

AN EXAMINATION OF A MODEL OF THE TERM STRUCTURE
OF INTEREST RATES: KENNEDY'S MODEL

by

KONSTANTINOS SAKELLARIS

B.Sc. , The University of Athens, 1998

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

Department of Mathematics
Mathematical Finance Programme

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

April 2000

copyright © Konstantinos Sakellaris, 2000

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Mathematics

The University of British Columbia
Vancouver, Canada

Date April 28, 2000

Abstract

We examine the term structure model proposed by Kennedy (1994). The model assumes that the interest rates can be described as a Gaussian Random Field. We find conditions that the drift and covariance function of the forward rates have to satisfy. We then try to calibrate the model using a number of approaches. Finally the proposed covariance function by Kennedy (1997) is tested.

Table of Contents

Abstract	ii
List of Tables	v
List of Figures	vi
Acknowledgements	vii
SECTION 1 Introduction	1
SECTION 2 Kennedy's Approach	2
SECTION 3 Comments on Kennedy's Model	3
SECTION 4 Data	4
SECTION 5 Working with Kennedy's Model	5
5.1 Estimating the Drift Function	6
5.1.a Formulating the Problem	6
5.1.b Estimating the Drift Function from Equation (5.1.3)	8
5.2 Estimating the Covariance Function	9
5.2.a Estimating $Cov(p_k^{T_i}, p_s^{T_j})$	9
5.2.b Formulating the Problem	11
5.2.c Estimating the Covariance Function from Equation (5.2.6)	14
5.2.d Approach A	15
5.2.e Approach B	17
5.2.f Approach B assuming that the covariance function has the form mentioned by Kennedy (1997)	19
SECTION 6 Conclusion	20
References	22
Appendix A Graphs of the Treasury Yield data	24
Appendix B Estimation of the Drift Function	29
Appendix C Indicative plots of $A_{m,i,j}$ for certain i, j	32
Appendix D Indicative plots of $h_{i,j}(m)$ and $A_{m,i,j}$ for certain i, j	37

Appendix E	Statistical properties of the fitting error of $h_{i,j}(m)$	42
Appendix F	Indicative plots of $A_{m,i,j}$ for certain m	47
Appendix G	Calculation of $f_{k,m,i,j}$ for various forms of the covariance function	52
Appendix H	Plots of $G_1(m, i, j)$ when using $c_1(k, u, z)$ as the covariance function	59
Appendix I	Plots of $G_2(m, i, j)$ when using $c_2(k, u, z)$ as the covariance function	64
Appendix J	Plots of $G(m, i, j)$ when using $c_3(k, u, z)$ as the covariance function	67
Appendix K	Definition of Gaussian Random Field.	70

List of Tables

Table 1 5

Table 2 8

List of Figures

Graphs of the Treasury Yield data	24
Figure 1	29
Figure 2	29
Figure 3	30
Figure 4	30
Figure 5	31
Indicative plots of $A_{m,i,j}$ for certain i, j	32
Indicative plots of $h_{i,j}(m)$ and $A_{m,i,j}$ for certain i, j	37
Statistical properties of the fitting error of $h_{i,j}(m)$	42
Indicative plots of $A_{m,i,j}$ for certain m	47
Plots of $G_1(m, i, j)$ when using $c_1(k, u, z)$ as the covariance function	59
Plots of $G_2(m, i, j)$ when using $c_2(k, u, z)$ as the covariance function	64
Plots of $G(m, i, j)$ when using $c_3(k, u, z)$ as the covariance function	67

Acknowledgements

I'd like to thank my supervisor Dr. Ulrich Haussmann for his useful comments on my thesis and for all his help and support in the past one and a half years I've been in Vancouver.

I'd like to dedicate this work to my parents, who have always supported me, to my sister and to Smagdalen.

1. Introduction

Interest rates and their dynamics provide probably the most computationally difficult part of the modern financial theory. The modern fixed income market includes not only bonds but all kinds of derivative securities since they are used in time discounting. Interest rates are also important at the corporate level since most investment decisions are based on some expectations regarding alternative opportunities and the cost of capital - both depending on interest rates.

The classical approach to modelling interest rates and their associated bond prices is to specify the dynamics of a finite number of underlying processes, typically just the spot rate or the spot rate together with bond prices or interest rates of various durations. Then, using no-arbitrage conditions or equilibrium considerations, bond prices of all maturities are computed in terms of these basic processes.

The interest rate models can be divided into two classes. The first one involves modelling the short rate, by assuming it follows a stochastic differential equation. Here we can find three categories:

(i) Simple Gaussian models: In this category we include the models of Vasicek (1977), Merton (1973) and their generalisations like the Hull & White (1990) model. These models can be considered "simple" and the log-Gaussian distribution of bond prices makes it very easy to price derivative securities. The Hull & White model specifically is very flexible, in that any initial yield curve, and any initial term structure of volatility can be fitted. However the estimation of these models from data is not practical. The yield curve is not some nice smooth curve known at all positive real points; in practice, it is only known at a limited set of maturities, with dubious accuracy of measurement. Any procedure which requires repeated differentiation of this "curve" (like the procedure followed for these model's calibration) cannot be expected to work. Even if one could obtain estimates of the models' functions from the data, there is no reason why we should get consistent estimates if we performed the same analysis of the term-structure as it appears one week later. Additionally the undesirable property of these Gaussian models is that the interest rates may be negative.

(ii) Squared Gaussian models: These models escape the negative interest rates by modifying the variance structure. Cox, Ingersoll & Ross (1981) introduced such processes as models for the spot rate. Again a number of generalisations exist like the one proposed by Hull & White (1990). These models are particularly tractable because, like the Vasicek model, the yield curve is affine in the spot rate. All of the objections raised to the estimation of the Gaussian models apply equally here.

(iii) Multi-factor models: The single-factor models discussed so far are often criticised on the grounds that the long rate is a deterministic function of the spot rate, and that the prices of bonds of different maturities are perfectly correlated. These assumptions can't be verified by empirical evidence which on the contrary suggest that multi-factor models do significantly better than single-factor models. Models belonging in this class were proposed by Longstaff & Schwartz (1991), Duffie & Kan (1996), Constantinides (1992), Jamshidian (1988). By passing to multi-factor models, one should get an improved fit, but there is a heavy price to pay; if one wants to calculate prices of, say, options on bonds, the PDE to be solved is higher-dimensional, and will thus take much longer to be solved. Perhaps even more importantly, the factors used have to correspond to some observable variables. Longstaff &

Schwartz use the spot rate together with the volatility of the spot rate for example. Perhaps the best known model in this category is the model introduced by Brennan & Schwartz (1979), which takes as the variables of the (two-factor) model the spot rate and the long rate. Their pricing equation has to be solved numerically but their analysis of the Canadian Government Bonds gives impressive results. In this context, it is worth mentioning a result of Dybvig, Ingersoll & Ross (1996) who prove that the long rate is non-decreasing with respect to time. This makes one a little wary about a model which supposes that the long rate moves as a diffusion.

We would like to conclude the review of this class by saying that empirical studies tend to reject the hypothesis that the short rate is Markovian, let alone a diffusion [Ait-Sahalia 1997]. This is another objection to short-rate based models.

The second class of models, initiated by Ho & Lee (1986) and extended by Heath, Jarrow & Morton (1992) (HJM), involves modelling directly the forward rate processes for each maturity T . Specifically the HJM approach has as a starting point the observed yield curve and leads to the imposing of a relationship between the drifts and volatilities of forward rates on one hand, and the price volatility functions of discount bond prices on the other. In other words there is no such thing as the HJM model; rather there exists a whole class of models, each characterised by a specific functional form for the volatilities.

One important point about the HJM process is that it is intrinsically non-Markovian. From the implementation point of view this in turn means that a non-trivial HJM process cannot be mapped onto a recombining tree; therefore bushy trees or Monte Carlo paths are the tools available to the practitioner for pricing and hedging options.

Generally speaking, the complexity of this approach makes it very difficult to obtain simultaneously non-negativity of interest rates and a simple formula for bond pricing, properties desired for a term structure model. On the other hand the ability to input the initial yield curve directly and to vary the volatilities of the forward rates are attractions. A generalisation of this approach was proposed by Kennedy, whose model allowed the incorporation of an infinite number of sources of randomness in the evolution of the term structure.

Finally a number of models exist which do not belong in any of the two above classes, like the one proposed by Heath & Jara (1998) which is based on Futures Prices. These models though are relatively new and haven't been examined thoroughly enough so we are not going to review them.

The organisation of the thesis is as follows. Section 2 presents Kennedy's approach. In Section 3 we make some comments on Kennedy's model. Section 4 presents the data used for the empirical estimation of the model. In Section 5 we try to estimate the drift and covariance function and we derive a condition that the covariance function should satisfy. Section 6 concludes the examination of this model by summarizing our results.

2. Kennedy's approach

Heath, Jarrow & Morton (1992) approached the problem of modelling the term structure of interest rates by specifying the dynamics of all instantaneous forward rates $\{F_{s,t} : 0 \leq s \leq t\}$ where:

$$(2.1) \quad P_{s,t} = e^{-\int_s^t F_{s,u} du}$$

represents the price at time s of a bond paying one unit at time $t \geq s$. Their model assumes that the $F_{s,t}$ satisfy stochastic differential equations of the form

$$(2.2) \quad dF_{s,t} = \alpha(s,t)ds + \sum_{i=1}^m \sigma_i(s,t)dW_s^i$$

where $W^1, \dots, W^m$ are independent Brownian motions and $\alpha(s,t)$ and $\sigma_i(s,t)$ are processes adapted to the natural filtrations of the Brownian motions.

Kennedy (1994) while following the approach of modelling the instantaneous forward rates, considered the case where $\{F_{s,t} : 0 \leq s \leq t\}$ is a Gaussian random field¹, so that all finite dimensional distributions are normal. He dealt with the situation where, under the equilibrium or martingale measure, the random field has independent increments in the s -direction which is the direction of the evolution of “real” time. This framework includes the Heath, Jarrow and Morton model in the case when the coefficients $\alpha(s,t)$ and $\sigma_i(s,t)$ in the underlying stochastic differential equations (2.2) are not random and so the rates $\{F_{s,t}\}$ are Gaussian.

Kennedy assumes that the instantaneous forward interest rate $F_{s,t}$ for date t at time s , $0 \leq s \leq t$, are given by:

$$(2.3) \quad F_{s,t} = \mu_{s,t} + X_{s,t}$$

with $0 \leq s \leq t$ where $X_{s,t}$ is a centered continuous Gaussian random field and $\mu_{s,t}$ is the deterministic and continuous drift function.

He assumes that the initial term structure $\{\mu_{0,t}, t \geq 0\}$ is specified.

The covariance of $X_{s,t}$ is specified by :

$$(2.4) \quad Cov(X_{s_1,t_1}, X_{s_2,t_2}) = c(s_1 \wedge s_2, t_1, t_2)$$

where $0 \leq s_i \leq t_i$ and $i = 1, 2$. The function c is given and satisfies $c(0, t_1, t_2) \equiv 0$. It is implicit that $c(s_1 \wedge s_2, t_1, t_2)$ is symmetric in t_1 and t_2 and is nonnegative definite in (s_1, t_1) and (s_2, t_2) .

3. Comments on Kennedy's model

Kennedy's model includes infinite-factor models, such as those based on the Brownian sheet, but also finite-factor models such as that of Jamshidian (1989) (where bond prices of all maturities are just functions of the spot rate). It also includes the HJM model when the drift and volatilities are deterministic.

Kennedy's model allows the possibility of infinite-dimensional cases, when, for example, the forward rate surface is generated by a Brownian sheet. This fact has significant empirical implications. The Gaussian models can be described by their drift and covariance functions. If we combine this with Kennedy's Theorem 1.1, then we just need to estimate the covariance function in order to fully describe the model and price derivatives. We don't need to specify from before the number of factors that are driving the term structure, as in most of the other multi-factor models (like Brennan & Schwartz, Longstaff & Schwartz, etc). Once the

¹We define Gaussian Random Field in Appendix K.

covariance function has been estimated we can easily find the number of factor correlations that are significantly different from unity.

The Gaussian model with independent increments has the important advantage that it is quite tractable and yields a simple characterisation of the martingale measure. In addition, prices of derivatives, such as interest rate caps, have an intuitively appealing form. On the other hand though, the Gaussian model can't explain the heavier (than log-normal) tail behaviour that is exhibited by the observed distributions of market prices. The existence of these "fat tails" is related to the fact that the coefficients of the model, which are taken to be constant, are in fact random. This could be considered as an important drawback of the Kennedy model. For a generalisation of the Kennedy model to include non-Gaussian random fields we refer to Goldstein (1997).

Finally a drawback of the random field interest rate models is that they allow (but not necessarily impose) a multi-humped yield curve. This is rarely found empirically.

4. Data

The first, and most important step, in examining this model is to try and estimate the drift and the covariance function. In order to estimate them we are going to use interest rate data. Choosing the data though is not as simple a task as it may seem. This is perhaps not surprising in view of the nature and quality of data on term structure. Prices of coupon bearing and zero coupon bonds, LIBOR rates, index linked bonds together with options and futures on such things all provide information on interest rates.

The data that we used were provided by DATASTREAM. As we mentioned above, in the following we are going to try to calibrate the model using interest rate data. The best choice for interest rate data would be the US Treasury data. The existence though of a large number of methods to strip the Treasury bonds of their coupons (in order to calculate the zero curve), providing us with an equally large number of possible zero curves, made us choose another time series. So we worked with the "Treasury constant maturity" series. The description of this series, as provided by the Federal Reserve, is : "Yields on Treasury securities at "constant maturities" are interpolated by the U.S. Treasury from the daily yield curve. This curve is based on the closing market bid yields on actively traded Treasury securities in the over-the-counter market. These market yields are calculated from composites of quotations obtained by the Federal Reserve bank. The constant maturity yield values are read from the yield curve at fixed maturities, currently 3 and 6 months and 1, 2, 3, 5, 7, 10, 20 and 30 years". This series offers the advantage of not being dependant on which method we will use to calculate the yield curve. The only possible objection then is with the method that the Federal Reserve uses to calculate this zero curve! The above series begins in 01/01/1991 but for the calibration we used data from 01/01/1995 up until 12/10/1999. You can find the graphs of the treasuries in Appendix A².

In the following we are going to work with logs of treasury bond prices and not yields. We should therefore transform the yields into prices. This can be done easily through the

²In the following we will assume that $T_1 = 63$, in other words 3 months include 63 working days.

definition of the yield³:

$$Y_{t,T} = -\frac{\log P_{t,T}}{T-t}$$

where T is the maturity of the bond and t the time that has passed since it's issuing. Since the yields we have are $Y_{s,s+T}$ at each date s for some maturities T , we can easily get the corresponding prices $P_{s,s+T}$.

Treasury price logs	Test Statistic for data	Unit Root Rejected	Test Stat. for differ. data	Unit Root Rejected
3-month	-1.6673	No	-6.6969	Yes
6-month	-1.7918	No	-6.7725	Yes
1-year	-1.6122	No	-5.9662	Yes
2-year	-1.9245	No	-5.8454	Yes
3-year	-1.8749	No	-5.7466	Yes
5-year	-1.6318	No	-5.3576	Yes
7-year	-1.5013	No	-5.5792	Yes
10-year	-1.3089	No	-5.6666	Yes
20-year	-1.2108	No	-5.6015	Yes
30-year	-1.1183	No	-5.6382	Yes

Table 1 : The critical values for the test are -2.57 at the 10% level and -2.86 at the 5% level.

Examining the prices of the treasuries we notice two things. First of all their rates seem to have a big drop around October 1998. This is mainly because of the Asian crisis. The other thing that we notice is that the bond price series (and their logs) may be non-stationary. By applying the Dickey-Fuller unit root test to the data, the unit root hypothesis can't be rejected and therefore the series can not be assumed stationary. Then by applying the Dickey-Fuller test to the first differences of the series at the 5% and 10% level, we find that can reject the hypothesis that the first differences have a unit root with 95% confidence⁴. Therefore in the rest of the paper we can safely assume that the first differences of the bond price series (and their logs) are stationary.

The reader can find the results of the tests in Table 1 (we got the same results performing the Phillips-Perron test; in the above results we have assumed that there is no trend in the prices but we get similar results if we assume trend).

5. Working with Kennedy's model

From here on we are going to work in discrete time and we are going to try to find a formula which will help us estimate the drift and covariance functions. So the time s will take one

³We should note here that we treated the yield as annual although we weren't able to confirm that with DATASTREAM who didn't have this information.

⁴We remind the reader that the finding of a unit root in a time series indicates nonstationarity. The unit root hypothesis can be rejected if the T-TEST statistic is smaller than the critical value. We used SHAZAM to do the unit root tests.

of the (integer) values 0, 1, 2, ..., 986. For convenience let $M = 986$ be the last "date" of our data.

5.1 Estimating the drift function

5.1.a Formulating the problem

The price at time s of a bond paying one unit at time $t \geq s$ is :

$$(5.1.1) \quad P_{s,t} = e^{-\int_s^t F_{s,u} du}$$

The data we have restricts us to knowing just the prices $P_{s,s+T}$ where T is one of the maturities of the bond and $s = 0, 1, 2, \dots, M$. Let us pick a specific maturity T . Then, if we apply equation (2.3) in (5.1.1) we get:

$$P_{s,s+T} = e^{-\int_s^{s+T} \mu_{s,u} du - \int_s^{s+T} X_{s,u} du}$$

and equivalently

$$(5.1.2) \quad \log(P_{s,s+T}) = -\int_s^{s+T} \mu_{s,u} du - \int_s^{s+T} X_{s,u} du$$

Taking expectations:

$$E[\log(P_{s,s+T})] = E[-\int_s^{s+T} \mu_{s,u} du - \int_s^{s+T} X_{s,u} du] \Leftrightarrow$$

$$E[\log(P_{s,s+T})] = -\int_s^{s+T} E[\mu_{s,u}] du - \int_s^{s+T} E[X_{s,u}] du \Leftrightarrow$$

$$E[\log(P_{s,s+T})] = -\int_s^{s+T} \mu_{s,u} du$$

The last equivalency comes from Kennedy's assumption that $X_{s,t}$ is a centered Gaussian random field (so it has zero mean) and $\mu_{s,t}$ is deterministic.

As we mentioned above the logs of the bond data are not stationary, so the value of $E[\log(P_{s,s+T})]$ will depend on the time s . Therefore we can't estimate this expectation (for each s) by a sample mean. The first differences, though, are stationary and therefore we can work with these:

$$\begin{aligned} E[\log(P_{s+1,s+1+T}) - \log(P_{s,s+T})] &= E[\log(P_{s+1,s+1+T})] - E[\log(P_{s,s+T})] \\ &= -\int_{s+1}^{s+1+T} \mu_{s+1,u} du + \int_s^{s+T} \mu_{s,u} du \end{aligned}$$

Working recursively we have :

$$\int_{s+1}^{s+1+T} \mu_{s+1,u} du = \int_s^{s+T} \mu_{s,u} du - E[\log(P_{s+1,s+1+T}) - \log(P_{s,s+T})]$$

$$\begin{aligned}\int_{s+2}^{s+2+T} \mu_{s+k,u} du &= \int_{s+1}^{s+1+T} \mu_{s+1,u} du - E[\log(P_{s+1,s+1+T}) - \log(P_{s,s+T})] \\ &= \int_s^{s+T} \mu_{s,u} du - 2 \cdot E[\log(P_{s+1,s+1+T}) - \log(P_{s,s+T})]\end{aligned}$$

and therefore we get the general formula

$$\int_{s+k}^{s+k+T} \mu_{s+k,u} du = \int_s^{s+T} \mu_{s,u} du - k \cdot E[\log(P_{s+1,s+1+T}) - \log(P_{s,s+T})]$$

where $s = 0, 1, 2, \dots, M$ and $k = 0, 1, 2, \dots, M - s$.

We can then estimate $E[\log(P_{s+1,s+1+T}) - \log(P_{s,s+T})]$ from :

$$E[\log(P_{s+1,s+1+T}) - \log(P_{s,s+T})] = \frac{1}{M+1} \sum_{i=0}^M [\log(P_{i+1,i+1+T}) - \log(P_{i,i+T})] + \epsilon_T$$

where ϵ_T is the unbiased random error of the estimate.

Let $\Delta(T) = \frac{1}{M+1} \sum_{i=0}^M [\log(P_{i+1,i+1+T}) - \log(P_{i,i+T})]$. By replacing the last equation in the previous one we get :

$$\int_{s+k}^{s+k+T} \mu_{s+k,u} du = \int_s^{s+T} \mu_{s,u} du - k \cdot \Delta(T) + k \cdot \epsilon_T$$

and for $s = 0$ we have :

$$\int_k^{k+T} \mu_{k,u} du = \int_0^T \mu_{0,u} du - k \cdot \Delta(T) + k \cdot \epsilon_T$$

We know that $\mu_{0,u} = F_{0,u}$ and therefore $\int_0^T \mu_{0,u} du = -\log(P_{0,T})$. So we conclude that the drift function must satisfy:

$$(5.1.3) \quad \int_k^{k+T} \mu_{k,u} du = -\log(P_{0,T}) - k \cdot \Delta(T) + k \cdot \epsilon_T$$

where $k = 0, 1, 2, \dots, M$.

Notice in (5.1.3) that the right part of the equation is comprised of known parameters, except of course for the error ϵ_T . We can estimate ϵ_T as follows:

Let $p_k^T = \log(P_{k+1,k+1+T}) - \log(P_{k,k+T})$. Then

$$E[p_0^T] \simeq \frac{1}{M+1} \sum_{i=0}^M p_i^T + \epsilon_T$$

So

$$\begin{aligned}E[\epsilon_T^2] &= E \left[\frac{1}{M+1} \sum_{i=0}^M p_i^T - E[p_0^T] \right]^2 = \frac{1}{(M+1)^2} \cdot E \left[\sum_{i=0}^M (p_i^T - E[p_0^T]) \right]^2 \\ &= \frac{1}{(M+1)^2} \cdot \sum_{i,j=0}^M E[(p_i^T - E[p_0^T])(p_j^T - E[p_0^T])] = \frac{1}{(M+1)^2} \cdot \left\{ 2 \cdot \sum_{m=0}^M \sum_{i=0}^{M-m} \text{Cov}(p_i^T, p_{i+m}^T) \right\}\end{aligned}$$

$$- \sum_{m=0}^M Cov(p_0^T, p_0^T) \Big\} = \frac{2}{(M+1)^2} \cdot \sum_{m=1}^M (M-m+1) \cdot Cov(p_0^T, p_m^T) + \frac{1}{M+1} Cov(p_0^T, p_0^T)$$

We can easily estimate $Cov(p_i^T, p_{i+m}^T)$ (we refer the reader to section 5.2.a for that) and therefore get an estimate of the order of the error by calculating $\sqrt{E[\epsilon_T^2]}$:

Maturity	3m	6m	1y	2y	3y
$E[\epsilon_T^2]$	$O(10^{-11})$	$O(10^{-11})$	$O(10^{-10})$	$O(10^{-10})$	$O(10^{-9})$
$\sqrt{E[\epsilon_T^2]}$	$O(10^{-6})$	$O(10^{-6})$	$O(10^{-5})$	$O(10^{-5})$	$O(10^{-5})$
$ \Delta(T) $	$1.32 \cdot 10^{-6}$	$2.71 \cdot 10^{-6}$	$5.72 \cdot 10^{-6}$	$1.67 \cdot 10^{-5}$	$2.11 \cdot 10^{-5}$

Maturity	5y	7y	10y	20y	30y
$E[\epsilon_T^2]$	$O(10^{-9})$	$O(10^{-9})$	$O(10^{-8})$	$O(10^{-8})$	$O(10^{-8})$
$\sqrt{E[\epsilon_T^2]}$	$O(10^{-5})$	$O(10^{-5})$	$O(10^{-4})$	$O(10^{-4})$	$O(10^{-4})$
$\Delta(T)$	$3.6 \cdot 10^{-5}$	$5.7 \cdot 10^{-5}$	$5.2 \cdot 10^{-5}$	$1.2 \cdot 10^{-4}$	$8.2 \cdot 10^{-5}$

Table 2 : Estimate of the order of ϵ_T

We notice that the order of the value of ϵ_T and $\Delta(T)$ are close. One reason could be that the error in the estimation of $Cov(p_i^T, p_{i+m}^T)$ is **not** small, as we have assumed (this can actually partially explain the problems we had in estimating the covariance function of the model), although we can't be sure about that. Nevertheless in the following we will ignore ϵ_T .

5.1.b Estimating the drift function from equation (5.1.3)

We are going to work as following:

- We will fit functions to $\log(P_{0,T})$ and $\Delta(T)$, as presented in figures B1 and B2 respectively.
- We will then differentiate (5.1.3) with respect to the maturity T and get the drift function.

Fitting a function to $\log(P_{0,T})$ proves to be a simple task. Examining the goodness of fit of polynomials of degrees 1 to 6, we find that the cubic polynomial $\alpha T^3 + \beta T^2 + \gamma T + \delta$, presented in figure B3, is our optimal choice ⁵. If we use polynomials of degrees 4 or higher the fitting doesn't improve significantly.

Before moving on to the fitting of $\Delta(T)$, notice that:

$$\Delta(T) = \frac{1}{M+1} \sum_{i=0}^M [\log(P_{i+1,i+1+T}) - \log(P_{i,i+T})] = \frac{1}{M+1} \log \left(\frac{P_{M+1,M+1+T}}{P_{0,T}} \right)$$

and therefore $\Delta(0) = 0$.

⁵We found $\alpha = 1.27 \cdot 10^{-12}$, $\beta = -1.55 \cdot 10^{-8}$, $\gamma = -1.92 \cdot 10^{-4}$ and $\delta = -2.4 \cdot 10^{-3}$.

Examining the plot of $\Delta(T)$ for the 10 maturities we have available, testing a big number of functions and keeping in mind the above comment we concluded that $\Delta(T)$ is best described by the function $\Delta(T) = a(e^{b(T-c)^2} - e^{bc^2})$, as shown in figure B4⁶.

We can then rewrite (5.1.3) in the following way:

$$\int_k^{k+T} \mu_{k,u} du = -\alpha T^3 - \beta T^2 - \gamma T - \delta - k \cdot a(e^{b(T-c)^2} - e^{bc^2}) + \epsilon_T^1 + k \cdot \epsilon_T^2$$

where ϵ_T^1 is the fitting error of $\log(P_{0,T})$, ϵ_T^2 is the fitting error of $\Delta(T)$ and where $k = 0, 1, 2, \dots, M$.

Differentiating this equation with respect to T , and ignoring the errors, we can get an estimate of the drift function:

$$\bar{\mu}_{k,k+T} = -3\alpha T^2 - 2\beta T - \gamma - k \cdot 2ab(T-c)e^{b(T-c)^2}$$

with $k = 0, 1, 2, \dots, M$.

We should note here that $\bar{\mu}_{k,k+T}$ has the form we expected, since it can be easily shown that stationarity implies $\mu_{k,k+T} = f(T) + k \cdot g(T)$ for some functions f and g .

We plotted the graphs of $\bar{\mu}_{1,1+T}$, $\bar{\mu}_{400,400+T}$ and $\bar{\mu}_{1000,1000+T}$ in figure B5. We can see that the maturity T defines most of the drifts' behaviour, something perfectly logical, while the date k changes slightly the value of the drift, depending how far in the future we are looking (i.e. how big k is)⁷.

Finally, going back to Kennedy (1994) and looking at his Theorem 1.1 we can find a second equation that the drift function $\mu_{s,t}$ must satisfy in order for the discounted bond process to be a martingale :

$$(5.1.4) \quad \mu_{s,t} = \mu_{0,t} + \int_0^t c(s \wedge v, v, t) dv$$

for all $0 \leq s \leq t$. If (5.1.4) holds for all s, t then the real measure is the risk neutral measure. If it does not hold for all s, t then in an arbitrage free market there exists another measure under which the discounted bond-price process is a martingale.

5.2 Estimating the covariance function

5.2.a Estimating $Cov(p_k^{T_i}, p_s^{T_j})$

We are now going to try to estimate the covariance function c . In order to do that we will need to calculate $Cov(p_k^{T_i}, p_s^{T_j})$, where $p_k^T = \log(P_{k+1,k+1+T}) - \log(P_{k,k+T})$ with $k = 0, 1, 2, \dots, M-1$. $Cov(p_k^{T_i}, p_s^{T_j})$ are the partial autocovariances of all the series corresponding to each pair of maturities of our differenced-log-bond data.

Remark 1 : k, s can take any values between 0 and $M-1$. But the covariance function is symmetric and therefore $Cov(p_k^{T_i}, p_s^{T_j}) = Cov(p_s^{T_j}, p_k^{T_i})$ so we should calculate $Cov(p_k^{T_i}, p_s^{T_j})$ for $s \geq k$. It would then be easier if we wrote s in the form $s = k + m$

⁶We found $a = -1.3 \cdot 10^{-4}$, $b = -1.08 \cdot 10^{-7}$ and $c = 5019$.

⁷The reader shouldn't be troubled with the value $T = 5019$ which defines the drifts behaviour, because we believe it just reflects information extracted from our sample.

with $k = 0, 1, 2, \dots, M-1$ and $m = 0, 1, 2, \dots, M-1-k$. Then we want to estimate $Cov(p_k^{T_i}, p_{k+m}^{T_j})$ for $k = 0, 1, 2, \dots, M-1$ and $m = 0, 1, 2, \dots, M-1-k$. We should point out that stationarity implies that $Cov(p_k^{T_i}, p_{k+m}^{T_j})$ is independent of the value of k .

Remark 2 : We are going to estimate $Cov(p_k^{T_i}, p_{k+m}^{T_j})$ using the (unbiased) estimator $C_{i,j,m}$:

$$Cov(p_k^{T_i}, p_{k+m}^{T_j}) = C_{i,j,m} + \epsilon_{i,j,m}$$

where

$$C_{i,j,m} = \frac{1}{M-1-m} \sum_{l=0}^{M-1-m} (p_l^{T_i} - \Delta_i^{0,M-1-m}) \cdot (p_{l+m}^{T_j} - \Delta_j^{m,M-1})$$

$$\Delta_i^{a,b} = \frac{1}{b-a+1} \cdot \sum_{n=a}^b p_n^{T_i}$$

with $k = 0, 1, 2, \dots, N$, $m = 0, 1, 2, \dots, N-k$, $N < M-1$ and $\epsilon_{i,j,m}$ a random error which we assume to be small.

We should make some comments here regarding the estimation:

- (a) $C_{i,j,m}$ is independent of k since the series are stationary.
- (b) We placed a restriction on m by not allowing it to become bigger than $N-k$, where $N < M-1$ (this restriction immediately imposes another restriction on k to be less or equal than N or m may become negative). The reason is that the closer m gets to $M-1-k$, the smaller the sample we have for estimating $Cov(p_k^{T_i}, p_{k+m}^{T_j})$ is. Since we want the errors to be small we need a relatively “big” sample. Usually by “big” statisticians mean more than 50 – 60 data points, although this number is relevant to the application. In the following we assumed $N = M-1-100$ (and therefore the smallest sample we are going to use is 100).
- (c) The way we estimate the covariance, the sample size will differ depending on m . In one sense someone could argue that this is not consistent, but practically what we do is exploit as much information as we have available. Therefore the smaller m and k are the better our estimate is and consequently the smaller $\epsilon_{i,j,m}$ will be.
- (d) The indices i, j just specify with which series we are working.

5.2.b Formulating the problem

For $k = 0, 1, 2, \dots, N$ and $m = 0, 1, 2, \dots, N - k$ we have - using the definition of p_k^T , equation (5.1.2) and the fact that the drift function $\mu_{s,t}$ is deterministic - the following:

$$\begin{aligned}
Cov(p_k^{T_i}, p_{k+m}^{T_j}) &= \\
&= Cov\left(\int_{k+1}^{k+T_i+1} X_{k+1,u} du - \int_k^{k+T_i} X_{k,u} du, \int_{k+m+1}^{k+m+T_j+1} X_{k+m+1,u} du - \int_{k+m}^{k+m+T_j} X_{k+m,u} du\right) \\
&= \int_{k+1}^{k+T_i+1} \int_{k+m+1}^{k+m+T_j+1} Cov(X_{k+1,u}, X_{k+m+1,z}) dz du + \int_k^{k+T_i} \int_{k+m}^{k+m+T_j} Cov(X_{k,u}, X_{k+m,z}) dz du \\
&\quad - \int_k^{k+T_i} \int_{k+m+1}^{k+m+T_j+1} Cov(X_{k,u}, X_{k+m+1,z}) dz du - \int_{k+1}^{k+T_i+1} \int_{k+m}^{k+m+T_j} Cov(X_{k+1,u}, X_{k+m,z}) dz du \\
&= \int_{k+1}^{k+T_i+1} \int_{k+m+1}^{k+m+T_j+1} c((k+1) \wedge (k+m+1), u, z) dz du + \int_k^{k+T_i} \int_{k+m}^{k+m+T_j} c(k \wedge (k+m), u, z) dz du \\
&\quad - \int_k^{k+T_i} \int_{k+m+1}^{k+m+T_j+1} c(k \wedge (k+m+1), u, z) dz du - \int_{k+1}^{k+T_i+1} \int_{k+m}^{k+m+T_j} c((k+1) \wedge (k+m), u, z) dz du \\
&= \int_{k+1}^{k+T_i+1} \int_{k+m+1}^{k+m+T_j+1} c(k+1, u, z) dz du + \int_k^{k+T_i} \int_{k+m}^{k+m+T_j} c(k, u, z) dz du \\
&\quad - \int_k^{k+T_i} \int_{k+m+1}^{k+m+T_j+1} c(k, u, z) dz du - \int_{k+1}^{k+T_i+1} \int_{k+m}^{k+m+T_j} c(k+(1 \wedge m), u, z) dz du
\end{aligned}$$

with $k = 0, 1, 2, \dots, N$, $m = 0, 1, 2, \dots, N - k$. So we end up with the formula :

(5.2.1)

$$\begin{aligned}
Cov(p_k^{T_i}, p_{k+m}^{T_j}) &= \int_{k+m+1}^{k+m+T_j+1} \left(\int_{k+1}^{k+T_i+1} c(k+1, u, z) du - \int_k^{k+T_i} c(k, u, z) du \right) dz \\
&\quad - \int_{k+m}^{k+m+T_j} \left(\int_{k+1}^{k+T_i+1} c(k+1, u, z) du - \int_k^{k+T_i} c(k, u, z) du \right) dz \\
&\quad + \delta_{m0} \int_k^{k+T_j} \int_{k+1}^{k+T_i+1} (c(k+1, u, z) - c(k, u, z)) du dz
\end{aligned}$$

with $k = 0, 1, 2, \dots, N$, $m = 0, 1, 2, \dots, N - k$.

Let

- $K_k^i(z) = \int_{k+1}^{k+T_i+1} c(k+1, u, z) du - \int_k^{k+T_i} c(k, u, z) du$
- $F_k^{i,j}(m) = \int_{k+m}^{k+m+T_j} K_k^i(z) dz$ if $m \geq 1$

- $F_k^{i,j}(0) = \int_k^{k+T_j} \left(\int_{k+1}^{k+T_i+1} c(k, u, z) du - \int_k^{k+T_i} c(k, u, z) du \right) dz$

We can then rewrite (5.2.1) as :

$$Cov(p_k^{T_i}, p_{k+m}^{T_j}) = F_k^{i,j}(m+1) - F_k^{i,j}(m)$$

Working recursively we have :

$$(5.2.2) \quad F_k^{i,j}(m+1) = \sum_{a=0}^m Cov(p_k^{T_i}, p_{k+a}^{T_j}) + F_k^{i,j}(0)$$

Let $G_k^i(z) = \int_k^{k+T_i} c(k, u, z) du$. Then

$$K_k^i(z) = G_{k+1}^i(z) - G_k^i(z) \Leftrightarrow G_{k+1}^i(z) = \sum_{b=0}^k K_b^i(z) + G_0^i(z)$$

But $G_0^i(z) = 0$ since $c(0, u, z) = 0 \quad \forall u, z$ and therefore we have :

$$(5.2.3) \quad G_k^i(z) = \sum_{b=0}^{k-1} K_b^i(z)$$

with $k = 1, 2, \dots, N$. So for these k we have :

$$\begin{aligned} \int_k^{k+T_i} \int_{k+m+1}^{k+m+1+T_j} c(k, u, z) dz du &= \int_{k+m+1}^{k+m+1+T_j} G_k^i(z) dz = \int_{k+m+1}^{k+m+1+T_j} \sum_{b=0}^{k-1} K_b^i(z) dz \\ &= \sum_{b=0}^{k-1} \int_{k+m+1}^{k+m+1+T_j} K_b^i(z) dz = \sum_{b=0}^{k-1} F_b^{i,j}(m+1) = k \cdot \sum_{a=0}^m Cov(p_k^{T_i}, p_{k+a}^{T_j}) + \sum_{b=0}^{k-1} F_b^{i,j}(0) \end{aligned}$$

Remark 3 : Above we used the fact that the series is stationary and therefore

$$\sum_{b=0}^{k-1} \sum_{a=0}^m Cov(p_b^{T_i}, p_{b+a}^{T_j}) = k \cdot \sum_{a=0}^m Cov(p_k^{T_i}, p_{k+a}^{T_j})$$

So far we have:

$$(5.2.4) \quad \int_k^{k+T_i} \int_{k+m+1}^{k+m+1+T_j} c(k, u, z) dz du = k \cdot \sum_{a=0}^m Cov(p_k^{T_i}, p_{k+a}^{T_j}) + \sum_{b=0}^{k-1} F_b^{i,j}(0)$$

for $k = 1, 2, \dots, N$ and $m = 0, 1, 2, \dots, N - k$.

Remark 4 : We shouldn't be concerned with the fact that (5.2.4) holds for $k = 1, 2, \dots, N$ and not for $k = 0$. We didn't lose any information in the calculations. It is just that the left hand side of (5.2.4) contains $c(k, u, z)$, which is equal to 0 for $k = 0$.

Let

- $f_{k,m,i,j} = \int_k^{k+T_i} \int_{k+m}^{k+m+T_j} c(k, u, z) dz du$
- $V_{m,i,j} = \sum_{a=0}^m \text{Cov}(p_k^{T_i}, p_{k+a}^{T_j})$
- $B_{k,i,j} = \sum_{b=0}^{k-1} F_b^{i,j}(0) = (1 - \delta_{k1}) \cdot \sum_{b=1}^{k-1} (f_{b,1,j,i} - f_{b,0,j,i})$

for $k = 1, 2, \dots, N$, $m = 0, 1, 2, \dots, N - k$ and $i, j = 1, 2, \dots, 10$. In the above we kept in mind that $F_0^{i,j}(0) = 0$. Notice that $V_{m,i,j}$ does not depend on k since the series are stationary. Then by replacing the above variables in (5.2.4) and keeping in mind Remark 4 we get:

$$(5.2.5) \quad f_{k,m,i,j} = k \cdot V_{m,i,j} + B_{k,i,j}$$

$k = 0, 1, 2, \dots, N$ and $m = 0, 1, 2, \dots, N - k + 1$. The very interesting thing about (5.2.5) is that m, k are separated. Now working recursively with (5.2.5) we have :

$$\begin{aligned} f_{1,m,i,j} &= V_{m,i,j} \\ f_{2,m,i,j} &= 2 \cdot V_{m,i,j} + B_{2,i,j} = 2 \cdot V_{m,i,j} + (f_{1,1,j,i} - f_{1,0,j,i}) = 2 \cdot V_{m,i,j} + (V_{1,i,j} - V_{0,i,j}) \\ f_{3,m,i,j} &= 3 \cdot V_{m,i,j} + (V_{1,i,j} - V_{0,i,j}) + 2 \cdot (V_{1,i,j} - V_{0,i,j}) = 3 \cdot V_{m,i,j} + 3 \cdot (V_{1,i,j} - V_{0,i,j}) \end{aligned}$$

.

.

.

$$f_{k,m,i,j} = k \cdot V_{m,i,j} + \frac{k(k-1)}{2} \cdot (V_{1,i,j} - V_{0,i,j})$$

with $k = 0, 1, 2, \dots, N$ and $m = 0, 1, 2, \dots, N - k + 1$. Let:

$$v_{i,j} = \frac{1}{2} \cdot (V_{1,i,j} - V_{0,i,j})$$

Then we have the main result of this paragraph, the equation:

$$(5.2.6) \quad f_{k,m,i,j} = v_{i,j} \cdot k^2 + (V_{m,i,j} - v_{i,j}) \cdot k$$

for $k = 0, 1, 2, \dots, N$ and $m = 0, 1, 2, \dots, N - k$.

Equation (5.2.6) gives us a very interesting result. For each pair of maturities (i, j) we can see exactly how $f_{k,m,i,j}$ is dependant on k and on m . Moreover, through $V_{m,i,j}$, we have an explicit formula for $f_{k,m,i,j}$ which we can use to estimate the covariance function $c(k, u, z)$. We should remind the reader here that $V_{m,i,j}$ can be estimated by $A_{m,i,j} \equiv \sum_{a=0}^m C_{i,j,a}$ for $m = 0, 1, 2, \dots, N - k$; we refer the reader to section 5.2.a.

5.2.c Estimating the covariance function from equation (5.2.6)

At first glance equation (5.2.6) gives us a clear way of how to estimate the covariance function $c(k, u, z)$. The unknown covariance function is on the left hand side while a certain transformation of the observed bond prices lies on the right hand side.

If we examine the problem closer though we can see that it can't be solved that easily. The limits of the double integral are such that most numerical schemes will never work or even if they work the estimation errors will be huge. The main problem is that we have to look at the problem for each k separately. By picking specific k, i, j we get to solve a system where the number of unknowns and equations varies greatly depending on the choice of k, i, j , but with the unknowns being always more than the equations. We can increase the number of equations we have in our disposal by keeping constant two of the three parameters k, i, j and changing the third one but even then we won't get enough equations to solve our system. We should add here that this system is really hard to manipulate because of its specific form. Moreover all the numerical schemes and optimization algorithms we used failed.

After seeing that we couldn't attack the problem directly, we thought of combining this problem with the drift function problem, that is with equations (5.1.3) and (5.1.4). This way we get a few more equations from some restrictions we can impose to our system. Unfortunately on one hand the problem becomes much more complicated and on the other hand the size of it becomes "too big" (at least for the computers we were using). Moreover the equations we had still weren't enough for a big number of k so we had to stop.

Since we didn't manage to solve the problem using a direct numerical approach we tried two different approaches:

- (A) One approach is to fit a function to $V_{m,i,j}$, or more precisely to $A_{m,i,j}$. Then the right hand part of equation (5.2.6), namely $f_{k,m,i,j}$, will become a continuous function of k, m for each i, j . At that point we will have two options on how to continue. Either work numerically (now it will be easier to do so than before) or work backwards the algebra and try to find which function's double integral can give us the function fitted to $f_{k,m,i,j}$.
- (B) The other approach consists of making assumptions about the form of $c(k, u, z)$ and working through the algebra to find $f_{k,m,i,j}$. We can then compare the left with the right hand side of (5.2.6) and try to calibrate the parameters so that $f_{k,m,i,j}$ will have a form like (5.2.6) and -at the same time - give the best possible fit to the data.

The better approach of the two is (A) using some kind of numerical algorithm after fitting a function to $f_{k,m,i,j}$. If we don't want to do that, but want to work some more with the algebra then both approaches have their advantages and disadvantages. The algebraic (A) will most probably lead us to a covariance function that doesn't satisfy the necessary conditions for $c(k, u, z)$ (like the symmetry condition). Approach (B) on the other hand will lead us to a game of guesses regarding the form of $c(k, u, z)$ without knowing anything about it and hoping that we will get an equation like (5.2.6). On the other hand the advantage of these two algebraic approaches over the numerical (A) is that if we manage to solve the problem

using one of them, it's solution will probably be a more accurate one since it won't have an approximation error (or if it has it should be relatively small).

5.2.d Approach A

The reader can find graphs of $A_{m,i,j}$ in Appendix C for $m = 0, 1, 2, \dots, N$ and for various values of i, j . The reader can notice in the presented graphs that $A_{m,i,j}$ can have different forms for the various values of i, j . At the same time most of these graphs have some specific patterns:

- When $i \neq 1$ we have a sudden drop around $k = 650$.
- There is an upward spike around $k = 550$ for all i, j .

We can also get some interesting insight by examining the changes of the values of $A_{m,i,j}$ in one or two of the three indices while keeping the rest constant. For example we notice that if we keep i or j constant and examine the graphs of $A_{m,i,j}$ for various values of m and j , we can see similar patterns. This could help in understanding the behaviour of $A_{m,i,j}$ and therefore to realize which function would best fit $A_{m,i,j}$.

Unfortunately the dimensionality of the problem combined with the unusual behaviour of the $A_{m,i,j}$ curves for the different i, j 's make the problem really difficult. For the estimation we used a number of algorithms, mainly least-square based. The final results were derived though using the "lsqcurvefit" script of MATLAB 5.3 (which is based on a non-linear optimisation algorithm). We used this script despite the fact that during the course of the estimation we wrote some routines which performed better than "lsqcurvefit", because the difficulties we had in the estimation weren't so much in the algorithms used for it but more in the nature of the data.

We tried to fit a large number of functions to the data $A_{m,i,j}$, with the ones giving the best results being of the form:

$$g_{i,j}(m) = \alpha_{i,j} \cdot (m + \beta_{i,j})^{\gamma_{i,j}} + \delta_{i,j} \cdot \cos(\omega_{i,j} \cdot m + \varphi_{i,j}) + \eta_{i,j}$$

$$p_{i,j}(m) = \alpha_{i,j} \cdot m^4 + \beta_{i,j} \cdot m^3 + \gamma_{i,j} \cdot m^2 + \theta_{i,j} \cdot m + \delta_{i,j} \cdot \cos(\omega_{i,j} \cdot m + \varphi_{i,j}) + \eta_{i,j}$$

$$h_{i,j}(m) = \alpha_{i,j} \cdot e^{\sigma_{i,j} \cdot (m - \beta_{i,j})^2} + \gamma_{i,j} \cdot e^{\tau_{i,j} \cdot (m - \delta_{i,j})^2} + \theta_{i,j} \cdot \cos(\omega_{i,j} \cdot m + \varphi_{i,j}) + \eta_{i,j}$$

Of the above functions only $h_{i,j}$ with $\sigma_{i,j}, \tau_{i,j} < 0$ can give a realistic covariance function. The covariance function should give smaller values as m increases (i.e. when you examine the covariance of bonds priced at dates further apart) which is a property that the polynomials won't have. Therefore $p_{i,j}$ ⁸ is rejected since this polynomial suggests that the covariance function is a second degree polynomial in m . The first function, $g_{i,j}$, will be rejected for the same reason if $\gamma_{i,j}$ is bigger than two, which is the case for most (i, j) . Therefore $g_{i,j}$ is rejected too.

The reader can find in Appendix D, for comparison reasons, the graphs of $h_{i,j}$ for the same values of i, j as the ones plotted in Appendix C. It can be seen that $h_{i,j}$ can in general

⁸Picking a higher order polynomial improves only slightly our results.

replicate most of the patterns that $A_{m,i,j}$ follows. Unfortunately we weren't able to achieve a better fit than the one presented. We should also not that in some cases the errors are very big. The fitting could possibly be improved by adding some more terms in $h_{i,j}$ like another trigonometric function or an exponential, but it is our view that this would lead on one hand to a much more complex problem and on the other hand to over fitting.

Generally speaking we noticed that we had the biggest errors for $m > 650$ while the errors for small m were quite satisfactory - although almost never approaching (except in very few cases) the 5% (or even the 10%) statistical benchmark. The errors actually tended to be bigger than 30% of the actual values which suggests that the fitting - although being the best we managed to get - was still bad. The only relatively satisfactory fit was made for $1 \leq m \leq 300$ where the number of data points with errors less than 15% is "small". In Appendix E the reader can find the number of data points whose error exceeded the 5%, 10%, 15%, 20%, 25% and 30% benchmarks. The reader can also find there two statistical values we examined. These numbers, although they don't mean much in a non-linear optimisation procedure, can help us compare the goodness of the fit across the various values of i, j .

After having fitted the function $h_{i,j}(m)$ to the data we can use (5.2.6) to get an estimate of $f_{k,m,i,j}$. Recall that $f_{k,m,i,j} = \int_k^{k+T_i} \int_{k+m}^{k+m+T_j} c(k, u, z) dz du$. So now we have to figure a way to estimate the covariance function from this equation.

One way would be to work as we did with the drift function, by differentiating with respect to T_j and T_i . Unfortunately we can't do that here because we don't know how $f_{k,m,i,j}$ depends on T_j and T_i . This information is built in the coefficients of $h_{i,j}$. Another way would be to work numerically. Once again though we can't proceed any further in the estimation of the covariance function since in whatever numerical scheme we apply, we have to work for each k separately, which reduces greatly the amount of information we have (notice that we don't necessarily have $c(k, u, z) = c(l, u, z)$ if $k \neq l$). For this to work numerically we need more maturities, i.e. not 10 maturities but N maturities!

After seeing that we can't go any further in the estimation of the covariance function if we try to fit functions $h_{i,j}(m)$ to the $A_{m,i,j}$ data, we decided to work the other way i.e. fitting functions $h_m(T_i, T_j)$ to $A_{m,i,j}$. This method though proves to be even more difficult.

Plotting $A_{m,i,j}$ for various values of m we found it very difficult to find a function that could describe $A_{m,i,j}$'s behaviour. The plots $A_{m,i,j}$ are very similar for m up to 400, but for m bigger than that there are continuous changes in their form. The reader can find some of these graphs in Appendix F.

Even if we assume though that we can fit a function $h_m(T_i, T_j)$ to $A_{m,i,j}$, it would still be difficult to get the covariance function. Specifically we would have:

$$\int_k^{k+T_i} \int_{k+m}^{k+m+T_j} c(k, u, z) dz du \simeq h_m(T_i, T_j) \cdot k + \frac{k(k-1)}{2} \cdot (h_1(T_i, T_j) - h_0(T_i, T_j))$$

and after taking the partial derivatives with respect to T_i and T_j :

$$c(k, k+T_i, k+m+T_j) \simeq D_j D_i h_m(T_i, T_j) \cdot k + \frac{k(k-1)}{2} \cdot D_j D_i (h_1(T_i, T_j) - h_0(T_i, T_j))$$

From this equation we will get the values of $c(u \wedge v, u, v)$ for $u, v = 0, 1, 2, \dots, N + T_{10}$. So fitting a function to this surface will give us the covariance function. All this depends on the

fitting of $h_m(T_i, T_j)$, i.e. if the fitting errors are smalls and $h_m(T_i, T_j)$ being smooth (since we have to estimate $D_j D_i h_m(T_i, T_j)$). Unfortunately it is questionable how well we can fit $h_m(T_i, T_j)$ to $A_{m,i,j}$ when we have just 100 sample points for each m (which basically means that from 100 points we will get estimates for 886^2 points) and especially by examining the plots of $A_{m,i,j}$, which despite their smoothness for m up to 400, for larger m they become very coarse. Moreover the reader should keep in mind that the fitting should result in a symmetric covariance function, something that is not guaranteed by the above procedure.

5.2.e Approach B

The results of Approach A weren't satisfying so we decided to proceed with Approach B. The first step with this approach is to make an assumption on the form of the covariance function and then compare it with equation (5.2.6) to see if this assumption can hold. We list five of the functions we used in Appendix G, along with the calculation of their double integrals. As it can be seen none of these function satisfy (5.2.6).

The next thing we did was to see if any of those functions can approximately satisfy (5.2.6). Indeed we can find conditions so that two of the aforementioned functions can, approximately, satisfy (5.2.6). These are:

$$c_1(k, u, z) = k\alpha e^{-\beta|u-z|}, \beta > 0$$

$$c_2(k, u, z) = k \cdot \alpha \cdot [e^{-\beta(\gamma \cdot u - z + \delta)^2} + e^{-\beta(\gamma \cdot z - u + \delta)^2}] + k \cdot \zeta, \beta > 0$$

This can be seen in the following way (we are going to show that for $c_1(k, u, z)$ although it can be similarly shown for the various approximations of $c_2(k, u, z)$). We show in Appendix G that:

$$\int_k^{k+T_i} \int_{k+m}^{k+m+T_j} c(k, u, z) dz du = k \frac{\alpha}{\beta^2} e^{\beta(T_i-m)} [1 - e^{-\beta T_i}] [1 - e^{-\beta T_j}]$$

if $m \geq T_i$,

$$\int_k^{k+T_i} \int_{k+m}^{k+m+T_j} c(k, u, z) dz du = k \frac{\alpha}{\beta^2} \{ 2\beta \cdot (T_i - m) - [1 - e^{-\beta T_j} (1 - e^{\beta T_i})] \cdot e^{-\beta m} + e^{-\beta T_i} \cdot e^{\beta m} \}$$

if $T_i - T_j \leq m \leq T_i$ and

$$\int_k^{k+T_i} \int_{k+m}^{k+m+T_j} c(k, u, z) dz du = k \frac{\alpha}{\beta^2} \{ 2\beta \cdot T_j - (1 - e^{-\beta T_j}) \cdot e^{-\beta m} + e^{-\beta T_i} (1 - e^{-\beta T_j}) \cdot e^{\beta m} \}$$

if $m \leq T_i - T_j$.

Now let:

$$G_1(m, i, j) = \frac{\alpha}{\beta^2} e^{\beta(T_i-m)} [1 - e^{-\beta T_i}] [1 - e^{-\beta T_j}]$$

if $m \geq T_i$,

$$G_1(m, i, j) = \frac{\alpha}{\beta^2} \{ 2\beta \cdot (T_i - m) - [1 - e^{-\beta T_j} (1 - e^{\beta T_i})] \cdot e^{-\beta m} + e^{-\beta T_i} \cdot e^{\beta m} \}$$

if $T_i - T_j \leq m \leq T_i$ and

$$G_1(m, i, j) = \frac{\alpha}{\beta^2} \left\{ 2\beta \cdot T_j - (1 - e^{-\beta T_j}) \cdot e^{-\beta m} + e^{-\beta T_i} (1 - e^{-\beta T_j}) \cdot e^{\beta m} \right\}$$

if $m \leq T_i - T_j$.

Then $f_{k,m,i,j} = k \cdot G_1(m, i, j)$. Let $T_j \geq T_i$. For equation (5.2.6) to hold approximately, we need:

$$f_{k,m,i,j} \simeq v_{i,j} \cdot k^2 + (V_{m,i,j} - v_{i,j}) \cdot k = k(k-1) \cdot v_{i,j} + k \cdot V_{m,i,j}$$

$$\text{i.e. } V_{m,i,j} \gg (k-1) \cdot v_{i,j} \text{ or equivalently } f_{k,m,i,j} \gg k(k-1) \cdot v_{i,j}.$$

$$\text{i.e. } f_{k,m,i,j} \gg k(k-1) \cdot \frac{1}{2}(f_{1,1,i,j} - f_{1,0,i,j})$$

$$k \cdot G_1(m, i, j) \gg \frac{1}{2}k(k-1) \cdot (G_1(1, i, j) - G_1(0, i, j))$$

$$G_1(m, i, j) \gg \frac{1}{2}(k-1) \cdot \frac{\alpha}{\beta^2} \left\{ -2\beta + [1 - e^{-\beta T_j} (1 - e^{\beta T_i})] \cdot (1 - e^{-\beta}) + e^{-\beta T_i} \cdot (e^{\beta} - 1) \right\}$$

If $T_j < T_i$ we will use the third form of $G_1(m, i, j)$ (i.e. $G_1(m, i, j)$ for $m \leq T_i - T_j$) and working similarly we get:

$$G_1(m, i, j) \gg \frac{1}{2}(k-1) \cdot \frac{\alpha}{\beta^2} (1 - e^{-\beta T_j}) (1 + e^{-\beta(T_i-1)}) \cdot (1 - e^{-\beta})$$

Now we have to calibrate the coefficients α, β in a such a way so that $G_1(m, i, j) \simeq A_{m,i,j}$ (where $A_{m,i,j}$ is the estimator of $V_{m,i,j}$) and at the same time the above equation to hold for $k = 1, 2, \dots, N$. If we can find such coefficients we will have estimated the covariance function.

So now we are going to try to estimate the coefficients of $G_1(m, i, j)$ for $c_1(k, u, z)$ (for all cases) and $G_2(m, i, j)$ for $c_2(k, u, z)$ (for all approximations) and check if we can fit them to $A_{m,i,j}$. The functions $G_1(m, i, j)$ and $G_2(m, i, j)$ can be easily found by taking in mind the results in Appendix G (just go to the double integral of the appropriate covariance function and take $k = 1$). Now as far as the fitting is concerned:

- $c_1(k, u, z)$: The plots of $G_1(m, i, j)$ can be found in Appendix H. If we look at the second plot in Appendix H, and specifically $G_1(m, 3, 1)$, we can see the three different behaviours of $G_1(m, i, j)$ (in general). When $m \leq T_i - T_j$ (so in our case when $m \leq T_i - T_j = 189$) the element $2\beta \cdot T_j - (1 - e^{-\beta T_j}) \cdot e^{-\beta m} \cong 2\beta \cdot T_j - e^{-\beta m}$ dominates. Then for $T_i - T_j \leq m \leq T_i$ (in our graph $189 \leq m \leq 252$) $2\beta \cdot (T_i - m)$ takes over and finally, when $m \geq T_i$ the values of $G_1(m, i, j)$ are too small compared to it's values for smaller m , so it appears as if $G_1(m, i, j) = 0$. It is obvious that $G_1(m, i, j)$ can not be used to fit $A_{m,i,j}$.
- $c_2(k, u, z)$, approx. (i) : While calibrating this function we notice that we get the best results when β take values in the interval $[5 \cdot 10^{-7}, 8 \cdot 10^{-6}]$. These values however don't satisfy the restriction $\beta < 5 \cdot 10^{-8}$ so this approximation has to be rejected. The plots of $G_2(m, i, j)$ using this approximation can be found in Appendix I.
- $c_2(k, u, z)$, approx. (ii)/(iv) : These approximations are not going to work since they will be either exponentially decaying or exponentially increasing in m .

- $c_2(k, u, z)$, approx. (iii) : This approximation won't work either since in both I_C^1 and I_C^2 the two highest powers of the polynomial do not depend on the parameters we pick so the solution will soon blow up.

Theoretically speaking $c_2(k, u, z)$ has the desired property that at each date k the covariance between two bonds maturing at dates u and z will decay exponentially. It is obvious that the covariance function is not stationary. One other property of the above covariance function is that as time passes and k increases the covariance between bonds of different maturities will also increase. We don't find this very logical in general, but if we take two specific maturity dates u and z then $c_2(k, u, z)$ tells us that as we get closer to the maturity of the bonds the covariances between their prices (of those two specific bonds) increases, which sounds fairly logical.

As the reader can see, we weren't successful using this approach either. It seems that (5.2.6) is very hard to be solved. Although it does provide a useful condition that the covariance function of this model should satisfy, the actual estimation of the covariance function is still very hard. Something encouraging is that in both Approach A and B we found out that the best functions are of a similar form. Therefore maybe we should search for the covariance function in a set of functions similar to $k \cdot g(u - z, e^{-(f(u-z))^2})$.

5.2.f Approach B assuming that the covariance function has the form mentioned by Kennedy (1997).

In 1997 Kennedy wrote another paper in the Mathematical Finance journal extending his previous work from 1994. There he demonstrates that if further Markovian and stationarity assumptions are made, there is a unique covariance structure describing the equilibrium measure which depends on just three parameters, namely :

$$c_3(k, u, z) = \sigma^2 \cdot e^{\lambda(k-u \wedge z) - \mu \cdot |u-z|}$$

for some constants $\sigma, \lambda \geq 0$ and $\mu \geq \frac{\lambda}{2}$.

We tested this function and found out that it neither gives good empirical results nor satisfies equation (5.2.6) (notice that it's double integral has no k dependence). Specifically if we take it's double integral we get:

$$\int_k^{k+T_i} \int_{k+m}^{k+m+T_j} c_3(k, u, z) dz du = -\frac{\sigma^2}{\mu \cdot (\mu - \lambda)} e^{-\mu m} (e^{(\mu-\lambda)T_i} - 1)(e^{-\mu T_j} - 1)$$

if $m > T_i$,

$$\begin{aligned} \int_k^{k+T_i} \int_{k+m}^{k+m+T_j} c_3(k, u, z) dz du &= \frac{\sigma^2}{\mu \cdot (\mu - \lambda)} \cdot \left\{ (e^{-\lambda m} - e^{-\mu m})(1 - e^{-\mu T_j}) + e^{(\mu-\lambda)m} \cdot (e^{-\mu T_i} \right. \\ &\quad \left. - e^{-\mu m}) - e^{-\mu T_j} \cdot (e^{(\mu-\lambda)T_i} - e^{(\mu-\lambda)m}) \cdot e^{-\mu m} \right\} \\ &\quad + \frac{\sigma^2}{\lambda} \left(\frac{1}{\mu - \lambda} + \frac{1}{\mu} \right) (e^{-\lambda m} - e^{-\lambda T_i}) \end{aligned}$$

if $m \leq T_i \leq m + T_j$ and

$$\int_k^{k+T_i} \int_{k+m}^{k+m+T_j} c_3(k, u, z) dz du = \frac{\sigma^2}{\mu \cdot (\mu - \lambda)} \cdot \left\{ e^{-\mu m} (e^{-\mu T_j} - 1) - e^{(\mu - \lambda)m} \cdot e^{-\mu T_i} \cdot (e^{(\mu - \lambda)T_j} - 1) \right\} \\ + \frac{\sigma^2}{\lambda} \left(\frac{1}{\mu - \lambda} + \frac{1}{\mu} \right) \cdot e^{-\lambda m} (1 - e^{-\lambda T_j})$$

if $m + T_j < T_i$.

As we said $c_3(k, u, z)$ can't even approximately satisfy (5.2.6). But even empirically, if we examine it's graphs, some of which are presented in Appendix J, we can see that this function doesn't perform well. This can be seen clearly in the last graph of Appendix J.

We should remind the reader that the only assumption we made in the derivation of (5.2.6) was that the first differences of the log-bond prices are stationary. Therefore, combining the above with Kennedy's (1997) Theorem 3.3, leads us to the following result:

If the first differences of the (log) bond prices are stationary, the forward rates won't satisfy both the Markov and stationarity conditions.

This leads us to a very practical result. If somebody is using either Kennedy's model or a model which is included in Kennedy's, he should first check if the first differences of (log) bond prices he is using are stationary. If they are this means that the forward rates won't satisfy both the Markov and stationarity conditions. Therefore a certain number of these models will not perform well with this dataset, either because they may assume the forward rates being Markov and stationary or because they may lead to such forward rates after the calibration. In order to calibrate the model then, one suggestion would be to use Kennedy's model and estimate the unknown coefficients from the formulas presented in the previous sections.

6. Conclusion

In our days, finding a model which can describe the movements of interest rates satisfactory is a necessity for a big number of financial institutions and corporations. The model, apart from being able to reproduce the observed patterns, should be easy and fast to calibrate. Although Kennedy's model is very general, and therefore it should be able to reproduce the observed data, it's calibration seems to be more than a time consuming exercise. Our efforts to calibrate the model, presented in section 5, weren't successful. Assuming that the logs of the bond prices are a stationary series we showed that the covariance can be found as the solution of the following system of equations:

$$f_{k,m,i,j} = a_{i,j} \cdot k^2 + (A_{m,i,j} - a_{i,j}) \cdot k$$

$$c(k, u, z) = c(k, z, u)$$

for $k = 0, 1, 2, \dots, N$, $m = 0, 1, 2, \dots, N - k$, $i, j = 1, 2, \dots, 10$ where:

$$f_{k,m,i,j} = \int_k^{k+T_i} \int_{k+m}^{k+m+T_j} c(k, u, z) dz du$$

$$A_{m,i,j} = \sum_{a=0}^m Cov(p_k^{T_i}, p_{k+a}^{T_j})$$

$$p_k^T = \log(P_{k+1,k+1+T}) - \log(P_{k,k+T})$$

$$a_{i,j} = \frac{1}{2} \cdot (A_{1,i,j} - A_{0,i,j})$$

where $c(k, u, z)$ is the covariance function of the forward rates and $P_{k,t}$ is the the price at time s of a bond paying one unit at time $t \geq s$. This was the main result of the thesis.

We believe that the solution to the above problem will belong in a set of functions of the form $k \cdot g(u - z, e^{-(f(u-z))^2})$. This is where all our estimation efforts seem to point at.

Apart from the above a second result was derived from ours and Kennedy's (1997) results regarding how appropriate is the choice of a model included in Kennedy's framework, based on the stationarity of the first differences of bond data.

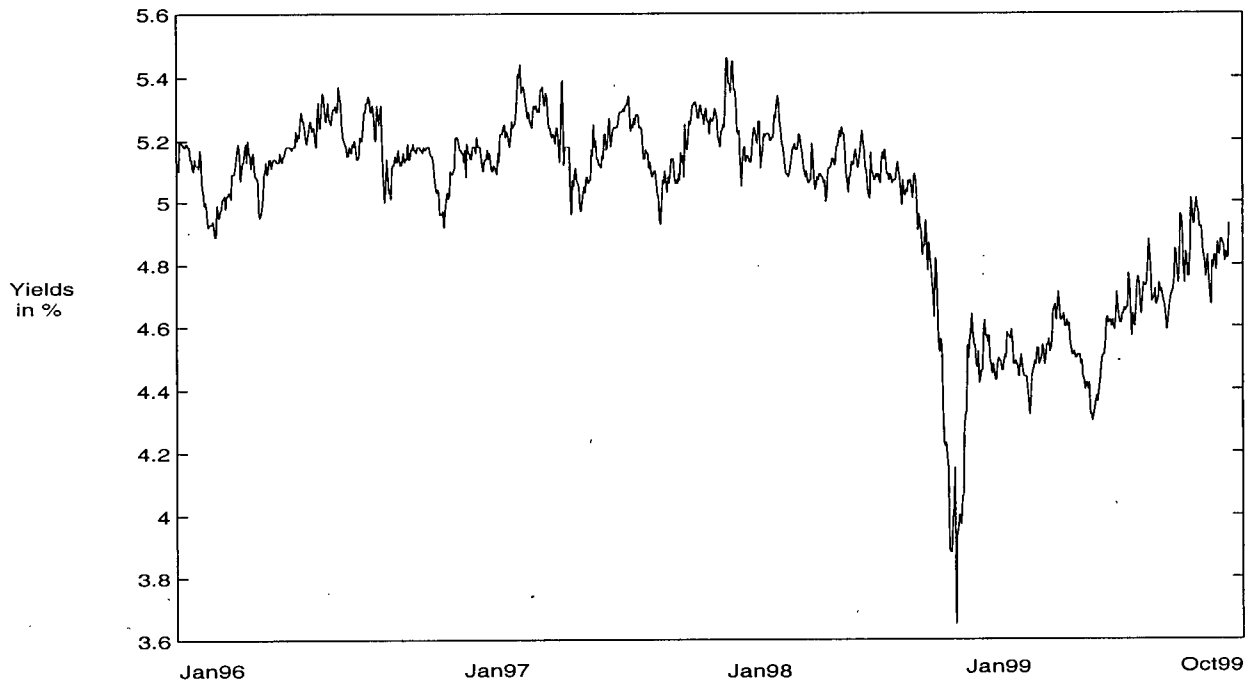
References

- [1] ADLER, R. (1981): "The Geometry of Random Fields", Wiley series in Probability and Mathematical Statistics
- [2] AIT-SAHALIA, Y. (1997): "Do Interest Rates Really Follow Continuous-Time Markov Diffusions?", working paper, University of Chicago.
- [3] BRENNAN M.J. and SCHWARTZ E.S. (1979): "A Continuous Time Approach to the Pricing of Bonds" J. Banking Fin., 3, 133-155.
- [4] CONSTANTINIDES, G.M. (1992): "A Theory of the Nominal Term Structure of Interest Rates" Rev. Fin. Studies, 5, 531-552.
- [5] CONT, R. (1999): "Modelling term structure dynamics: an infinite dimensional approach.", working paper, Ecole Polytechnique, Palaiseau, France.
- [6] COX J.C., INGERSOLL J.E. and ROSS S.A. (1981): "A Theory of the Term Structure of Interest Rates." Econometrica, 53, 385-408.
- [7] DUFFIE D. and KAN R. (1996): "A Yield Factor Model of Interest Rates." Math. Finance, 6, 379-406
- [8] DYBVIG P.H., INGERSOLL J.E. and ROSS S.A. (1996): "Long Forward and Zero Coupon Rates Can Never Fall" Journal of Business, 69, 1-26.
- [9] ENGELN-MULLGES, G. and UHLIG F. (1996): "Numerical Algorithms with Fortran." , Springer.
- [10] GOLDSTEIN, R. (1997): "The Term Structure of Interest Rates as a Random Field", working paper, Ohio State University.
- [11] HEATH, D. and JARA, D. (1998): "Term Structure Models Based on Future Prices" working paper, Carnegie Mellon University.
- [12] HEATH, D., JARROW, R.A. and MORTON, A. (1992): "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation." Econometrica, 60, 77-105.
- [13] HO, T.S.Y. and LEE S. (1986): "Term Structure Movements and the Pricing of Interest Rate Contingent Claims" Journal of Finance, 51, 1011-1029.
- [14] HULL J. and WHITE A. (1990): "Pricing Interest Rate Derivative Securities." Rev. Fin. Studies, 3, 573-592.
- [15] JAMSHIDIAN F. (1988): "The one factor Gaussian Interest Rate Model: Theory and implementation." working paper, Merrill Lynch Capital Markets.
- [16] JAMSHIDIAN F. (1989): "An Exact Bond Option Formula" J. Finance, 44, 205-209.

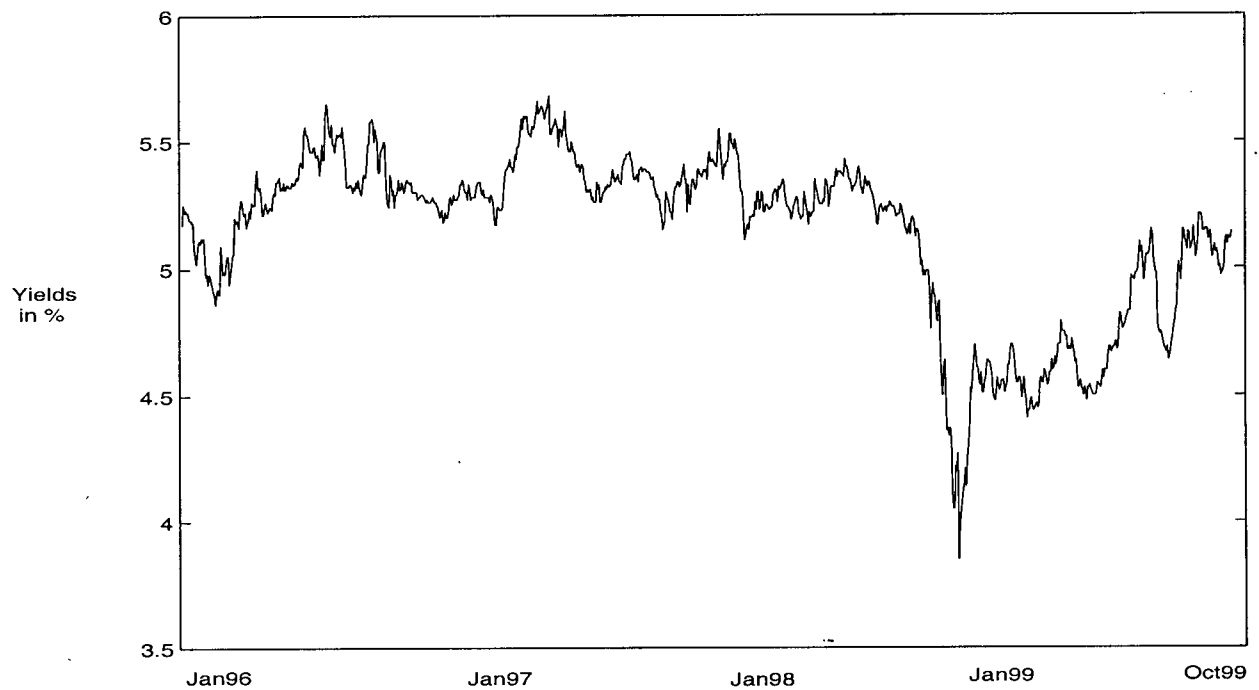
- [17] KENNEDY, D.P. (1994): "The Term Structure of interest rates as a gaussian random field." *Math. Finance*, 4, 247-258.
- [18] KENNEDY, D.P. (1997): "Characterising Gaussian Models of the Term Structure of Interest Rates." *Math. Finance*, 7, 107-118.
- [19] LONGSTAFF F.A. and SCHWARTZ E.S. (1991): "Interest Rate Volatility and the Term Structure: a Two Factor General Equilibrium Model." *J. Finance*, 47, 1259-1282.
- [20] MERTON, R.C. (1973): "Theory of Rational Option Pricing." *Bell J. Econ. Man. Sci.*, 4, 141-183.
- [21] ROGERS, L.C.G. (1995): "Which Model for Term-Structure of Interest Rates Should One Use?" *Math. Finance*, 93-116.
- [22] SQUIRE, W. (1970): "Modern Analytical Computational Methods in Science and Mathematics." , American Elsevier , NewYork.
- [23] VASICEK, O.A. (1977): "An Equilibrium Characterisation of the Term Structure." *J. Fin. Econ.*, 5, 177-188.

Appendix A

- 3-month Treasury yields



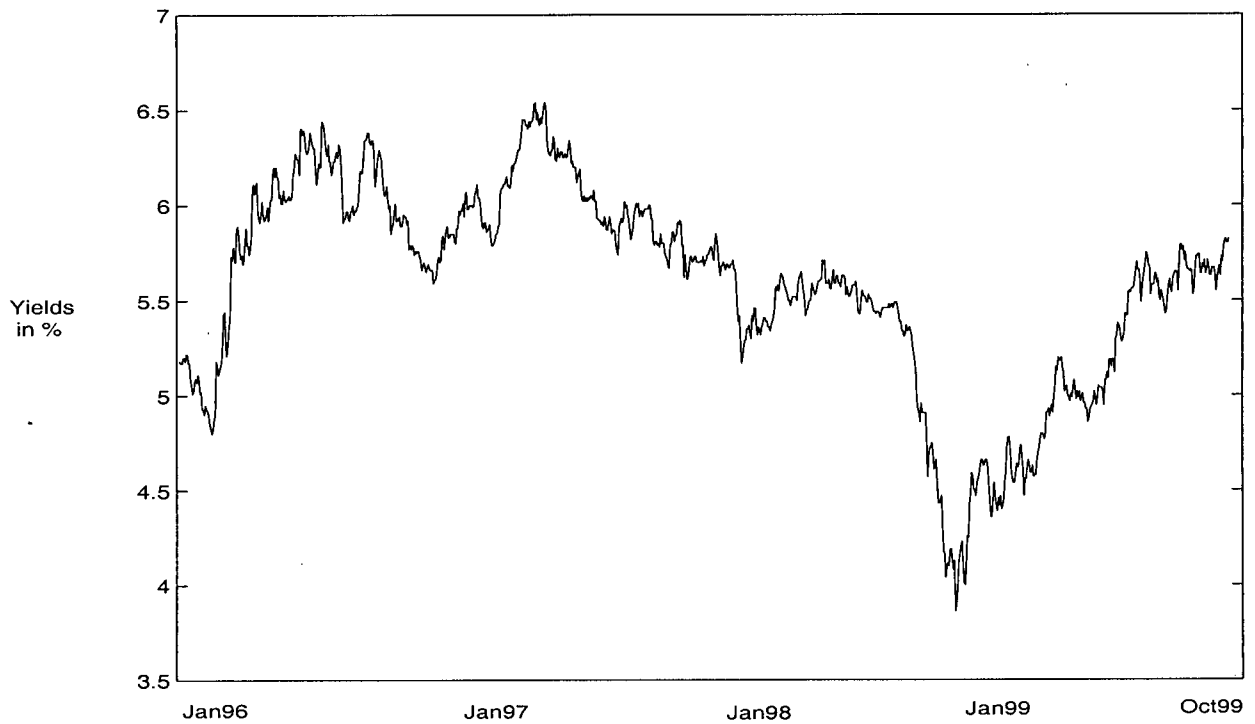
- 6-month Treasury yields



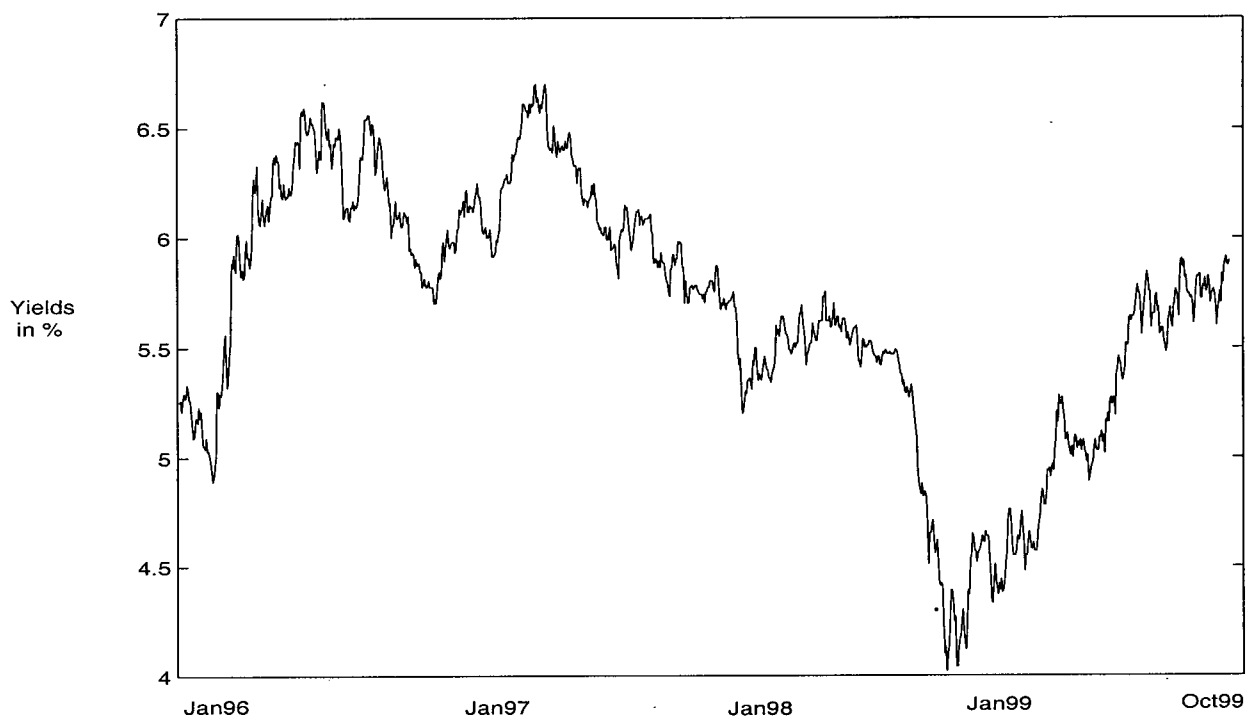
- 1-year Treasury yields



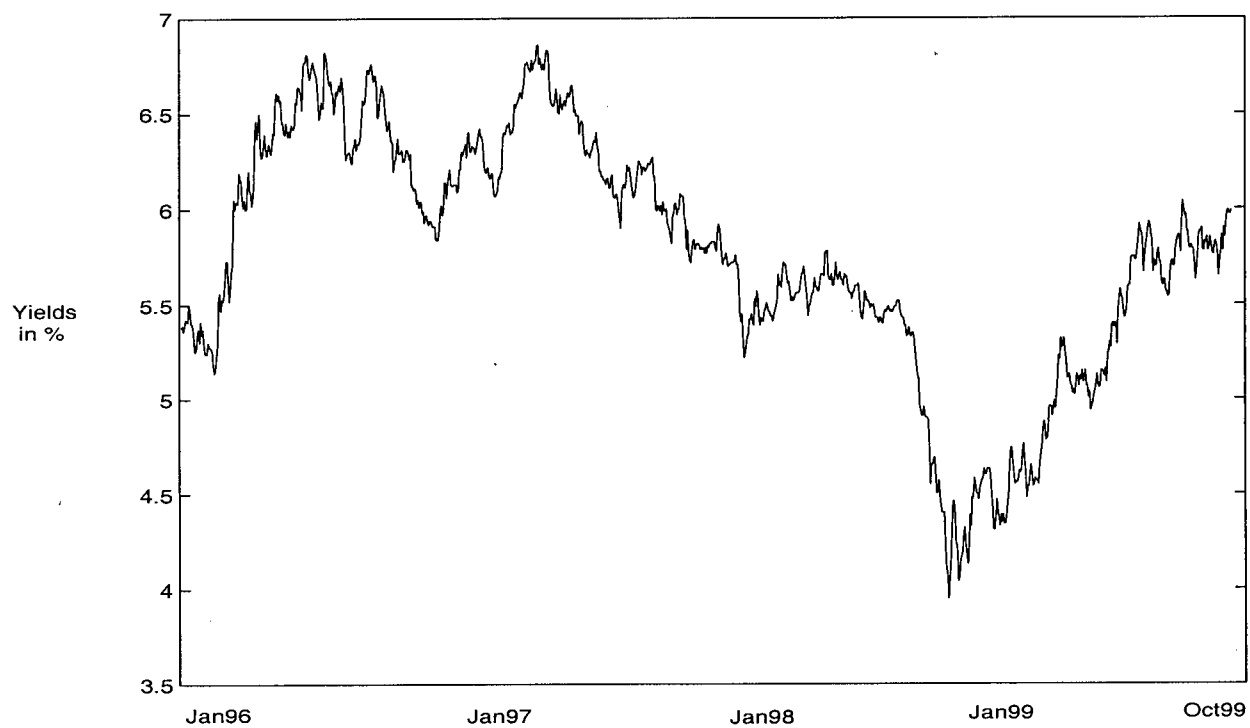
- 2-year Treasury yields



- 3-year Treasury yields



- 5-year Treasury yields



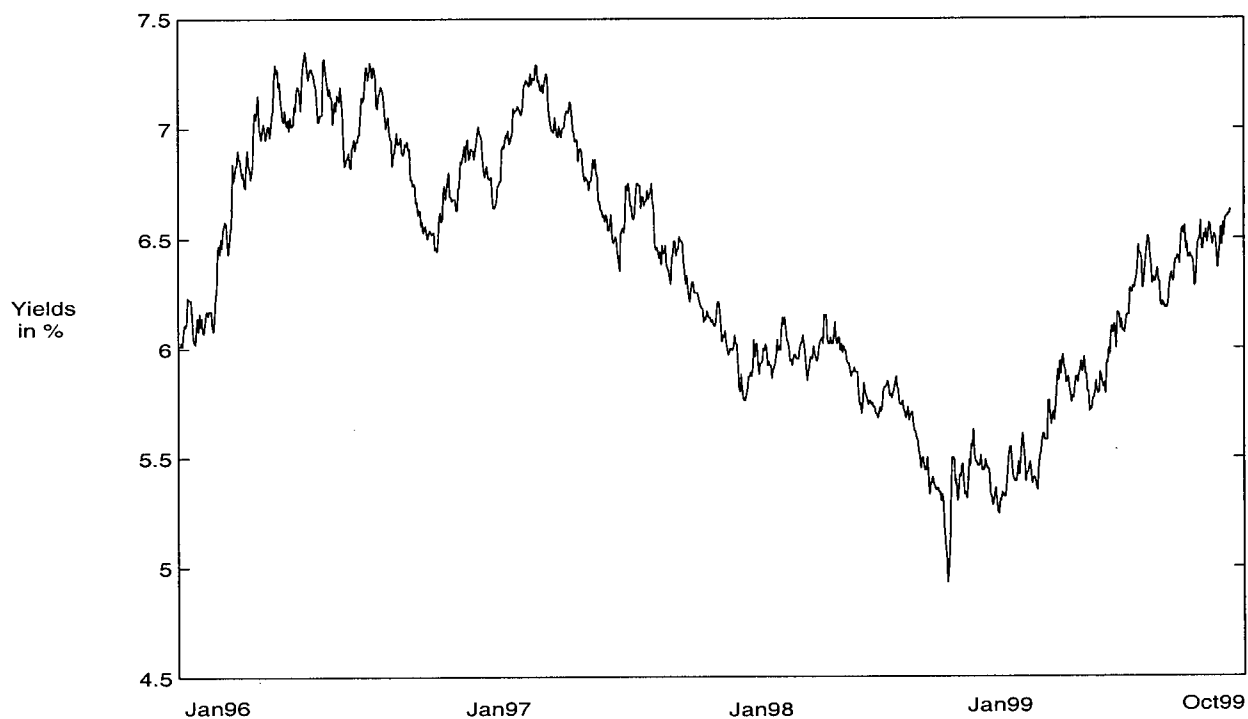
- 7-year Treasury yields



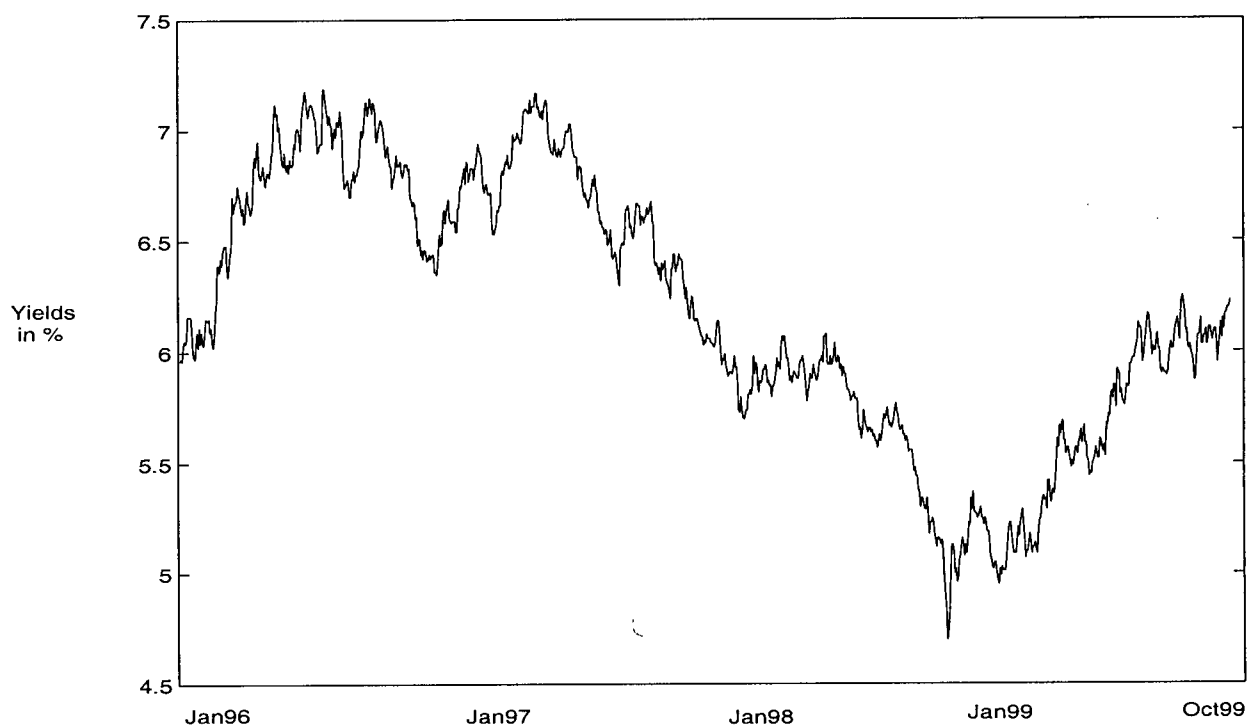
- 10-year Treasury yields



- 20-year Treasury yields



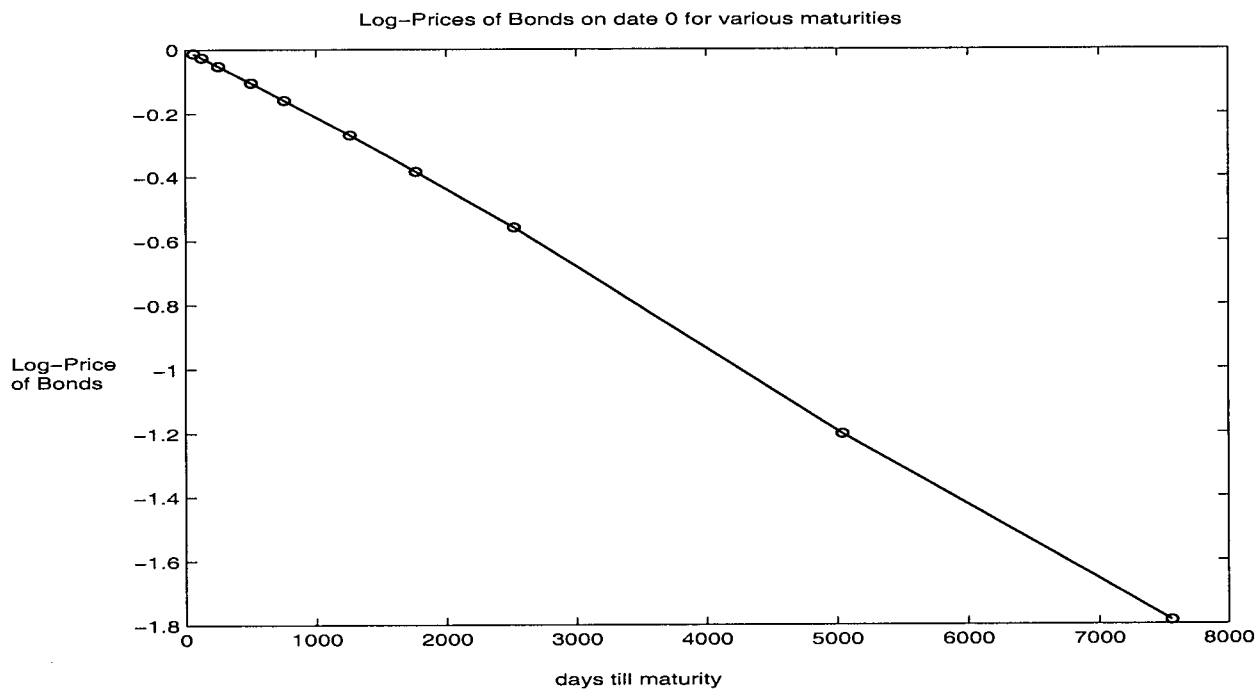
- 30-year Treasury yields



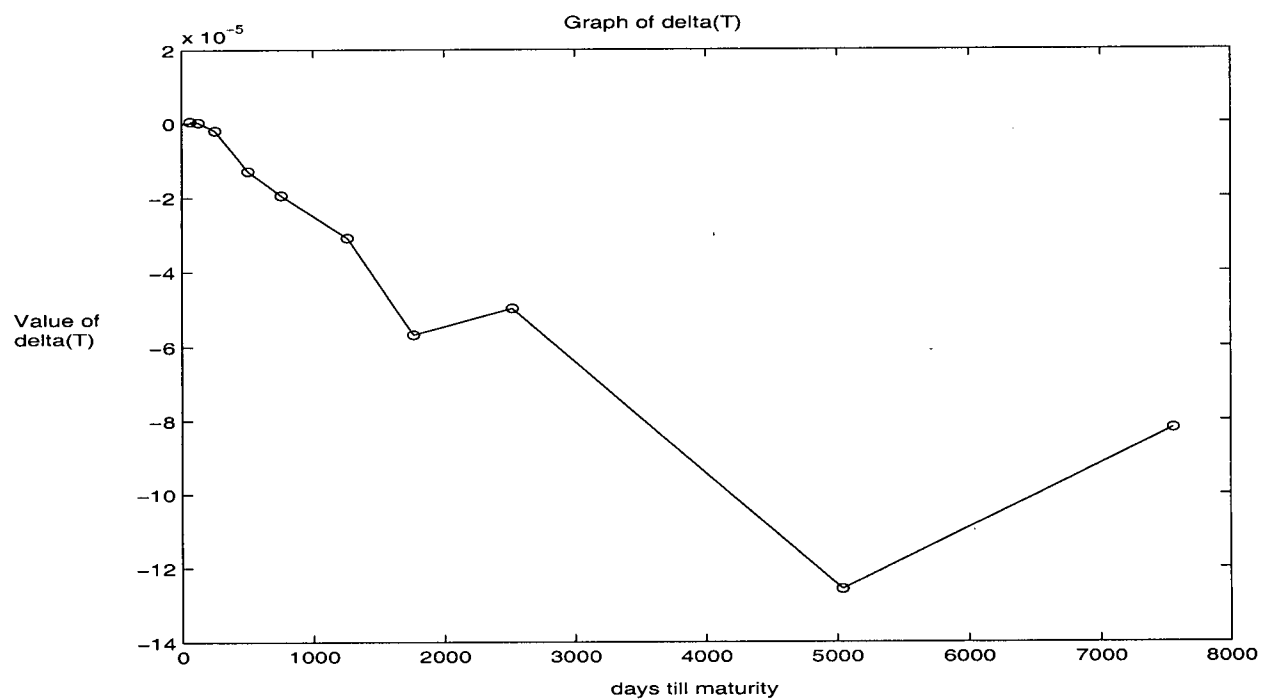
Appendix B

Estimation of the drift function.

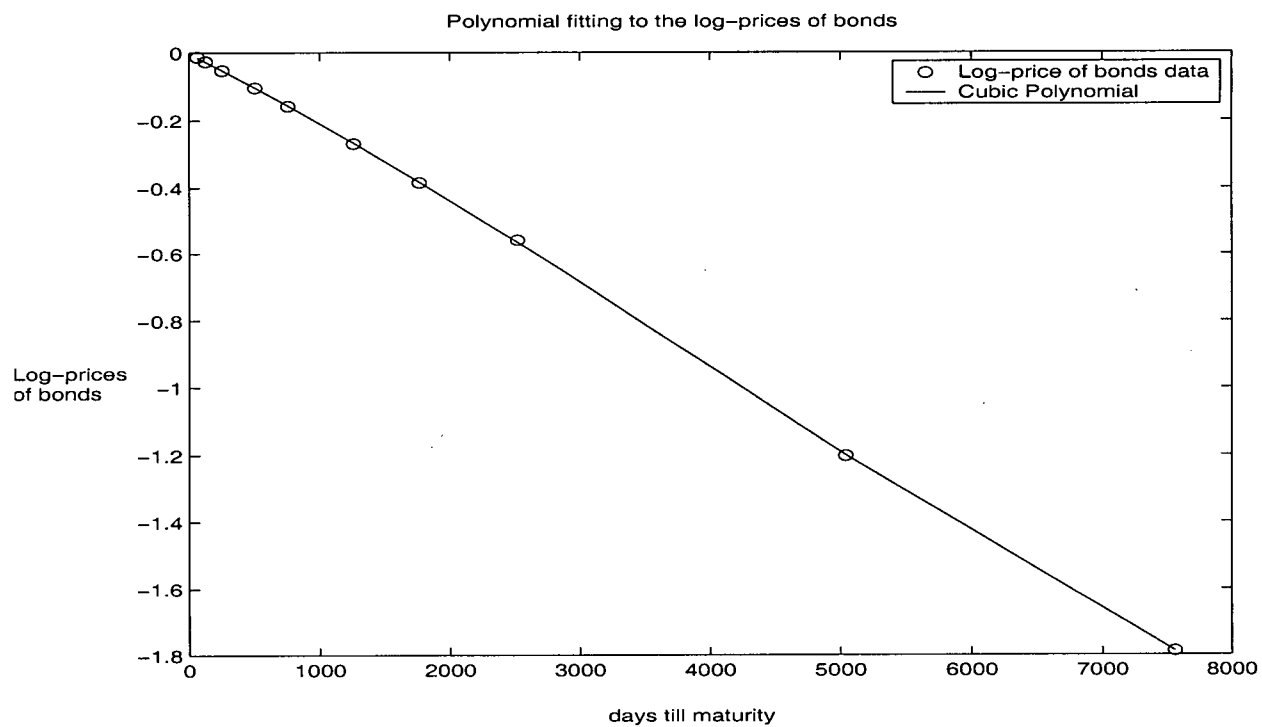
- Figure B1



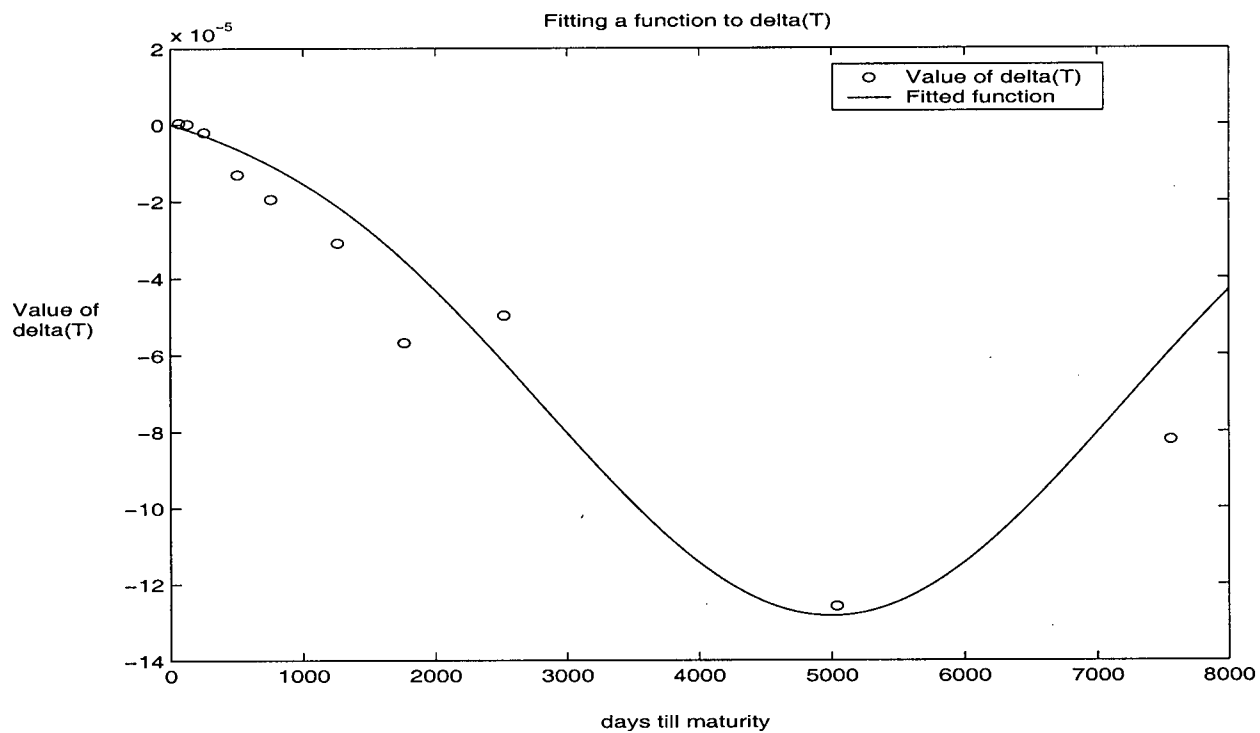
- Figure B2



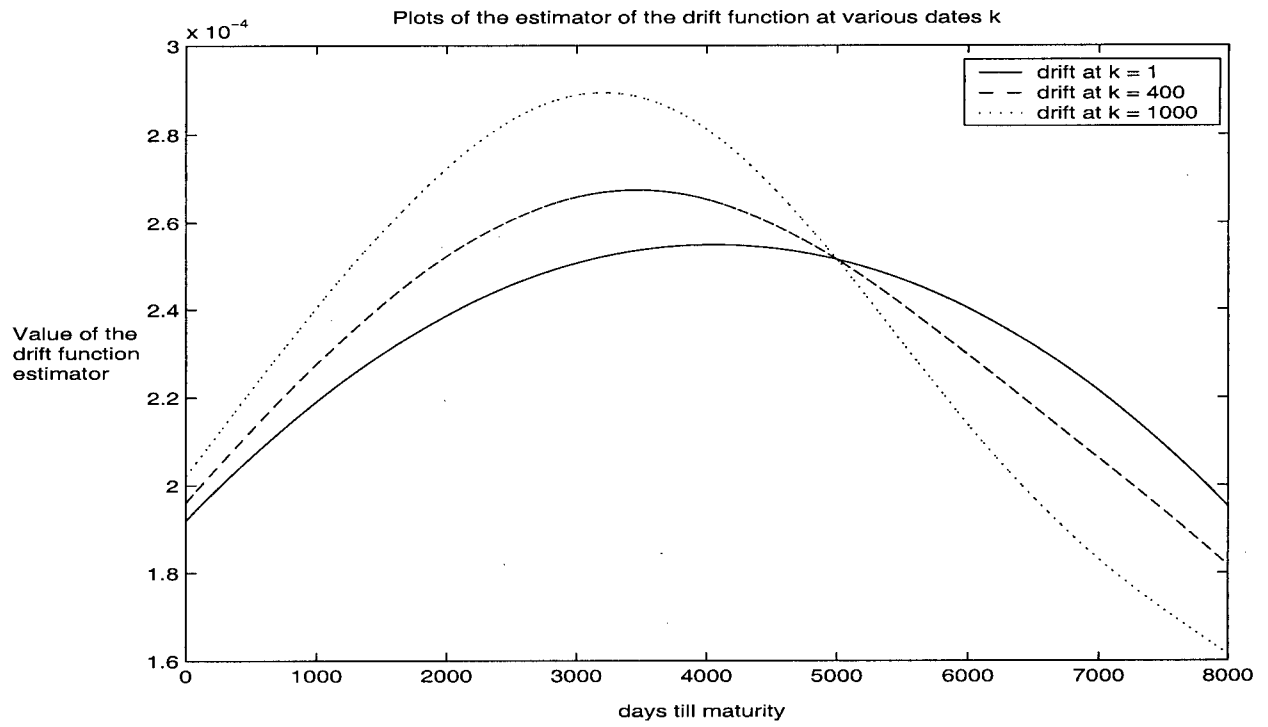
• Figure B3



• Figure B4

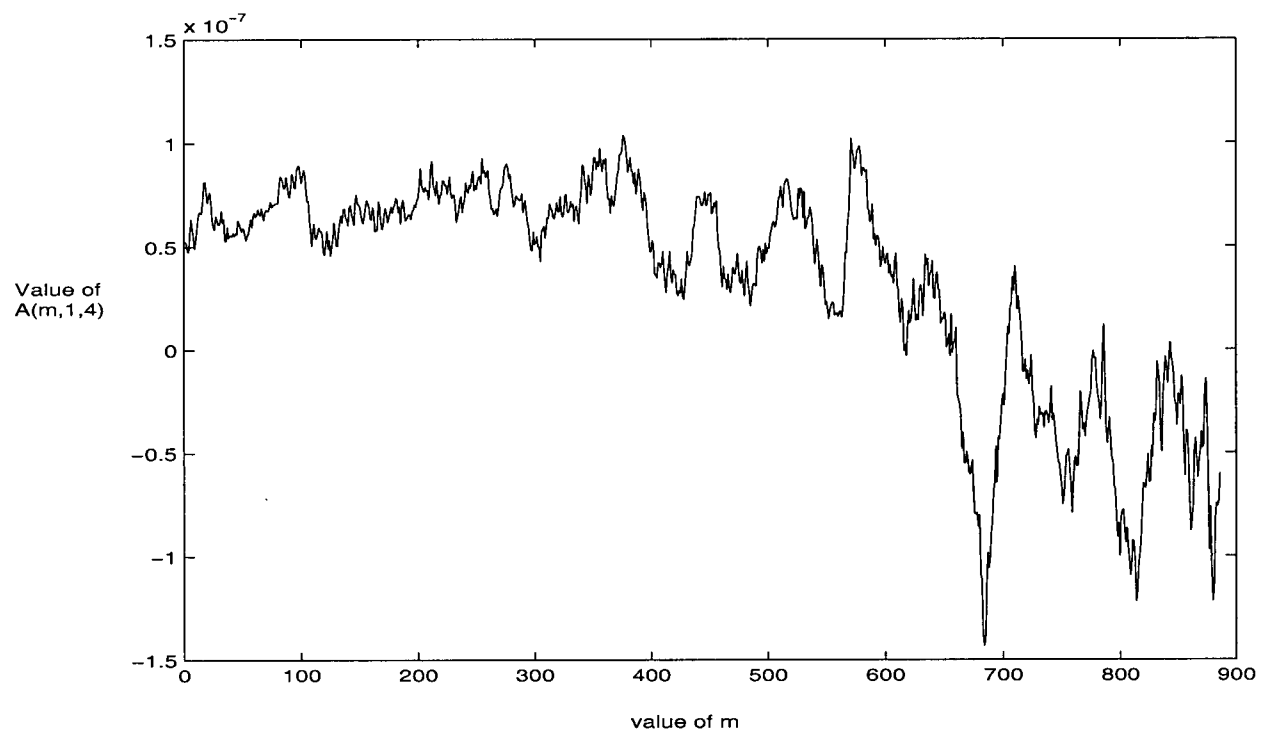
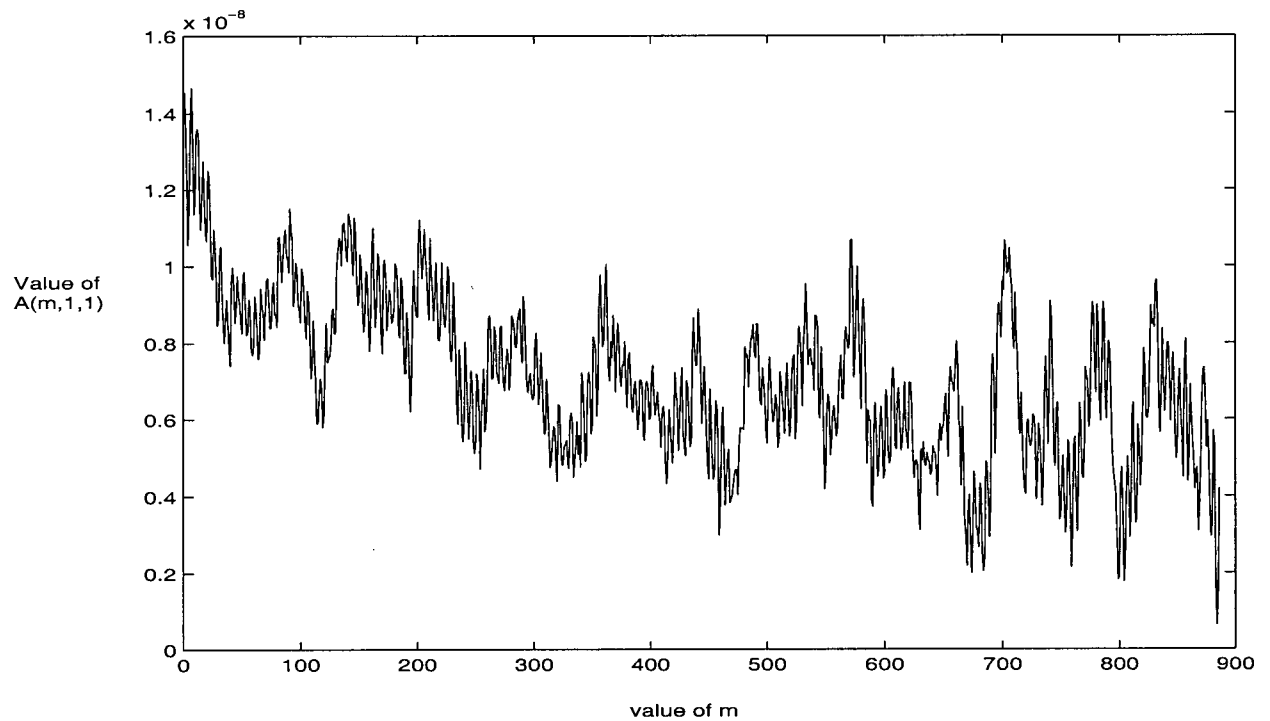


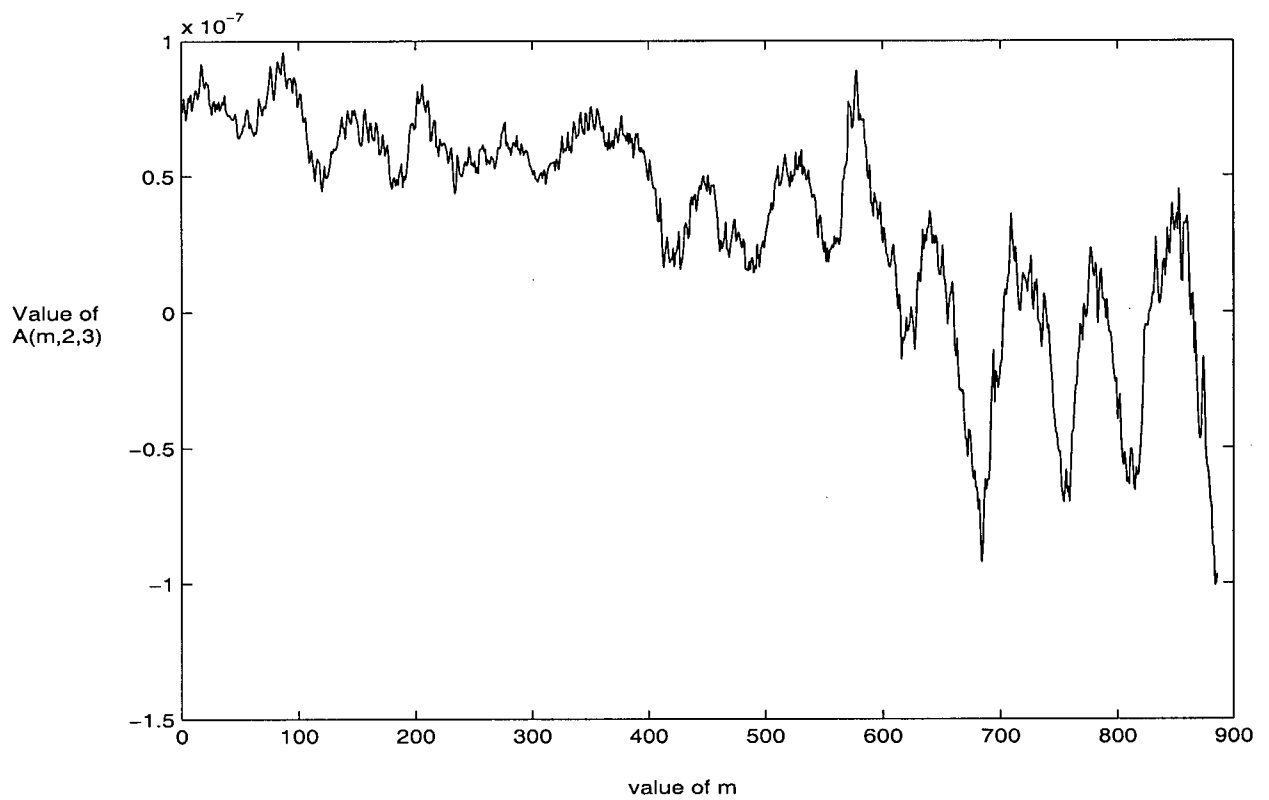
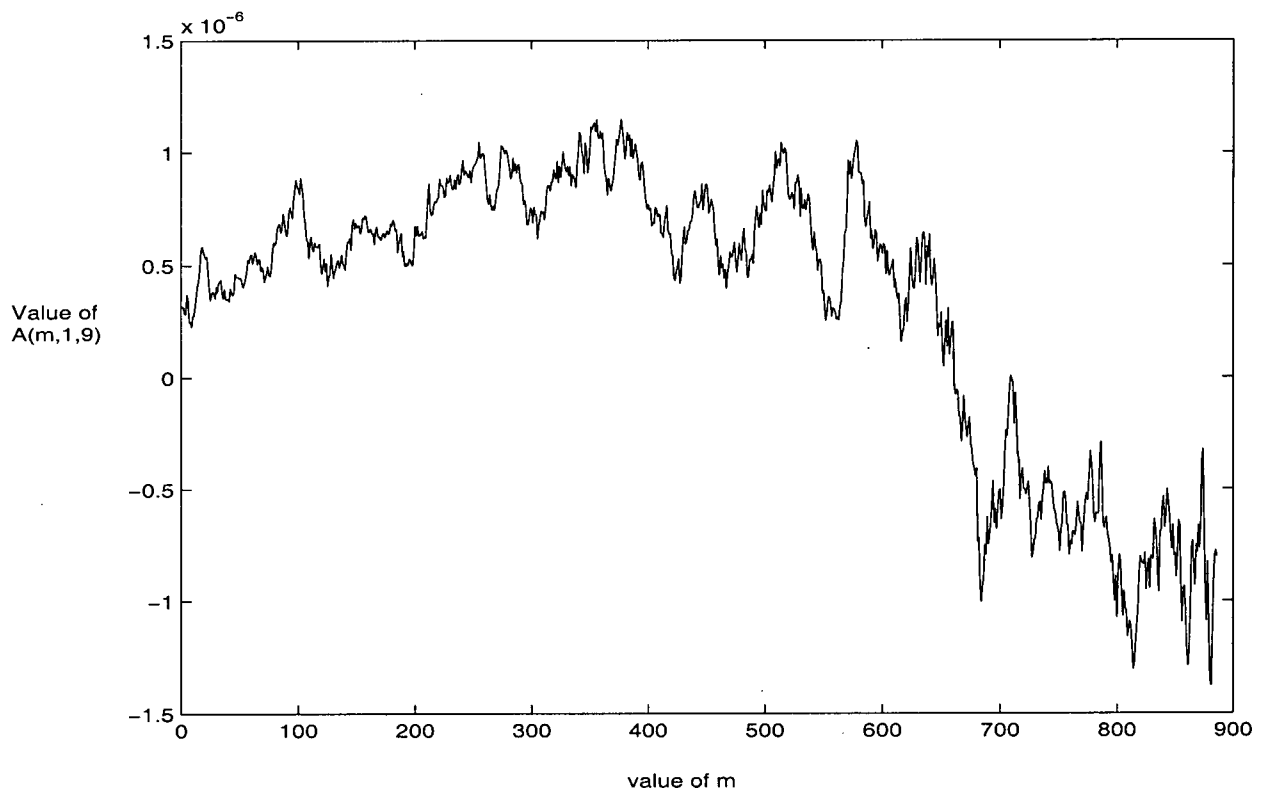
• Figure B5

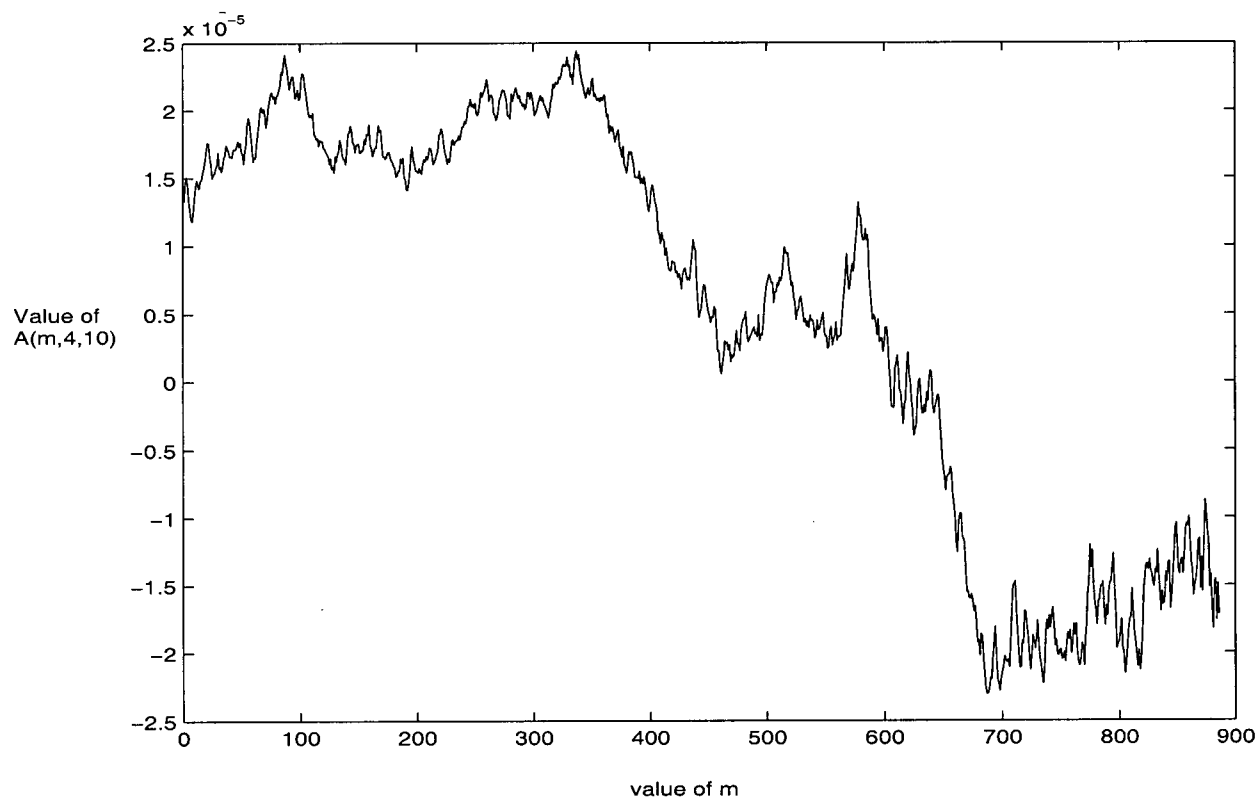
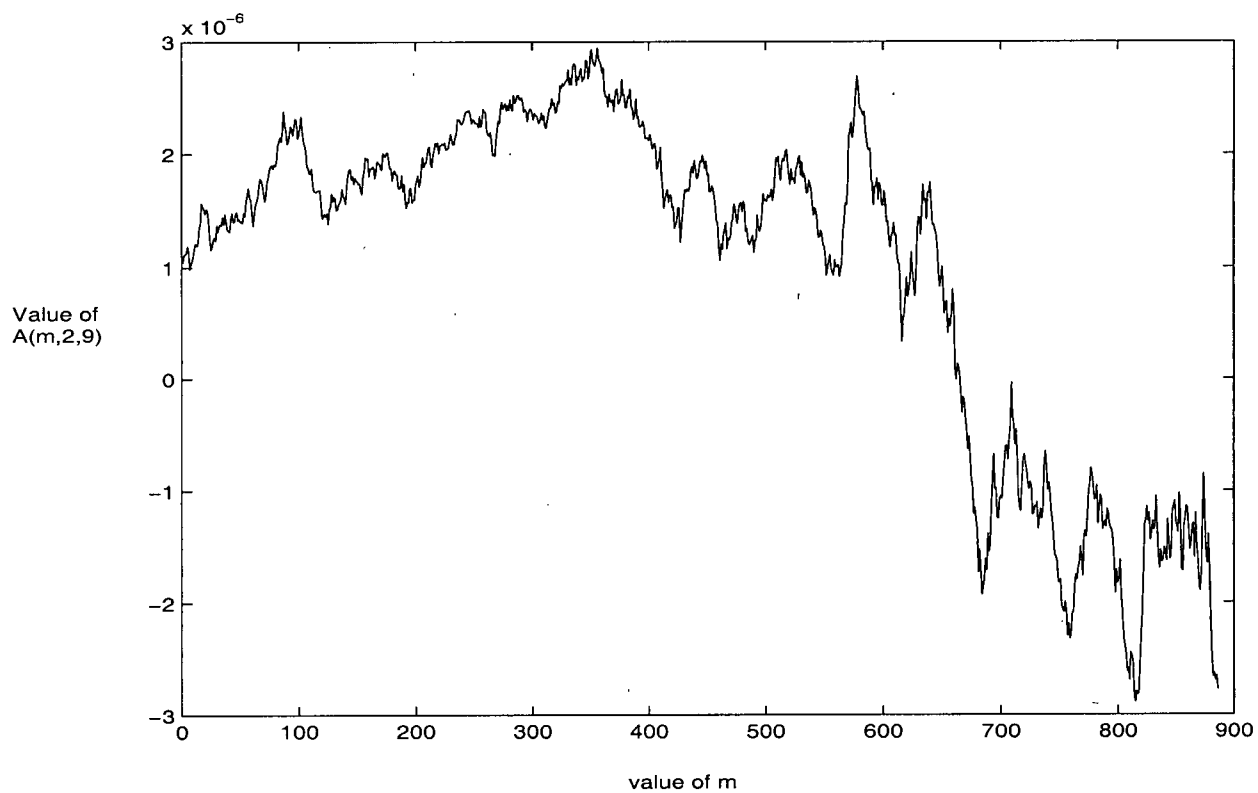


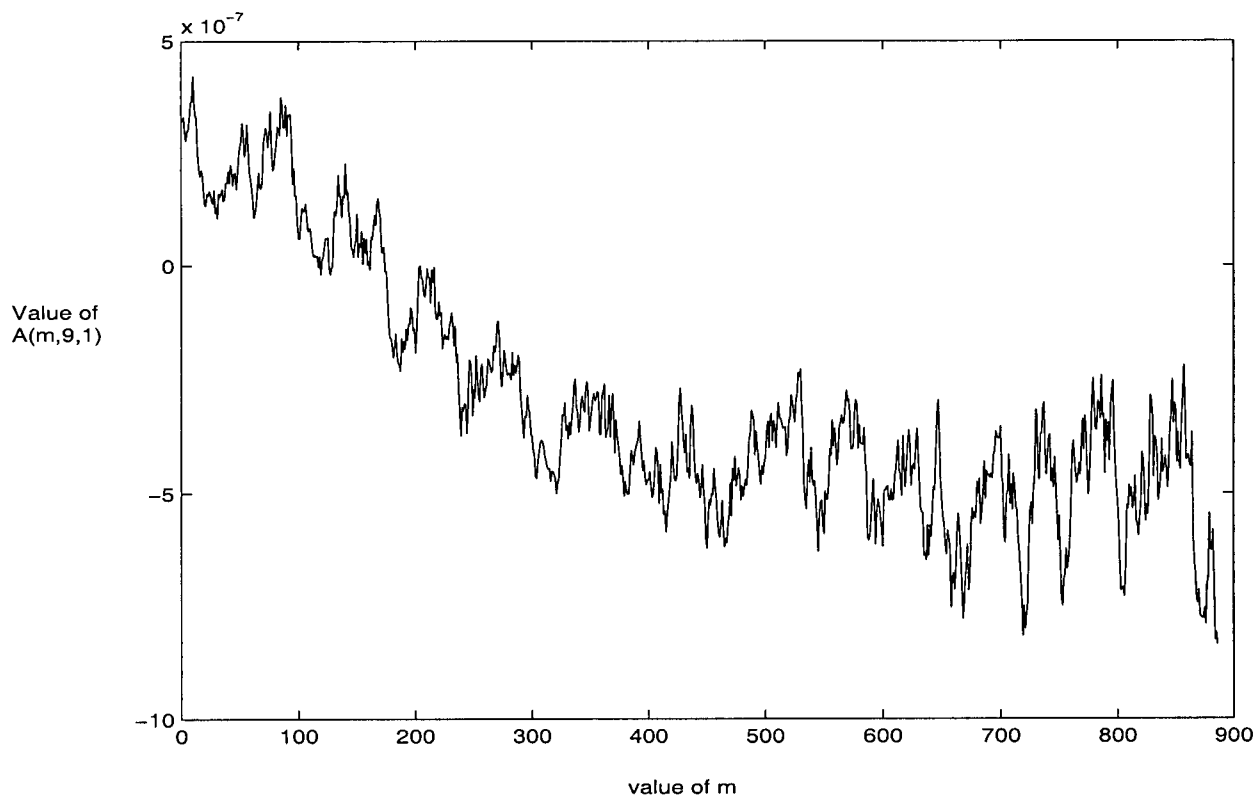
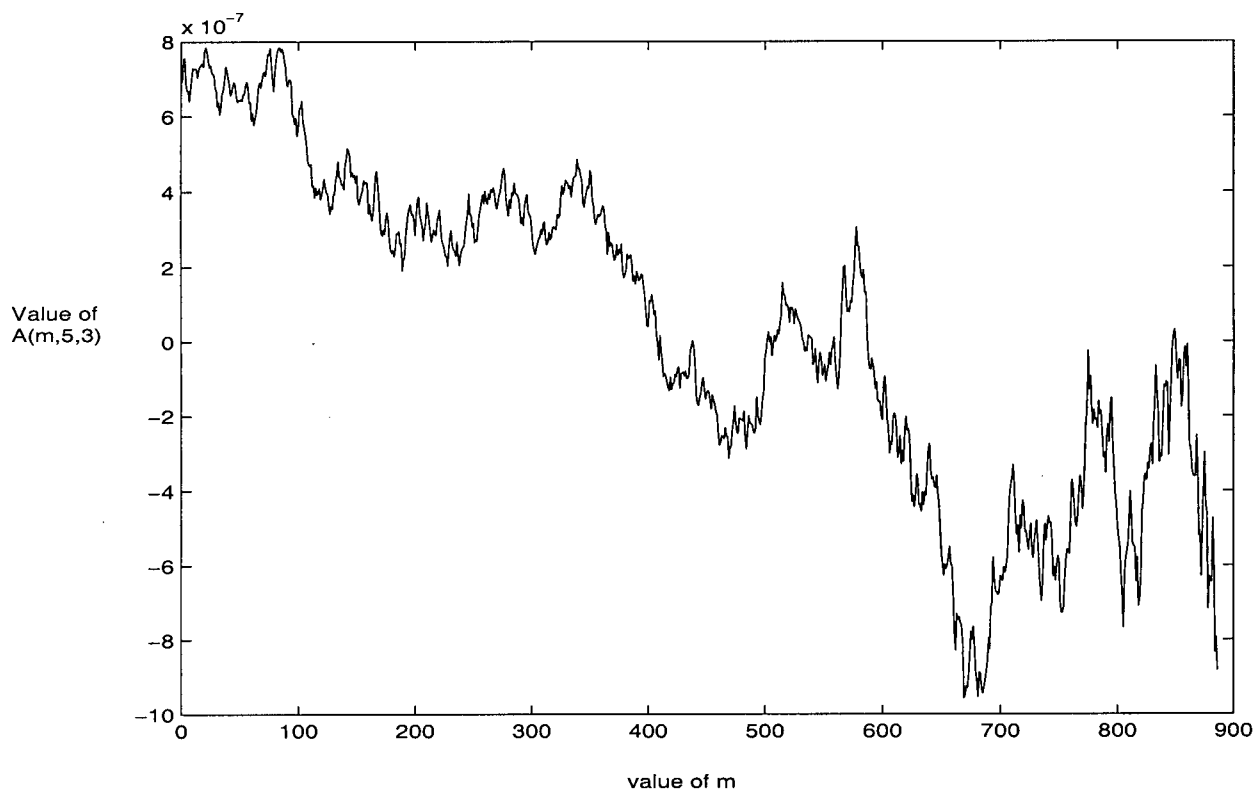
Appendix C

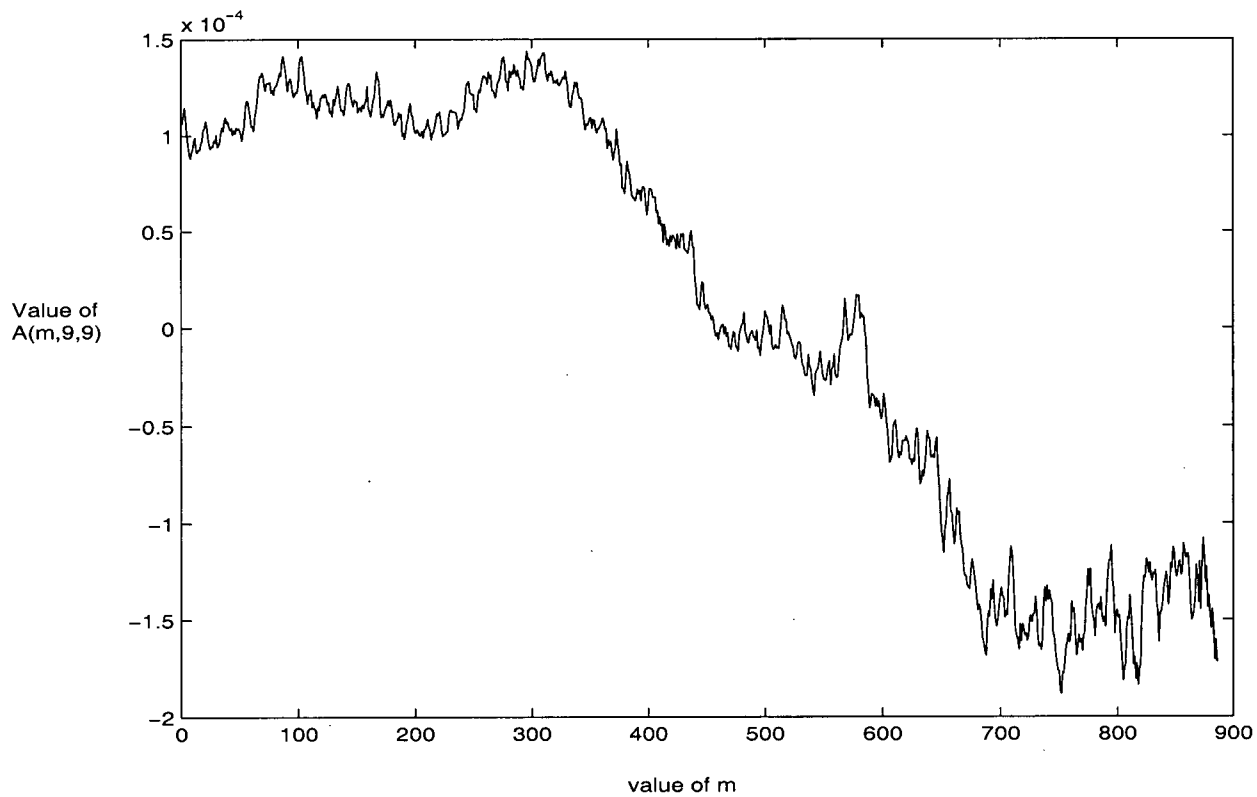
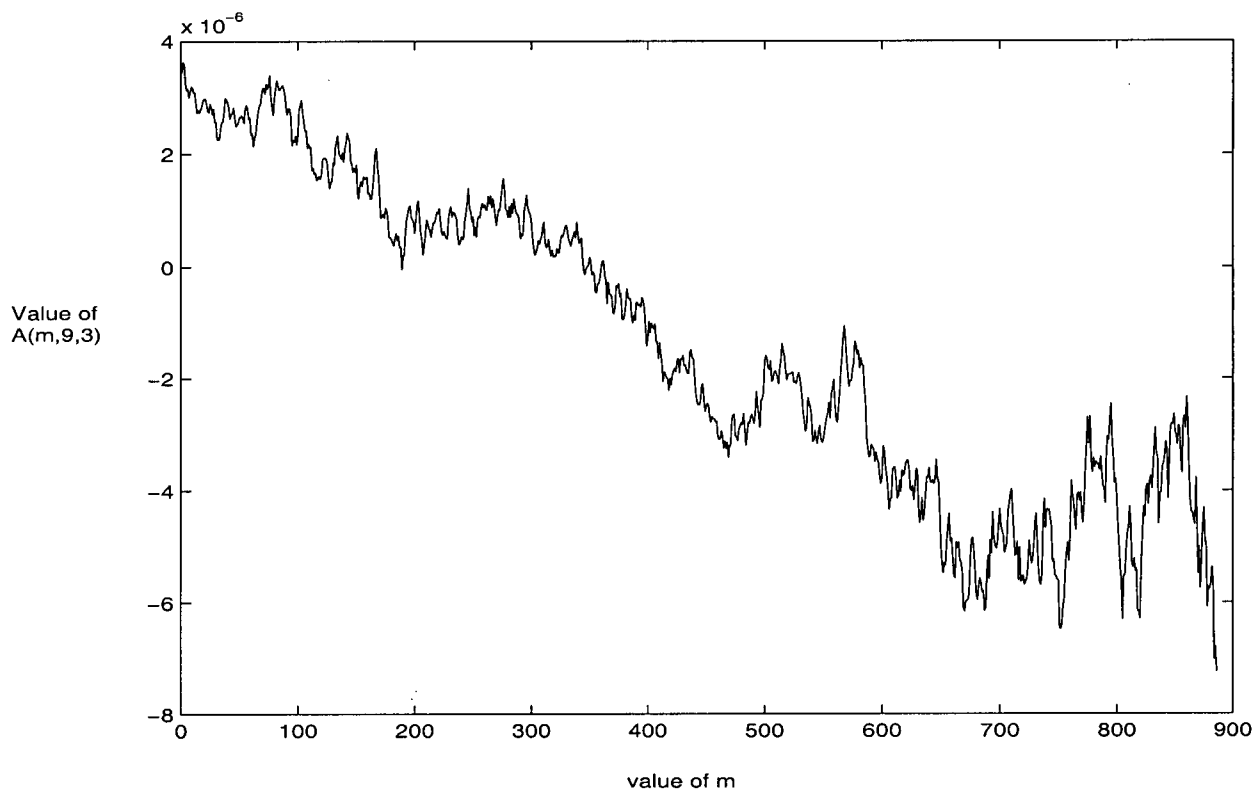
Indicative plots of $A_{m,i,j}$ for certain i, j . On the x-axis we present the values of m while on the y-axis the values of $A_{m,i,j}$ for those m .







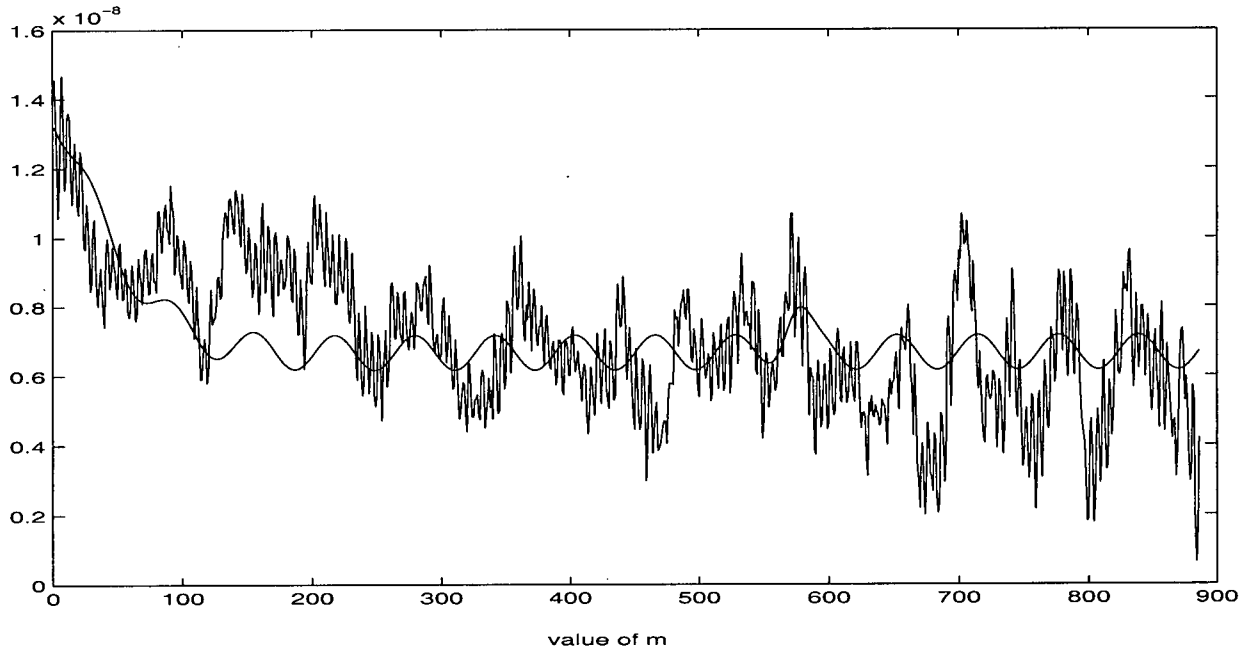




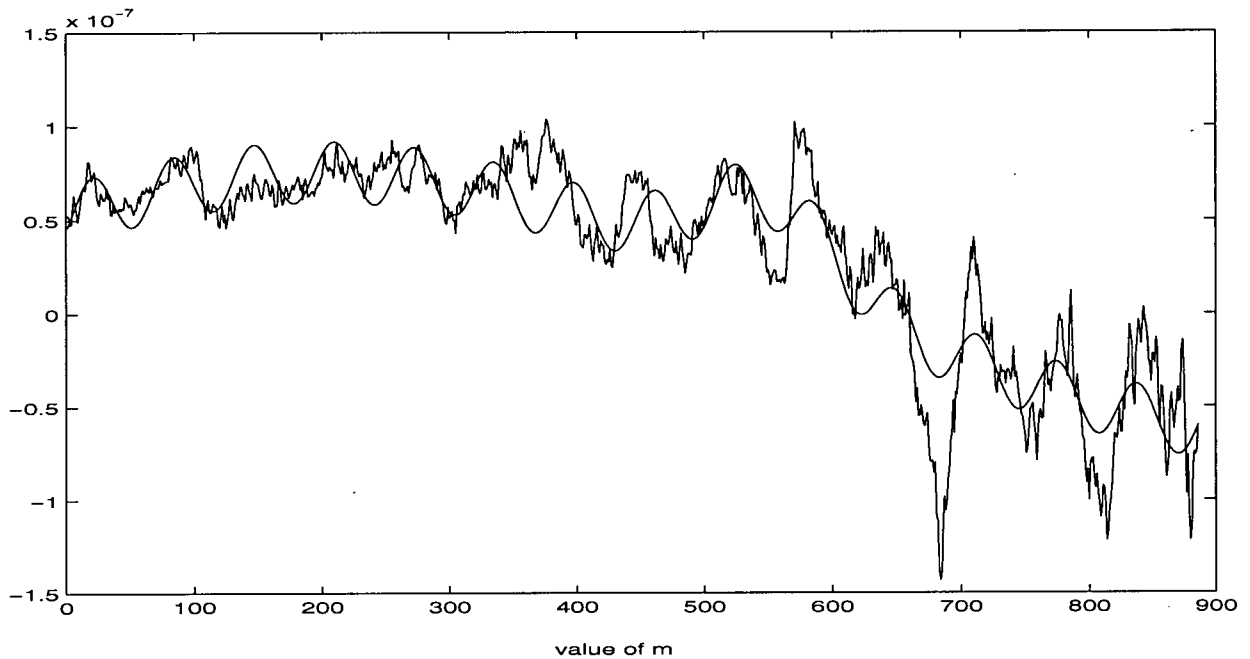
Appendix D

Indicative plots of $h_{i,j}(m)$ and $A_{m,i,j}$ for certain i,j . On the x-axis we present the values of m while on the y-axis the values of $h_{i,j}(m)$ and $A_{m,i,j}$ for those m . You can easily tell apart the graph of $h_{i,j}(m)$ from it's smoothness.

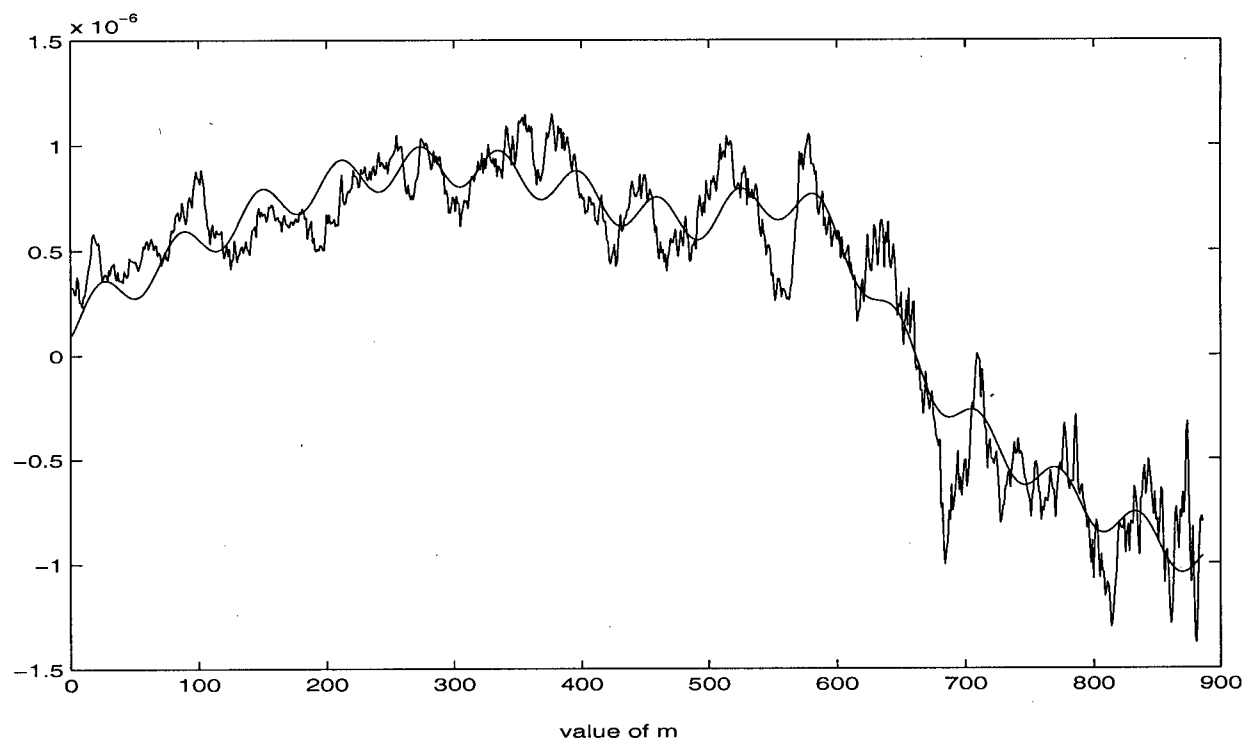
- $h_{1,1}(m)$ and $A_{m,1,1}$



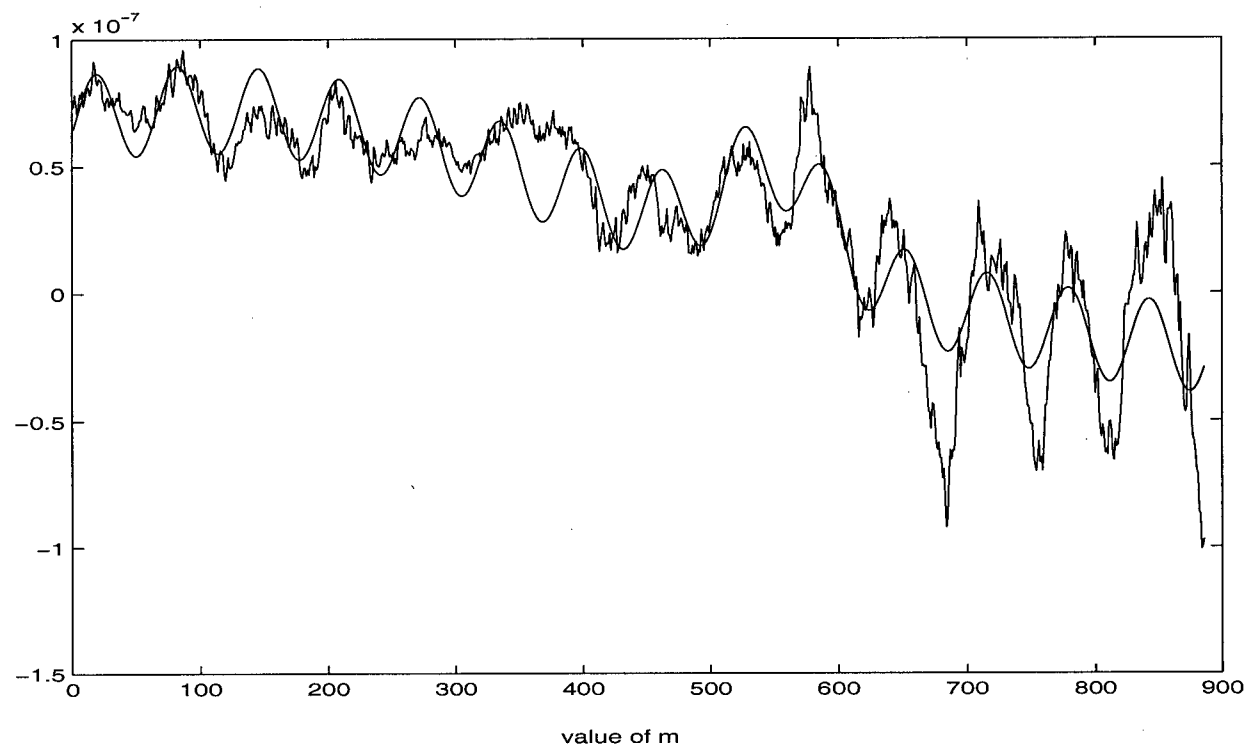
- $h_{1,4}(m)$ and $A_{m,1,4}$



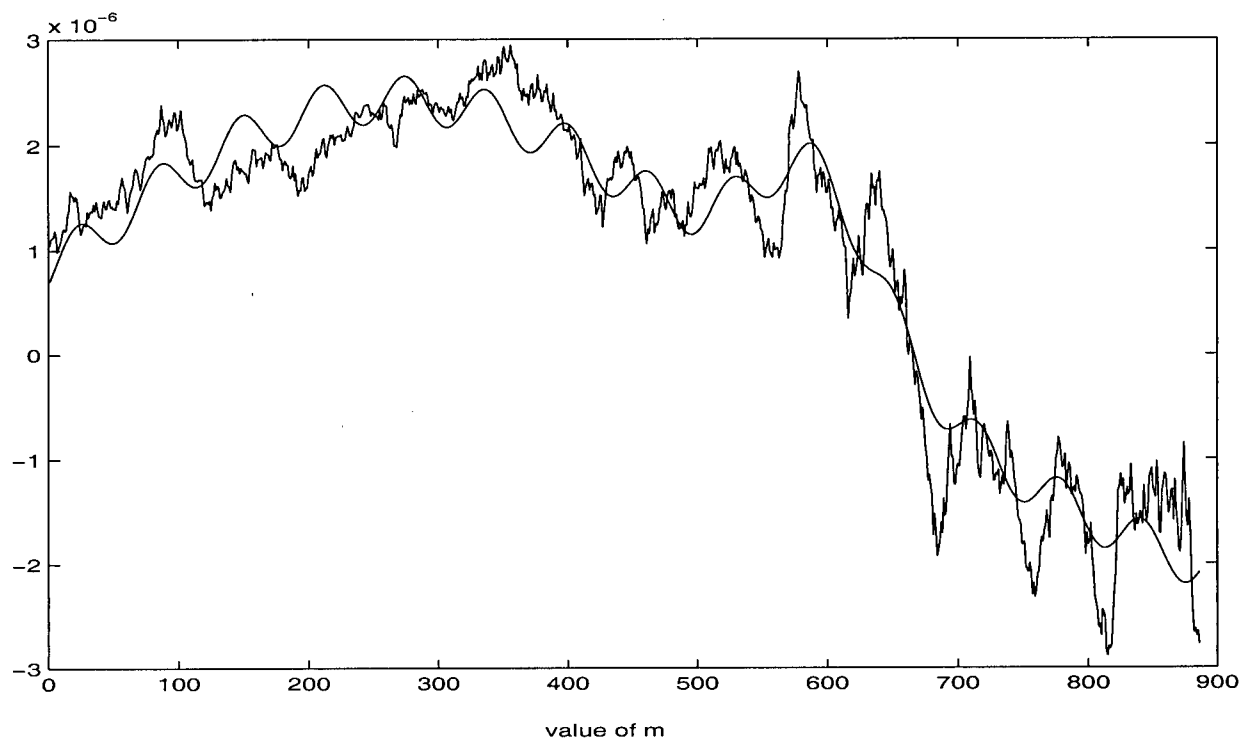
• $h_{1,9}(m)$ and $A_{m,1,9}$



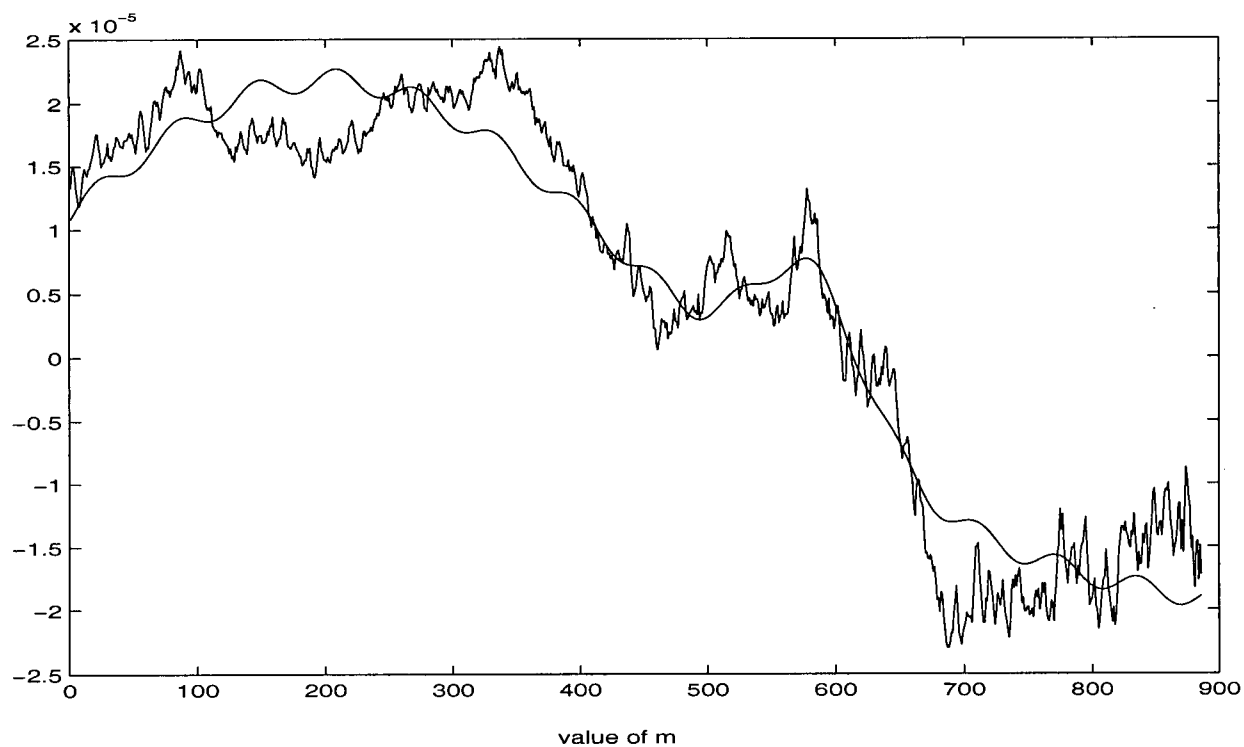
• $h_{2,3}(m)$ and $A_{m,2,3}$



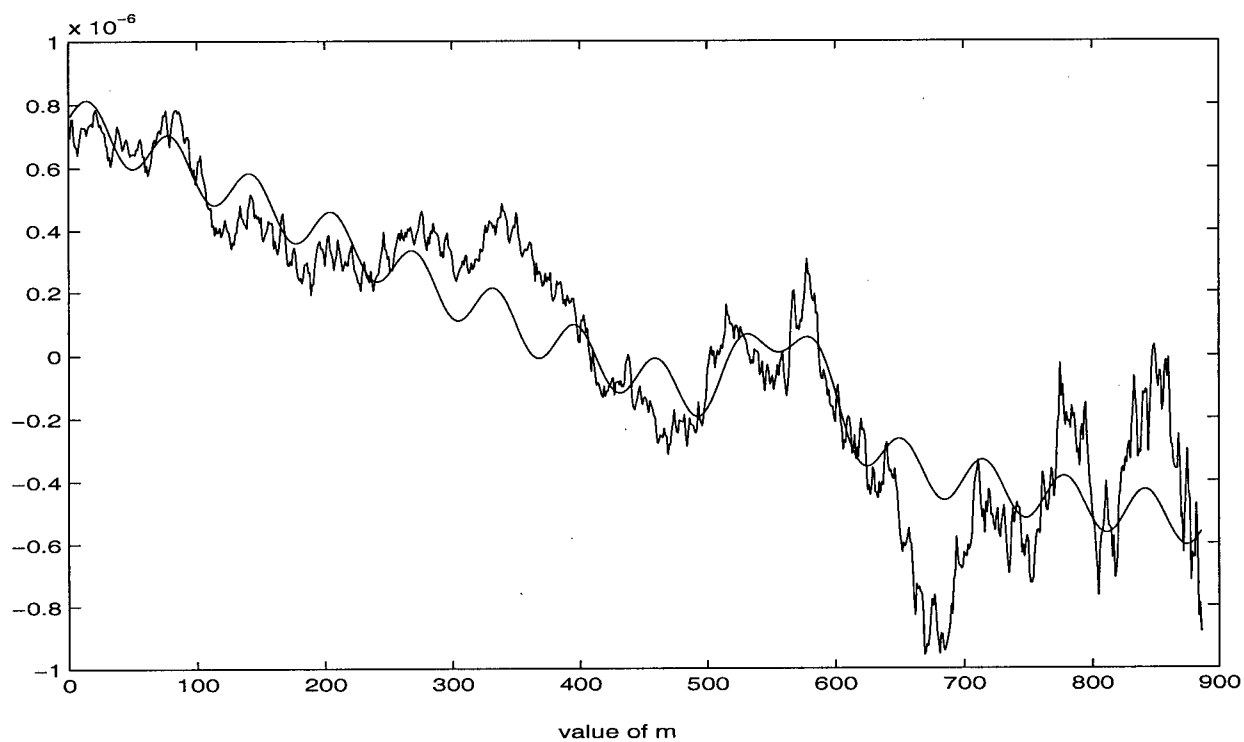
• $h_{2,9}(m)$ and $A_{m,2,9}$



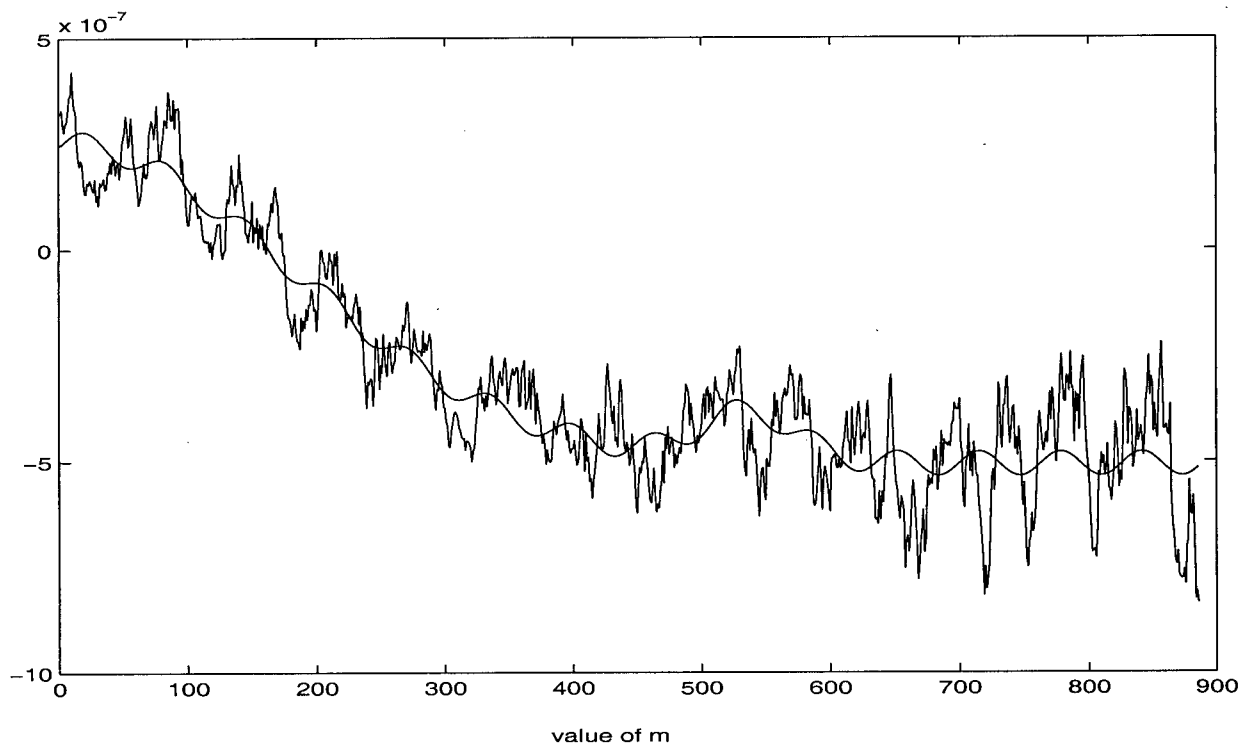
• $h_{4,10}(m)$ and $A_{m,4,10}$



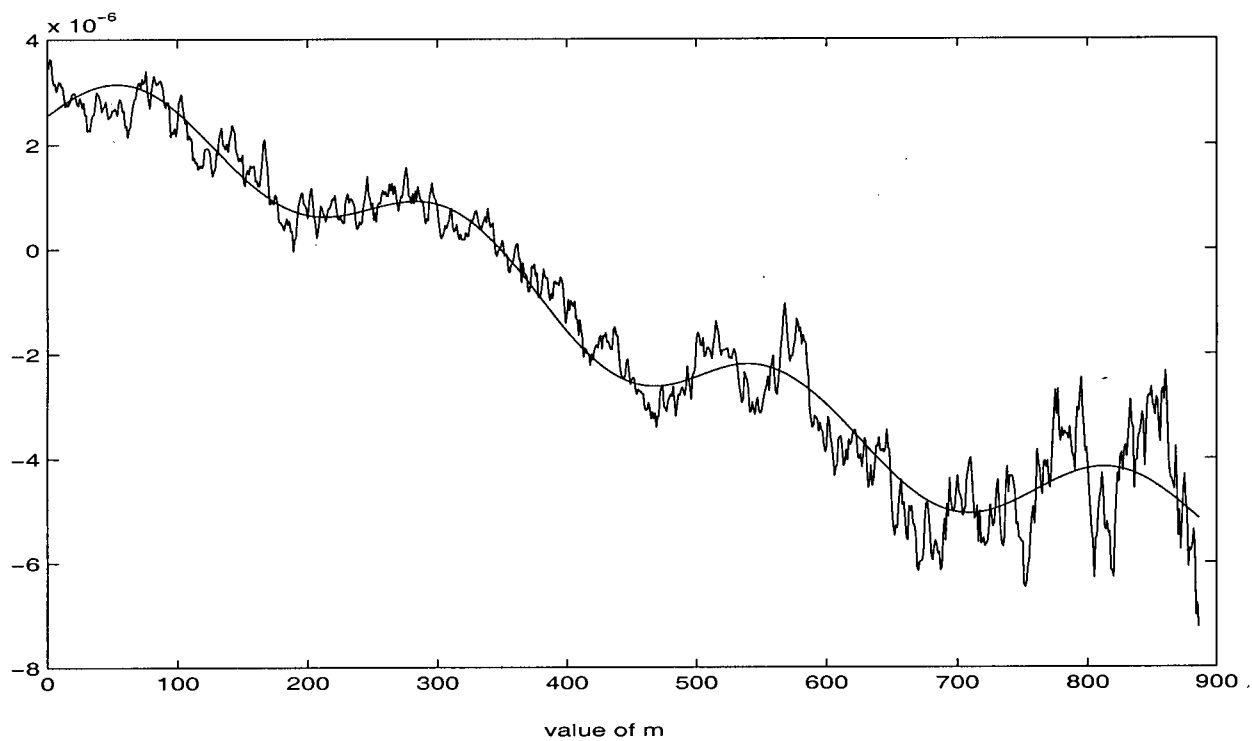
• $h_{5,3}(m)$ and $A_{m,5,3}$



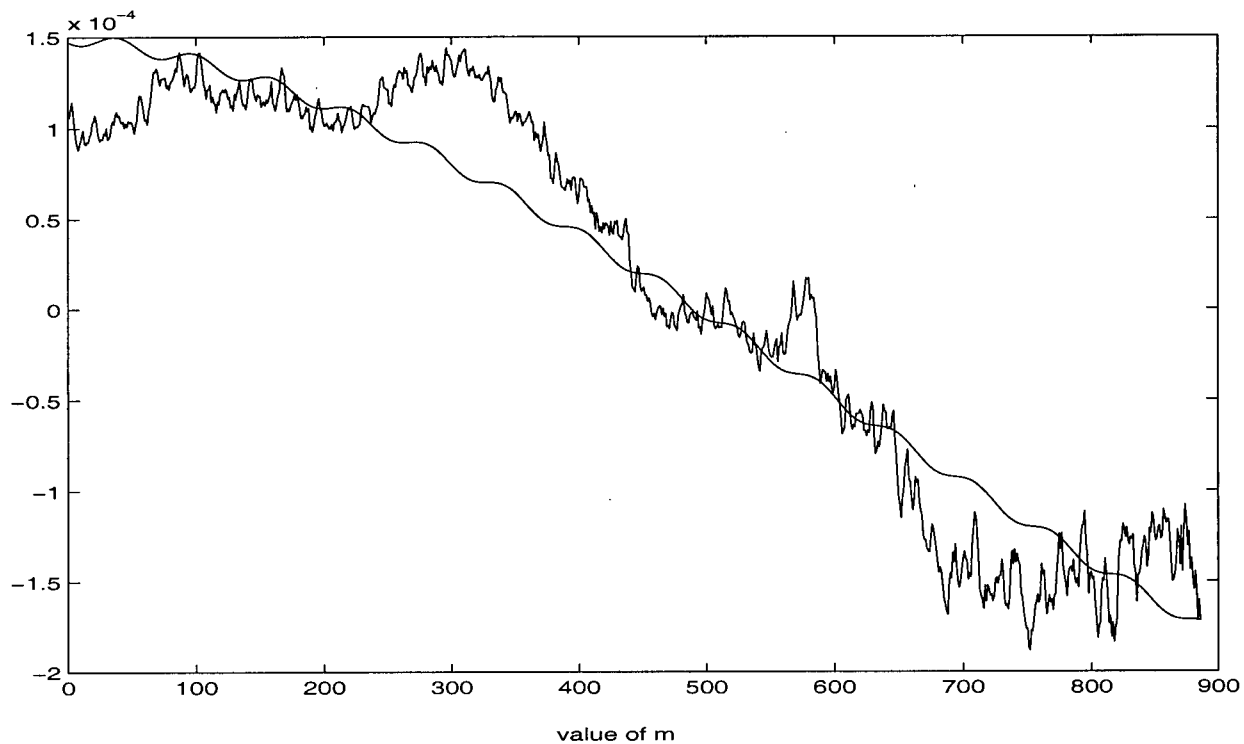
• $h_{9,1}(m)$ and $A_{m,9,1}$



• $h_{9,3}(m)$ and $A_{m,9,3}$



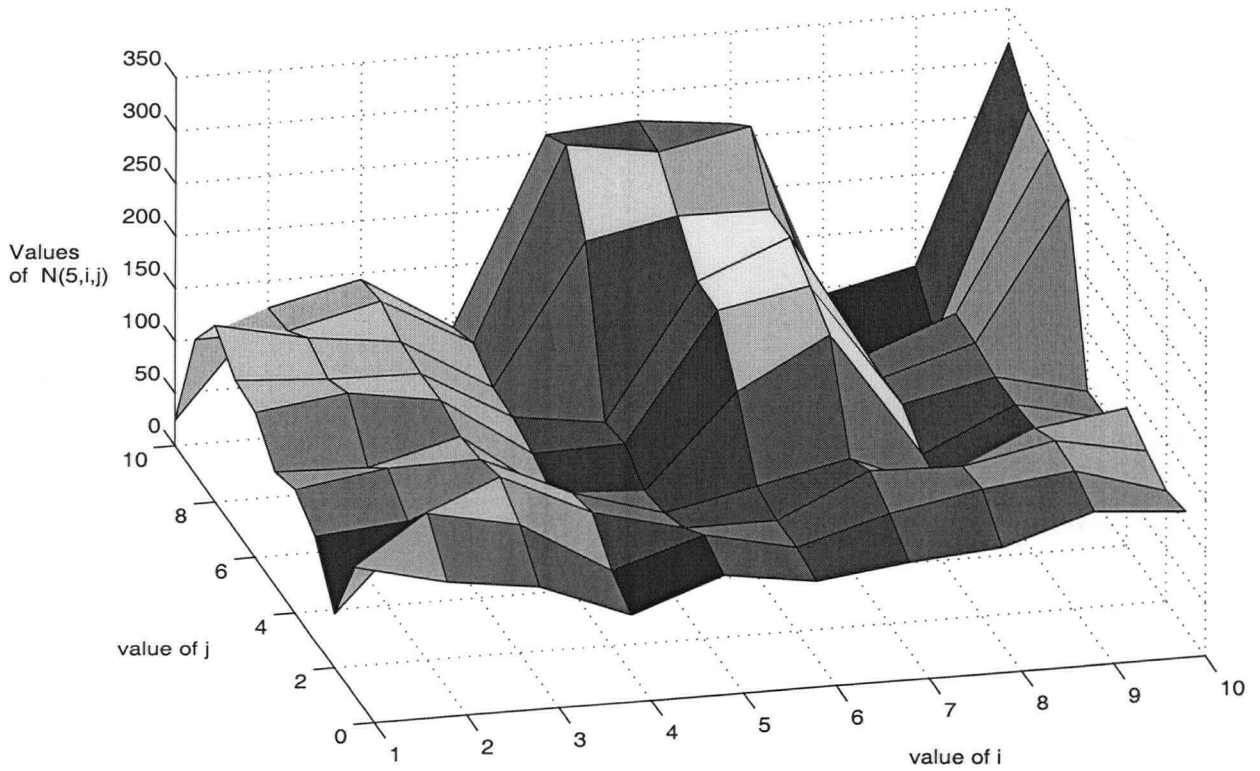
• $h_{9,9}(m)$ and $A_{m,9,9}$

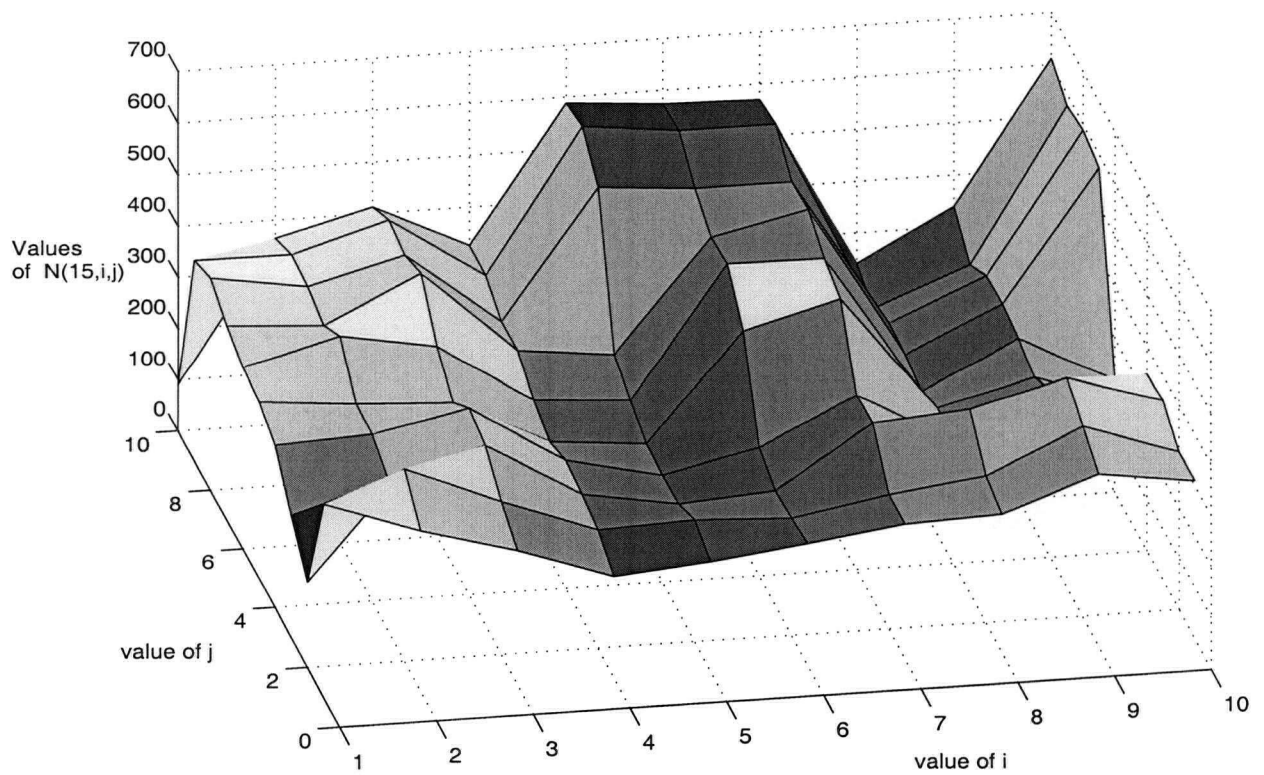
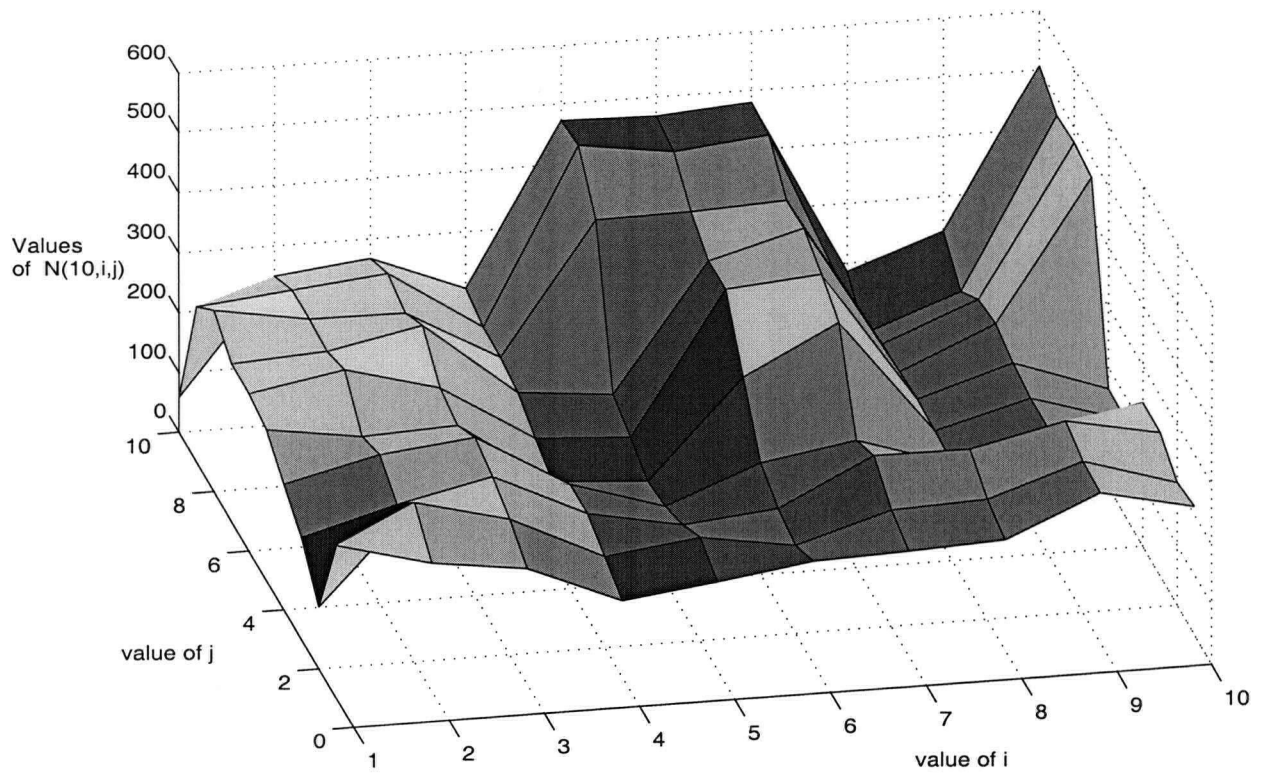


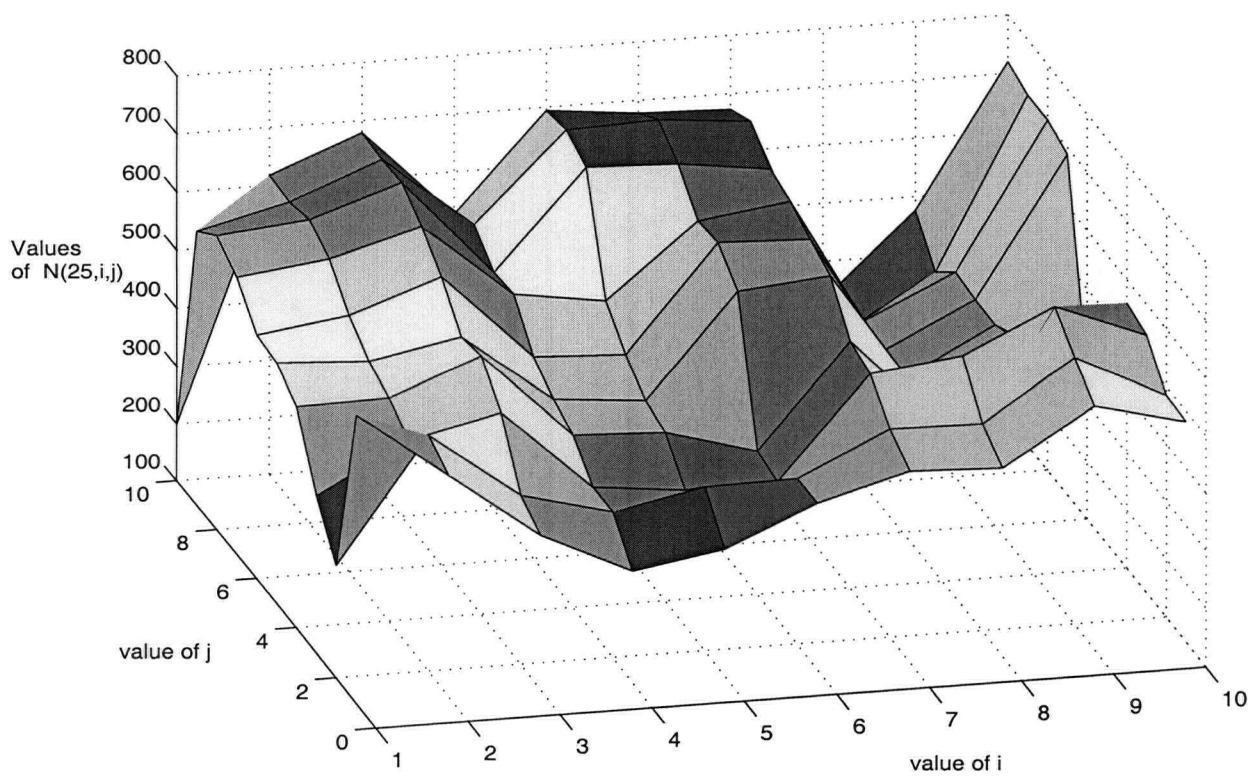
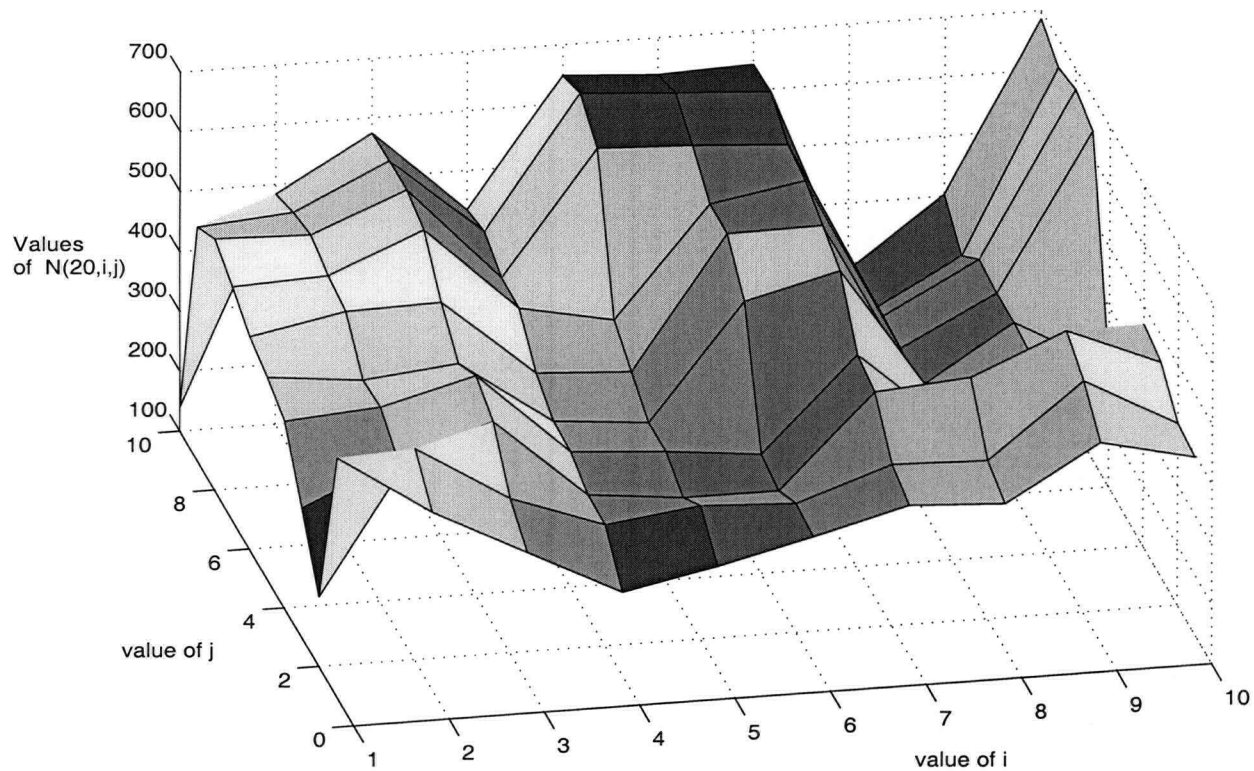
Appendix E

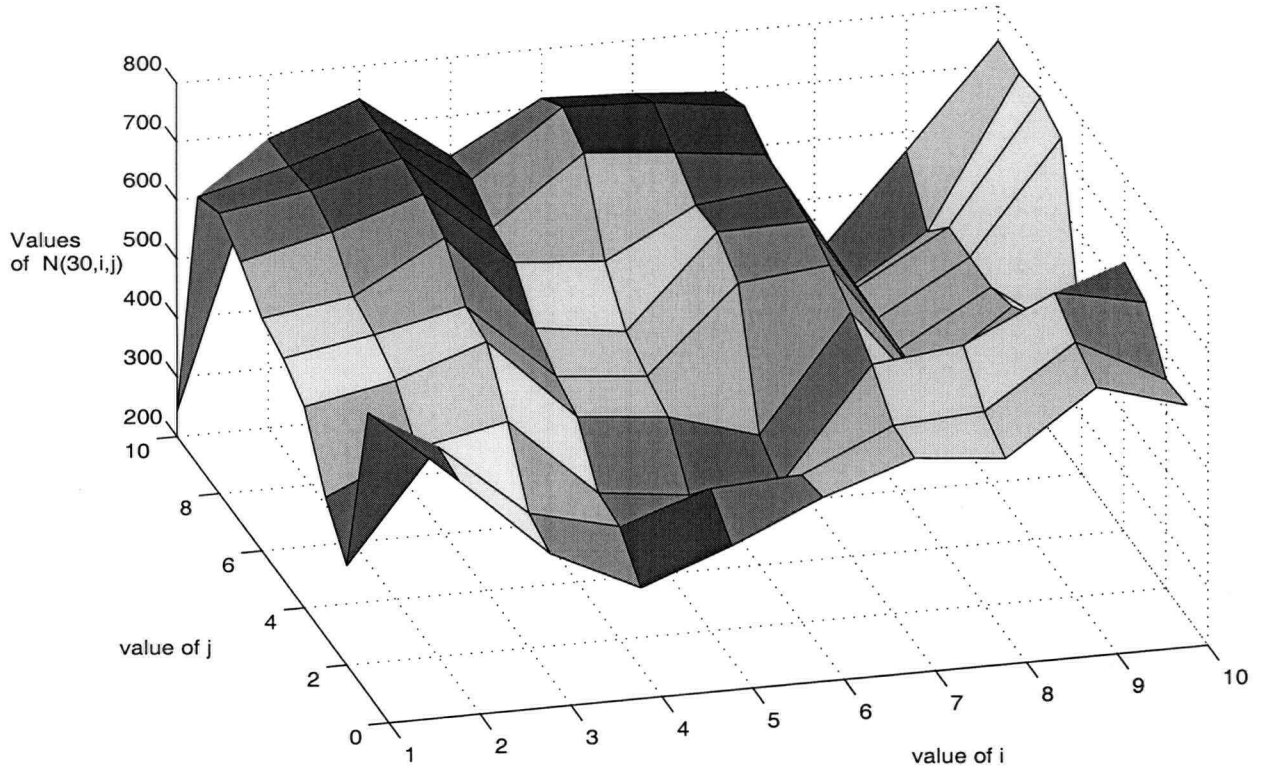
In this Appendix we present some (statistical) properties of the fitting error of $h_{i,j}(m)$.

- Let $N(p, i, j)$ be a function which measures the number of data-points whose fitting error is less than $p\%$, where as fitting error we consider the quotient $\frac{|A_{m,i,j} - h_{i,j}(m)|}{|A_{m,i,j}|}$. Then we can plot the following surfaces in order to see the levels at which the fitting errors fluctuate (we remind the reader that the total number of datapoints is 886). We can see from the plots that we have “good” fit for $j \geq 8$ and for $5 \leq i \leq 7$ and $i = 10$.

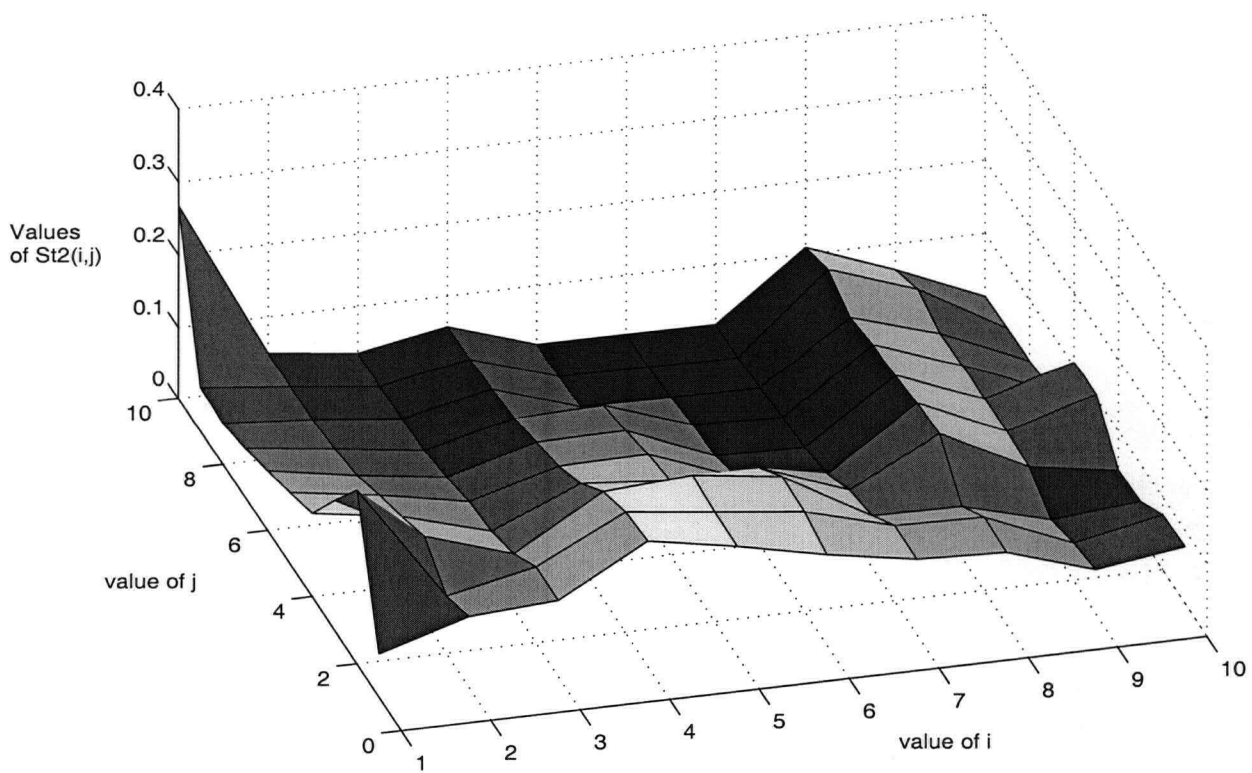
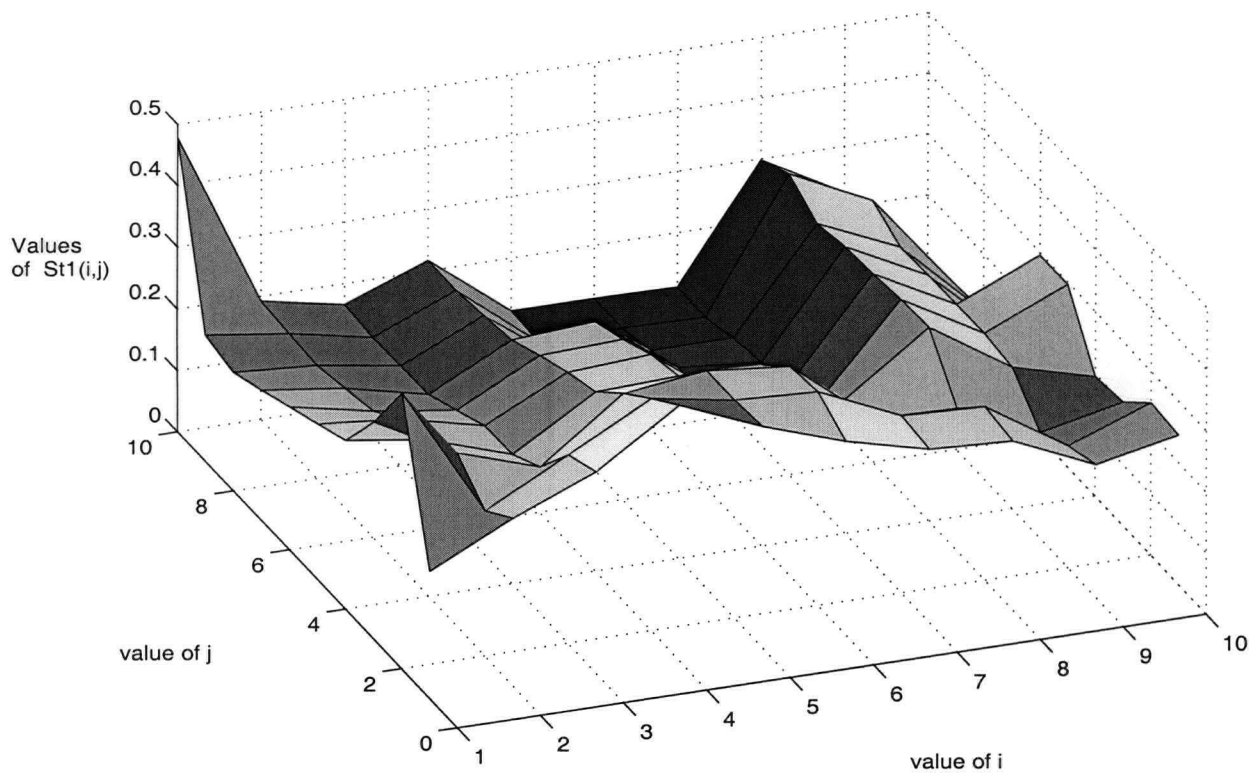






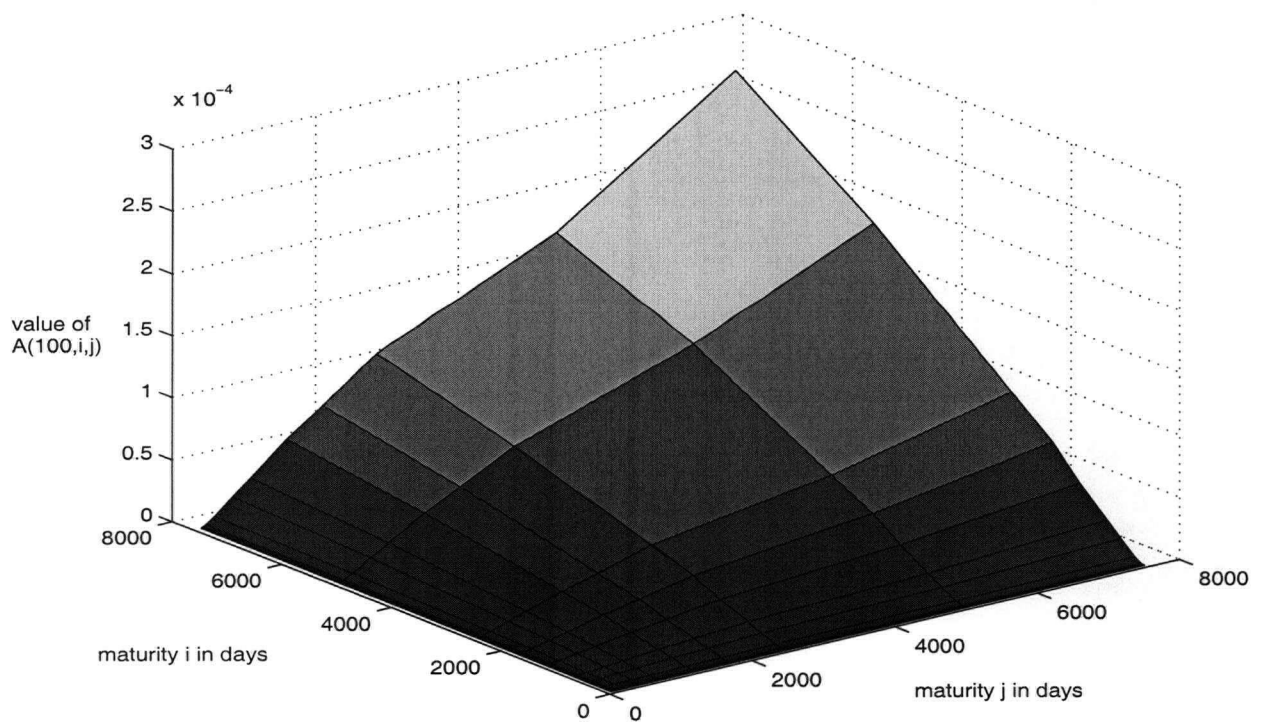
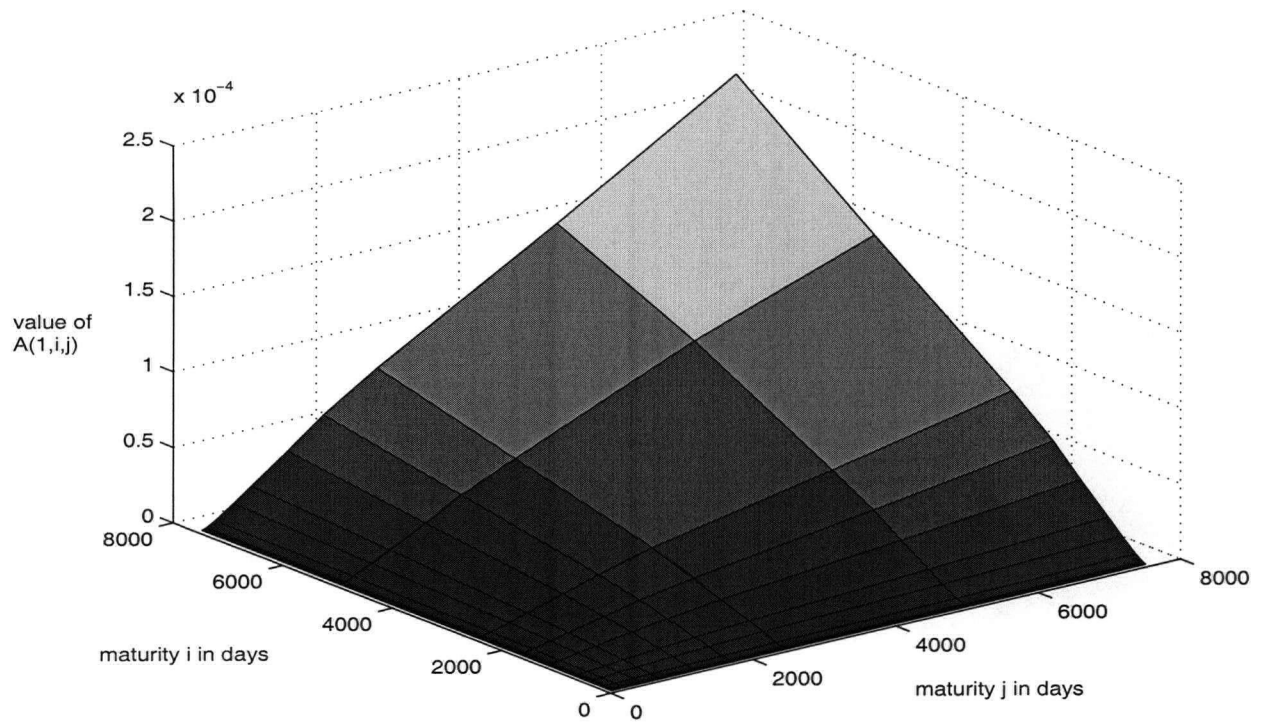


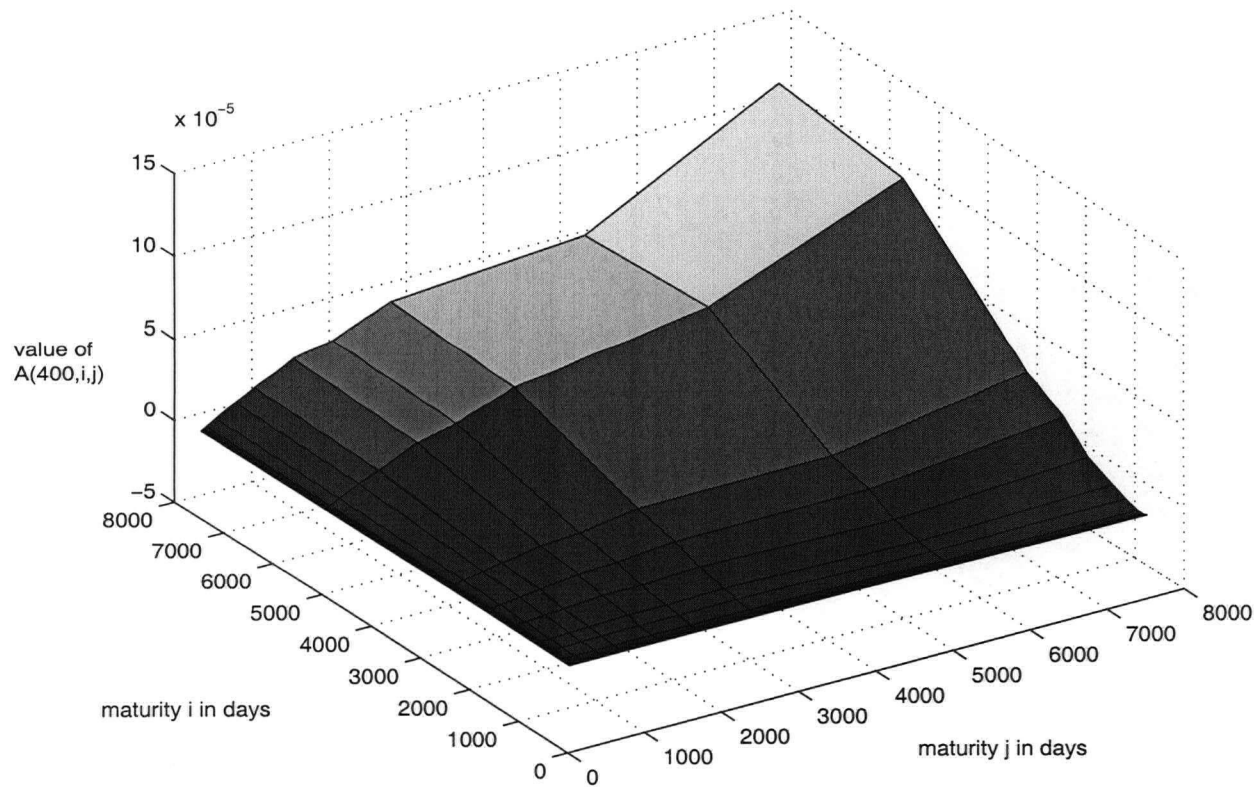
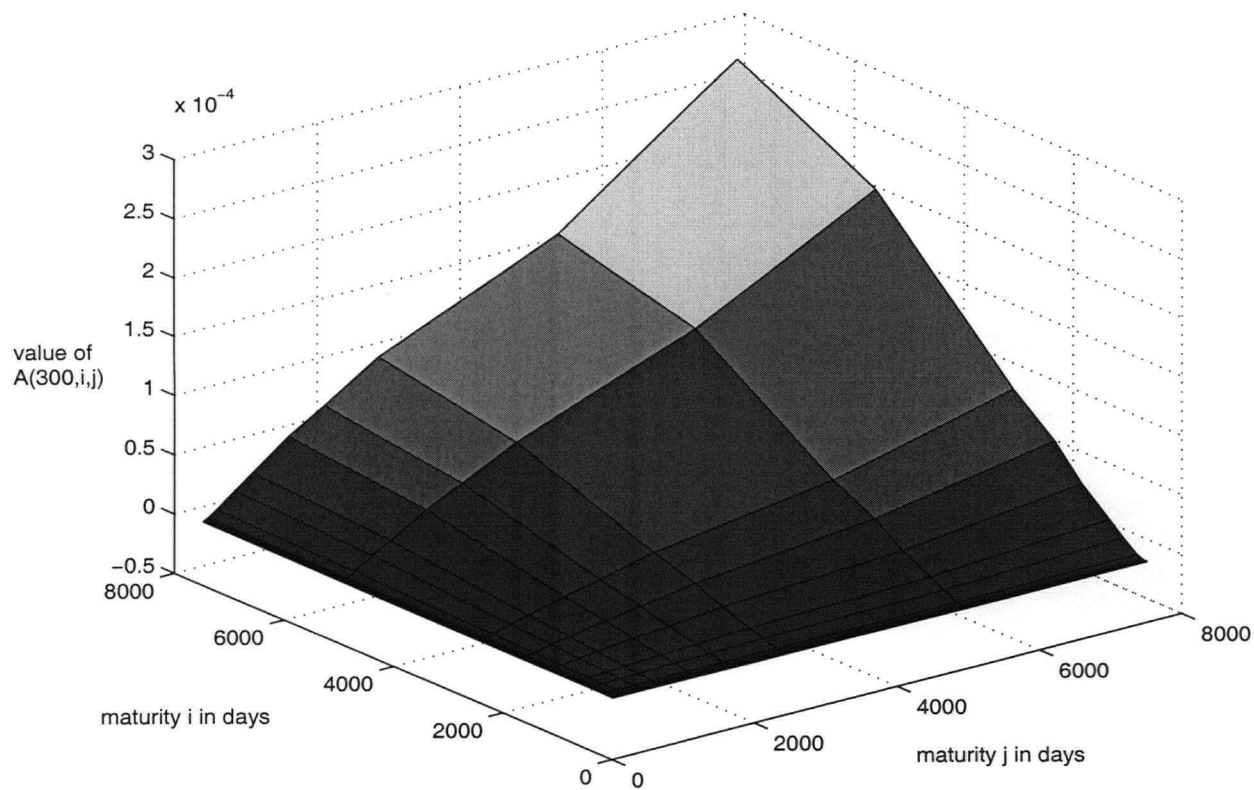
- Let $St1(i, j) = \frac{\sum |A_{m,i,j} - h_{i,j}(m)|}{\sum |A_{m,i,j}|}$ be the quotient of the sum of the absolute values of the fitting errors over the sum of the absolute values of $A_{m,i,j}$ for each pair of maturities i, j .
- Let also $St2(i, j) = \frac{\sum [A_{m,i,j} - h_{i,j}(m)]^2}{\sum (A_{m,i,j})^2}$ be the quotient of the sum of the squared absolute values of the fitting errors over the sum of the squared absolute values of $A_{m,i,j}$ for each pair of maturities i, j (this statistic is similar to the R^2 statistic of the least squares methodology).
- $St1$ and $St2$ can be used to give us an idea of how big the fitting errors are and for which pair of maturities we have the best and worst fitting. Plotting them we can get the following graphs:

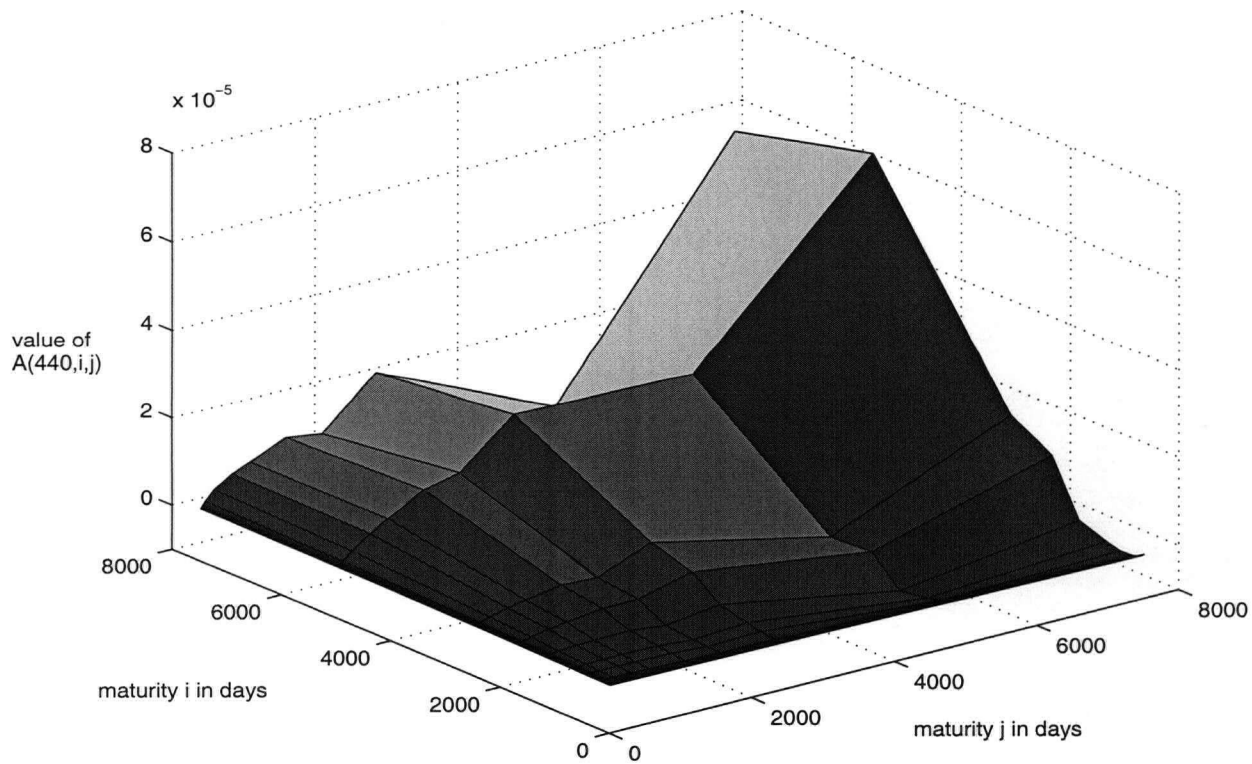
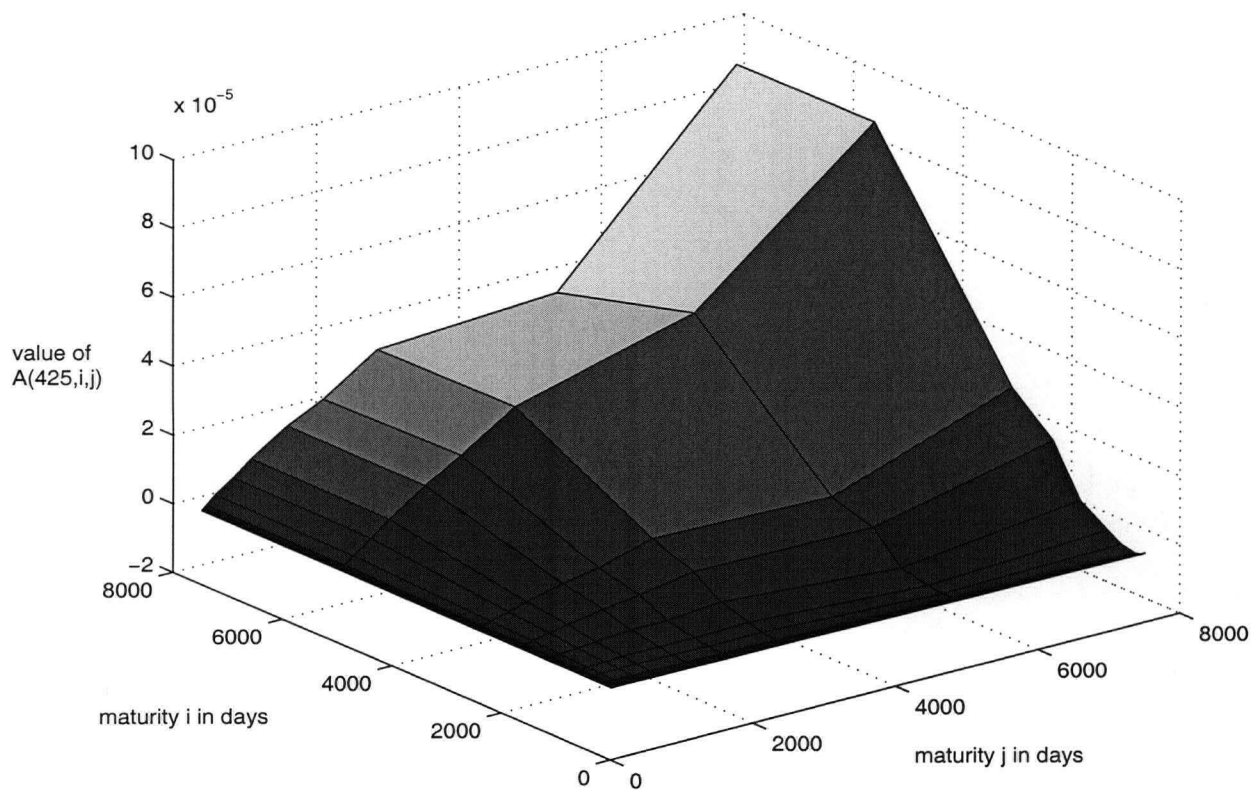


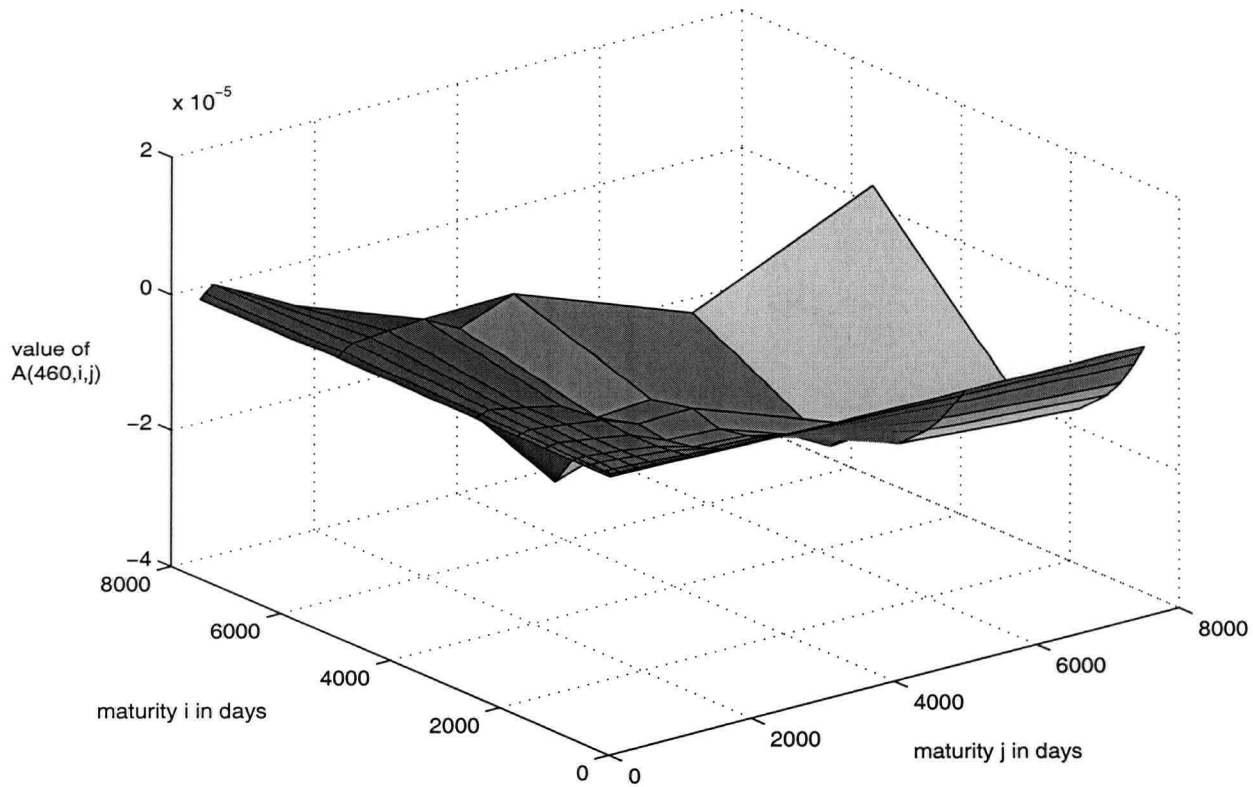
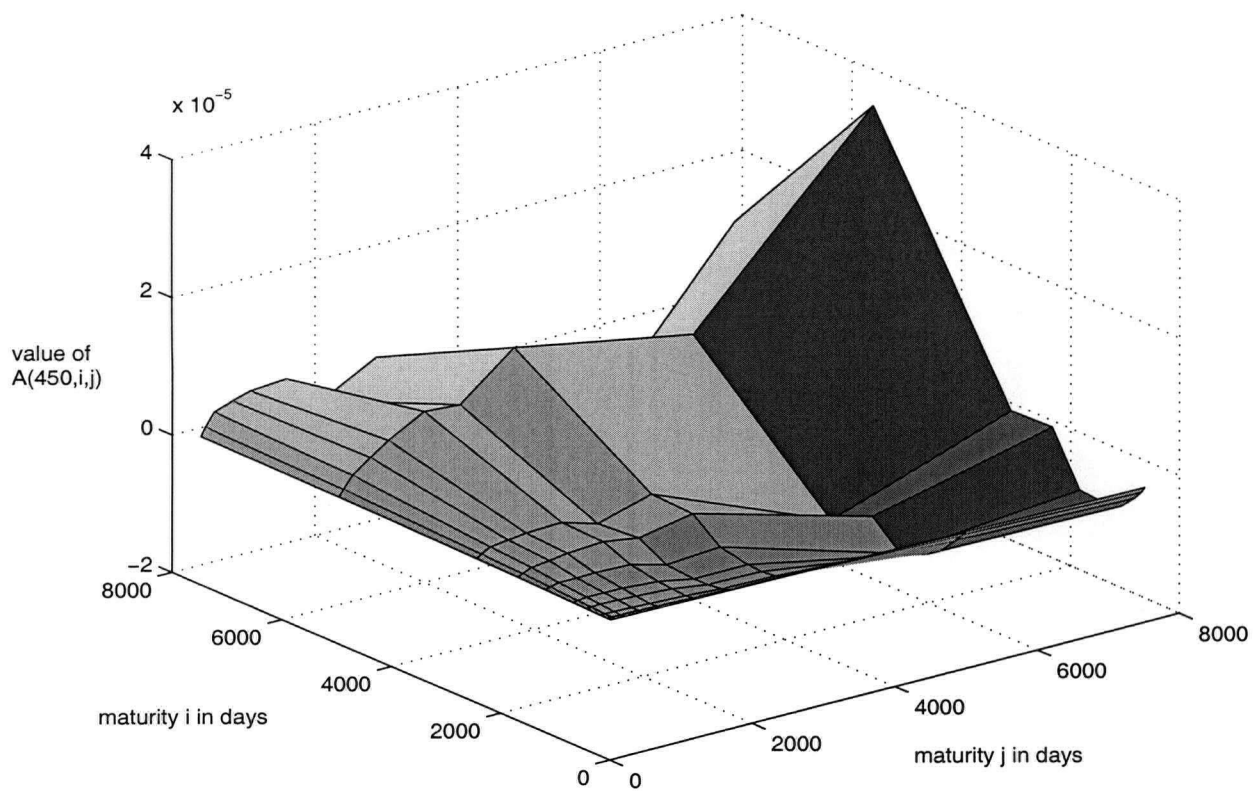
Appendix F

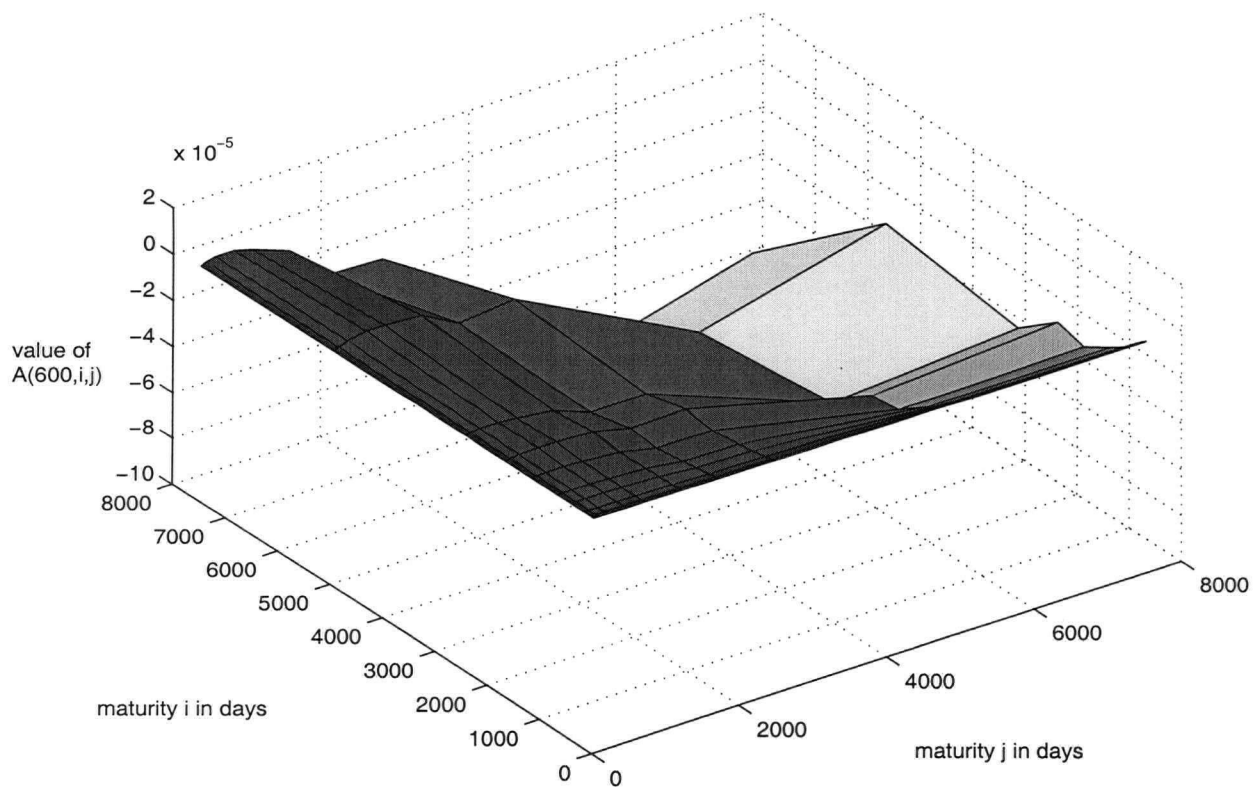
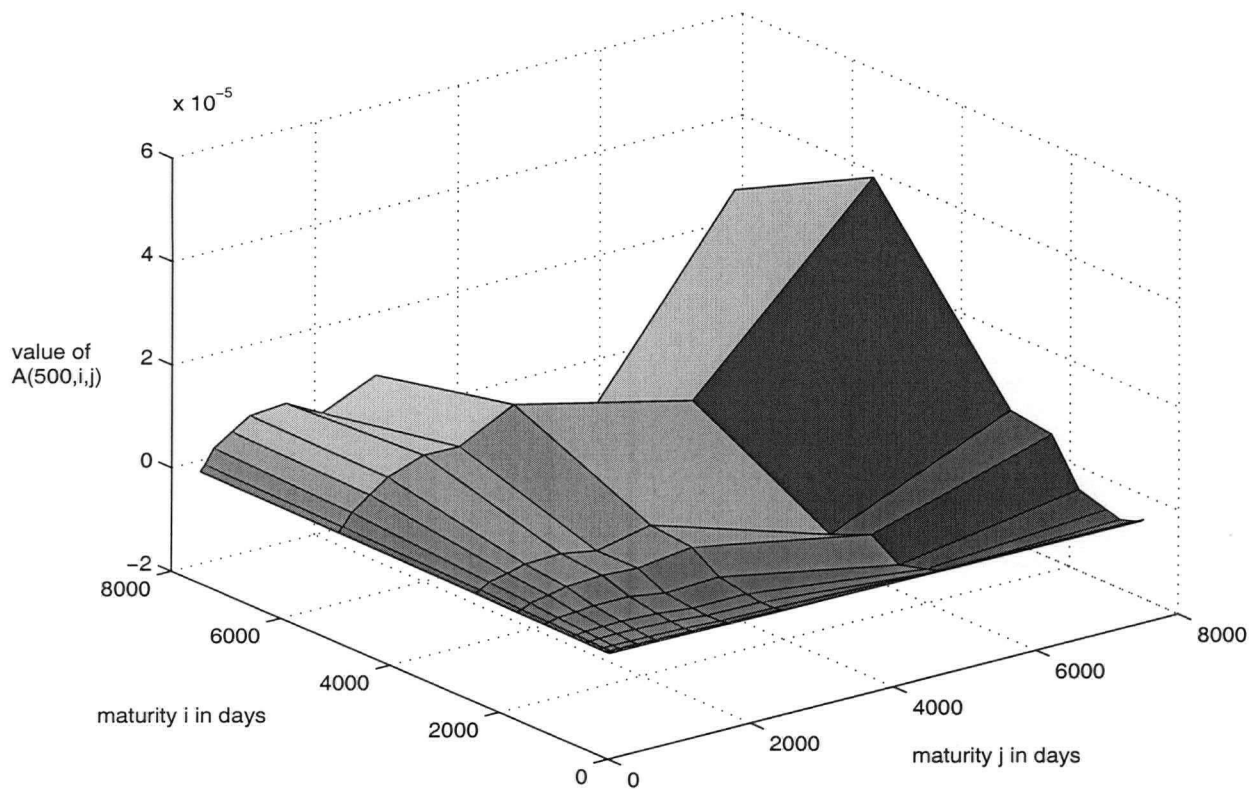
Indicative plots of $A_{m,i,j}$ for certain m .











Appendix G

Calculation of the integral $\int_k^{k+T_i} \int_{k+m}^{k+m+T_j} c(k, u, z) dz du = f_{k,m,i,j}$ for various forms of the covariance function $c(k, u, z)$. Let:

$$(A) \quad c(k, u, z) = \alpha \cdot (u + z + \beta)^2$$

Then

$$\begin{aligned} \int_k^{k+T_i} \int_{k+m}^{k+m+T_j} c(k, u, z) dz du &= \alpha \cdot T_i \cdot T_j \cdot \{4k^2 + (4m + 4\beta + 2 \cdot T_i + 2 \cdot T_j) \cdot k \\ &\quad + (m + \beta)^2 + (T_i + T_j)(m + \beta) + \frac{2T_j^2 + 2T_i^2 + T_j \cdot T_i}{6}\} \end{aligned}$$

which definitely doesn't satisfy (5.2.6) since the highest order k term is k^3 and not k^2 as we are supposed to have and also there is a big number of terms not multiplied with k .

$$(B) \quad c(k, u, z) = k\alpha e^{-\beta|u-z|}$$

Then:

$$\int_k^{k+T_i} \int_{k+m}^{k+m+T_j} c(k, u, z) dz du = k \frac{\alpha}{\beta^2} e^{\beta(T_i-m)} [1 - e^{-\beta T_i}] [1 - e^{-\beta T_j}]$$

if $m \geq T_i$,

$$\int_k^{k+T_i} \int_{k+m}^{k+m+T_j} c(k, u, z) dz du = k \frac{\alpha}{\beta^2} \{2\beta \cdot (T_i - m) - [1 - e^{-\beta T_j} (1 - e^{\beta T_i})] \cdot e^{-\beta m} + e^{-\beta T_i} \cdot e^{\beta m}\}$$

if $T_i - T_j \leq m \leq T_i$ and

$$\int_k^{k+T_i} \int_{k+m}^{k+m+T_j} c(k, u, z) dz du = k \frac{\alpha}{\beta^2} \{2\beta \cdot T_j - (1 - e^{-\beta T_j}) \cdot e^{-\beta m} + e^{-\beta T_i} (1 - e^{-\beta T_j}) \cdot e^{\beta m}\}$$

if $m \leq T_i - T_j$.

Again this covariance function doesn't satisfy (5.2.6) because the k^2 term doesn't appear in the right hand side. Notice also that the double integral is continuous at the points $m = T_i - T_j$ and $m = T_i$.

$$(C) \quad c(k, u, z) = k \cdot \alpha \cdot (e^{-\beta u} + e^{-\beta z})$$

Then

$$\int_k^{k+T_i} \int_{k+m}^{k+m+T_j} c(k, u, z) dz du = k \frac{\alpha}{\beta} [e^{-\beta(k+m+T_j)} - e^{-\beta(k+m)} + e^{-\beta(k+T_i)} - e^{-\beta k}]$$

which again doesn't satisfy (5.2.6), since k appears in the exponent of the exponential.

$$(D) \quad c(k, u, z) = k \cdot \alpha \cdot [e^{-\beta(\gamma u - z + \delta)^2} + e^{-\beta(\gamma z - u + \delta)^2}] + k \cdot \zeta, \quad \beta > 0$$

We have:

$$\begin{aligned}
I &= \int_k^{k+T_i} \int_{k+m}^{k+m+T_j} (e^{-\beta(\gamma \cdot u - z + \delta)^2} + e^{-\beta(\gamma \cdot z - u + \delta)^2}) dz du \\
&= \int_k^{k+T_i} \int_{k+m}^{k+m+T_j} e^{-\beta(\gamma \cdot u - z + \delta)^2} dz du + \int_k^{k+T_i} \int_{k+m}^{k+m+T_j} e^{-\beta(\gamma \cdot z - u + \delta)^2} dz du \\
(*) \quad &= -\frac{1}{\sqrt{\beta}} \int_k^{k+T_i} \left\{ \int_{a_1}^{a_2} e^{-x^2} dx \right\} du + \frac{1}{\gamma \sqrt{\beta}} \int_k^{k+T_i} \left\{ \int_{a_3}^{a_4} e^{-y^2} dy \right\} du
\end{aligned}$$

where we did the following substitutions:

$$x = (\gamma \cdot u - z + \delta) \sqrt{\beta} \quad \text{and} \quad y = (\gamma \cdot z - u + \delta) \sqrt{\beta}$$

The limits of the integrals then are:

$$\begin{aligned}
a_1 &= (\gamma \cdot u - k - m + \delta) \sqrt{\beta} \quad , \quad a_2 = (\gamma \cdot u - k - m - T_j + \delta) \sqrt{\beta} \\
a_3 &= (\gamma \cdot k + \gamma m - u + \delta) \sqrt{\beta} \quad , \quad a_4 = (\gamma \cdot k + \gamma m + \gamma \cdot T_j - u + \delta) \sqrt{\beta}
\end{aligned}$$

In order to calculate (*) we have to use some approximation of the integral $\int_a^b e^{-x^2} dx$. We used 4 different approximation for this integral, found in Squire (1970):

$$\begin{aligned}
\text{(i)} \quad & \int_a^b e^{-x^2} dx \simeq e^{-b^2} \left\{ b + \frac{2}{3}b^3 + \frac{4}{15}b^5 + \dots \right\} - e^{-a^2} \left\{ a + \frac{2}{3}a^3 + \frac{4}{15}a^5 + \dots \right\} \\
\text{(ii)} \quad & \int_a^b e^{-x^2} dx \simeq -e^{-b^2} \left\{ \frac{1}{2b} - \frac{1}{4b^3} + \frac{3}{8b^5} - \dots \right\} + e^{-a^2} \left\{ \frac{1}{2a} - \frac{1}{4a^3} + \frac{3}{8a^5} - \dots \right\} \\
\text{(iii)} \quad & \int_a^b e^{-x^2} dx \simeq \left\{ b - \frac{1}{3}b^3 + \frac{1}{10}b^5 - \dots \right\} - \left\{ a - \frac{1}{3}a^3 + \frac{1}{10}a^5 - \dots \right\} \\
\text{(iv)} \quad & \int_a^b e^{-x^2} dx \simeq \frac{e^{-b^2}}{b^2} - \frac{e^{-a^2}}{a^2}
\end{aligned}$$

approximation (i)

$$\begin{aligned}
(*) &= \frac{1}{\sqrt{\beta}} \left\{ - \int_k^{k+T_i} e^{-a_2^2} \left[a_2 + \frac{2}{3}a_2^3 + \frac{4}{15}a_2^5 + \dots \right] du + \int_k^{k+T_i} e^{-a_1^2} \left[a_1 + \frac{2}{3}a_1^3 + \frac{4}{15}a_1^5 + \dots \right] du \right. \\
&\quad \left. + \frac{1}{\gamma} \int_k^{k+T_i} e^{-a_4^2} \left[a_4 + \frac{2}{3}a_4^3 + \frac{4}{15}a_4^5 + \dots \right] du - \frac{1}{\gamma} \int_k^{k+T_i} e^{-a_3^2} \left[a_3 + \frac{2}{3}a_3^3 + \frac{4}{15}a_3^5 + \dots \right] du \right\} \\
&= -\frac{1}{\beta \cdot \gamma} \left\{ \int_{b_1}^{b_2} e^{-a^2} \left[a + \frac{2}{3}a^3 + \frac{4}{15}a^5 + \dots \right] da - \int_{b_3}^{b_4} e^{-a^2} \left[a + \frac{2}{3}a^3 + \frac{4}{15}a^5 + \dots \right] da \right. \\
&\quad \left. + \int_{b_5}^{b_6} e^{-a^2} \left[a + \frac{2}{3}a^3 + \frac{4}{15}a^5 + \dots \right] da - \int_{b_7}^{b_8} e^{-a^2} \left[a + \frac{2}{3}a^3 + \frac{4}{15}a^5 + \dots \right] da \right\}
\end{aligned}$$

where

$$\begin{aligned}
b_1 &= [(\gamma - 1)k - m - T_j + \delta]\sqrt{\beta} \quad , \quad b_2 = [(\gamma - 1)k - m + \gamma \cdot T_i - T_j + \delta]\sqrt{\beta} \\
b_3 &= [(\gamma - 1) \cdot k - m + \delta]\sqrt{\beta} \quad , \quad b_4 = [(\gamma - 1) \cdot k - m + \gamma \cdot T_i + \delta]\sqrt{\beta} \\
b_5 &= [(\gamma - 1)k + \gamma m + \gamma \cdot T_j + \delta]\sqrt{\beta} \quad , \quad b_6 = [(\gamma - 1)k + \gamma m + \gamma \cdot T_j - T_i + \delta]\sqrt{\beta} \\
b_7 &= [(\gamma - 1) \cdot k + \gamma m + \delta]\sqrt{\beta} \quad , \quad b_8 = [(\gamma - 1) \cdot k + \gamma m - T_i + \delta]\sqrt{\beta}
\end{aligned}$$

Notice here that unless $\gamma = 1$, k will be in the exponent of the exponential and therefore equation (5.2.6) won't be satisfied. Therefore we must have $\gamma = 1$. Then the limits of the integrals become:

$$\begin{aligned}
b_1 &= [-m - T_j + \delta]\sqrt{\beta} \quad , \quad b_2 = [-m + T_i - T_j + \delta]\sqrt{\beta} \\
b_3 &= [-m + \delta]\sqrt{\beta} \quad , \quad b_4 = [-m + T_i + \delta]\sqrt{\beta} \\
b_5 &= [m + T_j + \delta]\sqrt{\beta} \quad , \quad b_6 = [m + T_j - T_i + \delta]\sqrt{\beta} \\
b_7 &= [m + \delta]\sqrt{\beta} \quad , \quad b_8 = [m - T_i + \delta]\sqrt{\beta}
\end{aligned}$$

Moreover we can place the restriction $\delta \geq 0$ from the symmetry of the integrals and the restriction we just derived: $\gamma = 1$. A final remark is that if we assume $T_1 \cong 63$ (based on the fact that T_1 ranges around 60 to 65 depending on the number of non-working days in this period) we can find that:

$$\begin{aligned}
\min_{m,i}\{b_i\} &= \min_m b_1 = [\delta + \min(-m) + \min(-T_j)] \cdot \sqrt{\beta} \simeq (\delta - 8.5 \cdot 10^3)\sqrt{\beta} \\
\max_{m,i}\{b_i\} &= \max_m b_5 = [\delta + \max(m) + \max(T_j)] \cdot \sqrt{\beta} \simeq (\delta + 8.5 \cdot 10^3)\sqrt{\beta}
\end{aligned}$$

We can easily find that:

$$\begin{aligned}
I_1 &= \int_{b_1}^{b_2} e^{-a^2} a \cdot da = -\frac{1}{2}[e^{-b_2^2} - e^{-b_1^2}] \\
I_2 &= \int_{b_1}^{b_2} e^{-a^2} a^3 \cdot da = I_1 - \frac{1}{2}[e^{-b_2^2} b_2^2 - e^{-b_1^2} b_1^2]
\end{aligned}$$

Now let's make the assumption:

$$(D1) \quad |a + \frac{2}{3}a^3| \gg |\frac{4}{15}a^5|$$

Then keeping in mind that:

$$\int_{b_1}^{b_2} e^{-a^2} (a + \frac{2}{3}a^3) da = -\frac{5}{6}[e^{-b_2^2} - e^{-b_1^2}] - \frac{1}{3}[e^{-b_2^2}b_2^2 - e^{-b_1^2}b_1^2]$$

we finally find:

$$\begin{aligned} I_A = k \cdot (\alpha \cdot I + \zeta \cdot T_i \cdot T_j) &\simeq k \frac{\alpha}{\beta} \left\{ \frac{5}{6}[e^{-b_2^2} - e^{-b_1^2}] + \frac{1}{3}[e^{-b_2^2}b_2^2 - e^{-b_1^2}b_1^2] - \frac{5}{6}[e^{-b_4^2} - e^{-b_3^2}] \right. \\ &- \frac{1}{3}[e^{-b_4^2}b_4^2 - e^{-b_3^2}b_3^2] + \frac{5}{6}[e^{-b_6^2} - e^{-b_5^2}] + \frac{1}{3}[e^{-b_6^2}b_6^2 - e^{-b_5^2}b_5^2] - \frac{5}{6}[e^{-b_8^2} - e^{-b_7^2}] \\ &- \frac{1}{3}[e^{-b_8^2}b_8^2 - e^{-b_7^2}b_7^2] \\ &\left. + k \cdot \zeta \cdot T_i \cdot T_j \right\} \end{aligned}$$

A couple of points about approximation (D1):

- The infinite series $a + \frac{2}{3}a^3 + \frac{4}{15}a^5 + \dots$ is absolutely convergent since

$$\lim_{n \rightarrow \infty} \left| \frac{u_{n+1}}{u_n} \right| = \lim_{n \rightarrow \infty} \frac{2a^2}{2n+3} = 0 < 1$$

- Since $\delta \geq 0$ we have $a \geq -8.5 \cdot 10^3 \cdot \sqrt{\beta}$. Moreover the inequality (D1) can be written as $|a| + \frac{2}{3}|a^3| > \frac{4}{15}|a^5|$ since a is raised only in odd powers. We can then solve this inequality:

$$|a| + \frac{2}{3}|a^3| > \frac{4}{15}|a^5| \Leftrightarrow |a|(1 + \frac{2}{3}a^2 - \frac{4}{15}a^4) > 0$$

and find that we must have $|a| < 1.88544 \simeq 1.9$. Since we also have $a \geq -8.5 \cdot 10^3 \cdot \sqrt{\beta}$ we should then have $\beta < 5 \cdot 10^{-8}$

- The above solution satisfies the inequality $|a| + \frac{2}{3}|a^3| > \frac{4}{15}|a^5|$ but not necessarily the one we are actually interested in: $|a| + \frac{2}{3}|a^3| \gg \frac{4}{15}|a^5|$. For this we should solve an inequality of the form $|a| + \frac{2}{3}|a^3| \gg \frac{4\varepsilon}{15}|a^5|$ where ε is a number specifying the accuracy of our approximation. If we pick $\varepsilon = 20$ (which can be interpreted as having around 5% error in our approximation) we get $|a| < 0.7$ and therefore $\beta < 7 \cdot 10^{-9}$.
- Now we can proceed to the estimation of the parameters keeping in mind the restrictions:

$$0 < \beta < 5 \cdot 10^{-8} \quad , \quad \gamma = 1 \quad , \quad \delta \geq 0$$

Note: The above restriction for β doesn't ensure small errors but only that the two first terms of the series will be larger than the third one. For small errors we should either make sure that β satisfies the more strict: $\beta < 7 \cdot 10^{-9}$ or add more terms to our approximation (which increases greatly the difficulty of our problem).

approximation (ii)

Working the same way we get:

$$\begin{aligned}
 I_B \simeq k \frac{\alpha}{8\beta} \{ & -2 \frac{e^{-b_2^2}}{b_2^2} + 2 \frac{e^{-b_1^2}}{b_1^2} + 3 \frac{e^{-b_2^2}}{b_2^4} - 3 \frac{e^{-b_1^2}}{b_1^4} + 2 \frac{e^{-b_4^2}}{b_4^2} - 2 \frac{e^{-b_3^2}}{b_3^2} \\
 & - 3 \frac{e^{-b_4^2}}{b_2^4} + 3 \frac{e^{-b_3^2}}{b_1^4} - 2 \frac{e^{-b_6^2}}{b_6^2} + 2 \frac{e^{-b_5^2}}{b_5^2} + 3 \frac{e^{-b_6^2}}{b_6^4} - 3 \frac{e^{-b_5^2}}{b_5^4} \\
 & + 2 \frac{e^{-b_8^2}}{b_8^2} - 2 \frac{e^{-b_7^2}}{b_7^2} - 3 \frac{e^{-b_8^2}}{b_8^4} + 3 \frac{e^{-b_7^2}}{b_7^4} \} + k \cdot \zeta \cdot T_i \cdot T_j
 \end{aligned}$$

where we used the fact that

$$\int_{b_1}^{b_2} e^{-a^2} \left(\frac{1}{2a} - \frac{1}{2a^3} \right) da = -\frac{1}{4} \left[e^{-a^2} \frac{1}{a^2} \right]_{b_1}^{b_2} + \frac{3}{8} \left[e^{-a^2} \frac{1}{a^4} \right] + \int_{b_1}^{b_2} e^{-a^2} \left\{ \frac{3}{2a^5} \right\} da$$

and made the assumption:

$$(D2) \quad b_i \neq 0, i = 1, \dots, 8$$

$$(D3) \quad \left| \frac{1}{2a} - \frac{1}{4a^3} \right| \gg \left| \frac{15}{8a^5} \right|$$

In order for (D3) to be valid (in the 5% error sense - as above) we must have $|a| > 2.985$ and therefore $\beta > 1.234 \cdot 10^{-7}$. We should note here that again the sum of the approximated series converges. In order for (D2) to be true we need to have

$$\delta > \max m + \max_j \{T_j\} = 8446$$

If $\delta < 8446$ then we can always find some values (m, i, j) so that the denominator will be equal to 0. For the above number we used that $m = 1, \dots, 886$ and assumed, as before, that $T_1 = 63$.

If we keep in mind the above restrictions they would lead to:

$$I_B \simeq k \frac{\alpha}{4\beta} \left\{ \frac{e^{-b_1^2}}{b_1^2} \right\} + k \cdot \zeta \cdot T_i \cdot T_j, \text{ if } b_1 > 1$$

$$I_B \simeq -k \frac{3\alpha}{8\beta} \left\{ \frac{e^{-b_1^2}}{b_1^4} \right\} + k \cdot \zeta \cdot T_i \cdot T_j, \text{ if } b_1 < 1$$

since the other functions would have much smaller values compared to the above ones.

approximation (iii)

Working the same way we get:

$$\begin{aligned}
 I_C^1 &\simeq k \frac{\alpha}{2\beta} \left\{ -b_2^2 + b_1^2 + \frac{1}{6}b_2^4 - \frac{1}{6}b_1^4 + b_4^2 - b_3^2 - \frac{1}{6}b_4^4 + \frac{1}{6}b_3^4 \right. \\
 &\quad \left. - b_6^2 + b_5^2 + \frac{1}{6}b_6^4 - \frac{1}{6}b_5^4 + b_8^2 - b_7^2 - \frac{1}{6}b_8^4 + \frac{1}{6}b_7^4 \right\} + k \cdot \zeta \cdot T_i \cdot T_j \\
 &= -k \cdot \frac{\alpha}{2\beta} \cdot T_i \cdot T_j \left[m^2 + (T_j - T_i)m + A + \frac{T_j^2 + T_i^2}{3} - \frac{T_j \cdot T_i}{2} \right] + k \cdot \zeta \cdot T_i \cdot T_j
 \end{aligned}$$

where $A = \delta^2 - \frac{1}{\beta}$ and we made the assumption:

$$(D3) \quad |a - \frac{1}{3}a^3| \gg |\frac{1}{10}a^5|$$

In order for this to be valid (in the 5% error sense - as above) we must have $|a| < 0.8$ and therefore $\beta < 9 \cdot 10^{-9}$. It is easy to see that again the sum of the series converges.

For a more accurate result we can also keep in our approximations the third term of the series which would then lead us to the following result:

$$\begin{aligned}
 I_C^2 &\simeq k \cdot \frac{\alpha}{\beta^2} \cdot T_i \cdot T_j \left[m^4 + 2(T_j - T_i)m^3 + 2(B + T_i^2 + T_j^2 + \frac{3}{2}T_i \cdot T_j)m^2 \right. \\
 &\quad \left. + (T_j - T_i)(2B + T_i^2 + T_j^2 - T_i \cdot T_j)m + C + B(\frac{2}{3}T_i^2 + \frac{2}{3}T_j^2 - T_i \cdot T_j) \right. \\
 &\quad \left. + \frac{T_j^4 + T_i^4}{5} - \frac{T_j \cdot T_i}{2}(T_i^2 + T_j^2) + \frac{2}{3} \cdot T_i^2 \cdot T_j^2 \right] + k \cdot \zeta \cdot T_i \cdot T_j
 \end{aligned}$$

where $B = 3\delta^2 - \frac{1}{\beta}$ and $C = (\delta^2 - \frac{1}{\beta})^2 + \frac{1}{\beta^2}$.

approximation (iv)

Finally working with this approximation we end up to:

$$\begin{aligned}
 I_D &\simeq k \frac{\alpha}{\beta} \left\{ -\frac{e^{-b_2^2}}{b_2} + \frac{e^{-b_1^2}}{b_1} - 2\frac{e^{-b_2^2}}{b_2^2} + 2\frac{e^{-b_1^2}}{b_1^2} + \frac{e^{-b_4^2}}{b_4} - \frac{e^{-b_3^2}}{b_3} + 2\frac{e^{-b_4^2}}{b_2^2} - 2\frac{e^{-b_3^2}}{b_1^2} \right. \\
 &\quad \left. - \frac{e^{-b_6^2}}{b_6} + \frac{e^{-b_5^2}}{b_5} - 2\frac{e^{-b_6^2}}{b_6^2} + 2\frac{e^{-b_5^2}}{b_5^2} + \frac{e^{-b_8^2}}{b_8} - \frac{e^{-b_7^2}}{b_7} + 2\frac{e^{-b_8^2}}{b_8^2} - 2\frac{e^{-b_7^2}}{b_7^2} \right\} + k \cdot \zeta \cdot T_i \cdot T_j
 \end{aligned}$$

where we used that:

$$\int_{b_1}^{b_2} e^{-a^2} \frac{1}{a^2} da \simeq \frac{e^{-b_2^2}}{b_2} - \frac{e^{-b_1^2}}{b_1} + 2\frac{e^{-b_2^2}}{b_2^2} - 2\frac{e^{-b_1^2}}{b_1^2}$$

and made the assumption (D2). The above formula is very similar to the one for approximation (ii) which leads us to the following result:

$$I_D \simeq k \frac{\alpha}{\beta} \left\{ \frac{e^{-b_1^2}}{b_1} \right\} + k \cdot \zeta \cdot T_i \cdot T_j, \text{ if } b_1 > 1$$

$$I_D \simeq k \frac{2\alpha}{\beta} \left\{ \frac{e^{-b_1^2}}{b_1^2} \right\} + k \cdot \zeta \cdot T_i \cdot T_j, \text{ if } b_1 < 1$$

The reader can see that all the above approximations don't satisfy (5.2.6), since we do not get the k^2 term.

$$(E) \quad c(k, u, z) = k \cdot \alpha \cdot [e^{-\beta(u-\delta)^2} + e^{-\beta(z-\delta)^2}], \quad \beta > 0$$

Working similarly to (D) we end up to a solution which doesn't satisfy (5.2.6), since k appears in the exponent of the exponential:

$$\begin{aligned} \int_k^{k+T_i} \int_{k+m}^{k+m+T_j} c(k, u, z) dz du &= k \frac{\alpha \cdot T_j}{\sqrt{\beta}} e^{-a_2^2} \left\{ a_2 + \frac{2}{3} a_2^3 + \frac{4}{15} a_2^5 + \dots \right\} - \frac{\alpha \cdot T_j}{\sqrt{\beta}} e^{-a_1^2} \left\{ a_1 + \frac{2}{3} a_1^3 \right. \\ &\quad \left. + \frac{4}{15} a_1^5 + \dots \right\} + \frac{\alpha \cdot T_i}{\sqrt{\beta}} e^{-a_4^2} \left\{ a_4 + \frac{2}{3} a_4^3 + \frac{4}{15} a_4^5 + \dots \right\} \\ &\quad - e^{-a_3^2} \left\{ a_3 + \frac{2}{3} a_3^3 + \frac{4}{15} a_3^5 + \dots \right\} \end{aligned}$$

where

$$\begin{aligned} a_1 &= (k - \delta) \sqrt{\beta}, \quad a_2 = (k + T_i - \delta) \sqrt{\beta} \\ a_3 &= (k + m - \delta) \sqrt{\beta}, \quad a_4 = (k + m + T_j - \delta) \sqrt{\beta} \end{aligned}$$

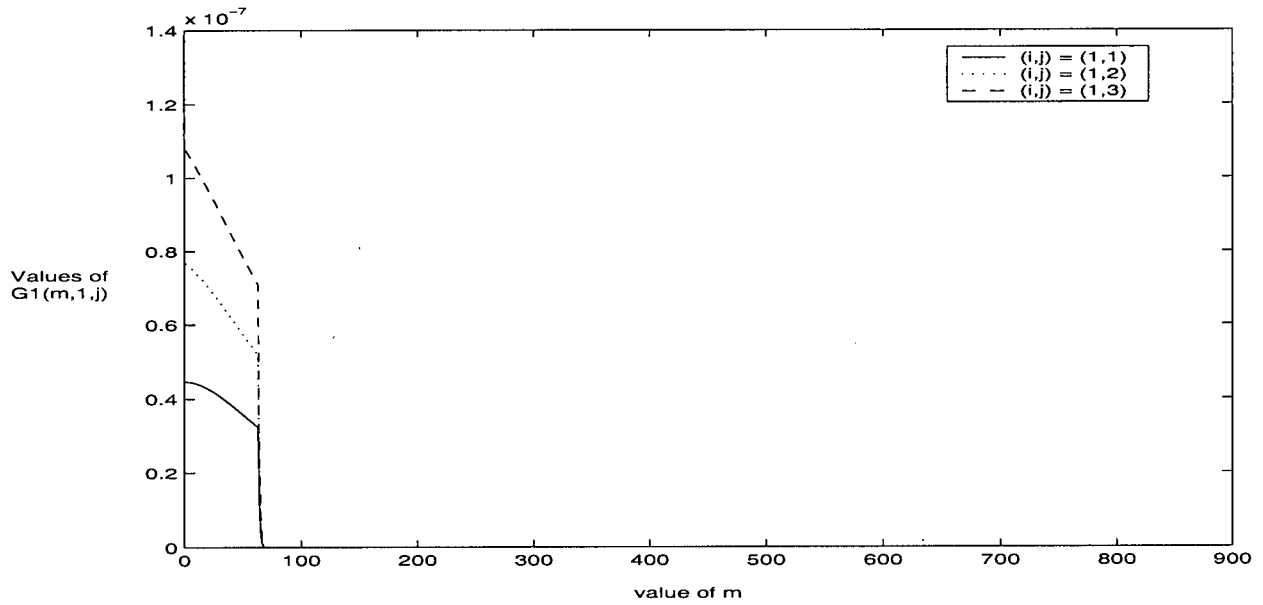
Note: The solution above is derived by using approximation (i). Using the other approximations lead to similar results - all of which are rejected.

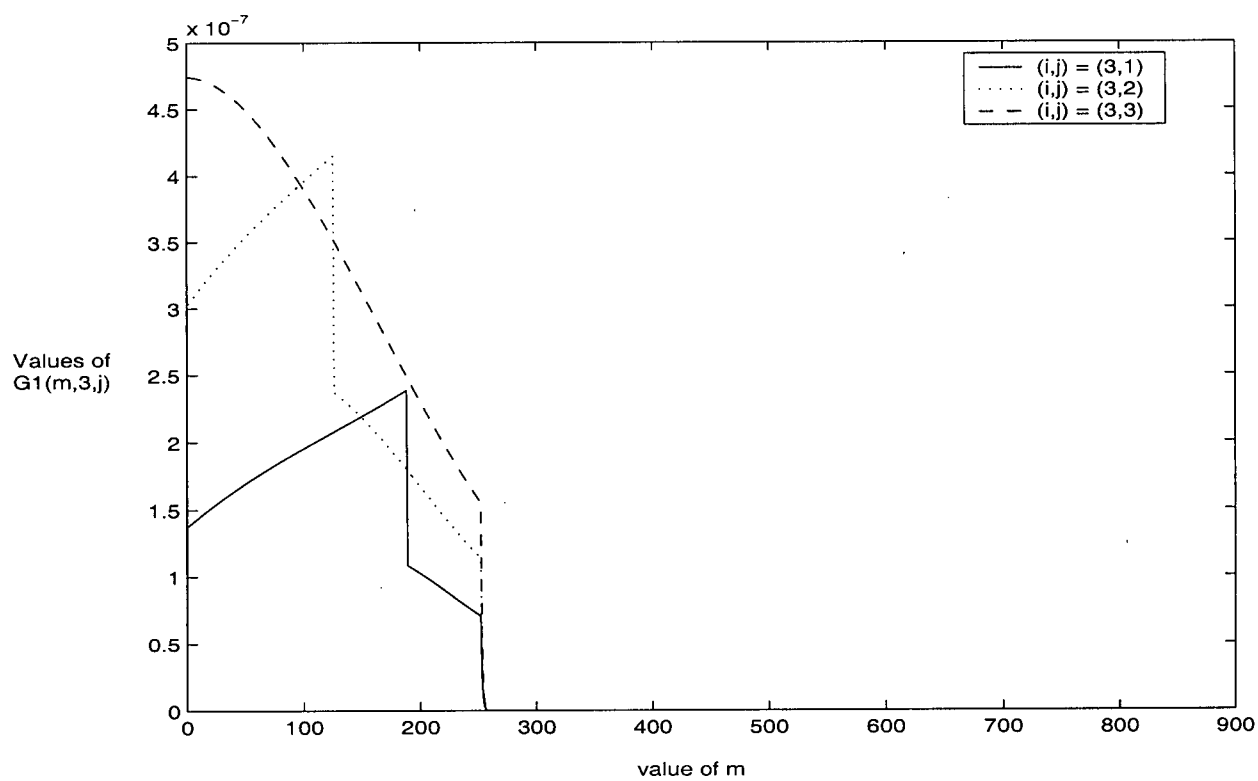
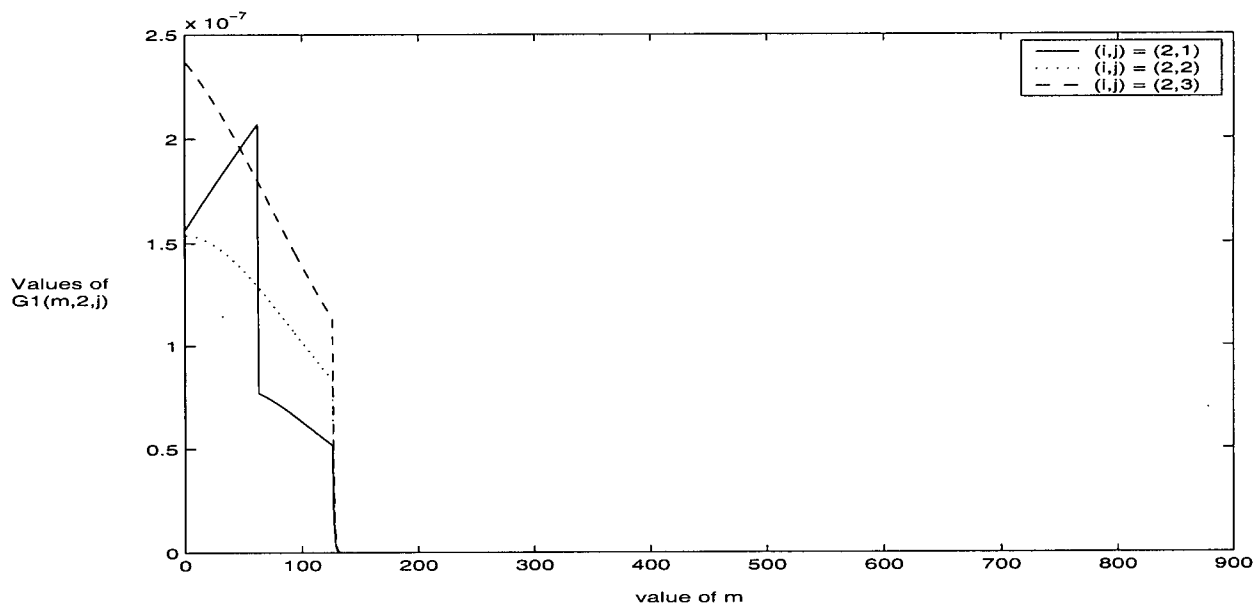
Appendix H

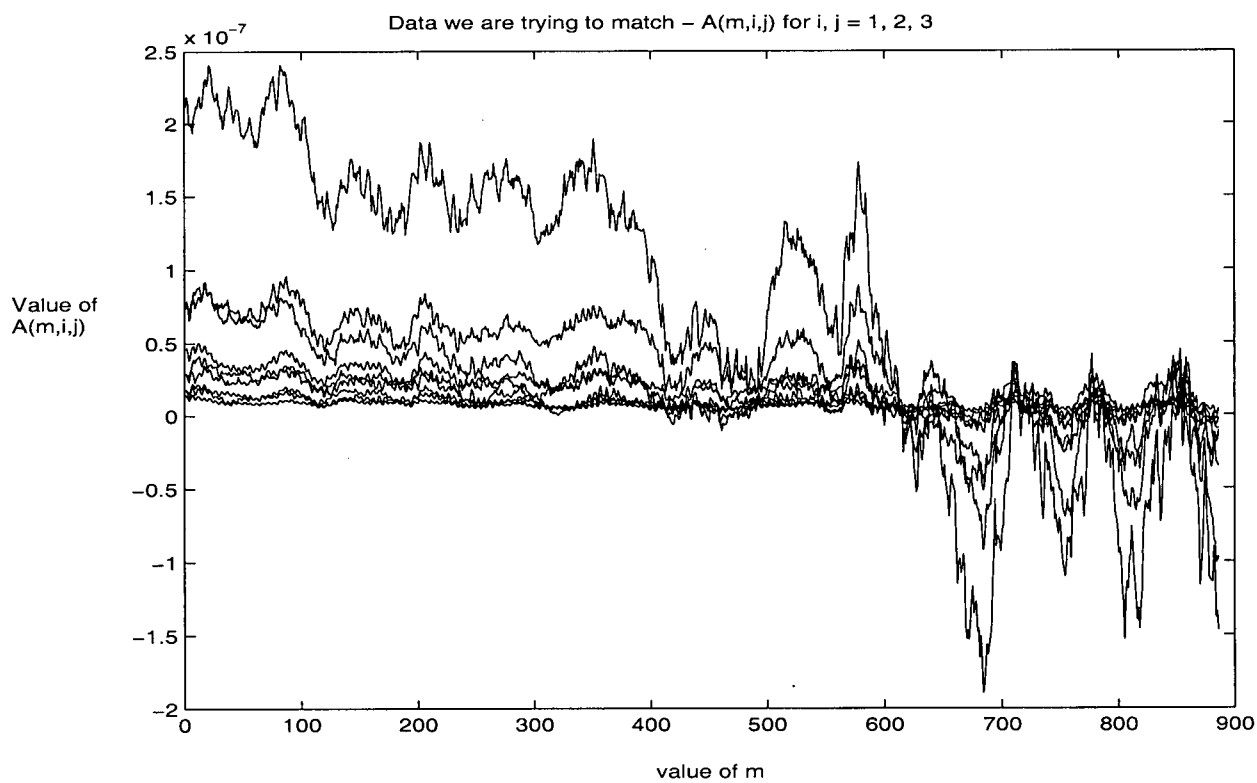
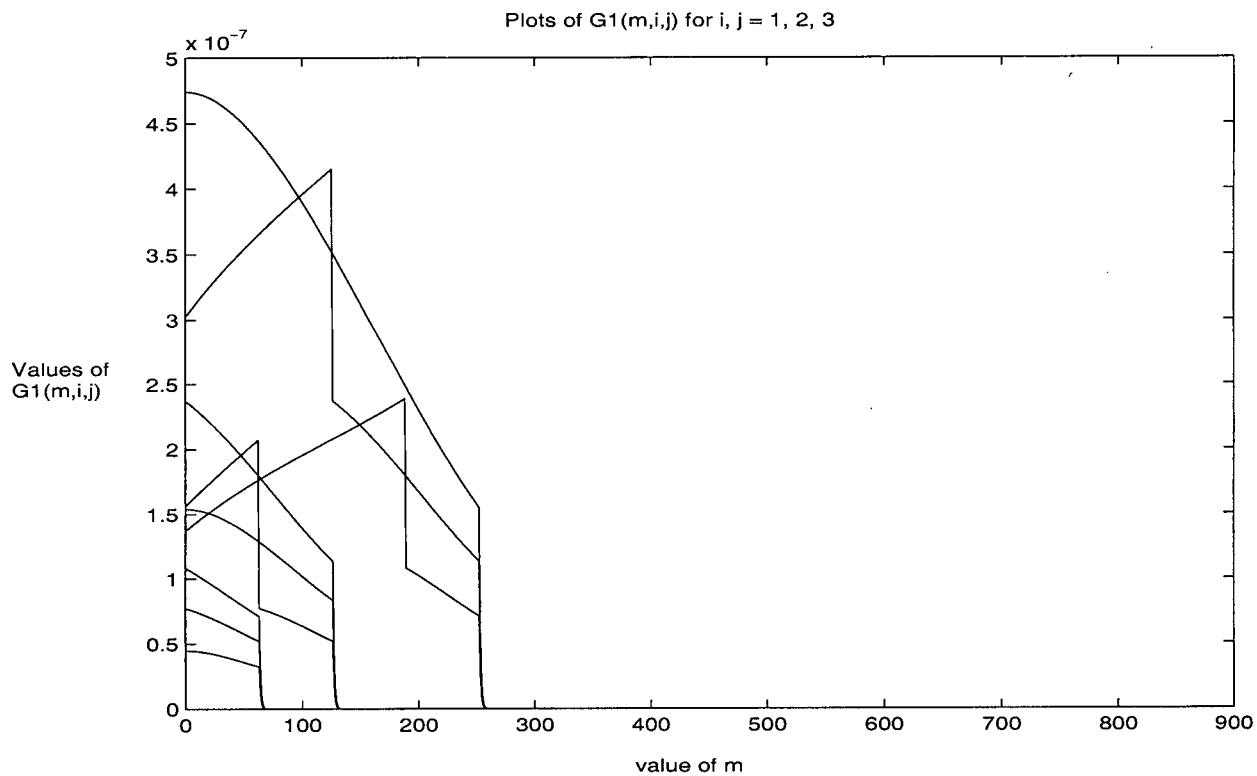
Plots of the function $G_1(m, i, j)$, the estimator of $A_{m,i,j}$ when using $c_1(k, u, z)$ as the covariance function where

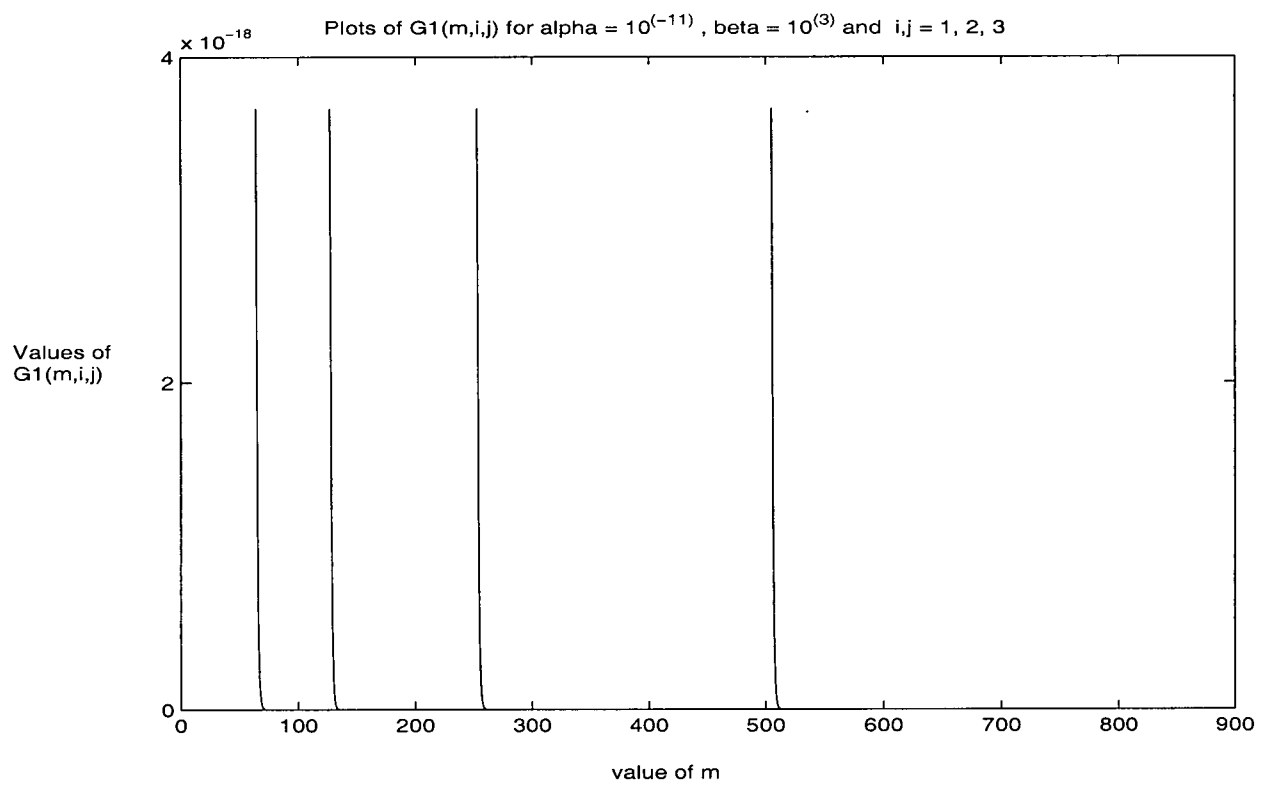
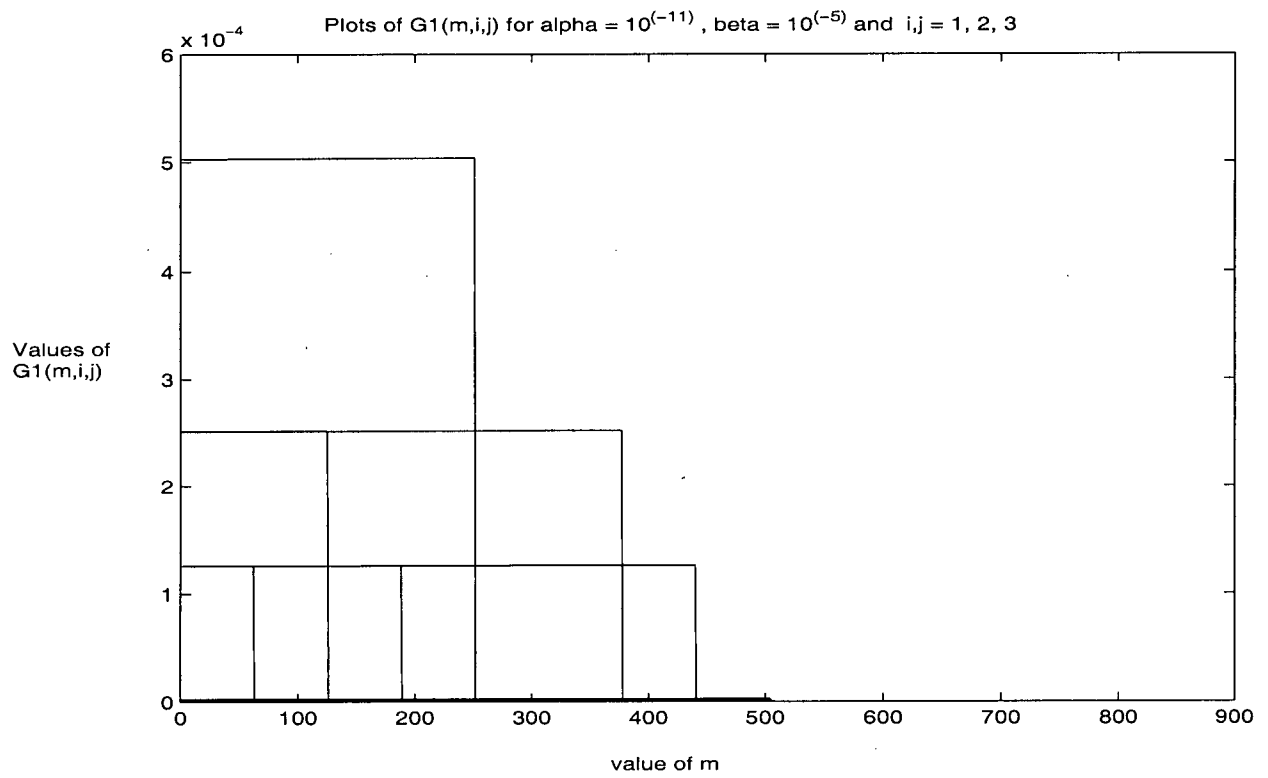
$$c_1(k, u, z) = k\alpha e^{-\beta|u-z|}, \quad \beta > 0$$

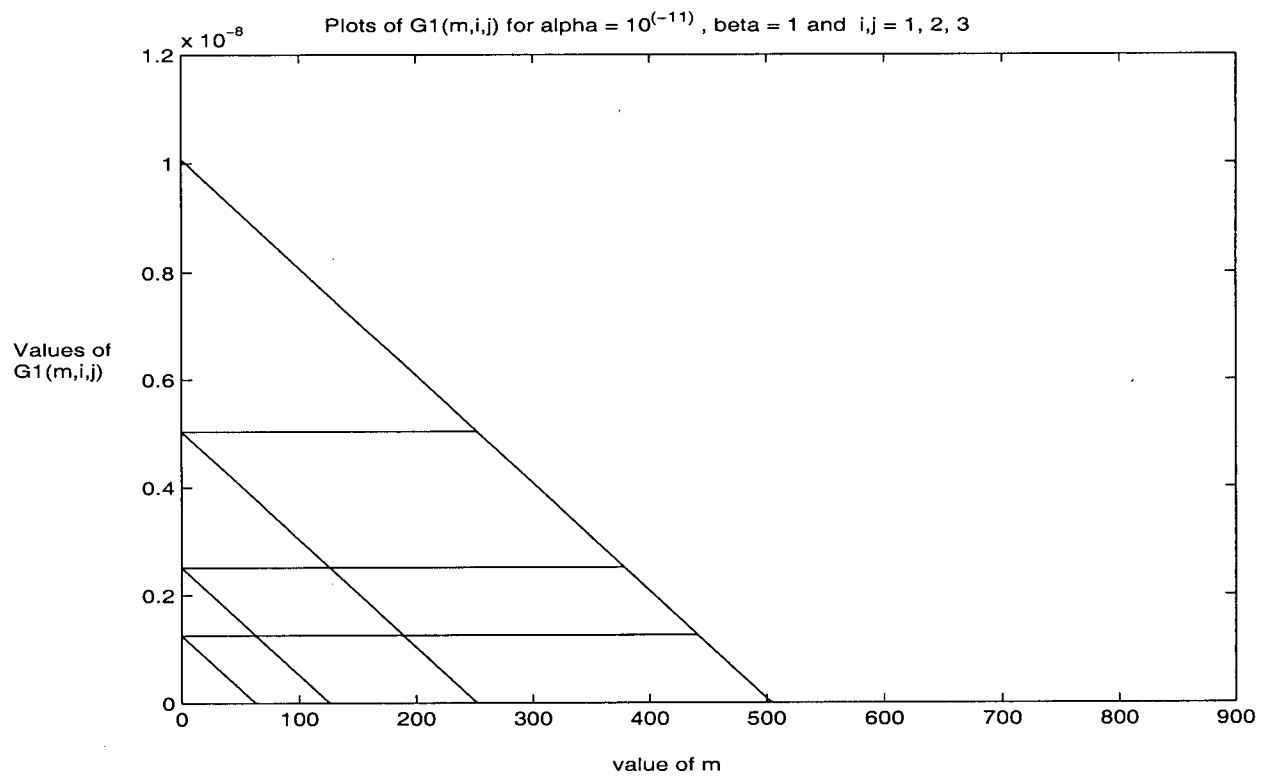
Using a non-linear optimization algorithm we found $\alpha = 1.32 \cdot 10^{-11}$ and $\beta = 0.08$ as the optimal values for the coefficients. When looking at the plots keep in mind that $G_1(m, i, j)$ is continuous. The jumps we see on the plots are not the caused because of discontinuities but because of a rapid change in the values of $G_1(m, i, j)$. The last three graphs of this section show the behaviour of $G_1(m, i, j)$ for non-optimal values of β (α doesn't concern us that much since it just a scaler) - we added them so that the reader can get a better view of this functions behaviour.











Appendix I

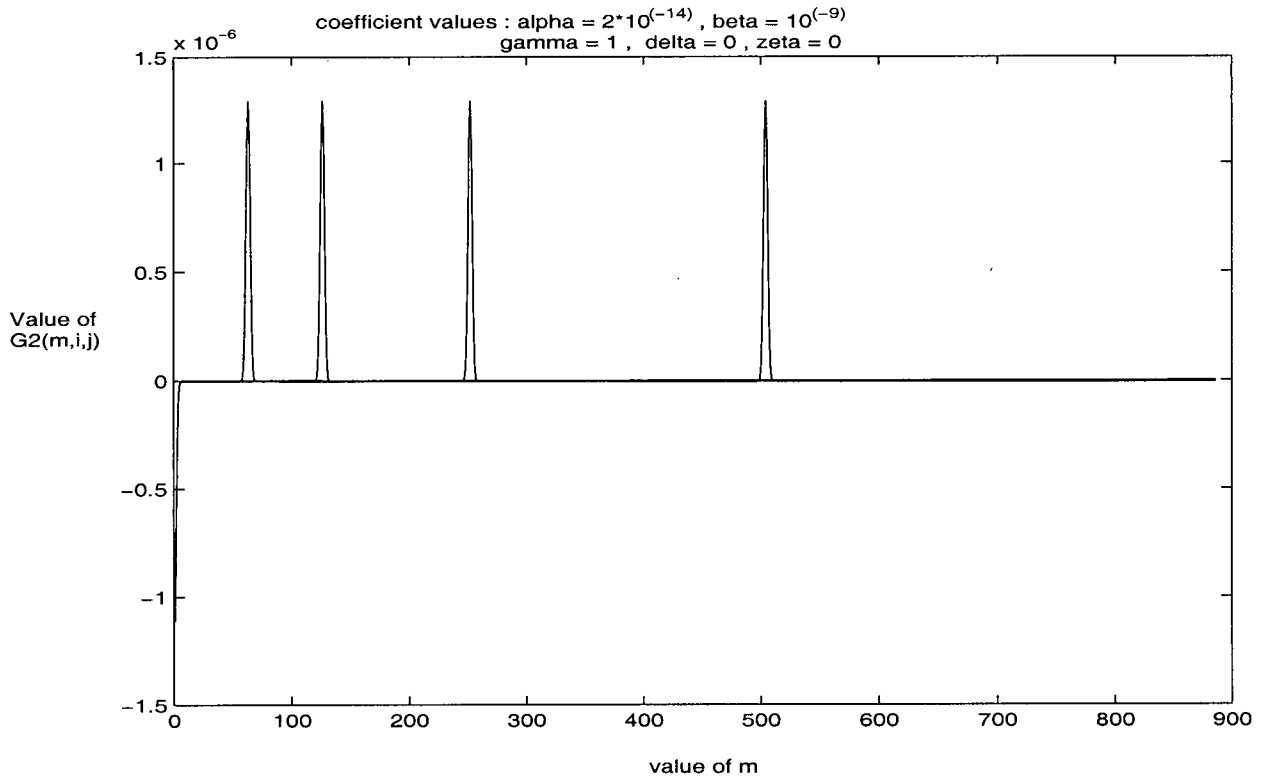
Plots of the function $G_2(m, i, j)$, the estimator of $A_{m,i,j}$ when using $c_2(k, u, z)$ as the covariance function where

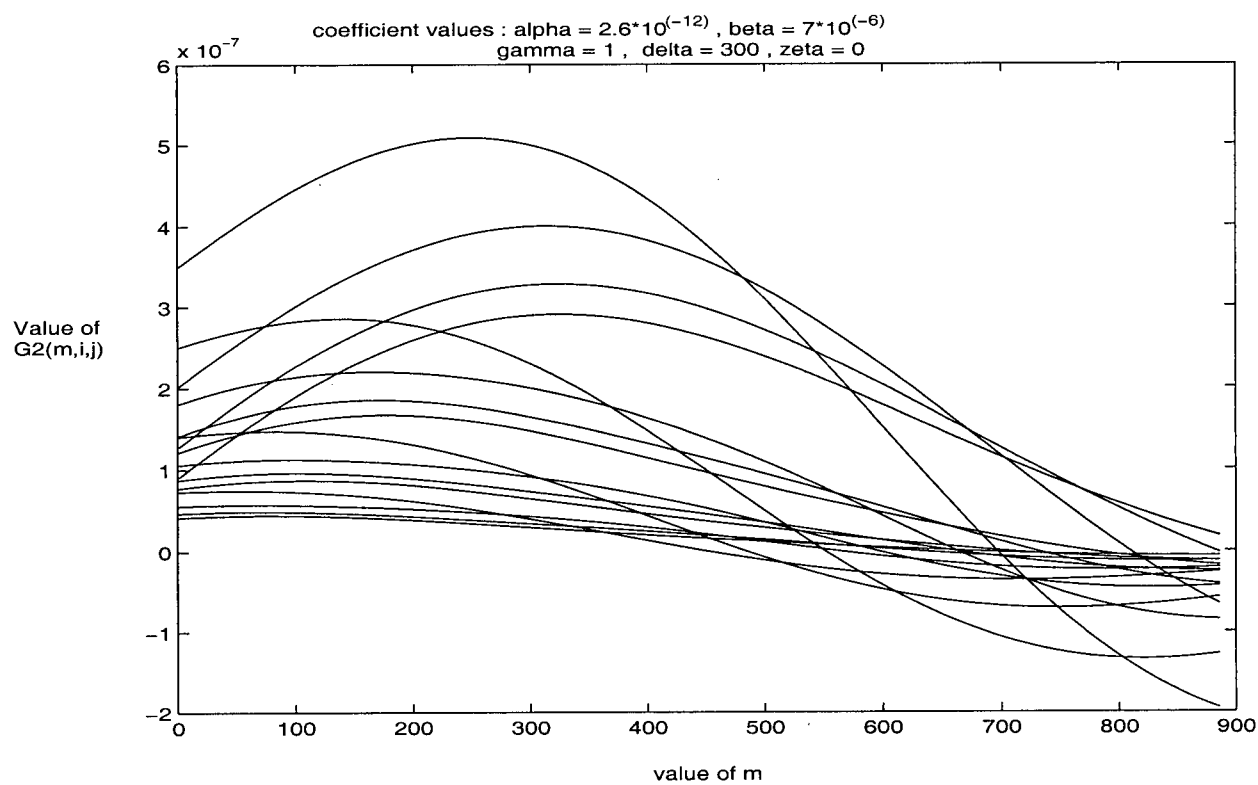
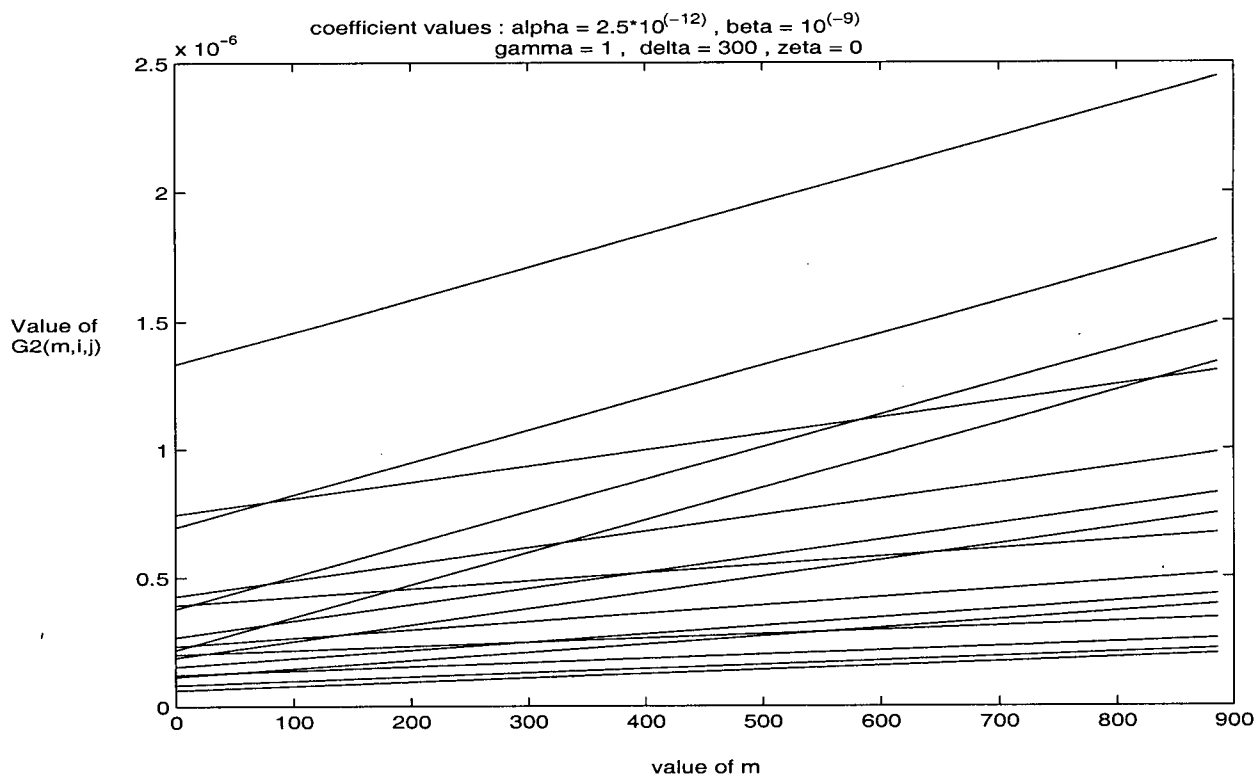
$$c_2(k, u, z) = k \cdot \alpha \cdot [e^{-\beta(\gamma \cdot u - z + \delta)^2} + e^{-\beta(\gamma \cdot z - u + \delta)^2}] + k \cdot \zeta, \quad \beta > 0$$

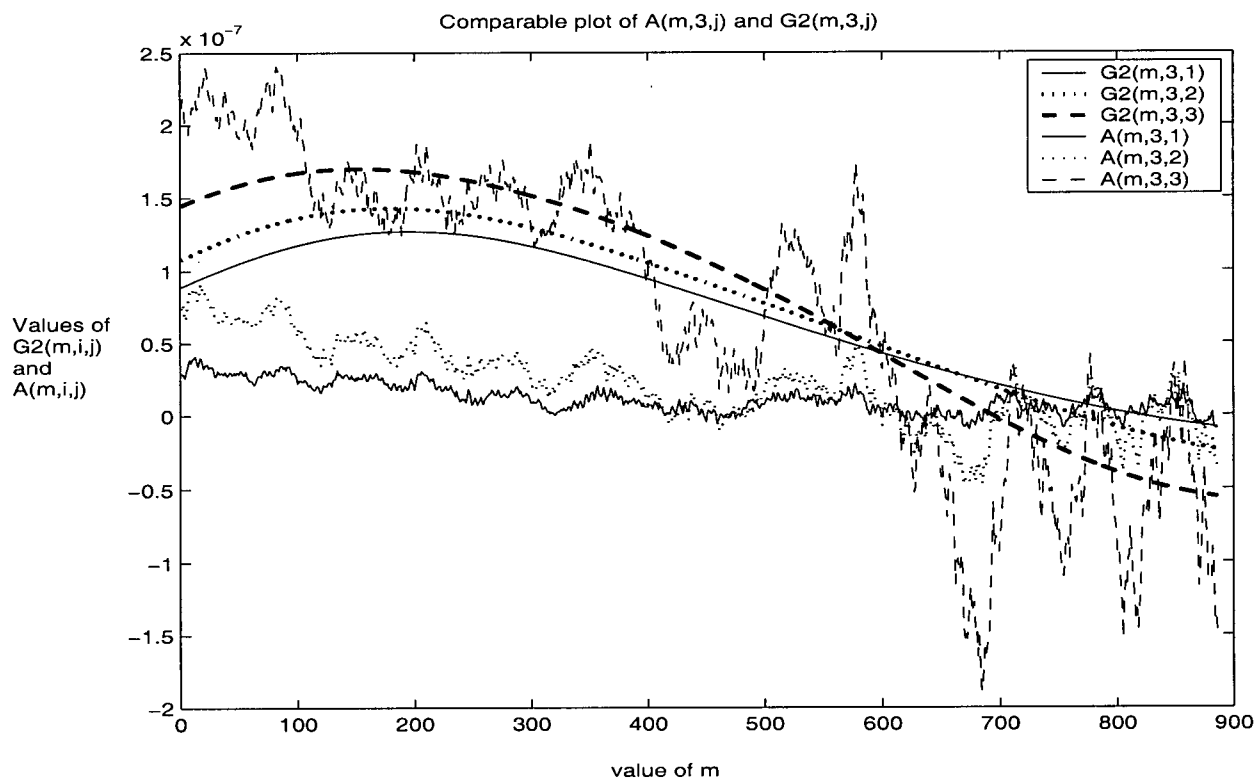
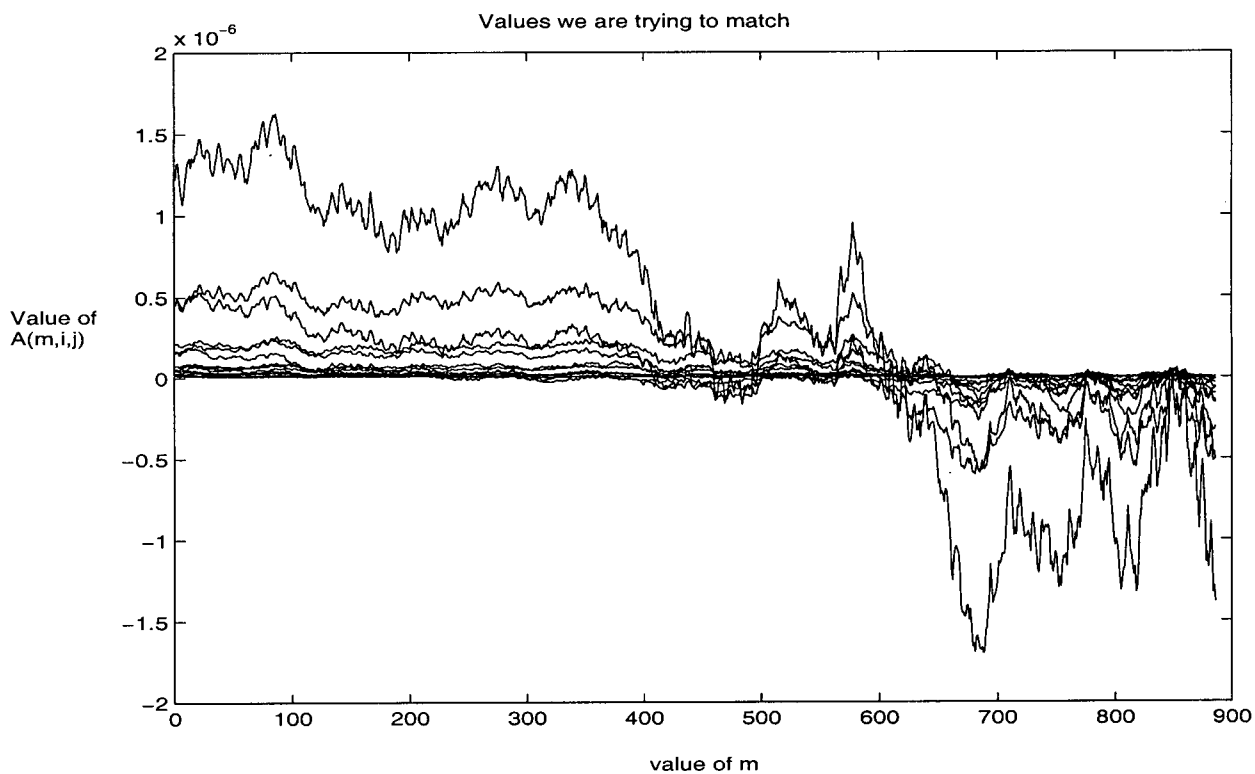
For the calculation of $G_2(m, i, j)$ we used approximation (i).

The first two graphs present the form of $G_2(m, i, j)$ if β is given a value close to the constraints mentioned in Appendix F. In the third graph we have $G_2(m, i, j)$ with the coefficients which make it fit best to $A_{m,i,j}$. In each graph there are 16 functions, the $G_2(m, i, j)$ for $i, j = 1, 2, 3, 4$.

We should note that in the first graph (where $\beta = 10^{-9}$, $\delta = 300$) each spike appears at one of the (fixed) bond maturities T_i . Moreover all 4 functions $G_2(m, i, j)$ which correspond to that specific i have exactly the same behaviour. In other words if $G_2(m, i, j)$ has the coefficients of the first graph then for each i the functions that correspond to the various j and this fixed i will be almost equal to 0 at all times except from $m = T_i$, where they will all have a spike (of the same height).





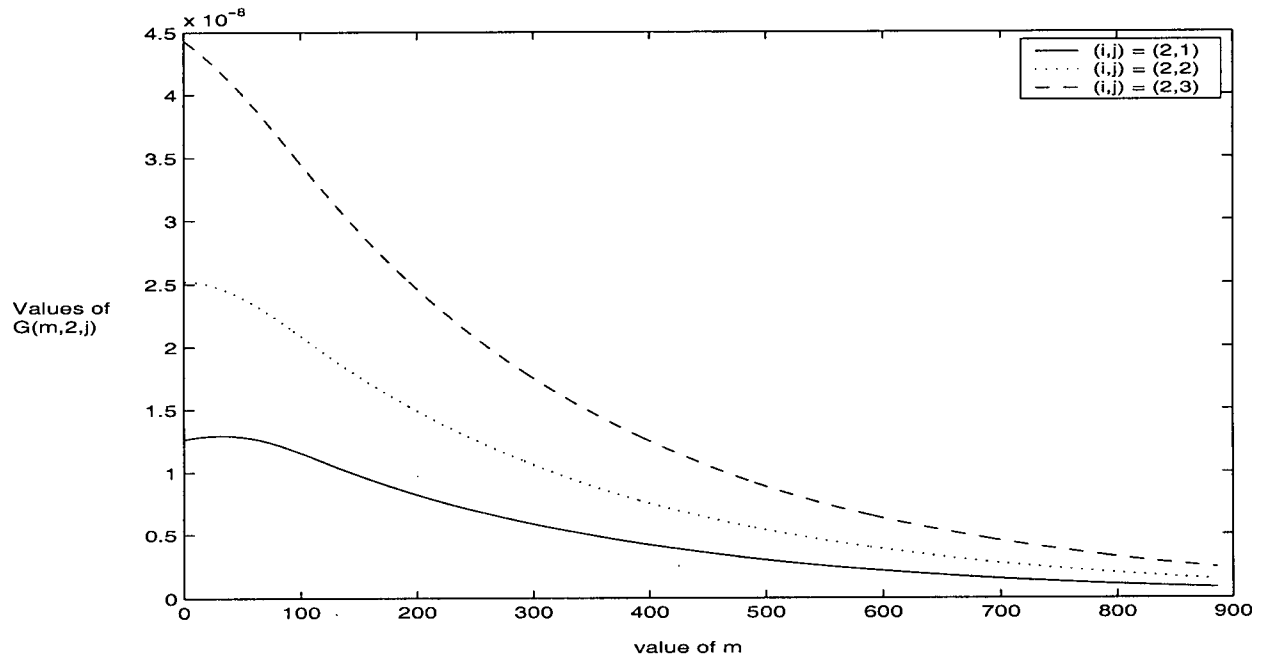
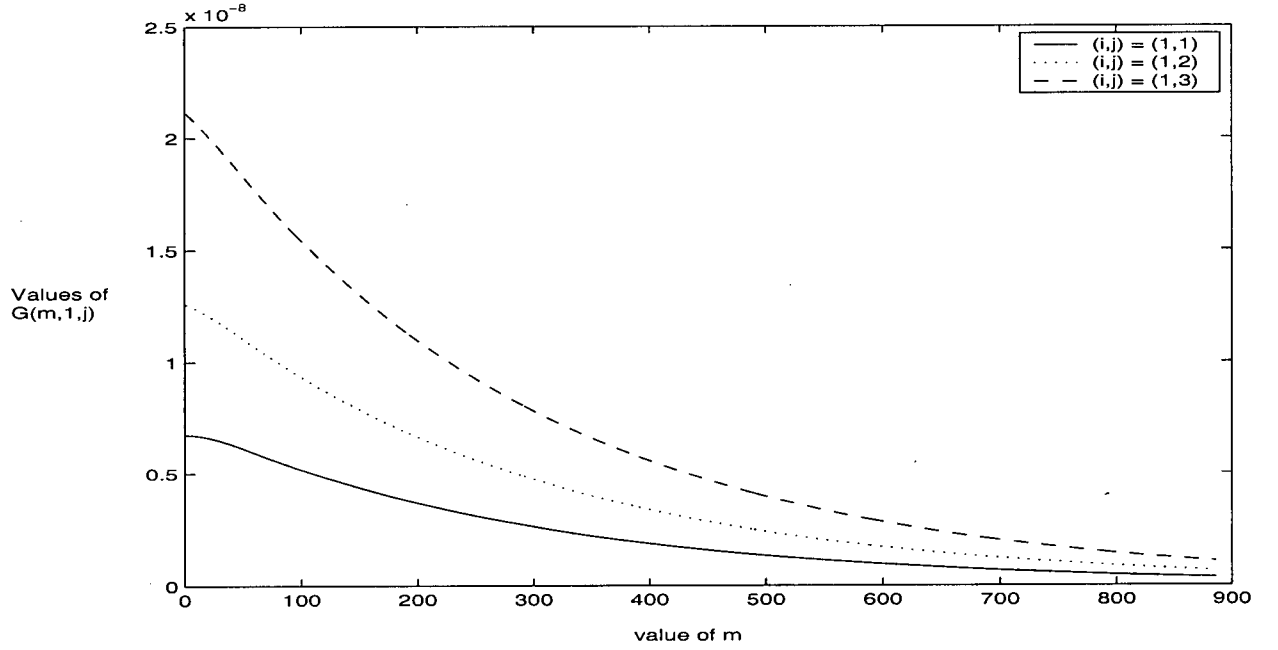


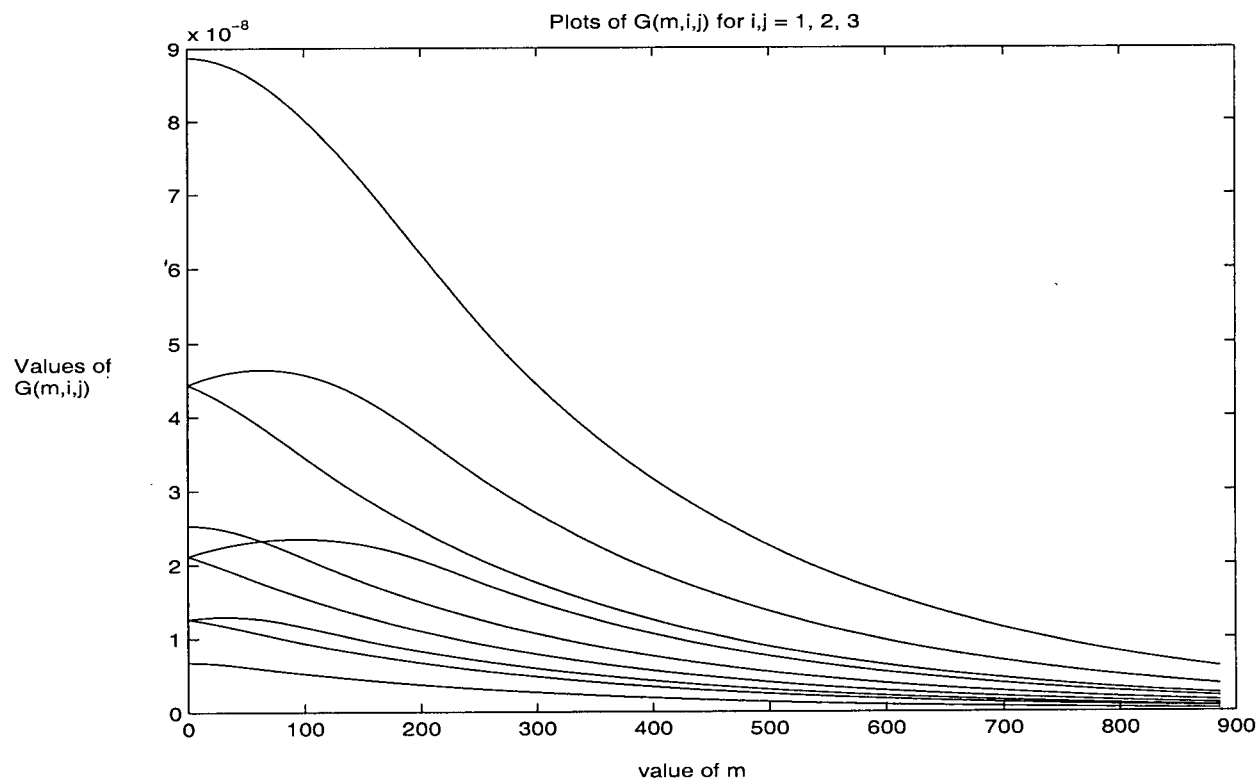
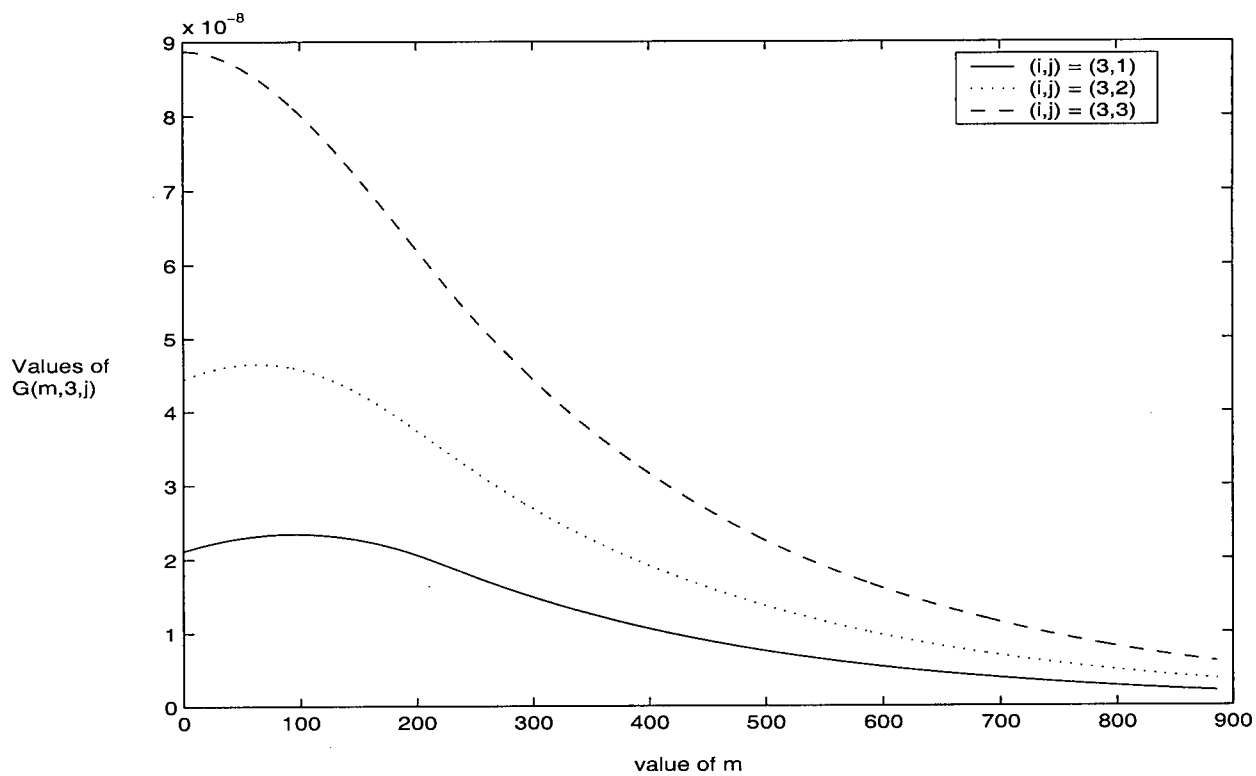
Appendix J

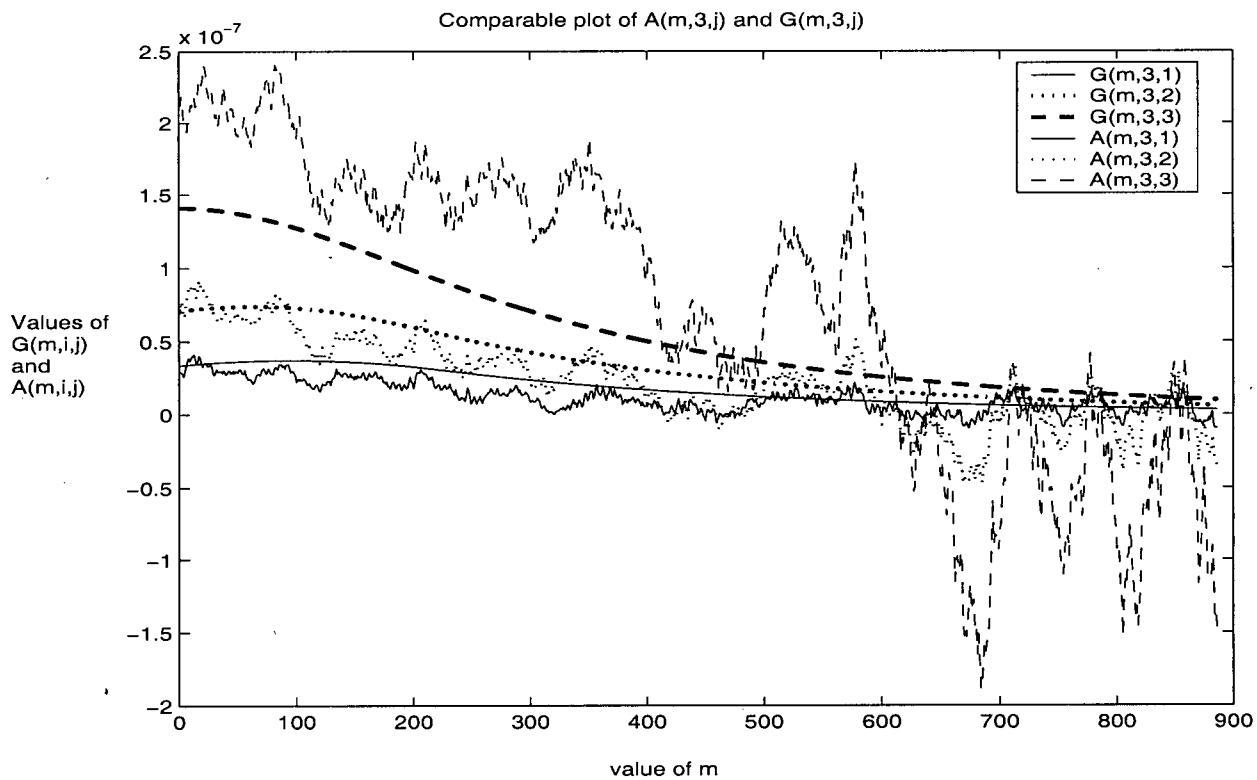
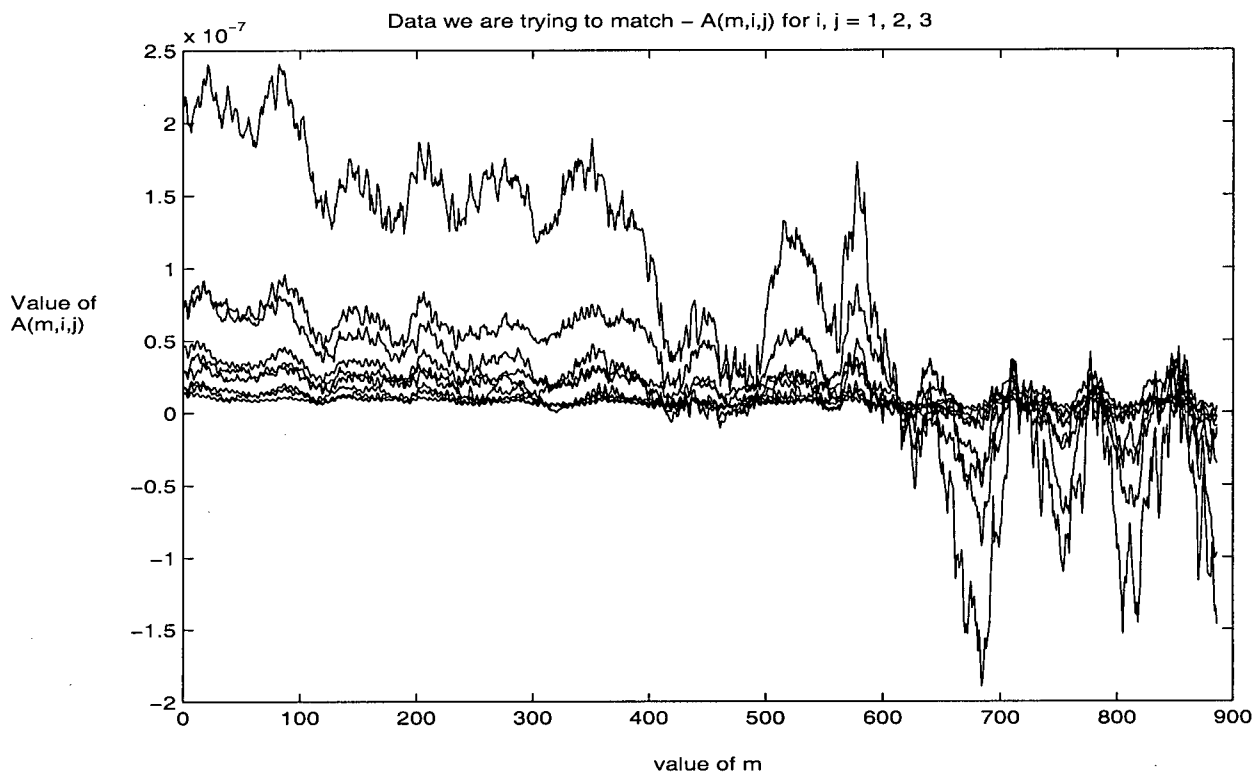
Plots of the function $G(m, i, j)$, the estimator of $A_{m,i,j}$, when using $c_3(k, u, z)$ as the covariance function where

$$c_3(k, u, z) = \sigma^2 \cdot e^{\lambda(k-u \wedge z) - \mu \cdot |u-z|}$$

Using a non-linear optimization algorithm we found $\sigma = 1.35 \cdot 10^{-6}$, $\mu = 0.0034$ and $\lambda = 2.17 \cdot 10^{-11}$ as the optimal values for the coefficients. We then got the following plots:







Appendix K

Definition of Gaussian Random Field, taken from ADLER (1981).

“Let $G^{2,1}$ denotes the set of all $\mathbb{R}^1$ -valued functions on $\mathbb{R}^2$ and $\wp^{2,1}$ the σ -field containing all sets of the form $\{ g \in G^{2,1}: g(t_j) \in B_j, j = 1, \dots, m \}$ where m is an arbitrary integer, the t_j are points of $\mathbb{R}^2$, and the B_j are open sets in $\mathbb{R}^1$.

Then we define a 2-dimensional random field to be a measurable mapping X from (Ω, F) into $(G^{2,1}, \wp^{2,1})$. We use the notation $X(t, \omega)$ to denote the value the function in $G^{2,1}$ corresponding to ω takes at the point t . for convenience we usually suppress the ω .

Given the existence of the probability measure P on $\wp$ we can immediately obtain from the definition a collection of measures $F_{t_1, \dots, t_n}$ on B^{n1} defined by

$$F_{t_1, \dots, t_n}\{B\} = P\{(X(t_1), \dots, X(t_n)) \in B\}$$

for any $B \in B^{n1}$. The collection of all such measures, or, equivalently, the corresponding distribution functions, is known as the family of finite-dimensional distributions for the field X .

We define Gaussian random field to be a random field possessing finite dimensional distributions all of which are multivariate Gaussian. From this definition and the above comments it is clear that all the finite dimensional distributions of a real valued Gaussian process, and hence the measures they induce on $\wp^{2,1}$ are completely determined once we specify the following two functions, known, respectively, as the mean and covariance functions:

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

$$R(s, t) = E[(X(s) - \mu(s))^T (X(t) - \mu(t))]$$

It can be easily seen from the form of the multivariate normal density that if a real valued Gaussian field has a constant mean and a covariance function that is dependant on $s - t$ only, then the field is homogeneous.”