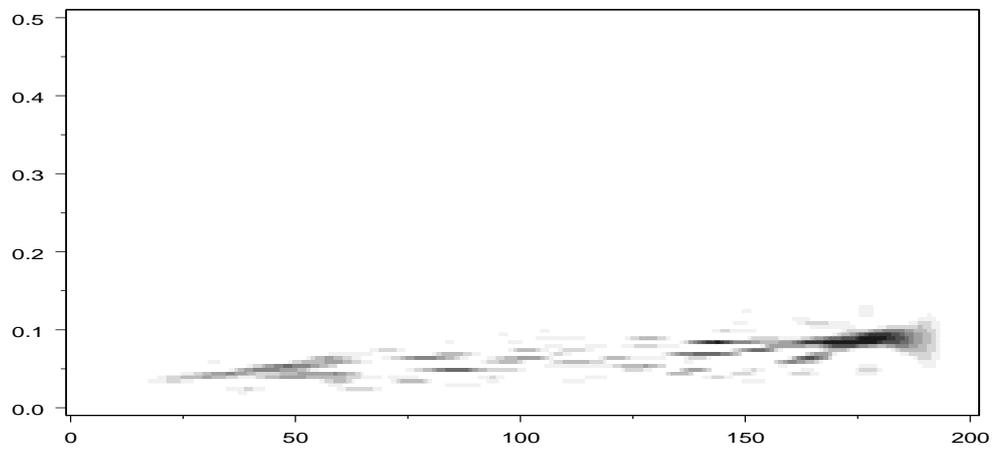
Figure 7: The WVD of Data in Example 2.2.
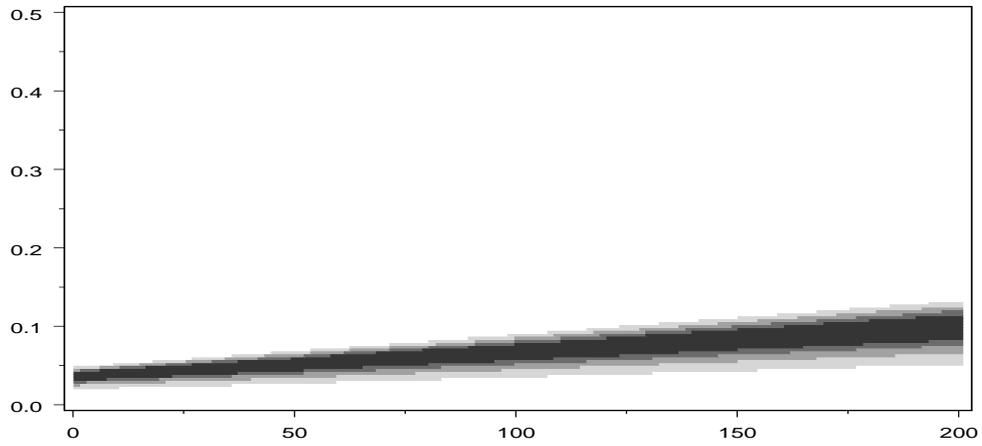
Figure 8: Instantaneous Spectrum of Data in Example 2.2 Using the True Parameters: $a = b = 0.00015$ and $\Lambda = 100$
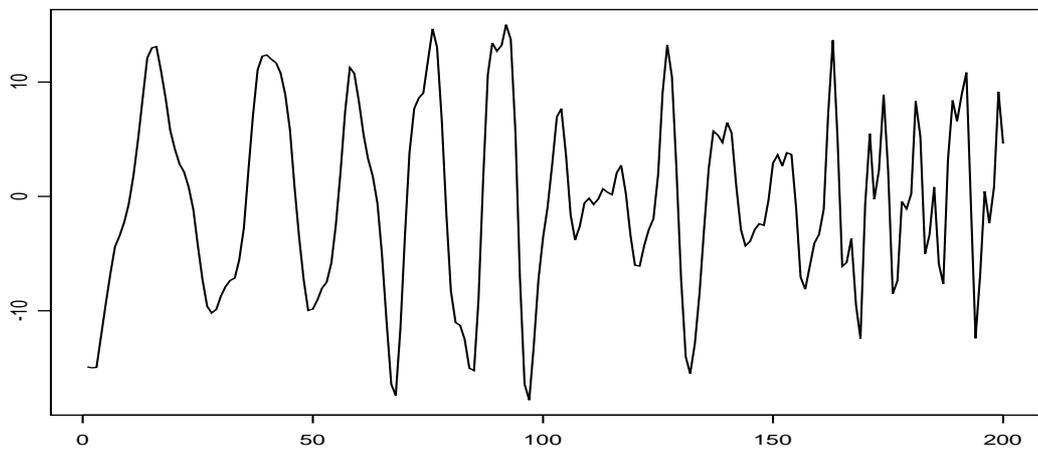
Figure 9: A Realization of Length 300 in Example 2.3.

34
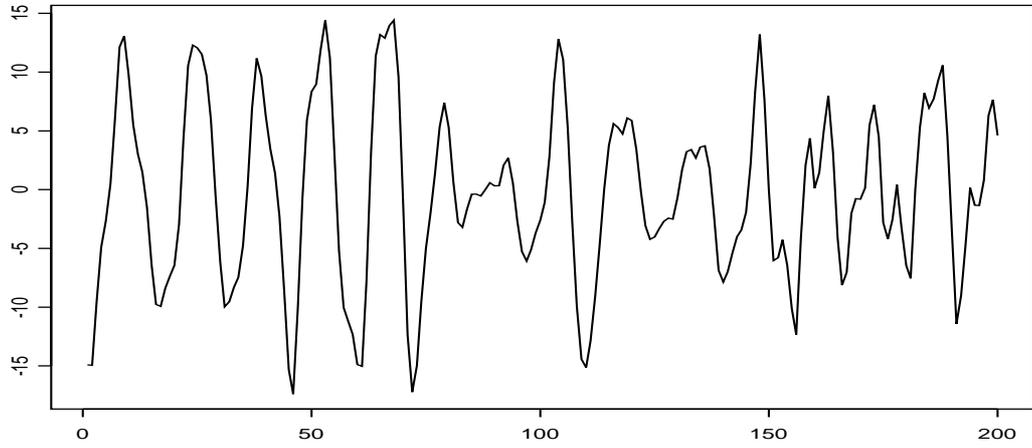
Figure 10: Dual of Data in Example 2.3.
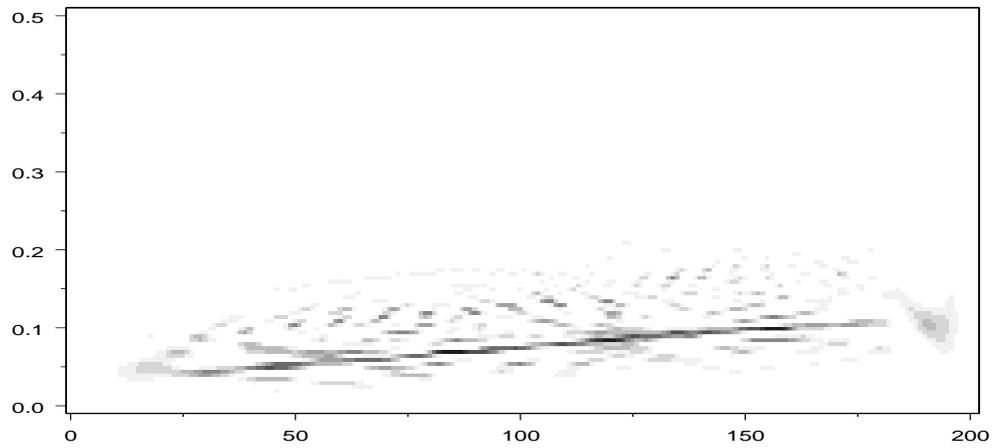
Figure 11: The WVD of the Data in Example 2.3.
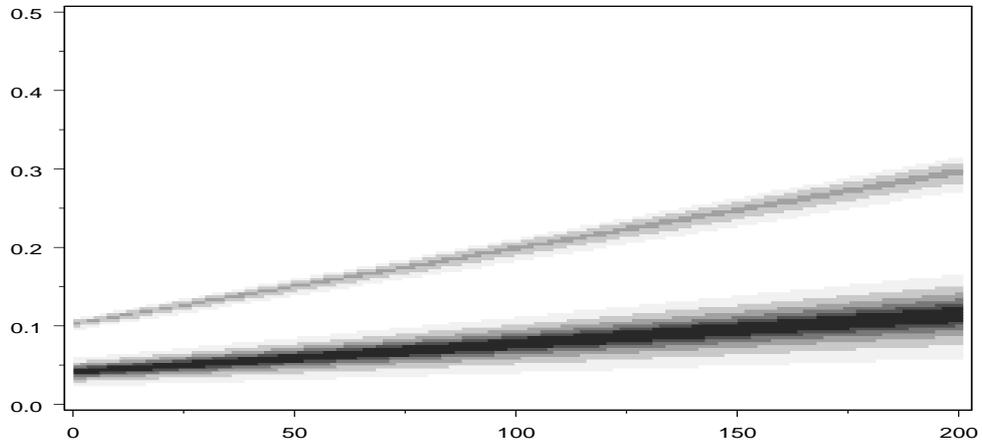
Figure 12: Instantaneous Spectrum of Data in Example 2.3.
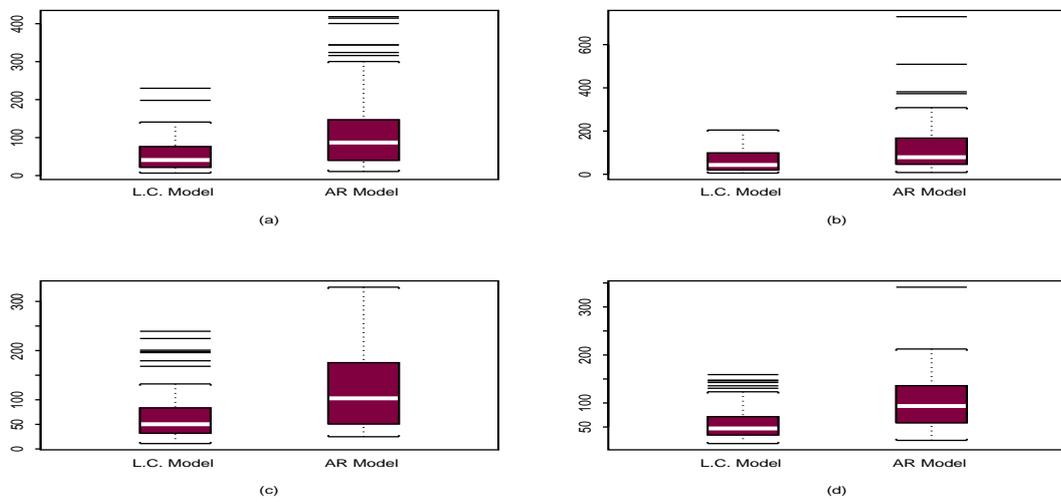
Figure 13: Comparison of Forecast Performance of the L-C Model and AR Model for 10% Validation of 4 Different Models
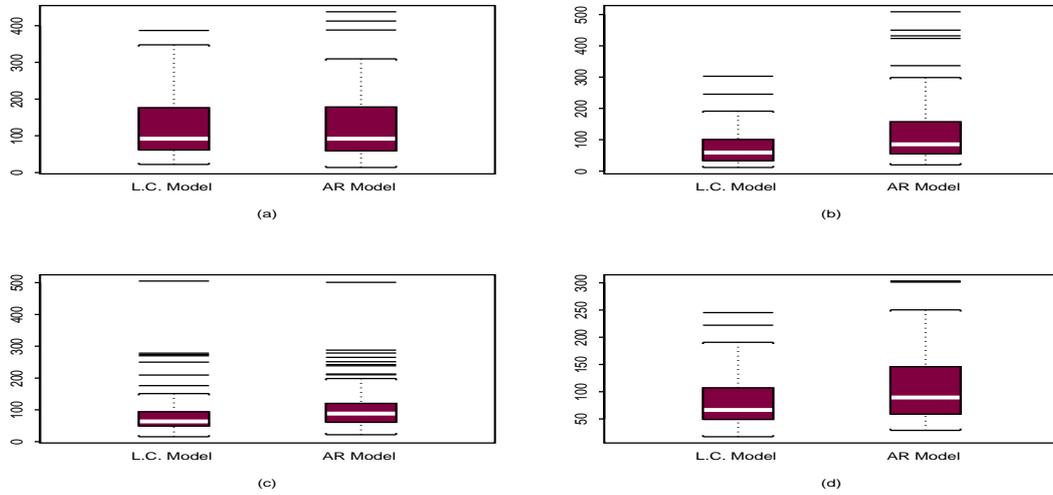
Figure 14: Comparison of Forecast Performance of the L-C Model and AR Model for 20% Validation of 4 Different Models
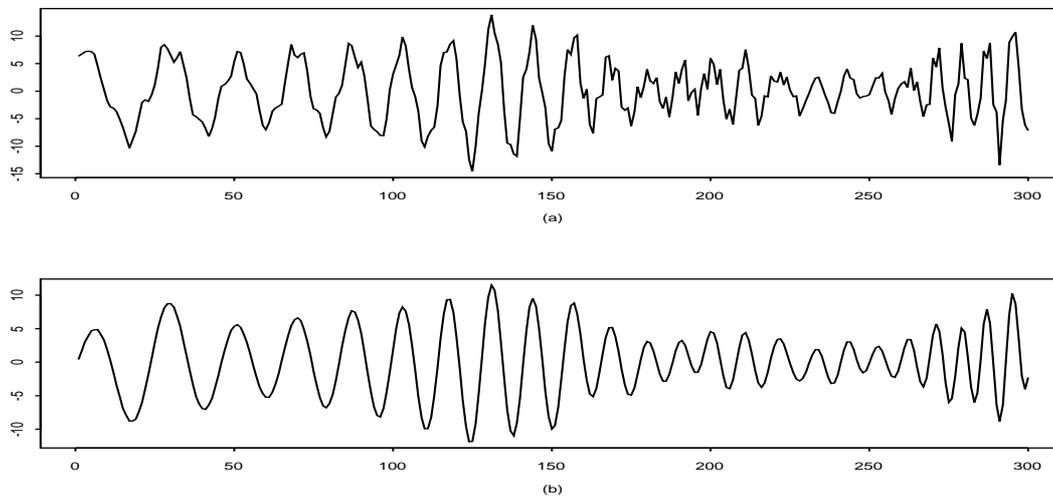
Figure 15: (a) The Original and (b) Filtered Data in Example 2.4

37
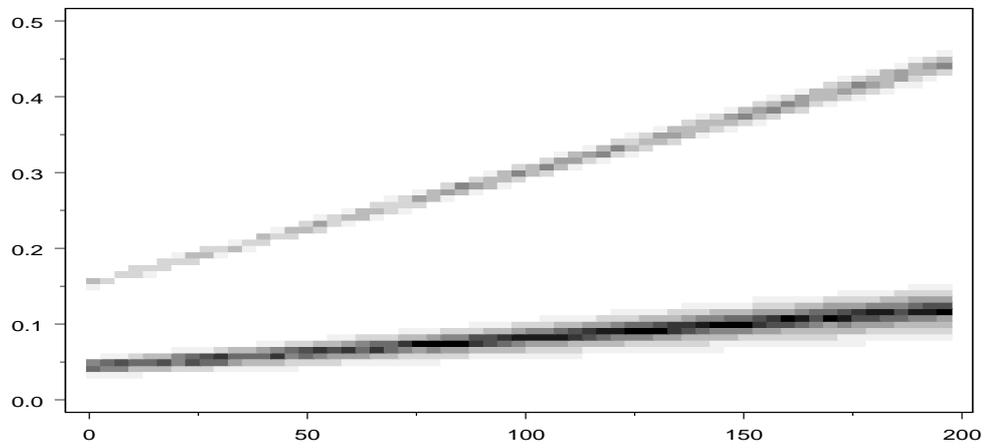
Figure 16: Instantaneous Spectrum of Data in Example 2.4
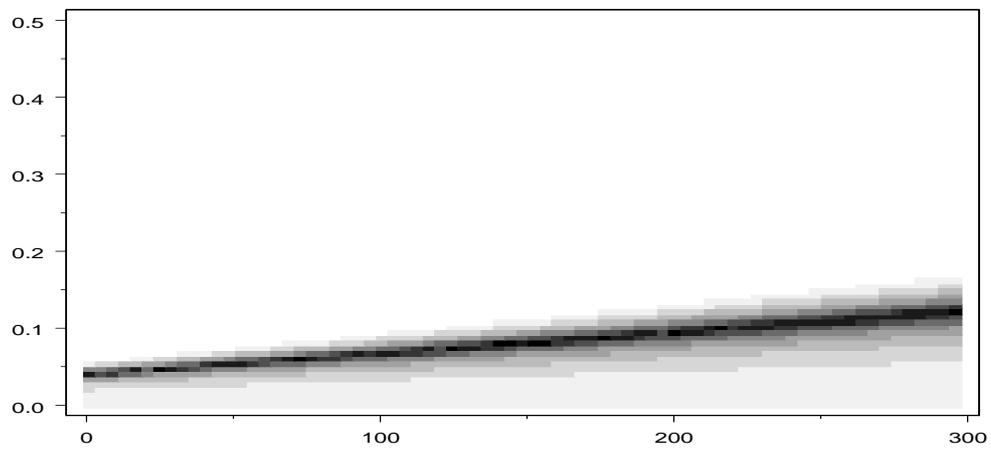
Figure 17: Instantaneous Spectrum of the Filtered Data in Example 2.4

38
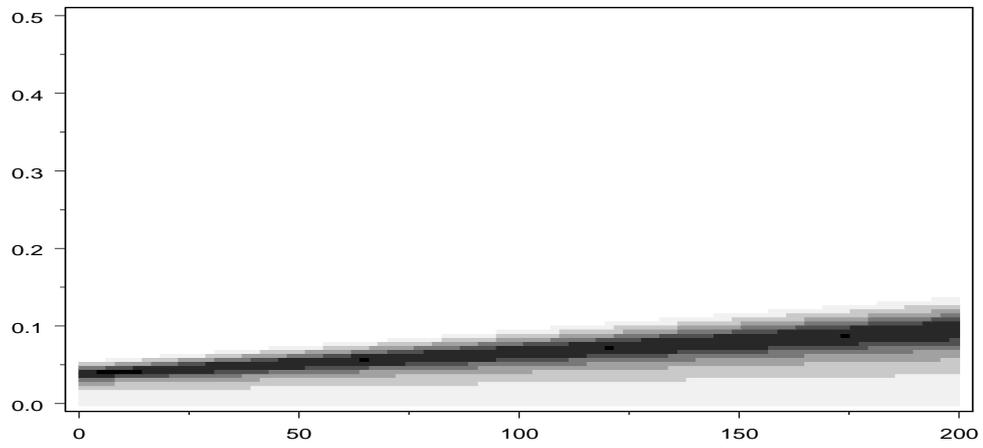
Figure 18: The $G(\lambda)$ Instantaneous Spectrum of the Data in Example 2.2.
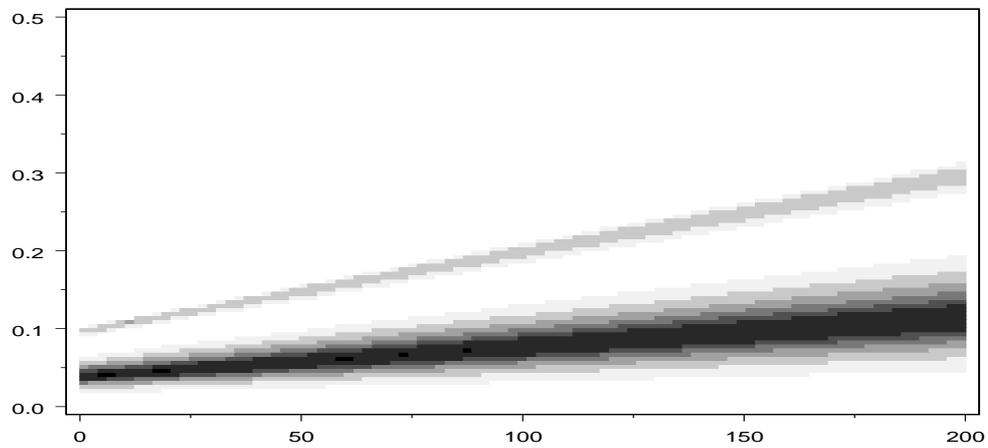
Figure 19: The $G(\lambda)$ Instantaneous Spectrum of the Data in Example 2.3.

39
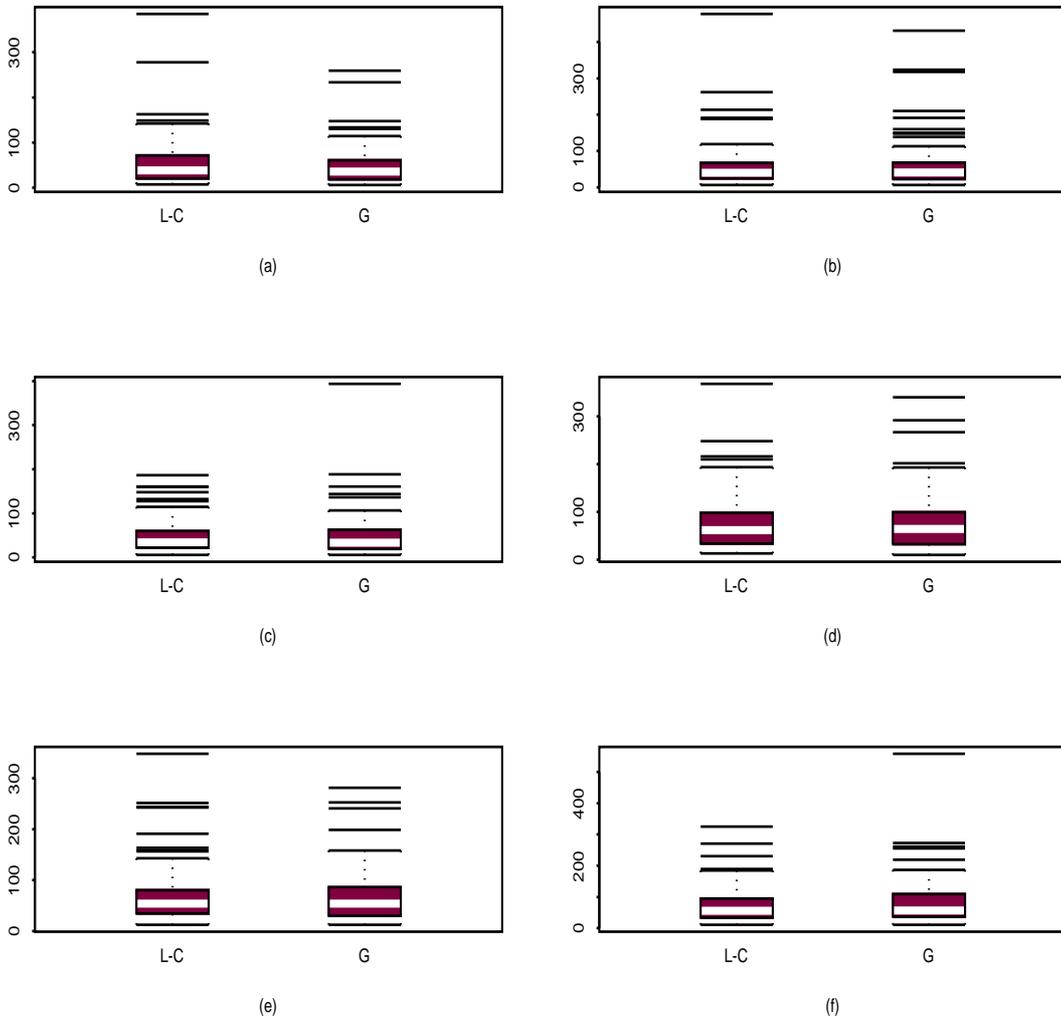
Figure 20: Comparison of Forecast MSE of the L-C Process and $G(\lambda)$ Process in the simulation study 1.
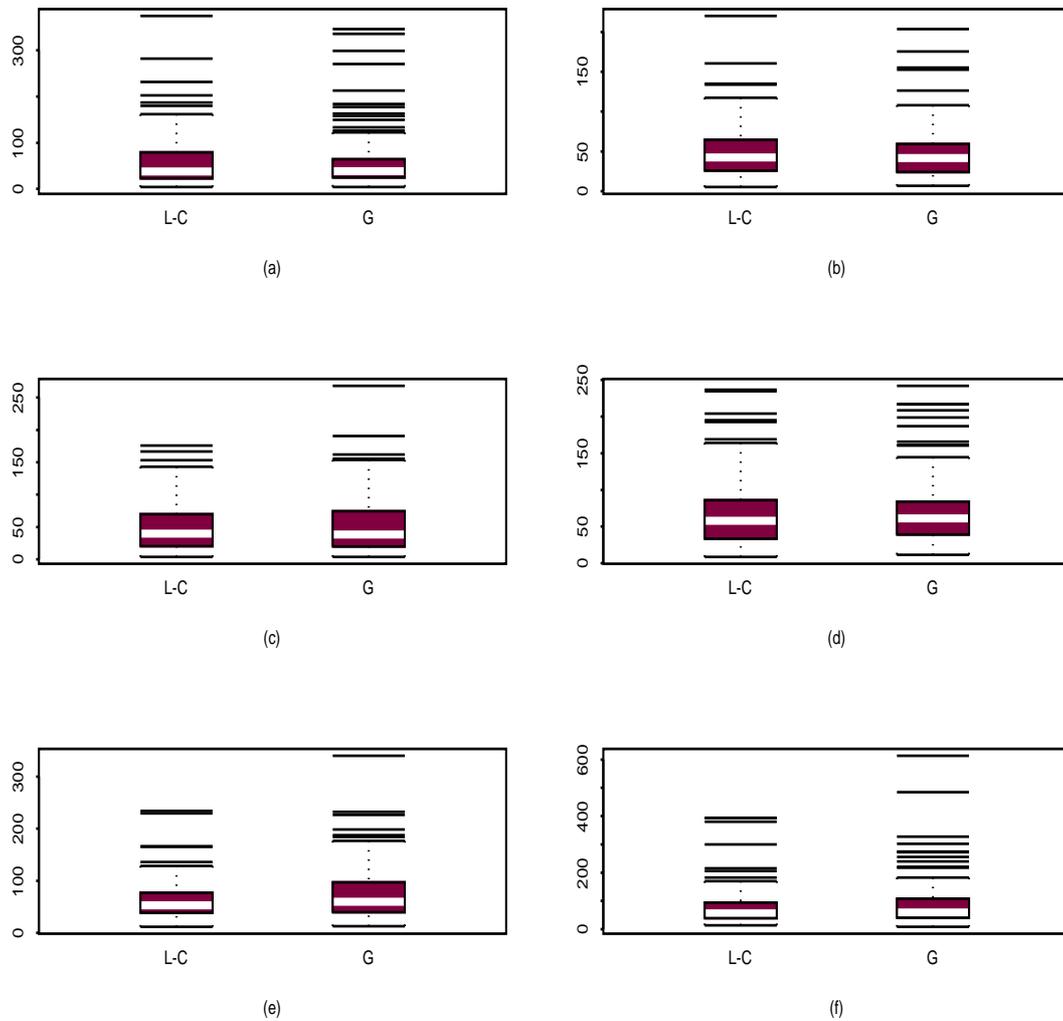
Figure 21: Comparison of Forecast MSE of L-C Process and $G(\lambda)$ Process in the simulation study 2.
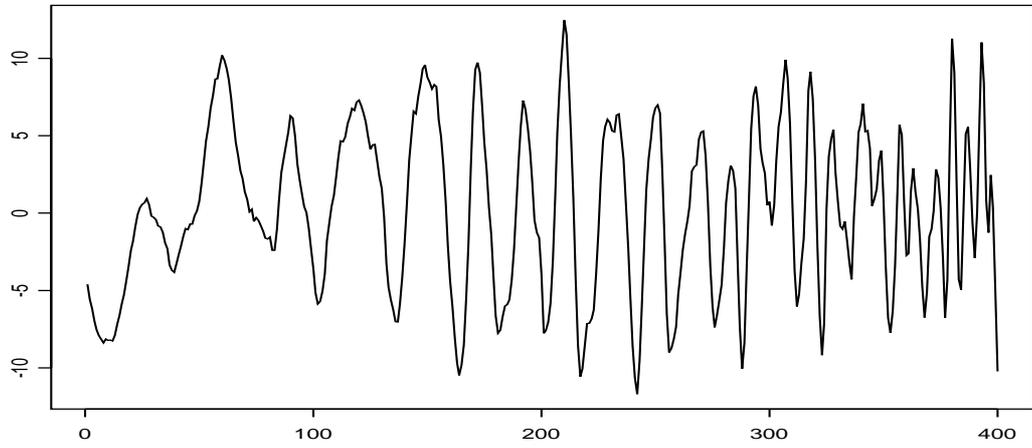
Figure 22: A Realization of Length 400 in Example 3.1.
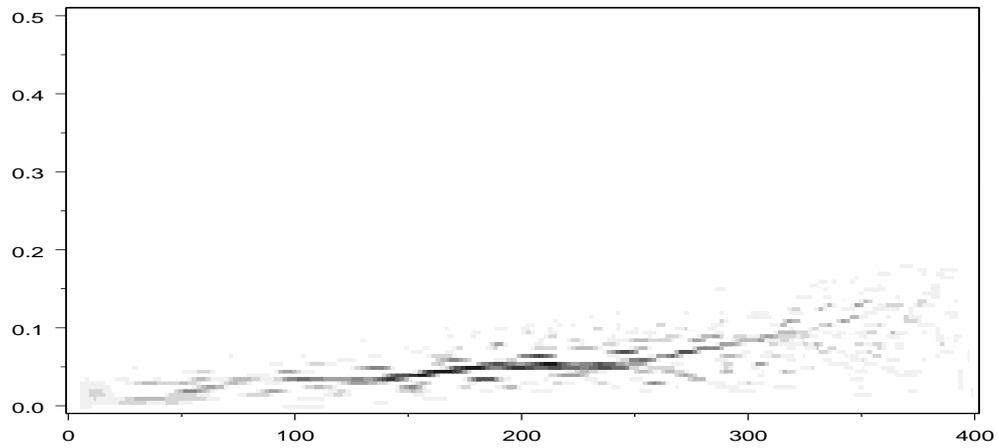
Figure 23: The WVD of the Data in Example 3.1.

42
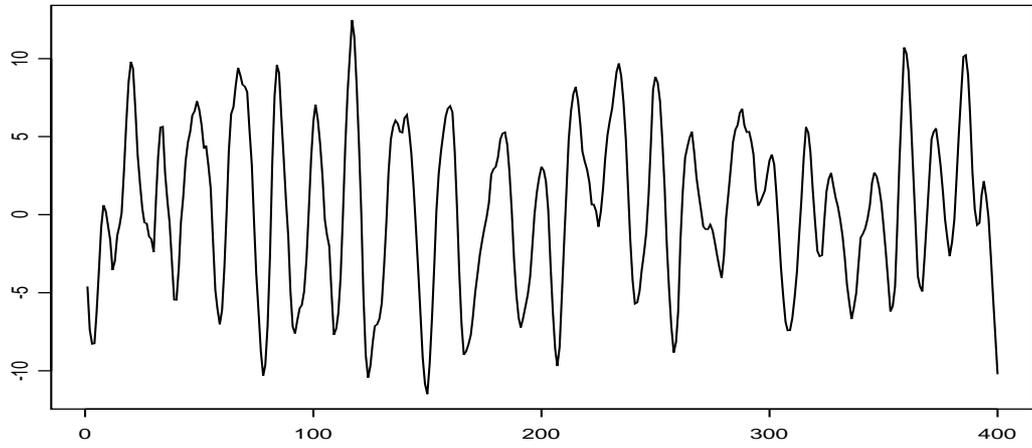
Figure 24: The Dual of the Data in Example 3.1.
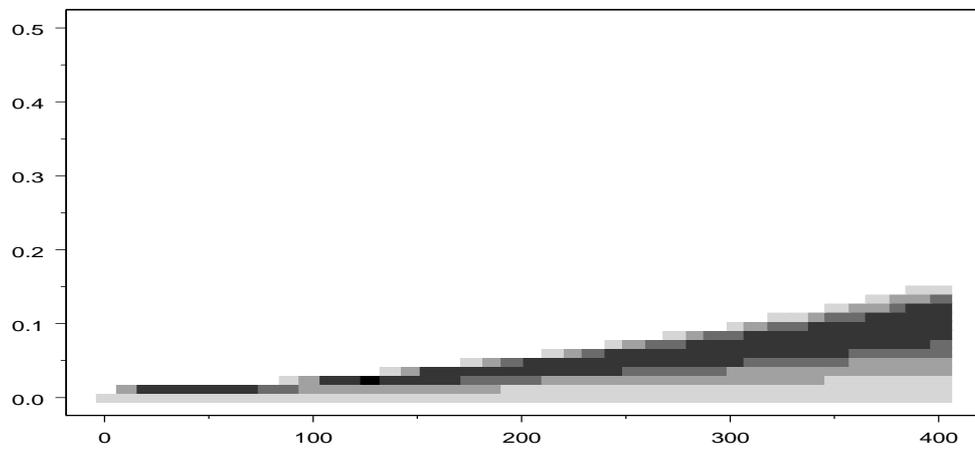
Figure 25: The Instantaneous Spectrum of the Data in Example 3.1.

43
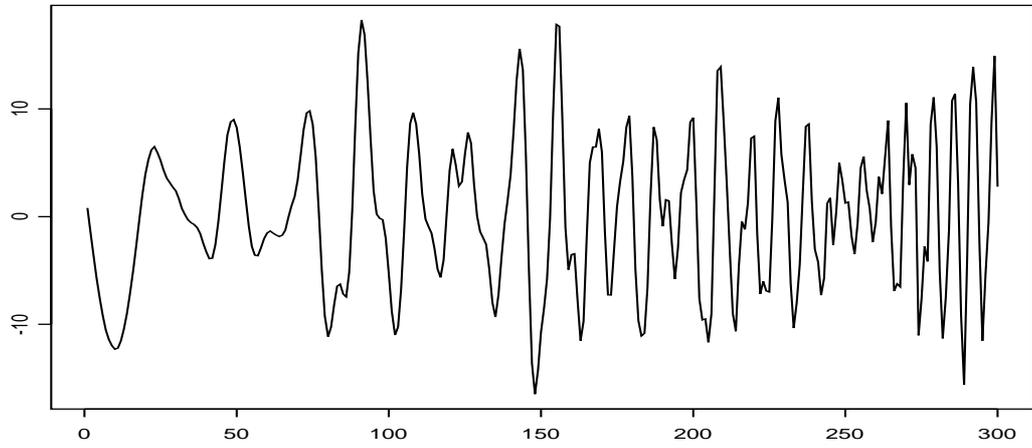
Figure 26: A Realization of Length 400 in Example 3.2.
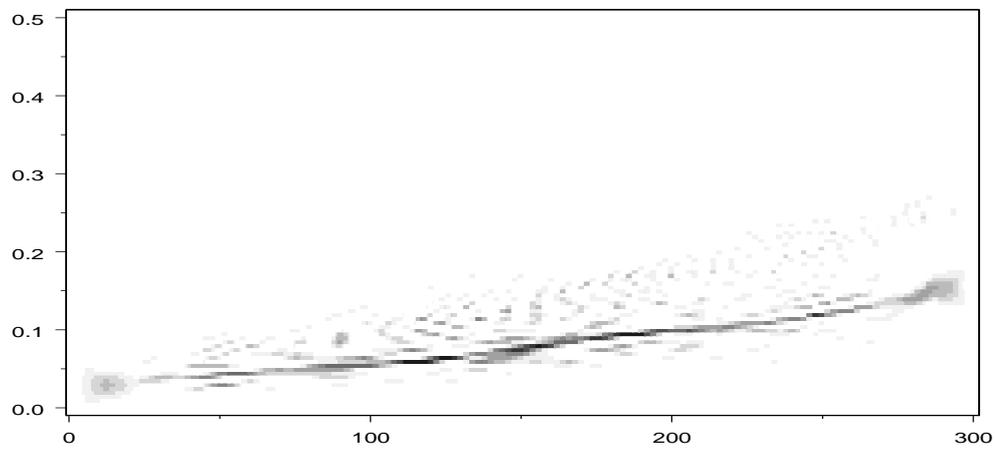
Figure 27: The WVD of the Data in Example 3.2.

44
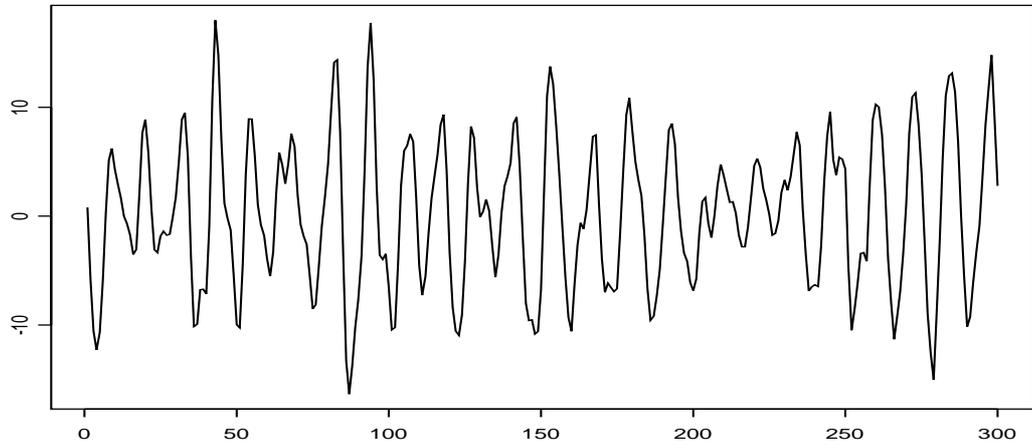
Figure 28: The Dual of the Data in Example 3.2.
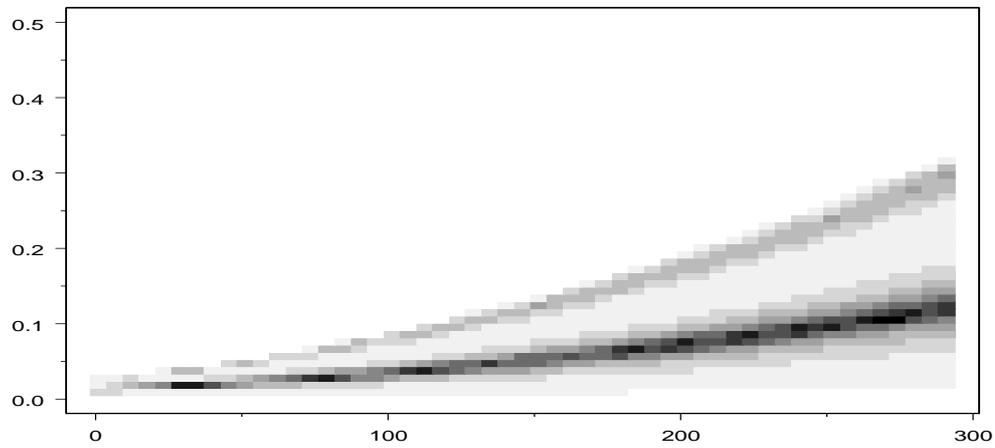
Figure 29: The Instantaneous Spectrum of the Data in Example 3.2.

45
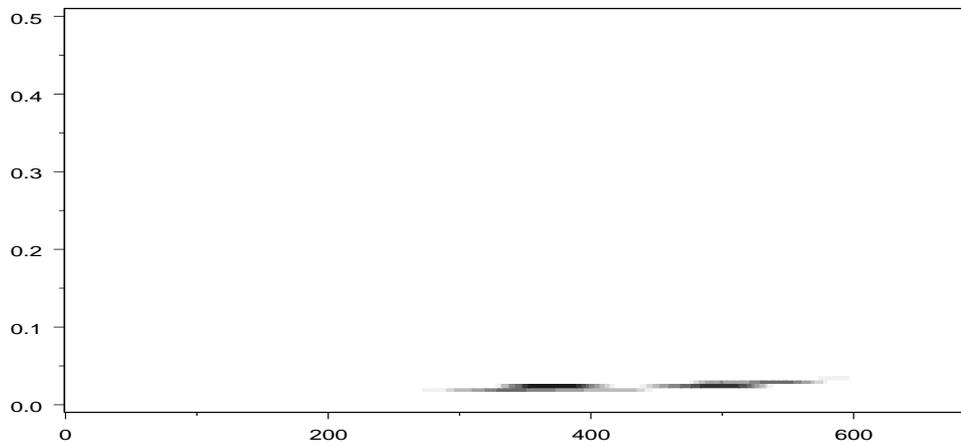
Figure 30: The WVD of MNTA Data.
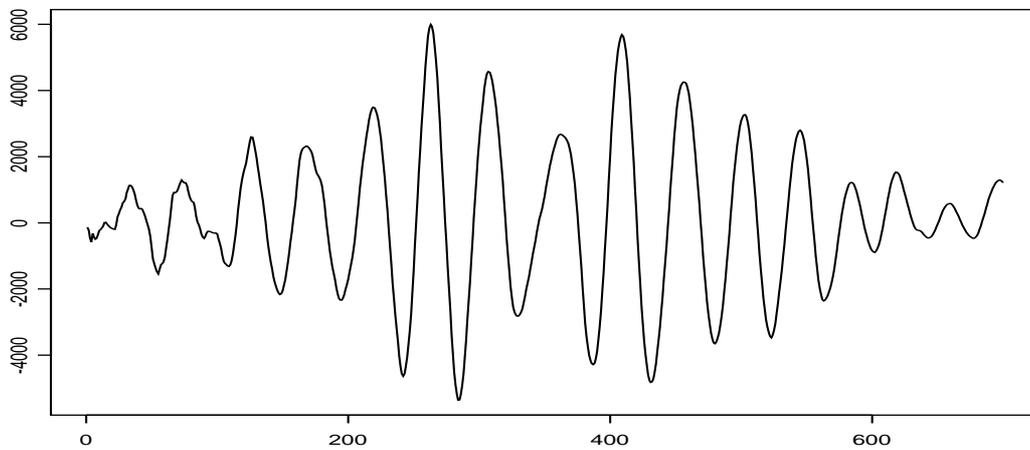
Figure 31: The Dual of MNTA Data Based on L-C Modeling.

46
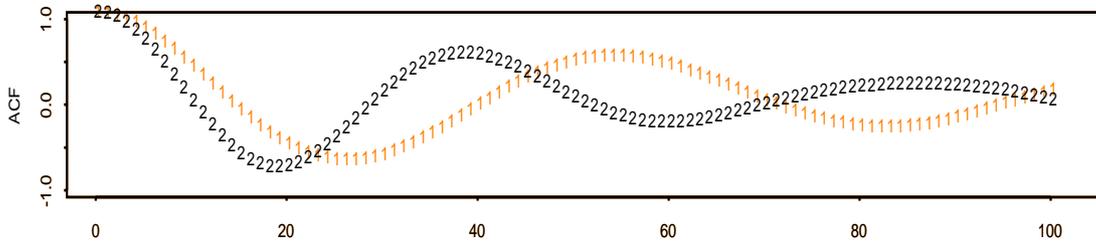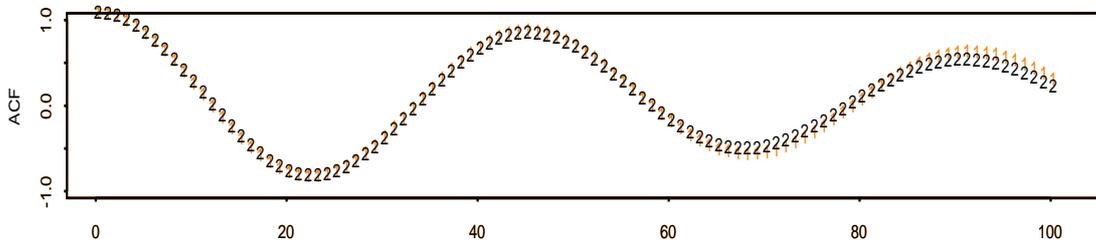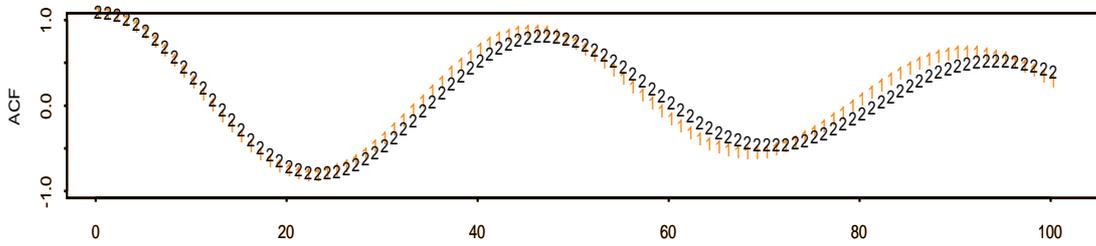
Figure 32: The Comparison of the ACF of (a) MNTA Data; (b) The Dual from L-C Modeling and (c) The Dual from $G(2)$ Modeling. Note '1' is the ACF of 1st half, and '2' is the ACF of the 2nd half.
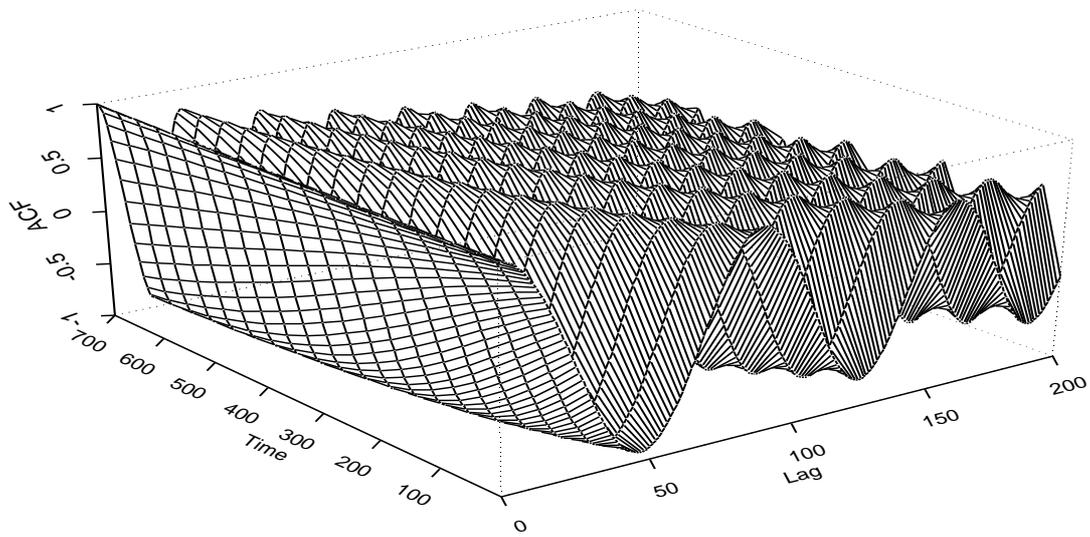
Figure 33: The L-C Instantaneus ACF of MNTA Data.
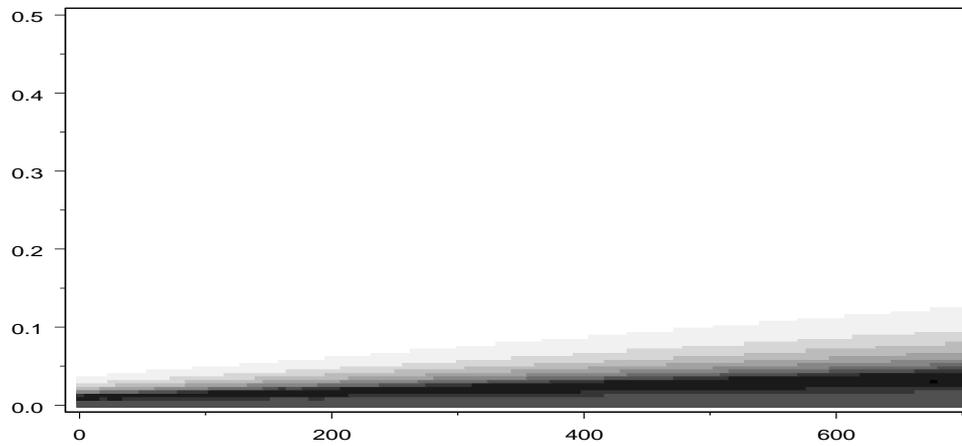
Figure 34: The L-C Instantaneus Spectrum of MNTA Data.

Figure 35: The $G(\lambda)$ Instantaneous Spectrum of MNTA Data.
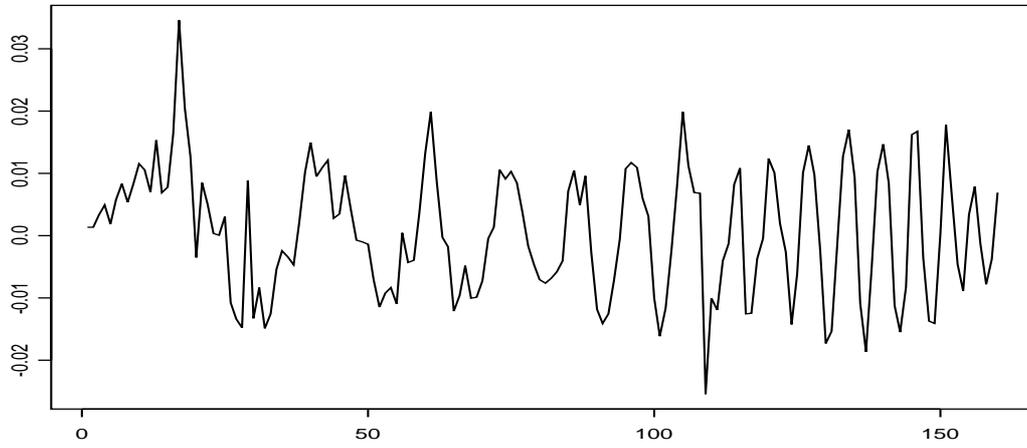
Figure 36: Whale Click Data.
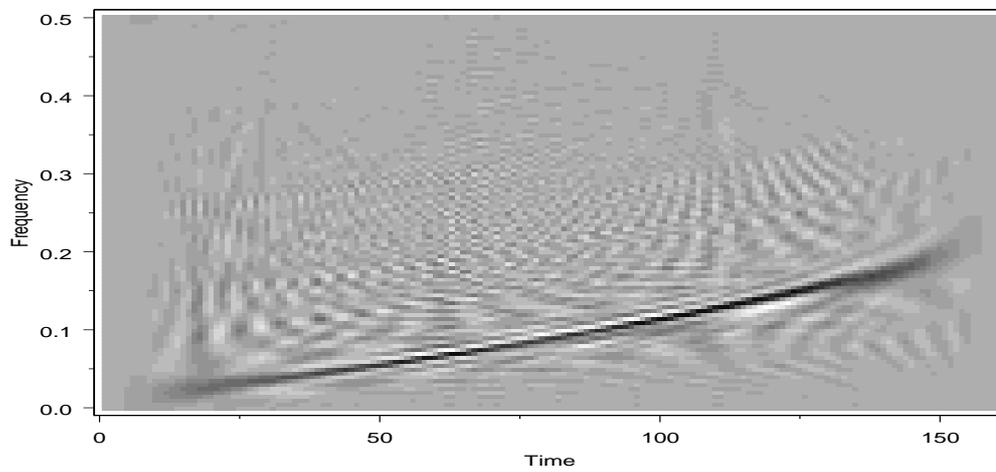

Figure 37: The WVD of Whale Click Data.
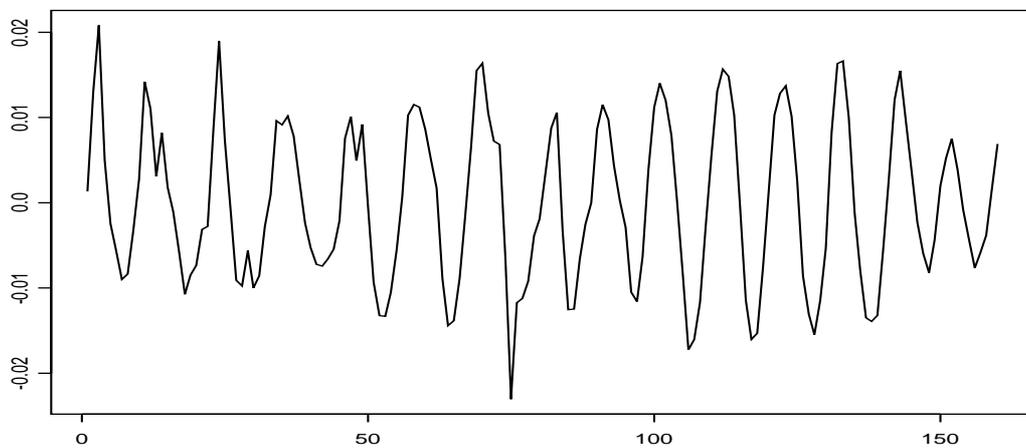
Figure 38: The Dual of the Whale Click Data.
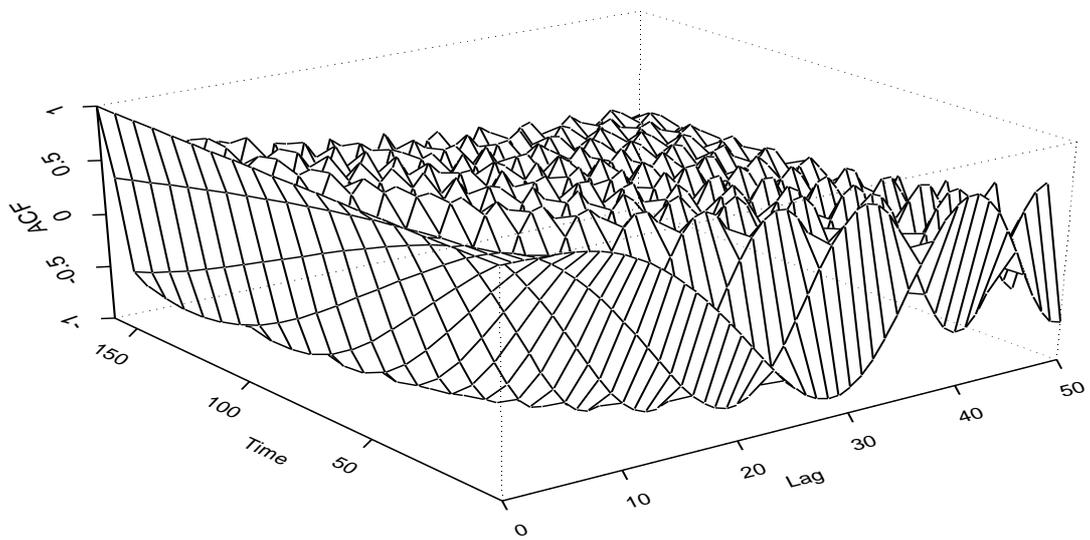
Figure 39: The Instantaneous ACF of the Whale Click Data.
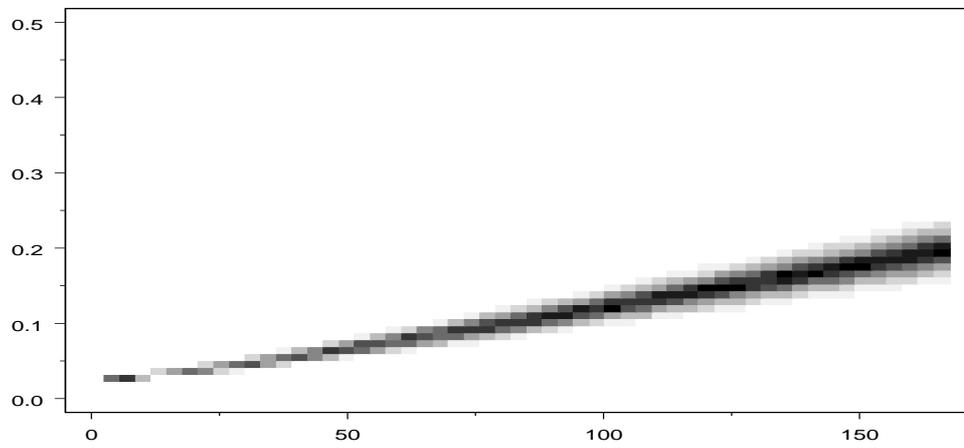
Figure 40: The L-C Instantaneous Spectrum of the Whale Click Data.