LUND UNIVERSITY

Experiment with an artificial time series

Åström, Karl Johan; Wensmark, Sture

1965

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Åström, K. J., & Wensmark, S. (1965). *Experiment with an artificial time series*. (IBM Technical Paper (TP); No. 18.146). IBM Nordic Laboratory.

Total number of authors:
2

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Download date: 19. Dec. 2025

EXPERIMENT WITH AN ARTIFICIAL TIME SERIES

K. J. Åström and S. Wensmark

TP 18.146
Technical Paper
April 20, 1965

IBM NORDIC LABORATORY
SWEDEN

EXPERIMENT WITH AN ARTIFICIAL TIME SERIES

K. J. Åström and S. Wensmark

SYNOPSIS

Results of simulation of the scheme for cross machine stretch estimation and prediction are given. In the simulation we use data with known statistical properties.

TP 18.146
Technical Paper
April 20, 1965

IBM NORDIC LABORATORY
SWEDEN

table of contents

INTRODUCTION

The purpose of this work was three-fold

- to test the identification program described in [7]
- to test the estimation scheme proposed in [8] and to perform a sensitivity study
- to judge the value of the asymptotic formulas for estimating the fluctuations in the sample covariance function due to finite observation time

We generated an artificial time series having essentially the same statistics as the series analysed in [8]. This artificial time series was processed in the same way as was done in [8] with the experimental data.

The following operations were performed

- calculation of the sample means and covariances from the time series
- approximation of the sample covariance function by the covariance function of a time series generated by a dynamical system forced by independent disturbances
- construction of an optimal estimator on the basis of the approximated dynamical model
- simulation of the optimal estimator

As the time series was artificially generated, its properties are known, and we have possibilities to check the errors at each step.

The report is organized as follows. In section 2 we describe the generation of the artificial time series.

The analysis of the generated time series is described in section 3. The analysis consists of the calculation of the sample means and covariances. The sample fluctuations due to the finite length of the time series is evaluated using Bartletts [1] formulas. The agreement with the asymptotic formulas is good. In section 4, we give the solution to the inverse problem using the identification technique described in [7]. The FORTRAN programs required to do this are found in [9]. In section 5, we present an optimal Wiener estimator and predictor for the time series. The estimator is optimal in the sense of minimum mean square estimates and is calculated from recursive equations given by Kalman [6]. The FORTRAN programs required to calculate the estimator are given in [9]. Finally in section 6 we simulate the estimator on the artificially generated time series. The estimates are compared with the true values. All numerical computations were done on an IBM 1401. Descriptions of the programs used as well as listings and outputs are given in [9].

GENERATION OF AN ARTIFICIAL TIME SERIES

An artificial time series $\{y(t),\ t = 1, 2, \ldots, N\}$ was generated using the recursive equations

$$x_1(t + 1) = x_2(t) \qquad (2.1)$$

$$x_2(t + 1) = -0.7x_1(t) + 1.5x_2(t) + 0.0548\ v(t)$$

$$y(t) = 4.0 + x_1(t) + 0.158\ w(t) \qquad (2.2)$$

where $\{v(t),\ t = 1, 2, \ldots\}$ and $\{w(t),\ t = 1, 2, \ldots\}$ are sequences of independent gaussian random variables with zero means and the covariances

$$\text{cov.}\{v(t),\ v(s)\} = \delta_{t,s} \qquad (2.3)$$

$$\text{cov}\ \{v(t),\ w(s)\} = 0 \qquad (2.4)$$

$$\text{cov}\ \{w(t),\ w(s)\} = \delta_{t,s} \qquad (2.5)$$

The random numbers v(t) and w(t) were generated as sums of 12 rectangular pseudo-random numbers obtained from a modified Fibonacci series. A random number generator GAUSSF was written as a FORTRAN function in SPS. A listing of this routine is given in [9]

When generating the series $\{y(t),\ t = 1, 2, \ldots\}$ the following initial values were chosen for x_1 and x_2

$$x_1(0) = x_2(0) = 4.0$$

In order to achieve steady state, 100 values of x_1 and x_2 were first generated from (2.1) before the generation of the y process was started. From there on, new values of x(t) and y(t) were generated sequantially from (2.1) and (2.2). At each step the triplet $\{t,\ x_1(t),\ y(t)\}$ was stored on tape. For further details we refer to the program description and listing in [9]

The generated time series is listed in [9]. A sample of the generated time series is also given in Figure 2.1.

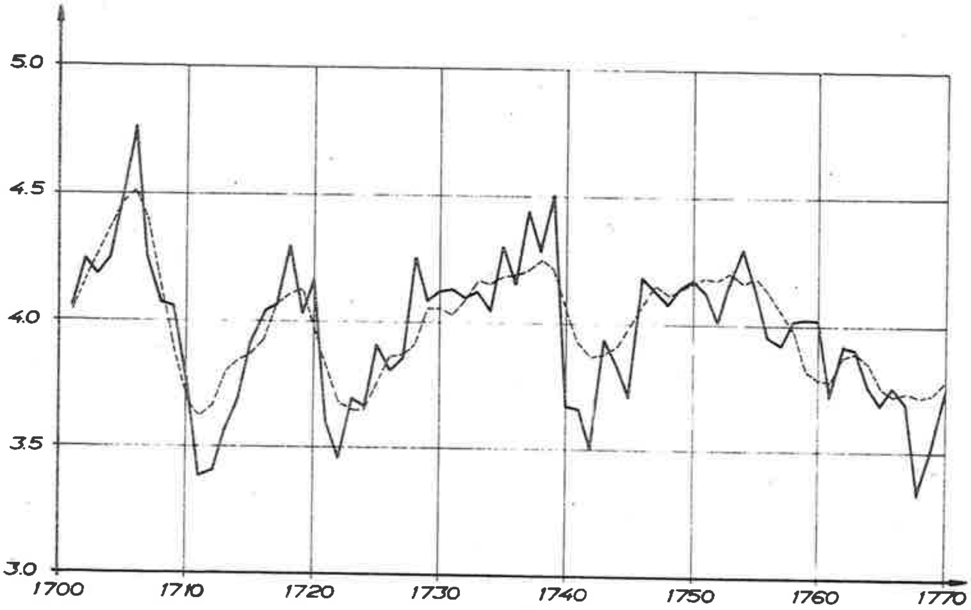

Figure 2.1

Sample of the generated time series $\{x_1(t), y(t), t = 1701 - 1770\}$. The solid line shows $x_1(t)$ and the dotted line $y(t)$.

The covariance function of the generated time series is given by

$$r(t) = \begin{cases} 0.026562496 + 0.158^2 = 0.051526496 & t = 0 \\ 0.023437495 & t = 1 \\ 0.016562495 & t = 2 \\ 1.5\, r(t-1) - 0.7\, r(t-2) & t \geq 2 \end{cases} \qquad (2.6).$$

The characteristic equation of (2.1) has the roots

$$\lambda_1 = 0.75 + 0.3708\, i$$

$$\lambda_2 = 0.75 - 0.3708\, i$$

ANALYSIS OF THE GENERATED TIME SERIES

Calculation of the sample means and sample covariances

We will now analyse the generated time series $\{y(t), t = 1, 2, \ldots\}$. The sample mean is defined by

$$m = \frac{1}{n} \sum_{s=1}^{n} y(s) \tag{3.1}$$

The sample means m were calculated for 10 samples of 500 values each. The results of the computations are shown in Table I.

Table I

Sample	Average of y
1-500	4.0068780
501-1000	3.9836884
1001-1500	3.9989198
1501-2000	3.9934112
2001-2500	4.0229672
2501-3000	3.9672824
3001-3500	4.0128164
3501-4000	3.9982452
4001-4500	3.9938372
4501-5000	4.0044870
1-5000	3.9982028

The sample covariance function c(t) is defined by

$$c(t) = \frac{1}{n-t} \sum_{s=1}^{n-t} [y(s)-m]\,[y(s+t)-m] \tag{3.2}$$

The sample covariances were also calculated for 10 samples of 500 values each. The FORTRAN programs used and the computer outputs are given in [9].

In Table II we have also given the average covariance function for the 10 samples, the true covariance function and an estimate of the sample fluctuations based on the 10 samples. In Figure 3.1 A-K we have graphed the results.

Table II

t	SAMPLE COVARIANCE FUNCTIONS OF ARTIFICAL TIME SERIES c(t) CALCULATED FROM SAMPLES OF 500 VALUES											AVERAGE OF D - X	COVARIANCE FUNCTION FROM SAMPLE OF 5000 VALUES	TRUE CONVARIANCE FUNCTION
	I	II	III	IV	V	VI	VII	VIII	IX	X				
	1 - 500	501-1000	1001-1500	1501-2000	2001-2500	2501-3000	3001-3500	3501-4000	4001-4500	4501-5000				
0	0.04469	0.05281	0.05450	0.06197	0.05764	0.04787	0.05439	0.04787	0.04840	0.05112	26.8	0.05213	0.05234	0.05153
1	0.01803	0.02621	0.02851	0.03359	0.02856	0.02202	0.02484	0.02245	0.02129	0.02331	19.6	0.02488	0.02511	0.02344
2	0.01262	0.01880	0.02123	0.02251	0.02152	0.01430	0.01788	0.01410	0.01561	0.01463	12.7	0.01732	0.01753	0.01657
3	0.00455	0.00742	0.01393	0.01317	0.01279	0.00865	0.01054	0.00819	0.00745	0.00623	10.2	0.00929	0.00951	0.00845
4	-0.00004	-0.00070	0.00312	0.00545	0.00160	0.00332	0.00250	0.00233	0.00211	-0.00191	4.6	0.00178	0.00207	0.00108
5	-0.00272	-0.01011	-0.00413	-0.00453	-0.00729	0.00079	-0.00469	-0.00060	-0.00360	-0.00792	10.9	-0.00448	-0.00422	-0.00430
6	-0.00564	-0.01189	-0.00554	-0.00912	-0.01133	-0.00306	-0.00548	-0.00677	-0.00857	-0.00936	8.0	-0.00768	-0.00749	-0.00721
7	-0.00409	-0.01489	-0.00932	-0.00952	-0.01234	-0.00248	-0.00942	-0.00658	-0.00643	-0.00724	13.7	-0.00823	-0.00804	-0.00781
8	-0.00316	-0.01097	-0.01203	-0.00786	-0.01280	-0.00465	-0.00661	-0.00684	-0.00783	-0.00332	11.8	-0.00761	-0.00743	-0.00667
9	-0.00058	-0.00555	-0.00878	-0.00603	-0.01137	-0.00120	-0.00258	-0.00447	-0.00540	-0.00333	11.0	-0.00493	-0.00477	-0.00454
10	0.00059	-0.00070	-0.00678	-0.00541	-0.00910	0.00185	-0.00197	-0.00251	-0.00064	0.00164	13.7	-0.00230	-0.00214	-0.00214
11	0.00273	0.00383	-0.00564	-0.00399	-0.00543	0.00293	-0.00258	-0.00296	0.00349	0.00335	16.0	-0.00043	-0.00026	-0.00003
12	0.00217	0.00625	-0.00291	-0.00595	-0.00298	0.00137	-0.00061	-0.00172	0.00523	0.00252	15.0	0.00034	0.00058	0.00145
13	0.00050	0.00666	-0.00503	-0.00512	-0.00255	0.00306	0.00053	-0.00161	0.00394	0.00045	14.5	0.00008	0.00035	0.00220
14	0.00051	0.00659	-0.00341	-0.00443	0.00376	0.00461	0.00206	-0.00174	0.00537	0.00174	14.1	0.00151	0.00168	0.00178
15	0.00237	0.00363	-0.00416	-0.00317	0.00318	0.00264	0.00087	-0.00249	0.00395	-0.00224	9.4	0.00046	0.00074	0.00113
16	0.00601	0.00121	-0.00410	-0.00024	0.00113	-0.00079	0.00386	-0.00037	0.00245	-0.00228	8.6	0.00069	0.00089	0.00045
17	0.00464	-0.00033	-0.00136	0.00202	0.00228	-0.00068	0.00558	0.00044	0.00004	-0.00614	10.9	0.00065	0.00069	0.00012
18	0.00150	-0.00146	0.00004	0.00504	0.00132	-0.00291	0.00978	0.00171	-0.00094	-0.00513	17.5	0.00090	0.00113	0.00050
19	-0.00171	-0.00255	0.00131	0.00547	-0.00151	-0.00365	0.00757	0.00493	-0.00415	-0.00296	18.2	0.00028	0.00045	0.00067
20	-0.00068	-0.00540	0.00317	0.00254	-0.00074	-0.00271	0.00774	0.00543	-0.00527	-0.00342	20.3	0.00007	0.00009	0.00066

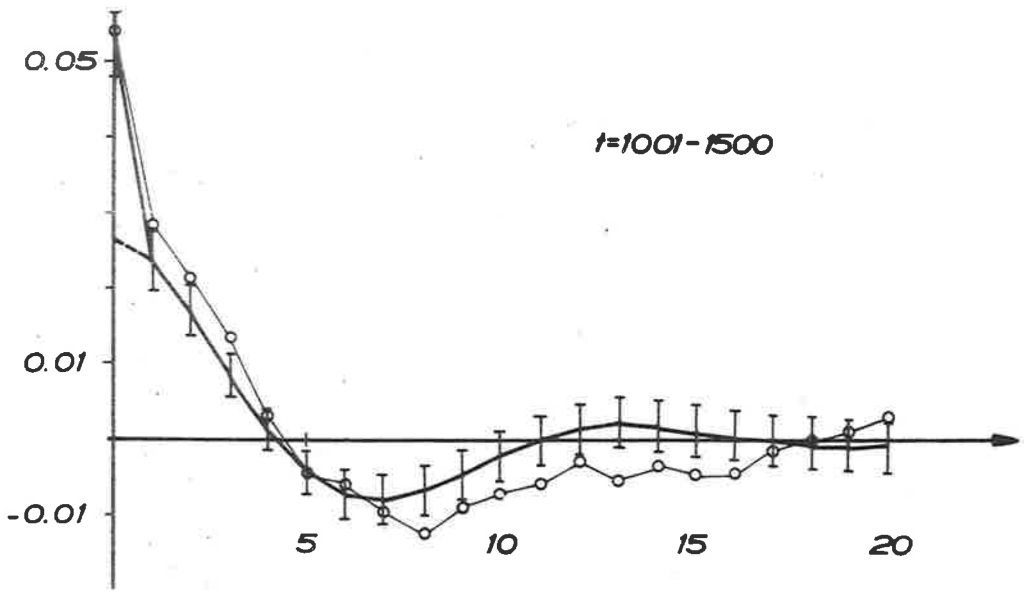

Figure 3.1 C

Sample covariance function calculated from 500 values of the artificial time series. The solid line shows the true covariance function.

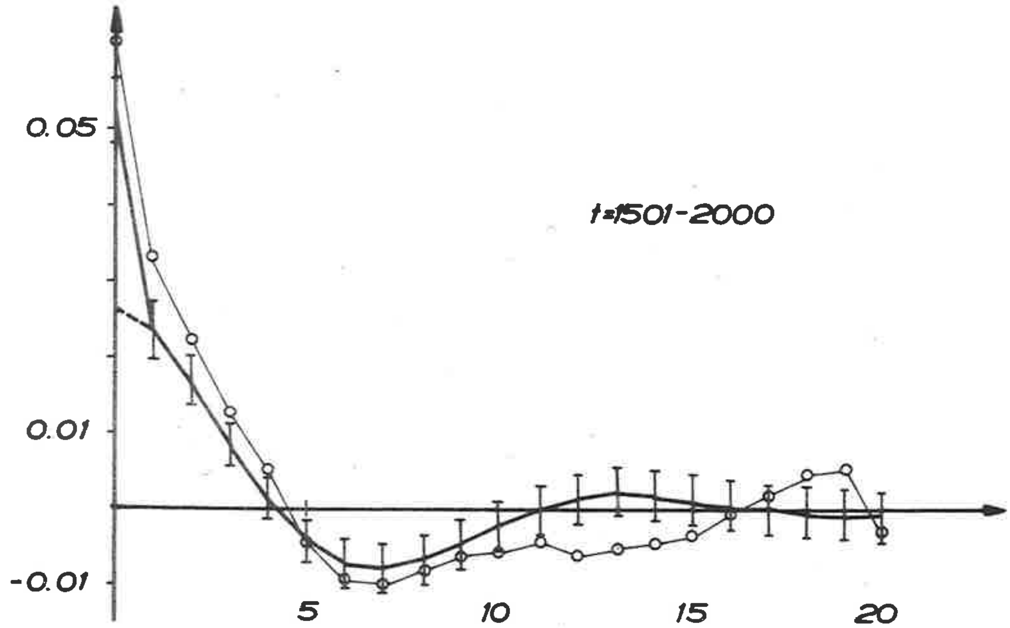

Figure 3.1 D

Sample covariance function calculated from 500 values of the artificial time series. The solid line shows the true covariance function.

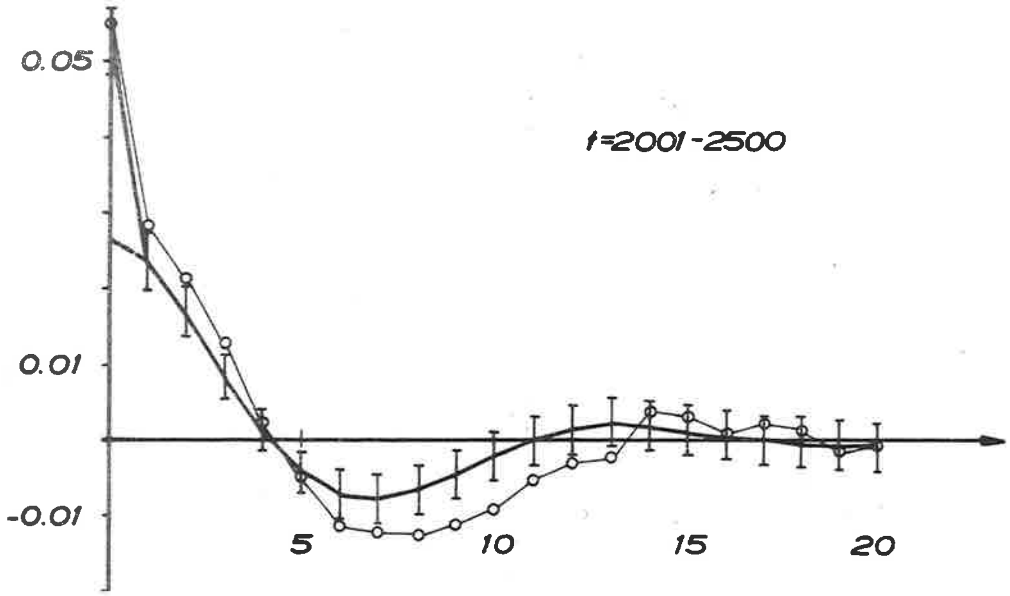

Figure 3.1 E

Sample covariance function calculated from 500 values of the artificial time series. The solid line shows the true covariance function.

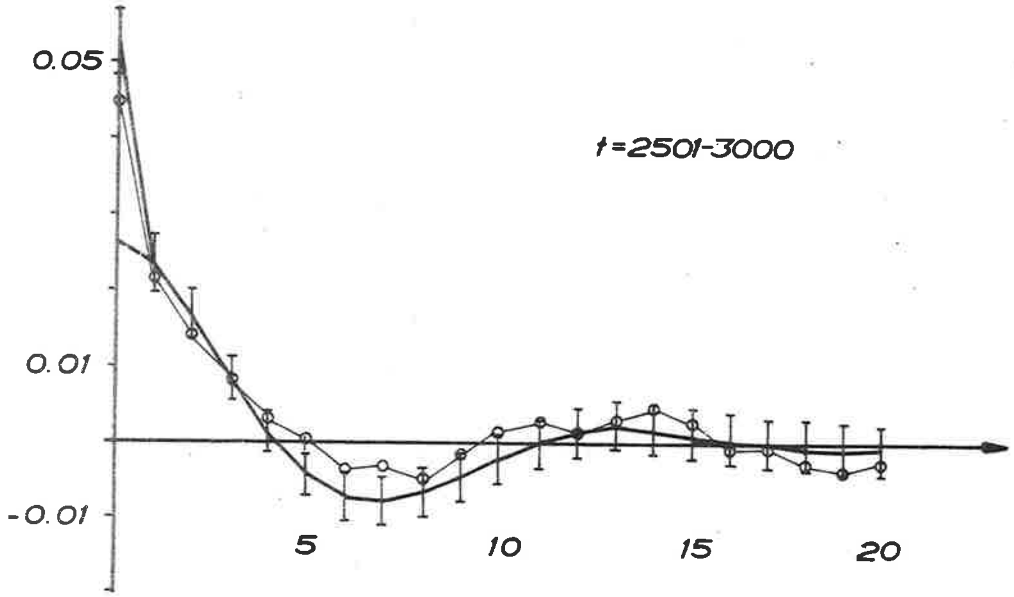

Figure 3.1 F

Sample covariance function calculated from 500 values of the artificial time series. The solid line shows the true covariance function.

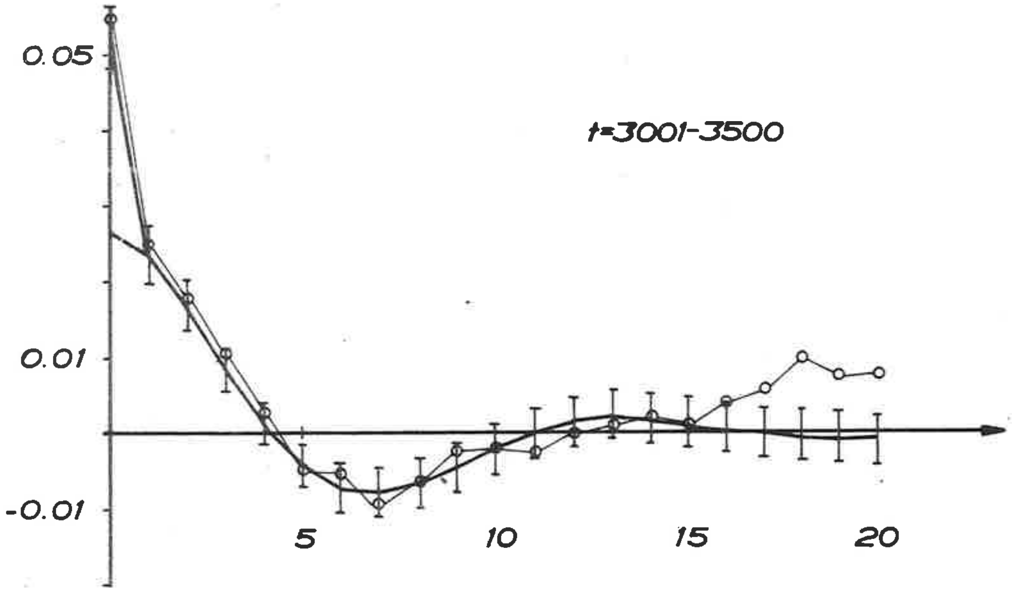

Figure 3.1 G

Sample covariance function calculated from 500 values of the artificial time series. The solid line shows the true covariance function.

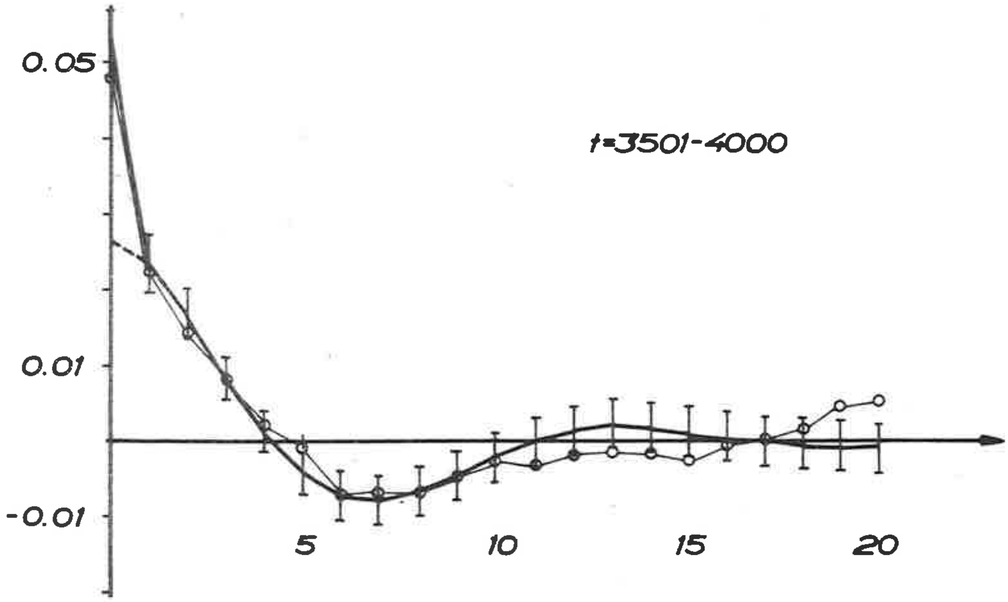

Figure 3.1 H

Sample covariance function calculated from 500 values of the artificial time series. The solid line shows the true covariance function.

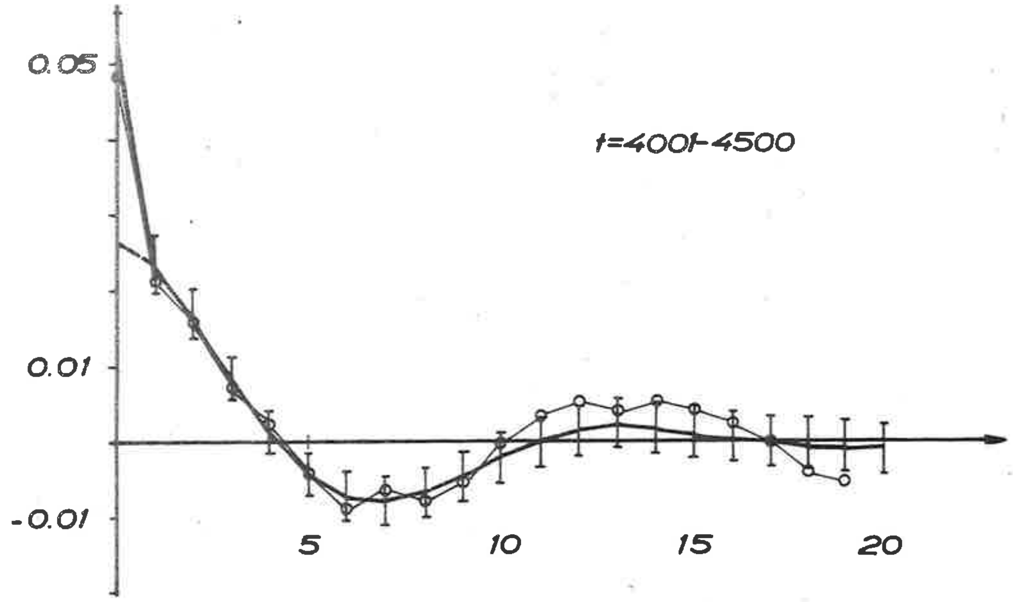

Figure 3.1 I

Sample covariance function calculated from 500 values of the artificial time series. The solid line shows the true covariance function.

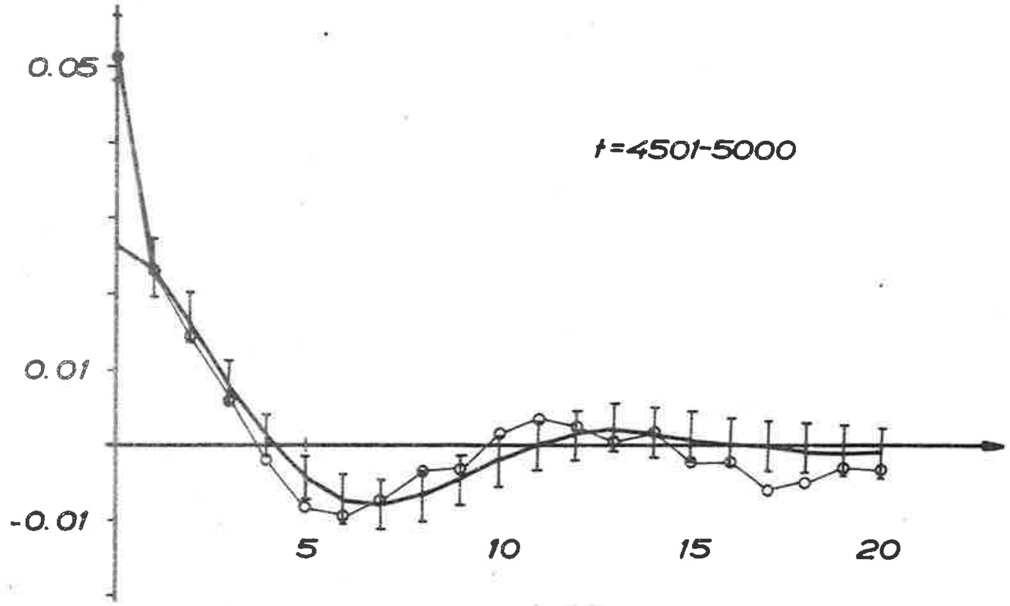

Figure 3.1 K

Sample covariance function calculated from 500 values of the artificial time series. The solid line shows the true covariance function.

Analysis of the sample fluctuations

It is seen from tables I and II that there are fluctuations in the sample means and covariances. These fluctuations occur because the sample means and covariances were calculated from a finite sample consisting of 500 points of the time series. Bartlett [1, chapter 9] has given formulas which expresses the sample fluctuations in terms of the true covariance function of the time series. The variance of the sample means is given by

$$\mathrm{var}\ (m - \mu) = \frac{1}{n} \left[r(0) + 2 \sum_{s=1}^{n-1} \left(1 - \frac{s}{n}\right) r(s) \right] \tag{3.3}$$

[1, p. 254] where μ is the average of the time series and $r(t)$ is its covariance function. Evaluating this formula for $r(t)$ given by (2.6) and $n = 500$ we get

$$\mathrm{var}\ (m - \mu) = 0.000200$$

which should be compared with the variance 0.000275 calculated from the ten sample means.

Now consider the sample covariance functions. If $\mu = 0$ we have

$$Ec(t) = r(t)$$

For example, if the process $\{y(t),\ t = 1, 2, \dots\}$ has zero mean, the sample covariances is an unbiased estimate of the covariance function. If the mean is not zero, we have to estimate it using (3.1) and then the sample covariance is a biased estimate of the covariance. The bias is given by

$$E\,[c(t) - r(t)] = - \frac{1}{n(n-s)} \sum_{i=-n+s+1}^{n-s+1} (n-s-|i|)\, r(s+i)$$

$$- \frac{1}{n(n-s)} \sum_{i=-s+1}^{s-1} (s-|i|)\, r(s+i) + \frac{s}{n^2(n-s)} \sum_{i=-n+1}^{n-1} (n-|i|)\, r(s+i) \tag{3.4}$$

We will also give a formula for the fluctuations in the covariance function. If the mean of the process is known, we have

$$\operatorname{cov}\left[c(t),\ c(s)\right] = \frac{1}{n-s}\Big\{ r(0)\, r(t-s) + r(s)\, r(t) + 2\sum_{i=1}^{n-t-1}\left(1-\frac{i}{n-t}\right)\left[r(i)\, r(t-s+i) + r(t+i)\, r(s-i)\right] + \sum_{i=1}^{t-s}\left[r(i)\, r(t-s-i) + r(t+i)\, r(s-i)\right]\Big\} \quad t \ge s \qquad (3.5)$$

Formulas (3. 4) and (3. 5) are given in [1, p. 254 - 255]. Notice that the bias of the sample covariance is asymptotically $O(\frac{1}{n})$ while the sample fluctuations are of the order $O(1/\sqrt{n})$.

The numerical values of var c(t) for the covariance function (2. 6) are listed in Table III for n = 50, 500 and 5000 respectively.

Notice that the numerical values show good agreement with the actual fluctuations of the sample covariances given in Table II. The numerical values of Table III were obtained using the FORTRAN program described in [9]. In this reference we have also given the original computer outputs. The results are illustrated in Figure 3. 2 A-C, where we have given the covariance function (2. 6) and 68% confidence intervals for the values of the sample covariances based on 50, 500 and 5000 values.

Table III Cov [c(t), c(t)] · 10^6

t	n=50	n=500	n=5000
0	190.6	19.50	1.954
1	154.1	15.52	1.553
2	103.5	10.07	1.005
3	76.70	7.255	0.7216
4	81.46	7.594	0.7543
5	90.19	8.268	0.8199
6	114.5	10.55	1.047
7	125.5	11.39	1.129
8	123.4	10.93	1.080
9	116.8	10.01	0.9871
10	114.0	9.503	0.9353
11	117.0	9.550	0.9377
12	122.2	9.762	0.9570
13	128.7	10.12	0.9903
14	133.1	10.23	0.9996
15	135.9	10.16	0.9913
16	138.6	10.07	0.9803
17	141.9	10.03	0.9743
18	146.3	10.06	0.9753
19	151.1	10.11	0.9784
20	156.3	10.17	0.9799

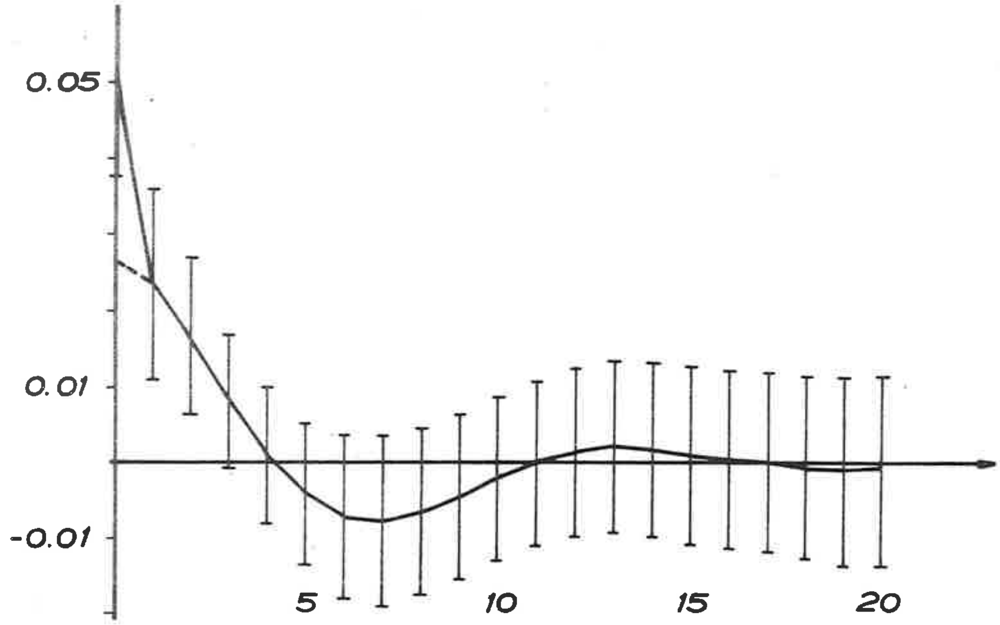

Figure 3.2 A

True covariance function and 68% confidence intervals for the values of sample covariances based on a sample of 50 values

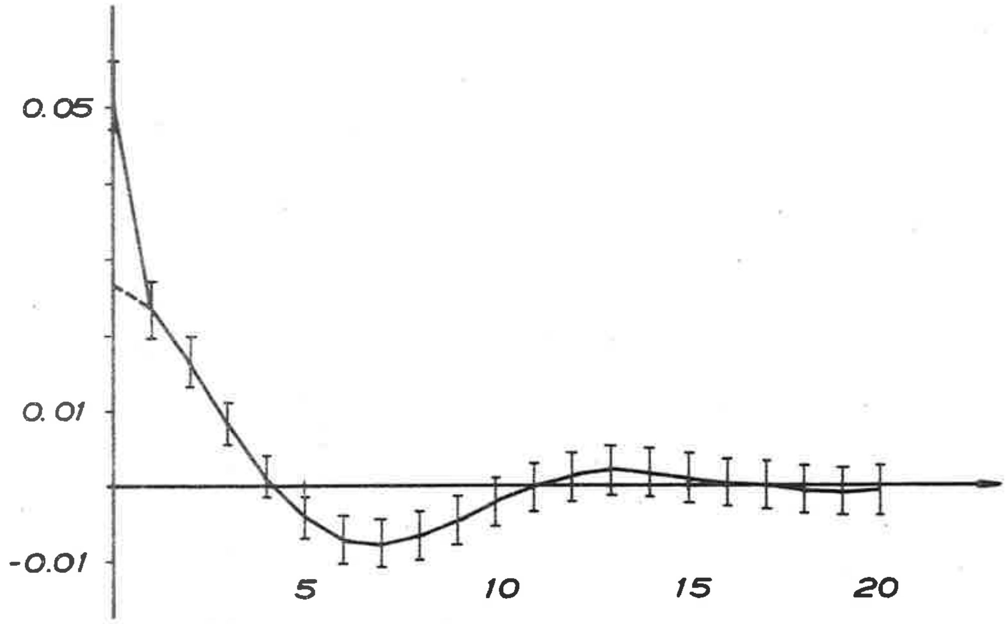

Figure 3.2 B

True covariance function and 68% confidence intervals of sample covariances based on samples of 500 values

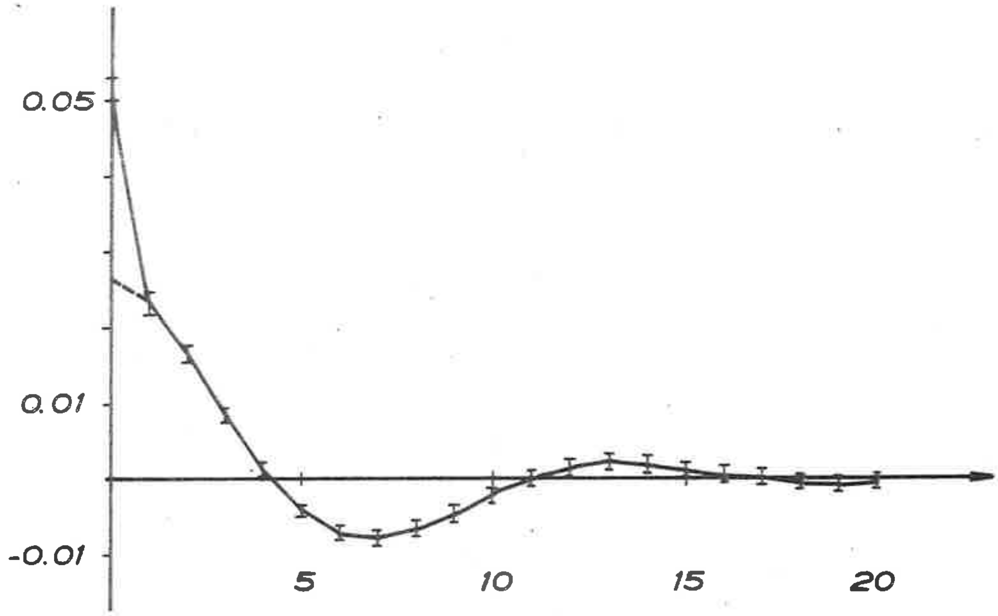

Figure 3. 2 C

True covariance function and 68% confidence intervals of sample covariances based on samples of 5000 values.

In Table IV we have also given the correlation matrix corresponding to the covariance matrix (3. 5).

Table IV

CORRELATION MATRIX

1.00	0.89	0.75	0.52	0.19	-0.14	-0.23	-0.24	-0.20	-0.14	-0.04	0.04	0.09	0.11	0.10	0.19	0.14	0.08	0.00	-0.01	-0.02
0.89	1.00	0.81	0.62	0.29	-0.05	-0.21	-0.27	-0.25	-0.19	-0.07	0.03	0.09	0.12	0.20	0.20	0.15	0.03	0.01	-0.01	-0.02
0.75	0.81	1.00	0.79	0.42	0.27	-0.02	-0.15	-0.25	-0.22	-0.12	-0.01	0.07	0.11	0.11	0.09	0.07	0.04	0.02	0.01	-0.01
0.52	0.62	0.79	1.00	0.82	0.52	0.30	0.05	-0.15	-0.20	-0.22	-0.08	-0.01	0.04	0.07	0.07	0.07	0.05	0.05	0.02	0.01
0.19	0.29	0.42	0.82	1.00	0.71	0.53	0.35	0.13	-0.08	-0.15	-0.20	-0.12	-0.06	-0.01	0.03	0.05	0.06	0.06	0.05	0.03
-0.14	-0.05	0.27	0.52	0.71	1.00	0.71	0.57	0.36	0.17	-0.06	-0.16	-0.20	-0.17	-0.11	-0.04	0.01	0.05	0.06	0.06	0.14
-0.23	-0.21	-0.02	0.30	0.53	0.71	1.00	0.67	0.68	0.34	0.15	-0.07	-0.16	-0.18	-0.15	-0.10	-0.04	0.02	0.05	0.06	0.08
-0.24	-0.27	-0.15	0.05	0.35	0.57	0.67	1.00	0.81	0.64	0.32	0.12	-0.09	-0.17	-0.18	-0.15	-0.09	-0.03	0.02	0.05	0.08
-0.20	-0.25	-0.25	-0.15	0.13	0.36	0.68	0.81	1.00	0.80	0.52	0.30	0.10	-0.10	-0.17	-0.18	-0.15	-0.09	-0.02	0.03	0.07
-0.14	-0.19	-0.22	-0.20	-0.08	0.17	0.34	0.64	0.80	1.00	0.72	0.53	0.30	0.11	-0.10	-0.17	-0.18	-0.15	-0.09	-0.02	0.03
-0.04	-0.07	-0.12	-0.22	-0.15	-0.06	0.15	0.32	0.52	0.72	1.00	0.73	0.54	0.31	0.11	-0.10	-0.17	-0.18	-0.15	-0.09	-0.03
0.04	0.03	-0.01	-0.08	-0.20	-0.16	-0.07	0.12	0.30	0.53	0.73	1.00	0.73	0.54	0.32	0.12	-0.09	-0.17	-0.18	-0.15	-0.10
0.09	0.09	0.07	-0.01	-0.12	-0.20	-0.16	-0.09	0.10	0.30	0.54	0.73	1.00	0.78	0.54	0.32	0.12	-0.09	-0.17	-0.18	-0.15
0.11	0.12	0.11	0.04	-0.06	-0.17	-0.18	-0.17	-0.10	0.11	0.31	0.54	0.78	1.00	0.80	0.53	0.32	0.12	-0.09	-0.17	-0.18
0.10	0.20	0.11	0.07	-0.01	-0.11	-0.15	-0.18	-0.17	-0.10	0.11	0.32	0.54	0.80	1.00	0.80	0.53	0.32	0.12	-0.09	-0.17
0.19	0.20	0.09	0.07	0.03	-0.04	-0.10	-0.15	-0.18	-0.17	-0.10	0.12	0.32	0.53	0.80	1.00	0.79	0.53	0.32	0.12	-0.09
0.14	0.15	0.07	0.07	0.05	0.01	-0.04	-0.09	-0.15	-0.18	-0.17	-0.09	0.12	0.32	0.53	0.79	1.00	0.79	0.53	0.32	0.12
0.08	0.03	0.04	0.05	0.06	0.05	0.02	-0.03	-0.09	-0.15	-0.18	-0.17	-0.09	0.12	0.32	0.53	0.79	1.00	0.79	0.54	0.32
0.00	0.01	0.02	0.05	0.06	0.06	0.05	0.02	-0.02	-0.09	-0.15	-0.18	-0.17	-0.09	0.12	0.32	0.53	0.79	1.00	0.79	0.54
-0.01	-0.01	0.01	0.02	0.05	0.06	0.06	0.05	0.03	-0.02	-0.09	-0.15	-0.18	-0.17	-0.09	0.12	0.32	0.54	0.79	1.00	0.79
-0.02	-0.02	-0.01	0.01	0.03	0.14	0.08	0.08	0.07	0.03	-0.03	-0.10	-0.15	-0.18	-0.17	-0.09	0.12	0.32	0.54	0.79	1.00

Notice that there are strong correlations between the sample fluctuations at neighboring points. This explains why we can get sample covariances with considerable fluctuations. Compare Figure 3.1. As an additional illustration of this we show in Figure 3.3 A-E some covariance functions calculated from samples of 50 values.

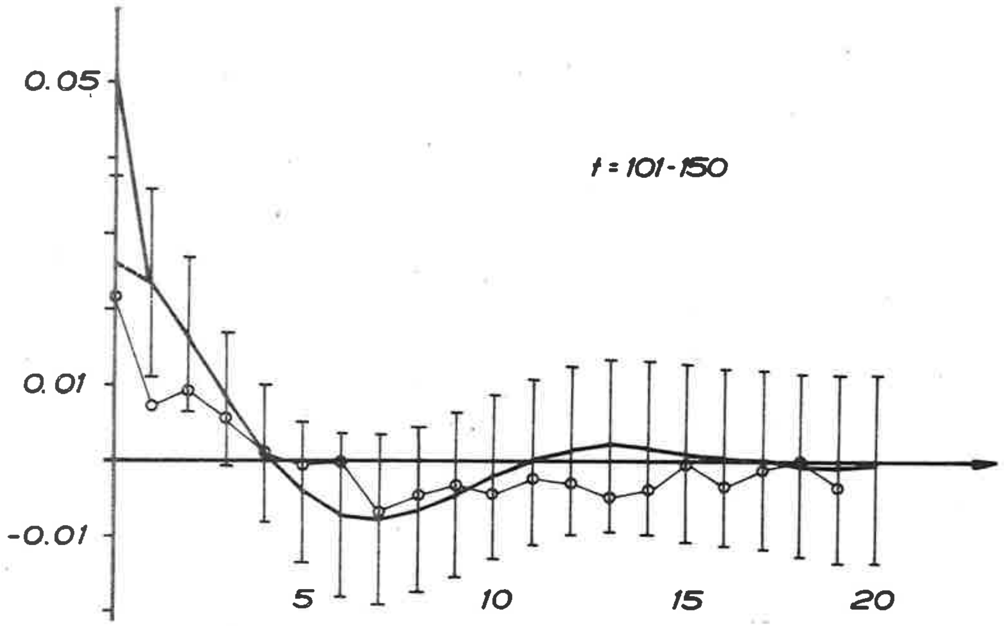

Figure 3.3 A

Sample covariance function calculated from 50 values of the artificial time series. The solid line shows the true covariance function.

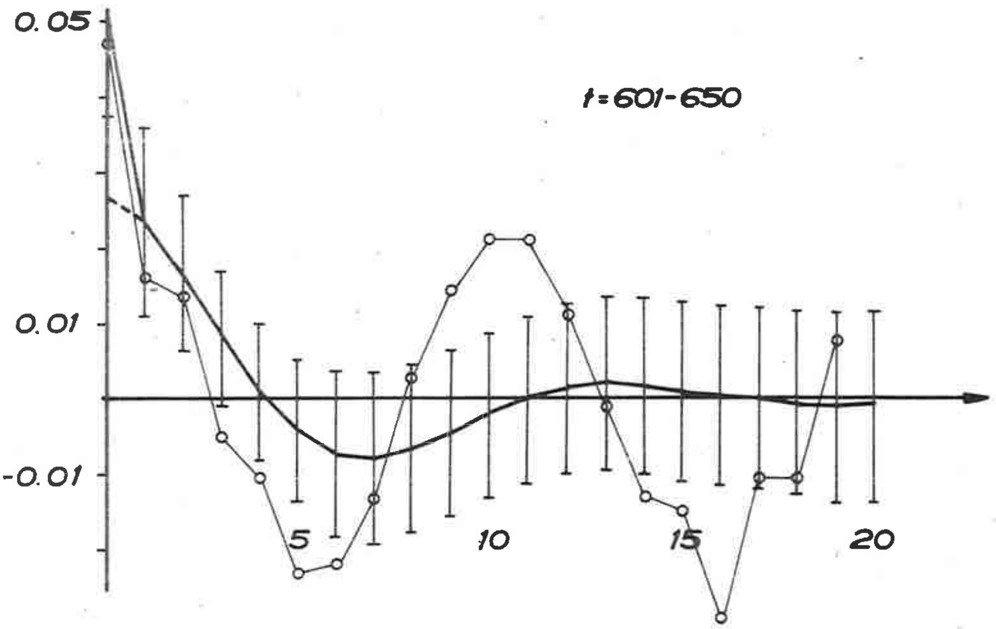

Figure 3.3 B

Sample covariance function calculated from 50 values of the artificial time series. The solid line shows the true covariance function.

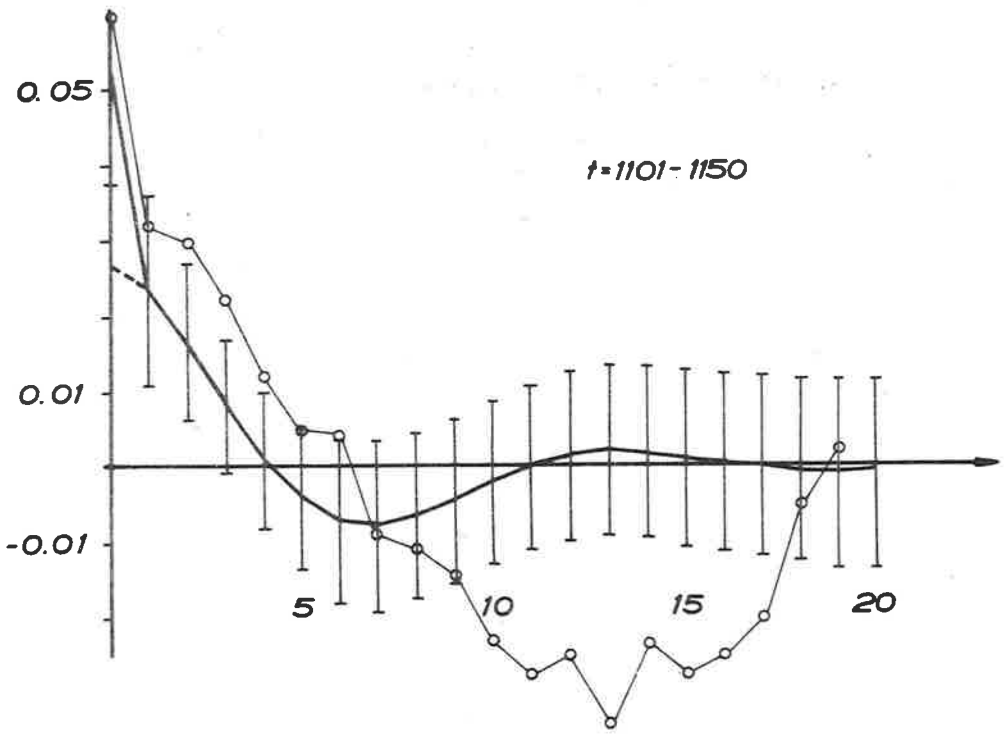

Figure 3.3 C

Sample covariance function calculated from 50 values of the artificial time series. The solid line shows the true covariance function.

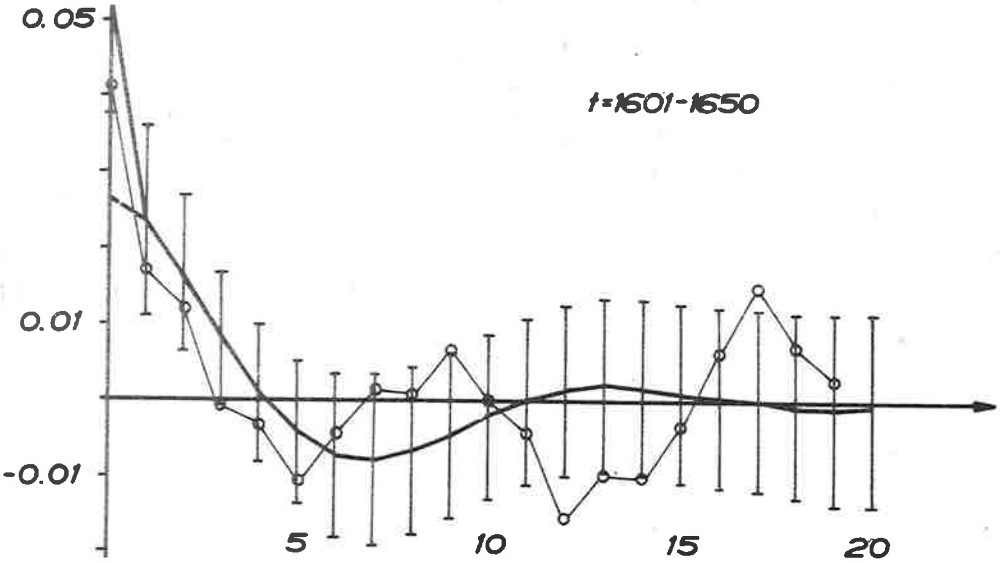

Figure 3. 3 D

Sample covariance function calculated from 50 values of the artificial time series. The solid line shows the true covariance function.

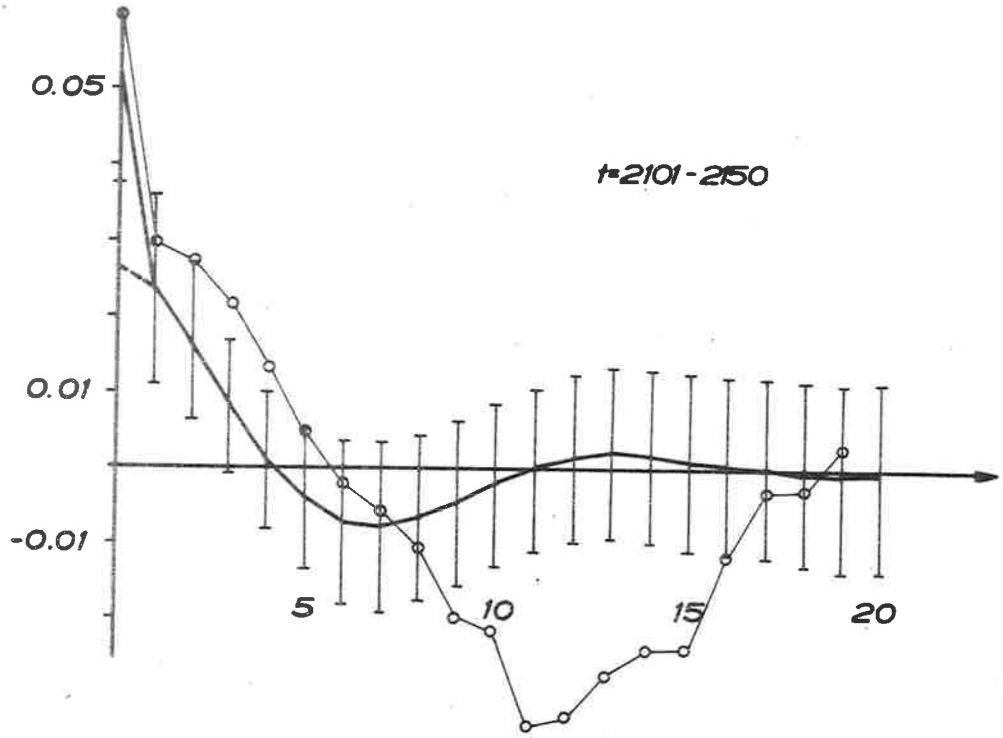

Figure 3. 3 E

Sample covariance function calculated from 50 values of the artificial time series. The solid line shows the true covariance function.

Let us finally point out that in order to evaluate the fluctuations we must know the statistical properties of the process. In this particular case, they are known, as we generated the series ourselves. This is, however, not the situation in most applications. One usually substitutes the estimated covariance function for the true covariance function when estimating the fluctuations. The error obtained when doing this is difficult to estimate.

IDENTIFICATION OF THE GENERATED TIME SERIES

We will now consider the inverse problem i. e., to construct a model which might have generated the observed time series. This problem was treated in [7] where a FORTRAN identification program also was given. Notice in particular that the inverse problem is not unique. In [7] we eliminated the non-uniqueness by choosing a standard for the model.

$$x(t+1) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ 0 & 0 & 0 & & 1 \\ a_k & a_{k-1} & a_{k-2} & & a_1 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ v(t) \end{bmatrix} \qquad (4.1)$$

$$y(t) = x_1(t) + w(t) \qquad (4.2)$$

where $\{v(t), t = 1, 2, \ldots\}$ and $\{w(t), t = 1, 2, \ldots\}$ are sequences of independent normal random variables with zero means and the variances

$$\text{var } v(t) = \sigma_1^2$$

$$\text{var } w(t) = \sigma_2^2$$

In order to facilitate the comparison, the model of section 2, for the generation of the time series, was chosen to be of this standard form. The input to the identification program is a sample covariance function. The identification program will then fit a covariance function generated by a model of type (4.1) - (4.2) to the sample covariance function. The program operates sequentially, starting with a first order model then proceeding to a second order model etc. At each step the identification program prints out the coefficients of the model $(a_1, a_2, \ldots a_k, \sigma_1^2, \sigma_2^2)$. The error norm i. e. the mean square error between the sample covariance function and the covariance function of the time series of the model is also computed at each step.

The error norm is to be compared with the uncertainty of the sample covariance functions. The uncertainty is introduced by the finite length of the sample as discussed in section 3.

The identification procedure was applied to the sample covariance function of the artificial time series. See Table II. The output from the identification program is in appendix 6. The results are summarized in Table V.

Table V

covariance function	a_1	a_2	Error norm			σ_1^2	σ_2^2
			1:st order system	2:nd order system	3:d order system		
1	1.129	-0.497	0.0037	0.0027	0.0007	0.00828	0.0192
2	1.442	-0.705	0.0104	0.0034	0.0015	0.00487	0.0187
3	1.415	-0.628	0.0087	0.0033	0.0037	0.00536	0.0184
4	1.460	-0.627	0.0076	0.0033	0.0012	0.00424	0.0261
5	1.606	-0.782	0.0112	0.0013	0.0013	0.00240	0.0248
6	0.820	-0.188	0.0029	0.0024	0.0018	0.00172	0.0138
7	1.263	-0.540	0.0062	0.0031	0.0031	0.00780	0.0208
8	1.270	-0.479	0.0051	0.0034	0.0019	0.00512	0.0226
9	1.270	-0.545	0.0063	0.0031	0.0014	0.00662	0.0193
10	1.390	-0.636	0.0062	0.0014	0.0097	0.00471	0.0227
aver	1.518	-0.703	0.0066	0.0007	0.0006	0.00287	0.0245

To evaluate the numerical values given in Table V, we recall that for a sample covariance function based on 500 values the sample fluctuations have a standard deviation of 0.0033. This means that the errors for first order approximations are too large, except for the first covariance function. For second order approximations the approximation errors are apparently of the same order of magnitude as the sampling fluctuations, which implies that a second order model is a reasonable choice. For the average covariance function the sample fluctuations have a standard deviation of 0.001, which implies that a second order model is reasonable also in this case.

The true values of the coefficients are

$a_1 = 1.500$

$a_2 = 0.700$

$\sigma_1^2 = 0.00300$

$\sigma_2^2 = 0.0250$

The determination of the coefficients based on the average covariance function will thus give the coefficients a_1 and a_2 with an accuracy of 1 %.

AN OPTIMAL ESTIMATOR BASED ON THE RESULTS OF THE IDENTIFICATION

We now consider an estimation problem relative to the generated time series. We regard $x(t)$ as the signal which should be estimated and we regard $y(t)$ as an observation of $x(t)$. Compare Figure 2.1. The problem is then: given a sequence of observed values $\{y(s), s = 1, \ldots, \}$ find the best minimum mean square estimate of $x(t)$. As the statistics of the time series is known, we can immediately construct a Kalman estimator and thereby obtain the desired results. In order to simulate the situation when the statistics are not known, we shall forget that we know the statistical properties and proceed in the following way:

observe a sample of the time series

construct a mathematical model of the process on the basis of the observed sample

construct an optimal Wiener filter for the model obtained in step 2

use the filter to estimate the variables $\{x(t), t = 1, 2, \ldots\}$ from the observed outputs $\{y(s), s = 1, \ldots, t\}$

This was in fact the procedure used in [8]. Four cases will be considered.

I optimal estimator based on true model

II estimator based on second order model estimated from the average of the ten sample covariance functions

III estimator based on second order model estimated from the sample covariance function nr 6 i. e. $\{y(t), t = 2501, 3000\}$

IV estimator based on first order model estimated from the average of the sample covariance functions.

Case I is used as a reference. The model based on sample number 6 was chosen because it showed the largest deviations from the true model. The estimator IV based on first order model is chosen in order to demonstrate the sensitivity of the estimates to variations in the order of the model.

The model for case I is given by equations (2.1) and (2.2). The models required for case II, III and IV are found in section 4. We have summarized the values of the model parameters for the different cases in Table VI.

Table VI

Case	a_1	a_2	σ_1^2	σ_2^2
I	1.5	-0.7	0.00300	0.0250
II	1.517	-0.703	0.00287	0.0245
III	0.820	-0.188	0.00172	0.0138
IV	0.658	—	0.0184	0.0196

Programs for the computation of optimal filters from a model of a time series were described in [8]. Listings of the programs used are found in [9].

The filtergains for the different cases are also shown in Figure 5.1. As seen from this figure there are considerable differences between the filtergains in the various cases.

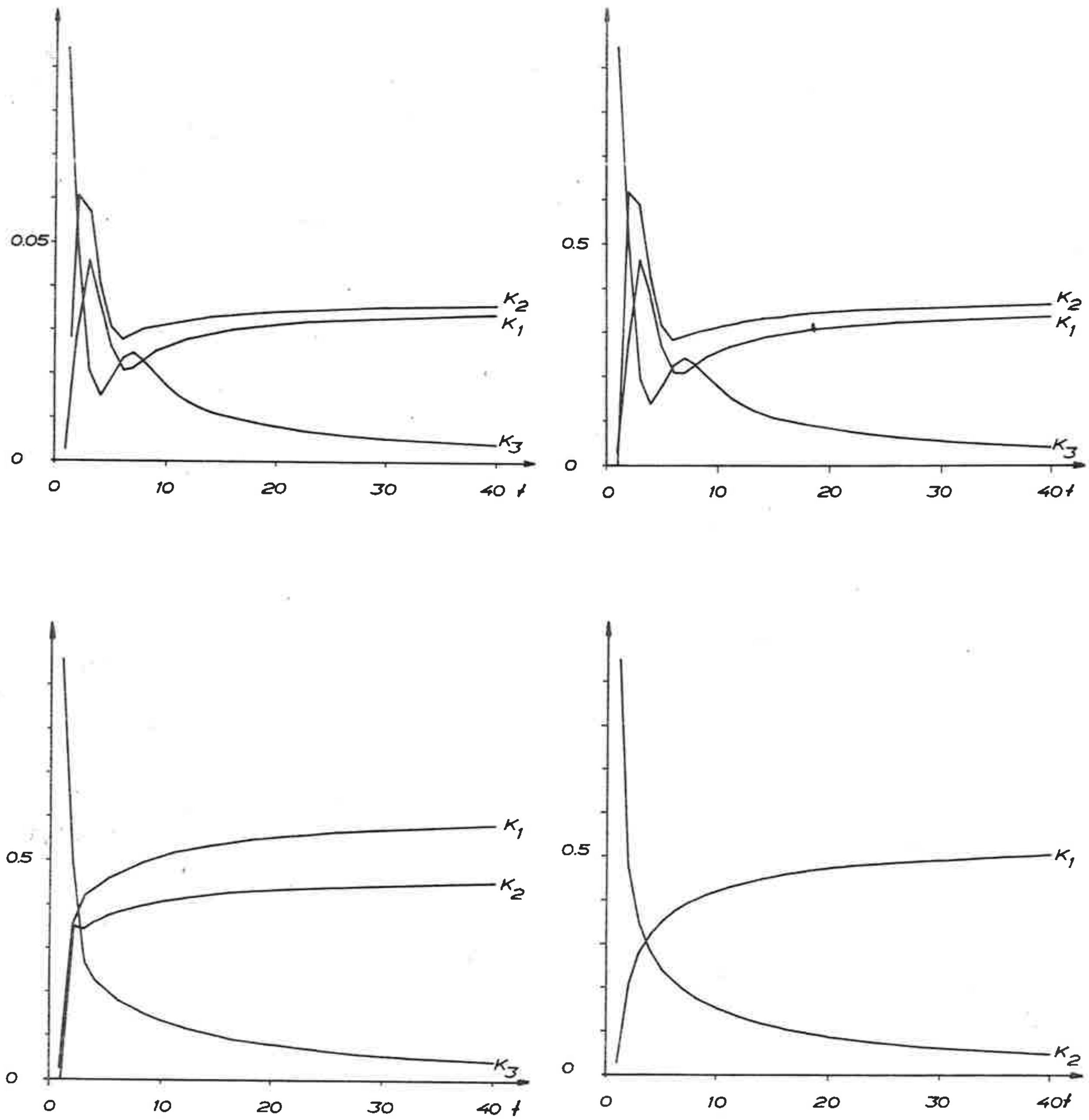

Figure 5.1

Filtergains for estimators I - IV

SIMULATION OF THE OPTIMAL ESTIMATOR

To test the four estimators described in the previous section, we used them to calculate estimates of x(t) on the basis of the "measured" y(t) from two samples of the time series $\{y(t),\ t = 1701,\ 1740\}$ and $\{y(t),\ t = 1731,\ 1770\}$. The results are summarized in Tables VII, VIII and in Figure 6.1.

As seen from the results, estimators I and II give practically the same estimates. The maximum difference is less than 0.02. In spite of the large differences in the filtergains the other estimators also produce good estimates. We can thus conclude that in this case the estimates are not too sensitive to variations in the filtergains.

Table VII

t	y	x_1	Estimate of x_1 I	II	III	IV
1701	4.04	4.02	4.04	4.04	4.04	4.04
1702	4.22	4.14	4.18	4.18	4.19	4.16
1703	4.17	4.25	4.19	4.19	4.17	4.16
1704	4.23	4.35	4.22	4.22	4.20	4.20
1705	4.46	4.44	4.32	4.32	4.36	4.35
1706	4.75	4.50	4.50	4.50	4.60	4.56
1707	4.24	4.40	4.38	4.38	4.33	4.34
1708	4.06	4.14	4.19	4.20	4.13	4.17
1709	4.04	3.88	4.08	4.08	4.07	4.10
1710	3.77	3.68	3.91	3.91	3.89	3.93
1711	3.38	3.61	3.67	3.66	3.59	3.65
1712	3.44	3.65	3.57	3.56	3.54	3.59
1713	3.56	3.78	3.60	3.59	3.62	3.64
1714	3.68	3.83	3.70	3.68	3.72	3.71
1715	3.91	3.86	3.86	3.85	3.89	3.87
1716	4.03	3.91	4.00	3.99	4.01	3.98
1717	4.06	4.06	4.08	4.07	4.05	4.03
1718	4.28	4.09	4.18	4.18	4.19	4.17
1719	4.02	4.11	4.12	4.12	4.07	4.06
1720	4.15	3.95	4.11	4.11	4.11	4.11
1721	3.60	3.81	3.88	3.89	3.78	3.81
1722	3.45	3.67	3.68	3.68	3.58	3.64
1723	3.69	3.65	3.67	3.66	3.69	3.72
1724	3.66	3.65	3.69	3.68	3.72	3.72
1725	3.91	3.75	3.81	3.81	3.87	3.86
1726	3.80	3.85	3.86	3.85	3.85	3.84
1727	3.85	3.86	3.90	3.89	3.86	3.86
1728	4.24	3.90	4.06	4.05	4.12	4.08
1729	4.07	4.04	4.08	4.08	4.08	4.06
1730	4.11	4.04	4.09	4.09	4.08	4.07
1731	4.11	4.01	4.08	4.08	4.09	4.08
1732	4.08	4.07	4.05	4.06	4.07	4.07
1733	4.11	4.15	4.04	4.05	4.08	4.08
1734	4.03	4.14	4.01	4.02	4.04	4.04
1735	4.28	4.16	4.09	4.09	4.18	4.16
1736	4.13	4.17	4.09	4.09	4.13	4.12
1737	4.42	4.19	4.19	4.20	4.29	4.27
1738	4.27	4.23	4.20	4.20	4.25	4.23
1739	4.48	4.19	4.27	4.28	4.36	4.34
1740	3.66	4.03	4.00	4.01	3.88	3.92

Table VIII

t	y	x_1	Estimate of x_1			
			I	II	III	IV
1731	4.11	4.01	4.11	4.11	4.11	4.11
1732	4.08	4.07	4.09	4.09	4.09	4.09
1733	4.11	4.15	4.10	4.10	4.10	4.10
1734	4.03	4.14	4.06	4.06	4.06	4.06
1735	4.28	4.16	4.16	4.16	4.21	4.19
1736	4.13	4.17	4.16	4.16	4.16	4.15
1737	4.42	4.19	4.28	4.28	4.32	4.30
1738	4.27	4.23	4.28	4.28	4.27	4.26
1739	4.48	4.19	4.36	4.36	4.39	4.38
1740	3.66	4.03	4.05	4.05	3.90	3.94
1741	3.65	3.91	3.84	3.84	3.74	3.80
1742	3.49	3.85	3.67	3.67	3.62	3.66
1743	3.92	3.87	3.78	3.77	3.86	3.86
1744	3.83	3.89	3.85	3.84	3.88	3.87
1745	3.70	3.97	3.85	3.84	3.79	3.80
1746	4.17	4.06	4.02	4.01	4.06	4.04
1747	4.11	4.14	4.10	4.10	4.11	4.08
1748	4.06	4.10	4.11	4.11	4.07	4.06
1749	4.13	4.13	4.12	4.12	4.10	4.09
1750	4.16	4.15	4.13	4.13	4.13	4.12
1751	4.11	4.16	4.11	4.11	4.11	4.11
1752	3.99	4.16	4.04	4.05	4.02	4.03
1753	4.16	4.19	4.07	4.07	4.11	4.11
1754	4.28	4.15	4.14	4.14	4.21	4.19
1755	4.13	4.17	4.13	4.13	4.15	4.14
1756	3.94	4.12	4.05	4.05	4.00	4.02
1757	3.91	4.04	3.97	3.98	3.94	3.96
1758	4.00	3.97	3.98	3.97	3.99	4.00
1759	4.01	3.81	3.99	3.99	4.02	4.01
1760	4.01	3.78	4.01	4.01	4.02	4.02
1761	3.71	3.77	3.91	3.90	3.83	3.85
1762	3.90	3.86	3.91	3.91	3.90	3.91
1763	3.89	3.88	3.92	3.92	3.92	3.92
1764	3.75	3.84	3.88	3.87	3.83	3.84
1765	3.67	3.75	3.82	3.82	3.75	3.78
1766	3.75	3.72	3.82	3.81	3.78	3.79
1767	3.68	3.73	3.80	3.79	3.75	3.77
1768	3.33	3.71	3.66	3.65	3.52	3.56
1769	3.50	3.72	3.63	3.62	3.56	3.59
1770	3.71	3.77	3.72	3.71	3.72	3.71

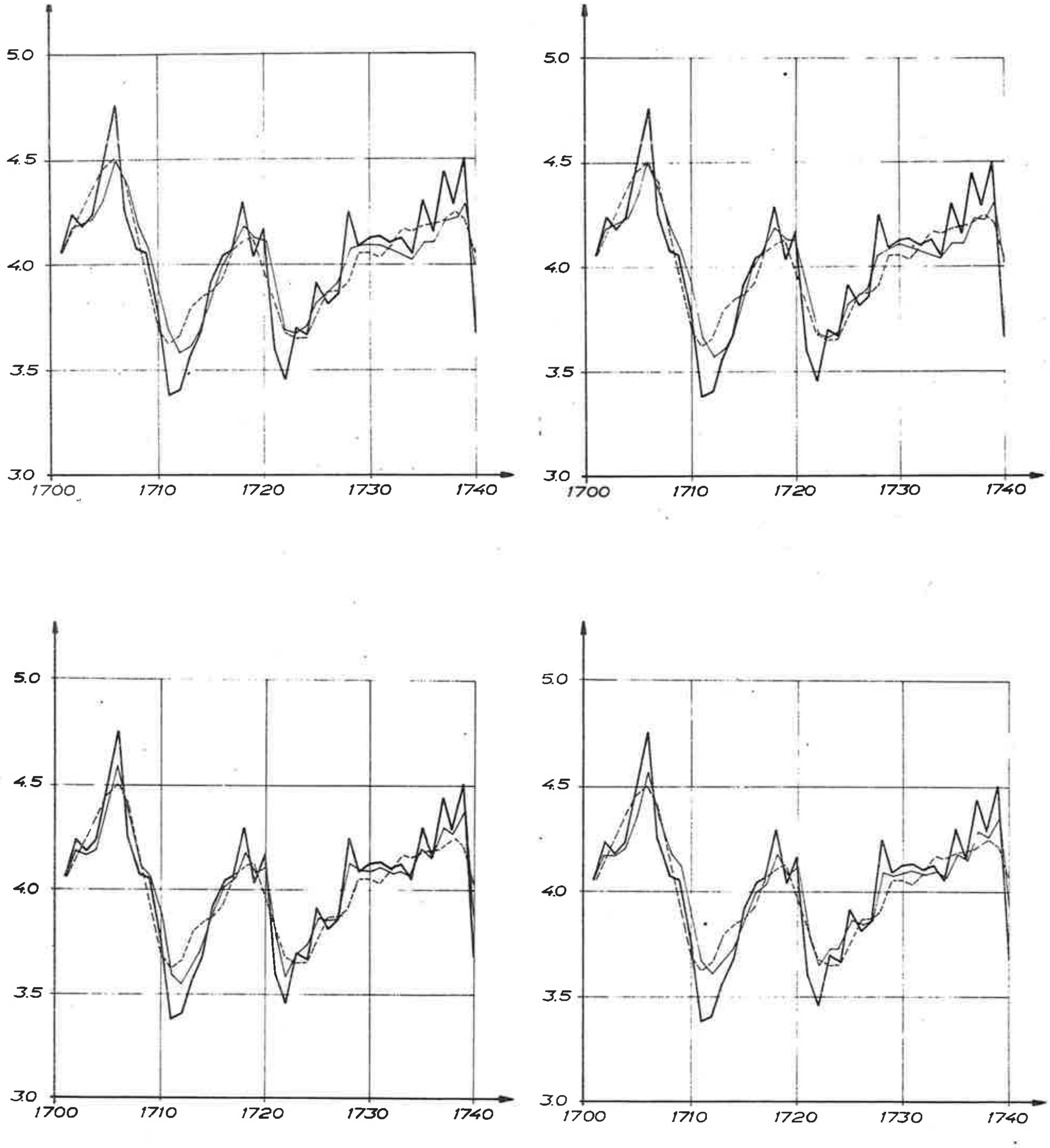

Figure 5.1

Results of simulations with estimators I - IV. The dotted line shows true values of $x_1(t)$, the thick solid line shows the measured values $y(t)$ and the thin solid line shows the estimate.

For the application described in [8] it is of interest to investigate the influence of the initial estimate which is normally set to $\hat{x}_3(0) = 4$. In Figure 6.2 we show the estimate obtained from the estimator based on the true model when the initial estimate is 2, 3 and 4 respectively.

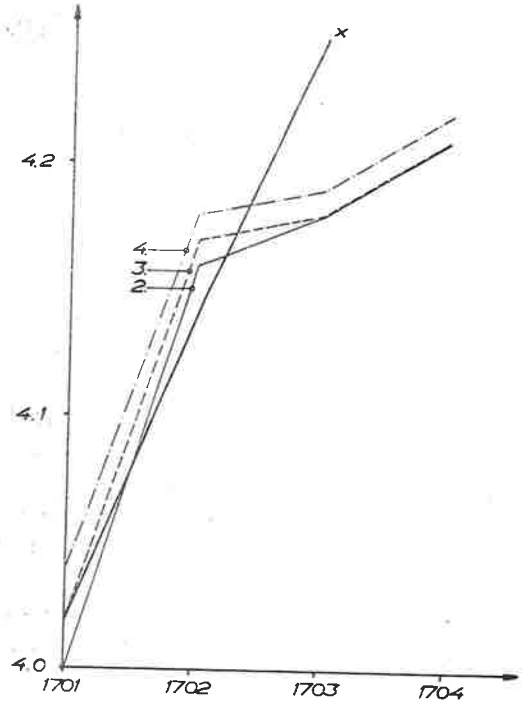

Figure 6.2
Effect of initial estimate

The corresponding numerical values are listed in Table VIIII

Table VIIII

t	y	x	Estimate of x Estimator based on true model		
			x(0) = 2	x(0) = 3	x(0) = 4
1701	4.04	4.02	4.00	4.02	4.04
1702	4.22	4.14	4.16	4.17	4.18
1703	4.17	4.25	4.18	4.18	4.19
1704	4.23	4.35	4.21	4.21	4.22
1705	4.46	4.44	4.31	4.31	4.32
1706	4.75	4.50	4.49	4.49	4.50
1707	4.24	4.40	4.37	4.37	4.38
1708	4.06	4.14	4.18	4.19	4.19
1709	4.04	3.88	4.07	4.07	4.08
1710	3.77	3.68	3.91	3.91	3.91

REFERENCES

[1] Bartlett, M.S. "Stochastic Processes", Cambridge University Press 1961.

[2] Blackman, R.B. Tukey, J.W. "The Measurement of Power Spectra" Dover, New York, 1950

[3] Grenander, U. Rosenblatt, M. "Statistical Analysis of Stationary Time Series", Almkvist & Wiksell, Stockholm J. Wiley, New York, 1956

[4] Hannan, E.J. "Time Series Analysis", Methuen, London, 1960

[5] Kalman, R.E. "Mathematical Description of Linear Dynamical Systems", SIAM Journal on Control 1 (1963) 152 - 192

[6] Kalman, R.E. "New Methods and Results in Linear Prediction and Filtering Theory", Technical Report RIAS 61-1 also in Bogdanoff, J.L. and Kozin, F ed. "Proceedings of the First Symposium on Engineering Applications of Random Function Theory and Probability", Wiley. New York, 1963

[7] Åström, K.J., Wensmark, S. "Numerical Identification of Stationary Time Series", Report IBM Nordic Laboratory, TP 18.134, May 15, 1964

[8] Åström, K.J., Wensmark, S., Tveit, O. "Quality Control at Billerud", Report IBM Nordic Laboratory TP IBM Confidential

[9] Åström, K.J., Wensmark, S. "Programs and Results of Experiments with Artificial Time Series", Internal Report 65 01 01 IBM Nordic Laboratory.

EXPERIMENT WITH AN ARTIFICIAL TIME SERIES

K. J. Åström
S. Wensmark

TP 18.146
April 20, 1965

IBM