Identification in the Presence of Drift

Banyasz, Csilla

1973

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

Link to publication

Citation for published version (APA):
Banyasz, C. (1973). *Identification in the Presence of Drift*. (Research Reports TFRT-3065). Department of Automatic Control, Lund Institute of Technology (LTH).

Total number of authors:
1

LUND UNIVERSITY

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

IDENTIFICATION IN THE PRESENCE OF DRIFT

CSILLA BANYASZ

Report 7343 December 1973
Lund Institute of Technology
Division of Automatic Control

IDENTIFICATION IN THE PRESENCE OF DRIFT

Cs.Banyasz

ABSTRACT

In this report the influence of drift of different types and magnitude on the ML and LS identification is examined.The drift effects are simulated, after identifying the results are analysed. It can be seen from the results that the presence of drift depending on the magnitude and type of it can have a significant influence on the parameter estimates. This means that it is necessary to analyse the measurements before the identification,in many cases more exact demands have to be created on the measuring conditions or prefiltering strategies must be applied.

TABLE OF CONTENTS

1. INTRODUCTION

During the identification from industrial data the problem appears usually whether the measurements correspond to the real values of signals or not, fulfil the requirements of the identification methods, otherwise what is the influence of inadequate data for the estimation.

Starting from these ideas the problem was approached in the following way: the input signals were PRBS and the output signals of the process were computed by simulation (supposing correlated noise at the output). Then different types of modifications were performed on the output data and after the identification of the input- and the modified values of output it was possible to draw some conclusions about the influence of drift of different character and magnitude on the identification.

We use the term drift in a very extended meaning, i.e. the following cases were considered:

1./ The values of output were changed in one or several points

2./ A constant level was added to the values of output

3./ A level changing linearly was added to the values of output

4./ A sinusoidal signal was added to the output.

The cases 1./ and 2./ are used as typical situations for industrial measurements and our aim is to examine the influence of the erroneous measurements on the estimates i.e. we do not filter the data before the identification.

The third case is generally considered as a typical drift effect. The fourth situation can appear in practice in the case of superposed signals.

In this paper we restrict ourselves to the examination of influence of drift in the output since we have used the off-line maximum likelihood method (ML) [1] for the identification and the input signal PRBS was given to the process by us. In the first step this method produces the least-squares (LS) estimates of process parameters so we can compare them with the ML ones for the different types of the drift effects. (I used the ML identification program in the program library of UNIVAC 1108 made by I.Gustavsson [2].)

The simulation model was the following:

$$y_m(t) = \frac{B(z^{-1})}{A(z^{-1})}\, u(t) + \lambda \frac{C(z^{-1})}{A(z^{-1})}\, e(t) + k\sigma_y \gamma(t)$$

See Fig.1.

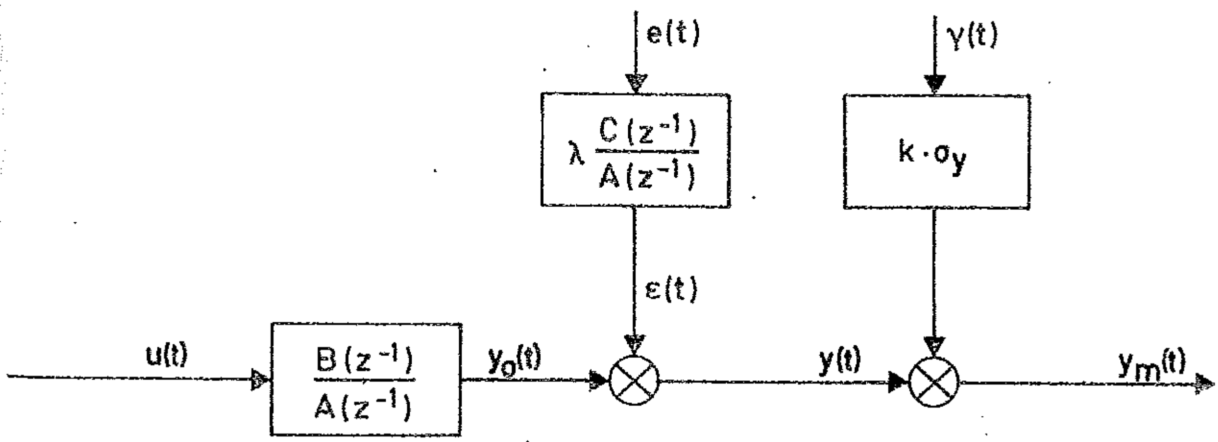

Fig.1. Structure for the simulation model

First we computed the values $y(t)$ and after this modified them.

The notations on the Fig.1. are:

$$A(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} = 1 - 1.5z^{-1} + 0.7z^{-2}$$
$$B(z^{-1}) = b_1 z^{-1} + b_2 z^{-2} = 1.0z^{-1} + 0.5z^{-2} \qquad (1)$$
$$C(z^{-1}) = 1 + c_1 z^{-1} + c_2 z^{-2} = 1 - 1.0z^{-1} + 0.2z^{-2}$$

$e(t)$ – white noise, $N(0, \lambda)$; $\lambda = 0.4$

σ_y – the standard deviation of $y(t)$

$\gamma(t)$ – means the time function of drift of different types
(in the 1st case $\gamma(t)$ is a sequence of impulse functions
in the 2nd case $\gamma(t)$ is a step function
in the 3rd case $\gamma(t)$ is a linear function
in the 4th case $\gamma(t)$ is a sinusoidal function).

k – is a relative number ; the value of $k\sigma_y$ characterizes the amplitude of $\gamma(t)$.(In the 3rd case $k\sigma_y$ means the slope of the linear function).

The number of samples was $N = 500$.

The results obtained by identification for the case without drift can be seen in Table I. and the time functions of the input,output, model output, model error, residual are shown on Fig.2.

Parameters	True values	LS	ML
a_1	-1.5	-1.449	-1.502
a_2	0.7	0.653	0.702
b_1	1.0	1.011	1.024
b_2	0.5	0.523	0.473
c_1	-1.0	0	-1.014
c_2	0.2	0	0.268
λ	0.4	0.552	0.391
$V(\hat{\theta})$	-	76.144	38.146
$\bar{y}$	0.6	-	-
σ_y	4.5	-	-
y_{max}	10.7	-	-
y_{min}	-12.6	-	-

Table I. Estimations from the data without drift.
N=500

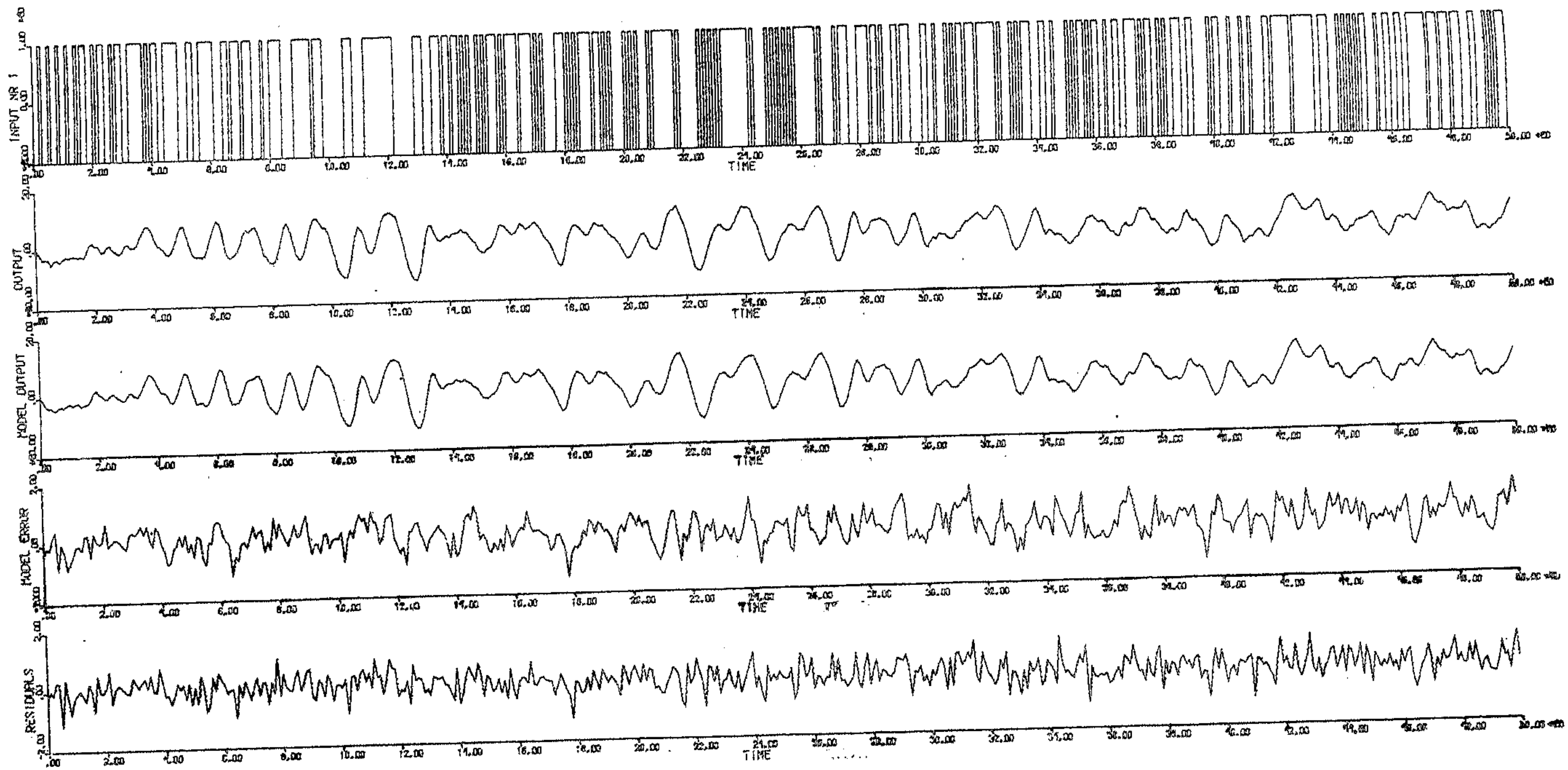

Fig.2. Results of the identification with the maximum likelihood method of data without drift

2. MEASURING ERRORS IN THE OUTPUT SIGNAL IN ONE OR SEVERAL POINTS

During the simulation of this situation we changed one or more output values in the following way:

$$y_m(t) = y(t) + k\sigma_y \gamma(t) \qquad (2)$$

where now $\gamma(t)$ is an impulse function series, i.e.

$$\gamma(t) = \begin{cases} 1 & \text{if} \quad t = 25, 75, 125, .., 475. \quad (\text{or } t=250\ ) \\ 0 & \text{otherwise} \end{cases}$$

See Fig.3.

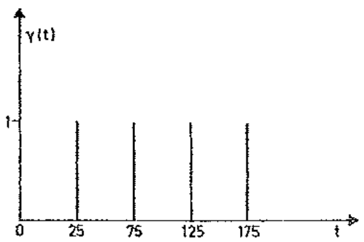

Fig.3. Measurement errors $\gamma(t)$

The identification was performed from the input and the modified output values. The ML estimates obtained in this way are shown in the Table II. together with the least-squares estimates. It can be seen from Table II. that the estimates of a_i, b_i are reasonably good but the estimates of c_i are increasingly worse if the k is increasing and the values of $\hat{c}_i$ tend to the values of the corresponding $\hat{a}_i$ when the number of the changed points is 1 and 10, as well. The reason for this can be seen easily because in this case the noise by the side of the signal $k\sigma_y \gamma(t)$ appearing in discrete time points (or only in one point) at the output is negligible and this situation can be identified only by such an identification model in which $C(z^{-1}) = A(z^{-1})$. (See Fig.4.)

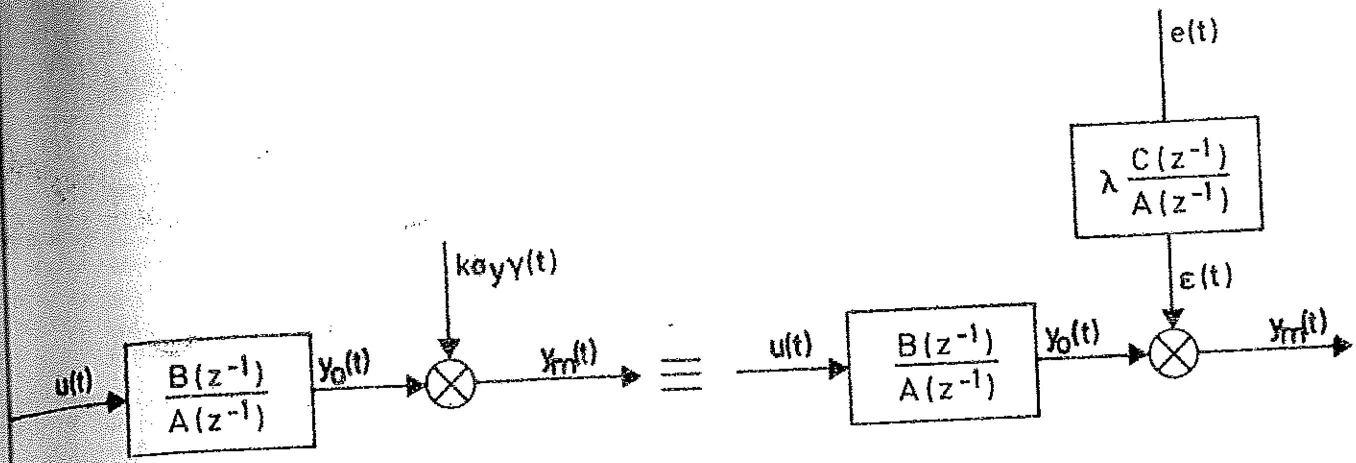

Fig.4. Resulting model structure, $C(z^{-1})=A(z^{-1})$, when measurement errors on data

It can also be seen from Table II. that these statements are not valid for the least-squares estimation because resulting from the behaviour of LS method it tries to smooth over the high jumpings in the output so the LS estimates of a_i and b_i are also spoiled. These establishments can also be observed on the time functions on Figs.5. and 6. both for the ML and LS estimations.

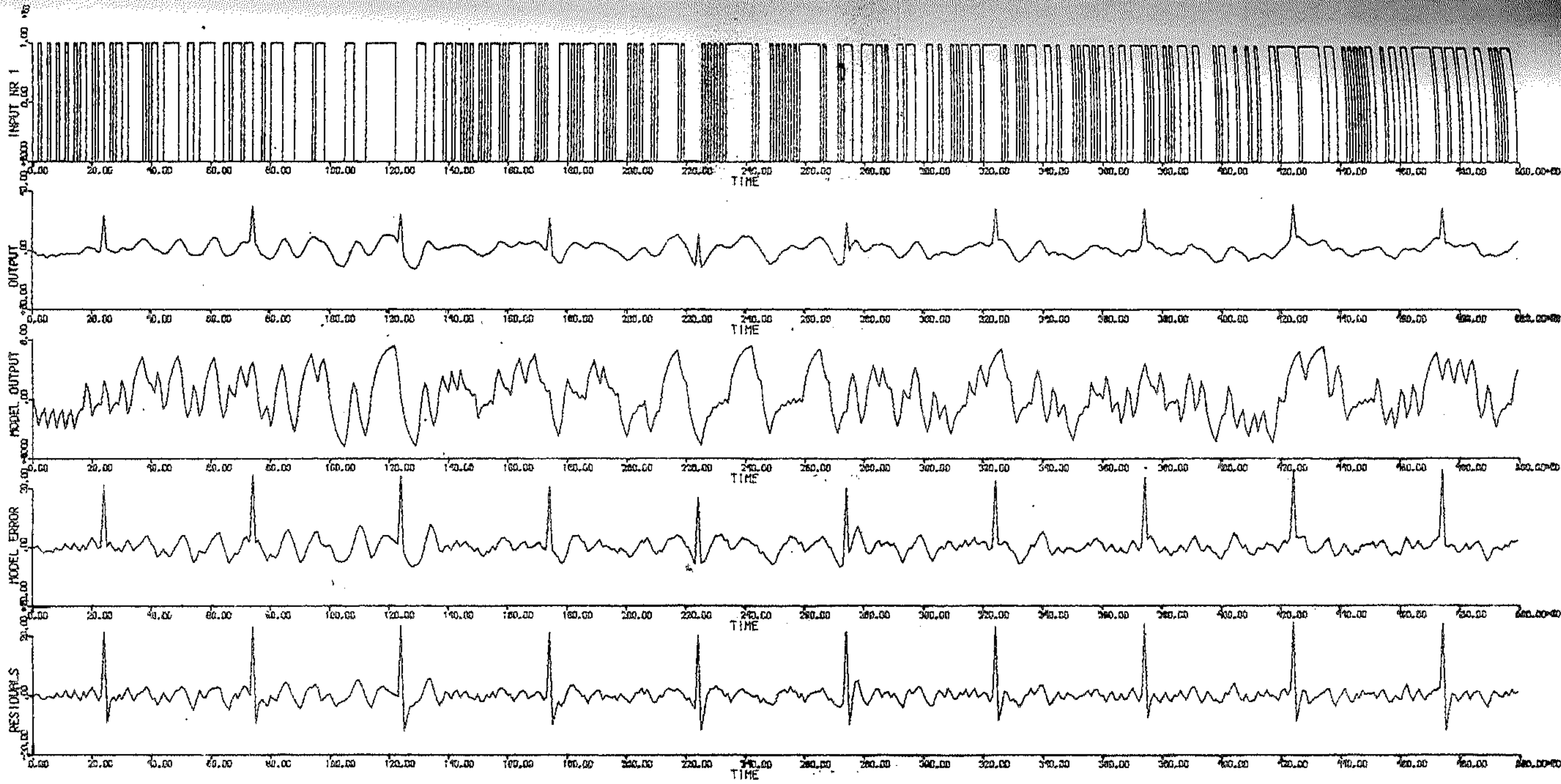

Fig.5. Results of identification with the least squares method on data with measurement errors

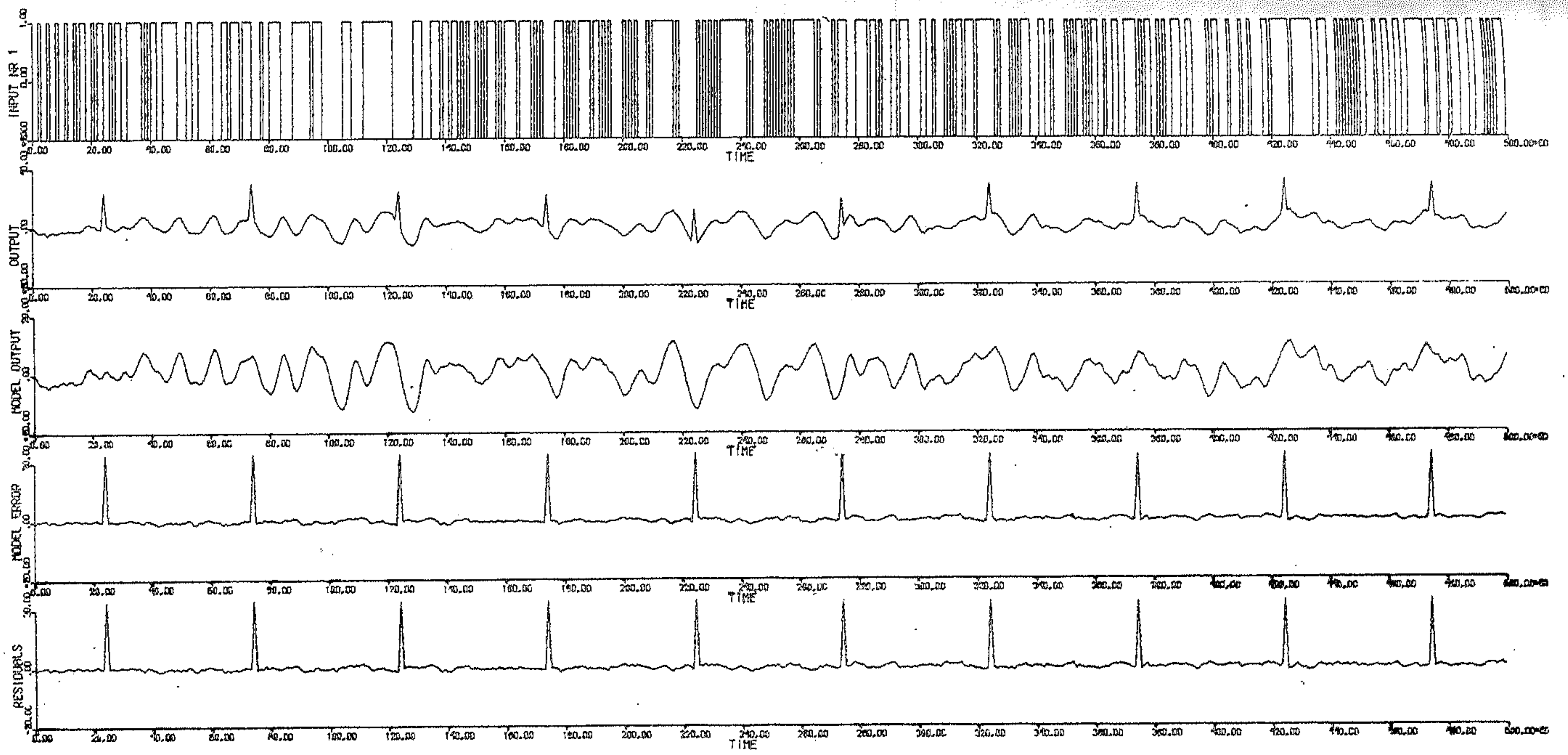

Fig.6. Results of identification with the maximum likelihood method on data with measurement errors

Parameters	True values	estimation from data without drift		number of the changed points: 1				number of the changed points: 10			
				k=3		k=5		k=3		k=5	
		LS	ML	LS	ML	LS	ML	LS	ML	LS	ML
a_1	-1.5	-1.449	-1.502	-1.232	-1.499	-0.990	-1.497	-0.645	-1.491	-0.429	-1.481
a_2	0.7	0.653	0.702	0.450	0.699	0.230	0.697	0.058	0.688	-0.180	0.675
b_1	1.0	1.011	1.024	1.104	1.048	1.155	1.054	1.230	1.106	1.323	1.151
b_2	0.5	0.523	0.473	0.674	0.437	0.899	0.422	1.353	0.442	1.636	0.440
c_1	-1.0	0.	-1.014	0.	-1.332	0.	-1.424	0.	-1.481	0.	-1.480
c_2	0.2	0.	0.268	0.	0.531	0.	0.615	0.	0.677	0.	0.672
λ	0.4	0.552	0.391	1.198	0.749	1.718	1.110	2.736	1.985	4.002	3.243
$V(\hat{\theta})$	-	76.144	38.146	359.07	140.14	737.67	308.23	1871.3	985.39	4003.7	2628.9
$\bar{y}$	0.63	0.63		0.65		0.67		0.90		1.08	
σ_y	4.54	4.54		4.53		4.57		4.99		5.65	
y_{max}	10.68	10.68		10.68		14.17		21.99		31.07	
y_{min}	-12.58	-12.58		-12.58		-12.58		-12.58		-12.58	

Table II Estimates in the case of measuring errors, N=500.

3. CONSTANT LEVEL ON THE OUTPUT

The simulated output is generated by the following equation:

$$y_m(t) = y(t) + k\sigma_y \gamma(t) \tag{3}$$

where now $\gamma(t)$ is a unit step signal. See Fig.7.

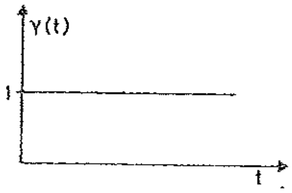

Fig.7. Disturbance $\gamma(t)$

The identified values of parameters can be seen in Table III. for different values of k. It is striking that the increasing value of k has the biggest influence for $\hat{a}_i$ and $\hat{c}_i$ .Analysing the results we can establish that the values of $\hat{a}_1, \hat{a}_2$ are formed in such way that the sum $1+\hat{a}_1+\hat{a}_2$ tends to 0, the values of $\hat{c}_i$ and the roots of the polynomial $\hat{C}(z^{-1})$ are in Figs.8/a and 8./b for different values of k where the roots seem to be moving to a certain point.

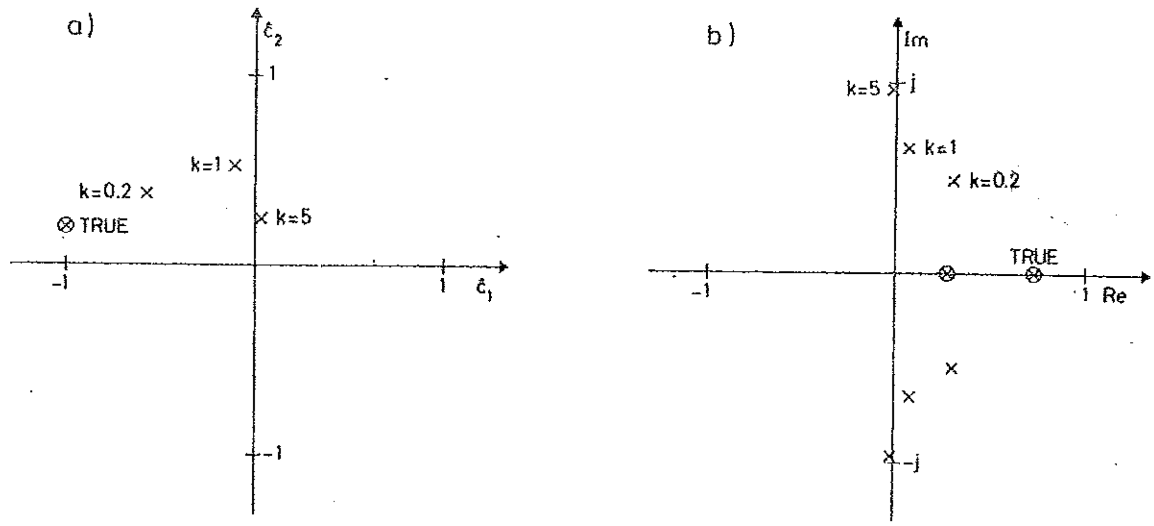

Fig.8. a) Estimates of $\hat{c}_1$ and $\hat{c}_2$ for different values of k

b) Estimates of the roots of $\hat{C}(z^{-1})$ for different values of k

Parameters	True values	estimations from data without drift		k=0.2		k=1.0		k=5.0	
		LS	ML	LS	ML	LS	ML	LS	ML
a_1	-1.5	-1.449	-1.502	-1.454	-1.494	-1.509	-1.486	-1.442	-1.388
a_2	0.7	0.653	0.702	0.645	0.686	0.609	0.596	0.454	0.403
b_1	1.0	1.011	1.024	1.024	1.021	1.034	1.028	1.029	1.039
b_2	0.5	0.523	0.473	0.529	0.499	0.473	0.557	0.511	0.639
c_1	-1.0	0	-1.014	0	-0.587	0	-0.125	0	0.014
c_2	0.2	0	0.268	0	0.371	0	0.453	0	0.241
λ	0.4	0.552	0.391	0.581	0.506	0.845	0.774	1.512	1.469
$V(\hat{\theta})$	-	76.144	38.146	84.487	63.978	167.76	149.85	571.69	539.52
$\bar{y}$	0.63	0.63		1.54		5.17		23.34	
σ_y	4.54	4.54		4.54		4.54		4.54	
y_{max}	10.68	10.68		11.59		15.22		33.39	
y_{min}	-12.58	-12.58		-11.67		-8.04		10.14	

Table III. Estimates in the case of constant drift.
N=500

In order to interpret this effect we have examined some different cases.

3.1.

First let us consider a simple case when the process model is:

$$y_m(t) = C(z^{-1})\, e(t) + k\sigma_y \gamma(t) \qquad (4)$$

which is identified by the model:

$$y_m(t) = \hat{C}(z^{-1})\,\varepsilon(t) \qquad (5)$$

where it is assumed that $e(t)$ is white noise.

Then

$$\varepsilon(t) = \frac{1}{\hat{C}(z^{-1})}\,y_m(t) = \frac{1}{\hat{C}(z^{-1})}\left(C(z^{-1})\,e(t) + k\sigma_y\,\gamma(t)\right) =$$
$$= \frac{C(z^{-1})}{\hat{C}(z^{-1})}\,e(t) + \frac{1}{\hat{C}(z^{-1})}\,k\sigma_y\,\gamma(t) \qquad (6)$$

Let us write down the loss function for a first order system:

$$F = E\left\{\varepsilon^2(t)\right\} = \frac{1}{2\pi j}\oint \frac{(1+cz^{-1})}{(1+\hat{c}z^{-1})}\,\frac{(1+cz)}{(1+\hat{c}z)}\,\frac{dz}{z} + \frac{1}{(1+\hat{c})^2}\,k^2\sigma_y^2 =$$
$$= \frac{1+c^2-2c\hat{c}}{1-\hat{c}^2} + k^2\sigma_y^2\,\frac{1}{(1+\hat{c})^2} \qquad (7)$$

The necessary condition for the extremum is:

$$\frac{\partial F}{\partial \hat{c}} = \frac{-2c(1-\hat{c}^2) - (1+c^2-2c\hat{c})(-2\hat{c})}{(1-\hat{c}^2)^2} - \frac{2k^2\sigma_y^2}{(1+\hat{c})^3} = 0 \qquad (8)$$

Exact analytical solution can not be given because of the complexity of this equation. Two special cases are examined, $k=0$ and $k\to\infty$.

In the case of $k=0$ we have

$$\frac{-2c + 2c\hat{c}^2 + 2\hat{c} + 2c^2\hat{c} - 4c\hat{c}^2}{(1-\hat{c}^2)^2} = 0 \qquad (9)$$

We get the equation

$$-c\hat{c}^2 + \hat{c}(1+c^2) - c = 0 \qquad (10)$$

if $\hat{c} \neq \pm 1$

Solutions of (10) are $\hat{c}=c$, $\hat{c} = \frac{1}{c}$ where the first one is admissible for the stable system in case without drift.

In the case of $k \to \infty$ the equality

$$-\frac{2\sigma_y^2}{(1+\hat{c})^3} = 0 \tag{11}$$

must be fulfilled and this is possible only in the case when $\hat{c} \to \infty$ but under the restrictions of stability $\hat{c}=1$. (During the identification the ML method allows only stable polynomial $\hat{C}(z^{-1})$.)

The solution $\hat{c}=1$ can also be obtained from (8) because if $\frac{\partial F}{\partial \hat{c}} = 0$ is fulfilled then we can write that:

$$\frac{-2c(1-\hat{c}^2)-(1+c^2-2c\hat{c})\ (-2\hat{c})}{(1-\hat{c}^2)^2} = \frac{2k^2\sigma_y^2}{(1+\hat{c})^3} \tag{12}$$

For $k \to \infty$ this can be true only when $\hat{c}=1$.

3.2.

Now let us consider a second order system for this simple model.

Then the loss function is:

$$F=E\left\{\varepsilon^2(t)\right\} = \frac{1}{2\pi j}\oint \frac{(1+c_1z^{-1}+c_2z^{-2})(1+c_1z+c_2z^2)}{(1+\hat{c}_1z^{-1}+\hat{c}_2z^{-2})(1+\hat{c}_1z+\hat{c}_2z^2)}\ \frac{dz}{z} + \frac{k^2\sigma_y^2}{(1+\hat{c}_1+\hat{c}_2)^2}$$

$$= \frac{(1+c_1^2+c_2^2)(1+\hat{c}_2)-2\hat{c}_1c_1(1+c_2)-2c_2(\hat{c}_1^2-\hat{c}_2-\hat{c}_2^2)}{(1-\hat{c}_2)\ ((1+\hat{c}_2)^2-\hat{c}_1^2)} + \frac{k^2\sigma_y^2}{(1+\hat{c}_1+\hat{c}_2)^2} \tag{13}$$

Equating the derivatives to zero we get the following equations:

$$\frac{\partial F}{\partial \hat{c}_1} =$$

$$= \frac{\left[-2c_1(1+c_2)+4c_2\hat{c}_1\right]\left[(1-\hat{c}_2)((1+\hat{c}_2)^2-\hat{c}_1^2)\right]}{\left\{(1-\hat{c}_2)\left[(1+\hat{c}_2)^2-\hat{c}_1^2\right]\right\}^2} -$$

$$- \frac{\left[(1+c_1^2+c_2^2)(1+\hat{c}_2)-2\hat{c}_1c_1(1+c_2)-2c_2(\hat{c}_1^2-\hat{c}_2-\hat{c}_2^2)\right]\left[-2\hat{c}_1(1-\hat{c}_2)\right]}{\left\{(1-\hat{c}_2)\left[(1+\hat{c}_2)^2-\hat{c}_1^2\right]\right\}^2} -$$

$$- \frac{2k^2\sigma_y^2}{(1+\hat{c}_1+\hat{c}_2)^3} = 0 \tag{14}$$

$$\frac{\partial F}{\partial \hat{c}_2} =$$

$$= \frac{\left[(1+c_1^2+c_2^2)-2c_2-4c_2\hat{c}_2\right]\left[(1-\hat{c}_2)((1+\hat{c}_2)^2-\hat{c}_1^2\right]}{\left\{(1-\hat{c}_2)\left[(1+\hat{c}_2)^2-\hat{c}_1^2\right]\right\}^2} -$$

$$- \frac{\left[(1+c_1^2+c_2^2)(1+\hat{c}_2)-2\hat{c}_1c_1(1+c_2)+2c_2(\hat{c}_1^2-\hat{c}_2-\hat{c}_2^2)\right]\left[-(1+\hat{c}_2)^2+2(1-\hat{c}_2)(1+\hat{c}_2)\right]}{\left\{(1-\hat{c}_2)\left[(1+\hat{c}_2)^2-\hat{c}_1^2\right]\right\}^2} -$$

$$- \frac{2k^2\sigma_y^2}{(1+\hat{c}_1+\hat{c}_2)^3} = 0 \qquad (15)$$

These are complicated functions of $\hat{c}_1$ and $\hat{c}_2$ and can not be handled analytically.

Examining the case $k \to \infty$ the equality

$$- \frac{2\sigma_y^2}{(1+\hat{c}_1+\hat{c}_2)^3} = 0 \qquad (16)$$

must be fulfilled. (this is valid both for (14) and (15)) which is true only if $\hat{c}_1$ and $\hat{c}_2 \to \infty$. But under the stability conditions the admissible domain for $\hat{c}_1, \hat{c}_2$ can be seen on Fig.9. (striped domain) so only the values $\hat{c}_1=2$, $\hat{c}_2=1$ are attainable as a maximum.

On the other hand this solution makes it possible to fulfil the conditions $\frac{\partial F}{\partial \hat{c}_1}=0$ and $\frac{\partial F}{\partial \hat{c}_2}=0$ in the case of $k \to \infty$, i.e. the first term of (14) and (15) is made infinite.

This solution concerning the roots of the polynomial $\hat{C}(z^{-1})$ means a double root on the z-plane in the point -1.

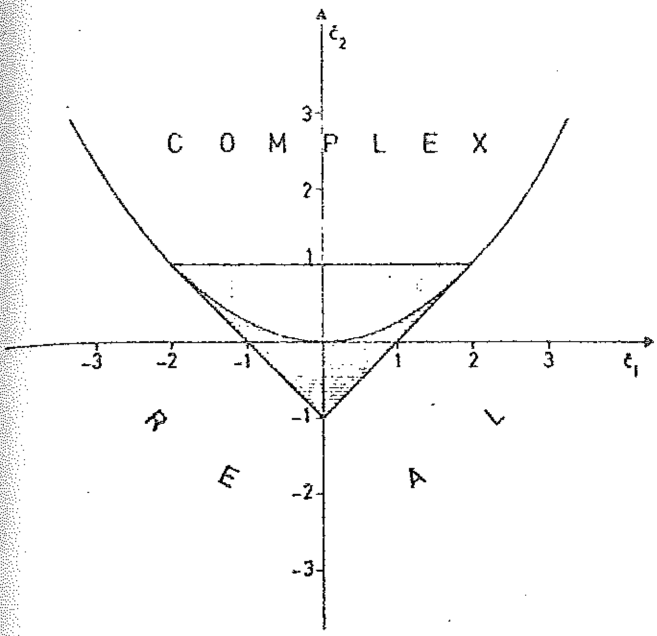

Fig.9. Admissible region for $\hat{c}_1$ and $\hat{c}_2$ (striped domain)

The parameter values obtained by the simulation and identification of model (4) can be seen in Table III./a for second order system. The changing of $\hat{c}_1$ and $\hat{c}_2$ and the roots of $\hat{C}(z^{-1})$ are also represented for different k on Fig.10. and these figures are similar to Fig.8./a and Fig.8./b.

Para-meters	True values	ML	
		k=1	k=5
a_1	0	-	-
a_2	0	-	-
b_1	0	-	-
b_2	0	-	-
c_1	-1.0	-0.218	0.657
c_2	0.2	0.513	0.622
λ	0.4	0.700	1.571
$V(\hat{\theta})$	-	124.786	619.113
$\bar{y}$	0.008	0.568	2.814
σ_y	0.563	0.563	0.563
y_{max}	-	2.23	4.47
y_{min}	-	-1.25	0.98

Table III./a. N=500

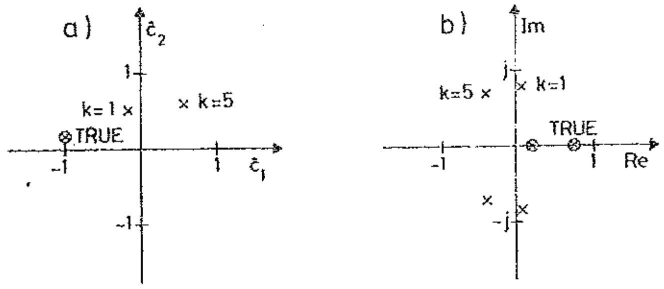

Fig.10. a) Estimates of $\hat{c}_1$ and $\hat{c}_2$ for different values of k
b) Estimates of the roots of $\hat{C}(z^{-1})$ for different values of k

But on the basis of previous examinations wecan not conclude that these statements are also valid for the general case because the model (4) does not include the parameters a_i.

3.3

Let us consider the following model:

$$y_m(t) = \frac{C(z^{-1})}{A(z^{-1})} e(t) + k\sigma_y \gamma(t) \tag{17}$$

which is identified by the model:

$$y_m(t) = \frac{\hat{C}(z^{-1})}{\hat{A}(z^{-1})} \varepsilon(t) \tag{18}$$

Then

$$\varepsilon(t) = \frac{\hat{A}(z^{-1})\ C(z^{-1})}{\hat{C}(z^{-1})\ A(z^{-1})} e(t) + \frac{\hat{A}(z^{-1})}{\hat{C}(z^{-1})} k\sigma_y \gamma(t) \tag{19}$$

and assuming that e(t) is white noise the loss function for a first order system is:

$$F=E\{\varepsilon^2(t)\} = \frac{1}{2\pi j} \oint \frac{(1+\hat{a}z^{-1})(1+cz^{-1})(1+\hat{a}z)(1+cz)}{(1+az^{-1})(1+\hat{c}z^{-1})(1+az)(1+\hat{c}z)} \frac{dz}{z} +$$
$$+ \frac{k^2\sigma_y^2(1+\hat{a})^2}{(1+\hat{c})^2} =$$

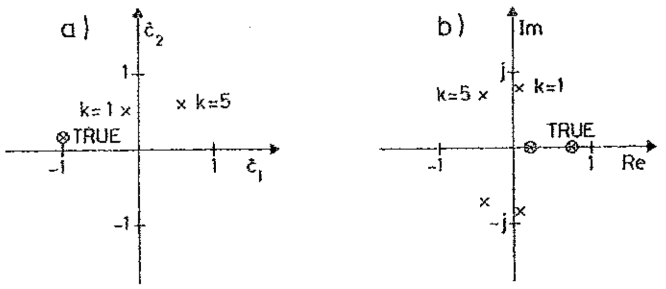

Fig.10. a) Estimates of $\hat{c}_1$ and $\hat{c}_2$ for different values of k
b) Estimates of the roots of $\hat{C}(z^{-1})$ for different values of k

But on the basis of previous examinations wecan not conclude that these statements are also valid for the general case because the model (4) does not include the parameters a_i.

3.3

Let us consider the following model:

$$y_m(t) = \frac{C(z^{-1})}{A(z^{-1})} e(t) + k\sigma_y \gamma(t) \tag{17}$$

which is identified by the model:

$$y_m(t) = \frac{\hat{C}(z^{-1})}{\hat{A}(z^{-1})} \varepsilon(t) \tag{18}$$

Then

$$\varepsilon(t) = \frac{\hat{A}(z^{-1})\; C(z^{-1})}{\hat{C}(z^{-1})\; A(z^{-1})} e(t) + \frac{\hat{A}(z^{-1})}{\hat{C}(z^{-1})} k\sigma_y \gamma(t) \tag{19}$$

and assuming that e(t) is white noise the loss function for a first order system is:

$$F = E\{\varepsilon^2(t)\} = \frac{1}{2\pi j} \oint \frac{(1+\hat{a}z^{-1})(1+cz^{-1})(1+\hat{a}z)(1+cz)}{(1+az^{-1})(1+\hat{c}z^{-1})(1+az)(1+\hat{c}z)} \frac{dz}{z} +$$
$$+ \frac{k^2\sigma_y^2(1+\hat{a})^2}{(1+\hat{c})^2} =$$

$$= \frac{(1+\hat{a}^2+2\hat{a}c+c^2+\hat{a}^2c^2)(1+\hat{a}^2c^2)-2(a+\hat{c})(\hat{a}+c)(1+\hat{a}c)+2\hat{a}c(a^2+2a\hat{c}+\hat{c}^2-a\hat{c}-a^2\hat{c}^2)}{(1-a\hat{c})(1+a^2\hat{c}^2-a^2-\hat{c}^2)} +$$

$$+ \frac{(1+\hat{a})^2}{(1+\hat{c})^2} \quad k^2 \sigma_y^2 \qquad (20)$$

Rewriting the equation (20) so that the numerator of the first term is marked by N, the denominator by D, i.e.

$$F = \frac{N}{D} + \frac{(1+\hat{a})^2}{(1+\hat{c})^2} k^2 \sigma_y^2 \qquad (21)$$

we can write the necessary conditions for the derivatives:

$$\frac{\partial F}{\partial \hat{a}} = \frac{N'_{\hat{a}} D - N D'_{\hat{a}}}{D^2} + \frac{2(1+\hat{a})}{(1+\hat{c})^2} k^2 \sigma_y^2 = 0 \qquad (22)$$

$$\frac{\partial F}{\partial \hat{c}} = \frac{N'_{\hat{c}} D - N D'_{\hat{c}}}{D^2} - \frac{2(1+\hat{a})^2}{(1+\hat{c})^3} k^2 \sigma_y^2 = 0 \qquad (22a)$$

where $N'_{\hat{a}}$ and $D'_{\hat{a}}$ mean the derivatives with respect to $\hat{a}$ of numerator and denominator, $N'_{\hat{c}}$, $D'_{\hat{c}}$ the derivatives of N and D with respect to $\hat{c}$. These equations are so complicated functions of $\hat{a}$ and $\hat{c}$ that they can not be handled analytically.

Let us take the case $k \to \infty$. Now there are two possibilities for satisfying equations (22),(22a) for $k \to \infty$. One of them is that $\hat{a}=-1$, the other one is that $\hat{c}=1$; this latter is equal to the previous one obtained for first order system taking account only stable solutions.

Examining the two terms of equations (22),(22a) the following equalities must be fulfilled:

$$\frac{N D'_{\hat{a}} - N'_{\hat{a}} D}{D^2} = \frac{2(1+\hat{a})}{(1+\hat{c})^2} k^2 \sigma_y^2 \qquad (23)$$

and

$$\frac{N'_{\hat{c}} D - N D'_{\hat{c}}}{D^2} = \frac{2(1+\hat{a})^2}{(1+\hat{c})^3} k^2 \sigma_y^2 \qquad (24)$$

from which we get that D must be 0 if $k \to \infty$.

Writing D in detail:

$$(1-a\hat{c})(1+a^2\hat{c}^2-a^2-\hat{c}^2) = 0 \quad (25)$$

After rearranging we get:

$$(1-a\hat{c})(1-a^2)(1-\hat{c}^2) = 0 \quad (26)$$

The solution $\hat{c}=-1$ must be excluded because this value would make the right sides of equations (23),(24) infinite even in the case of a small k. From the solutions $\hat{c}=1$, $\hat{c}=\frac{1}{a}$ the solution $\hat{c}=1$ is admissible taking account only stable solutions and this is justified by simulation, too. See Table III./b , Table III./c. and Figs. 11.,12. At the same time this value minimizes the right side terms of equations (23),(24).

Para-meters	True values	ML		
		k=0.2	k=1.0	k=5.0
a_1	-0.6	-0.926	-0.997	-1.000
b_1	0	-	-	-
c_1	-0.5	-0.792	-0.897	-0.758
λ	0.4	0.395	0.399	0.424
$V(\hat{\theta})$	-	39.152	39.981	45.103
$\bar{y}$	0.05	0.127	0.446	2.039
σ_y	0.39	0.398	0.398	0.398
y_{max}	-	1.395	1.714	3.307
y_{min}	-	-1.202	-0.883	0.709

Table III./b. N=500

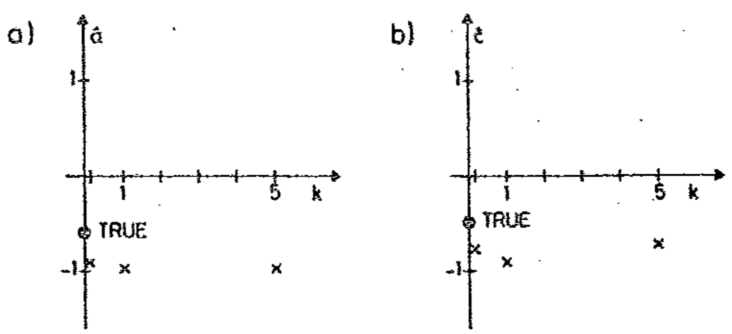

Fig.11. a) Estimates of $\hat{a}$ for different values of k

b) Estimates of $\hat{c}$ for different values of k

Parameters	True values	ML	
		k=1.0	k=5.0
a_1	0.5	-1.00056	-1.00048
b_1	0	-	-
c_1	-0.5	-0.9683	-0.8566
λ	0.4	0.605	0.678
$V(\hat{\Theta})$	-	91.643	115.144
$\tilde{y}$	0.0127	0.600	2.952
σ_y	0.58	0.58	0.58
y_{max}	-	2.227	4.578
y_{min}	-	-1.232	1.119

Table III./c. N=500

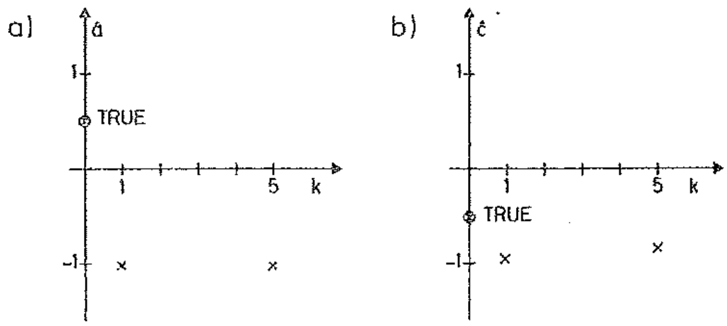

Fig.12. a) Estimates of $\hat{a}$ for different values of k

b) Estimates of $\hat{c}$ for different values of k

But these statements may still not be true in general case.

3.4.

Now let us consider the general case and try to develop the idea.

The simulation model:

$$y_m(t) = \frac{B(z^{-1})}{A(z^{-1})} u(t) + \frac{C(z^{-1})}{A(z^{-1})} e(t) + k \sigma_y \gamma(t) \tag{27}$$

The identification model is:

$$y_m(t) = \frac{\hat{B}(z^{-1})}{\hat{A}(z^{-1})} u(t) + \frac{\hat{C}(z^{-1})}{\hat{A}(z^{-1})} \varepsilon(t) \tag{28}$$

Then

$$\varepsilon(t) = \frac{A(z^{-1})\,\hat{B}(z^{-1}) - \hat{A}(z^{-1})\,B(z^{-1})}{A(z^{-1})\,\hat{C}(z^{-1})} u(t) + \frac{\hat{A}(z^{-1})\,C(z^{-1})}{A(z^{-1})\,\hat{C}(z^{-1})} e(t) +$$

$$+ \frac{\hat{A}(z^{-1})}{\hat{C}(z^{-1})} k \sigma_y \gamma(t) \tag{29}$$

The loss function can be written easily for first order system (assuming that $u(t)$ is white noise; for Table III. $u(t)$ was PRBS) but this simple case would not give new results. For higher order system, however, the relations would become too difficult. But it can be pursued that writing down the loss function and derivating it with respect to $\hat{a}_i$, $\hat{c}_i$ we get the following terms next to k:

$$\frac{\partial}{\partial \hat{a}_i} \longrightarrow \frac{2(1+\hat{a}_1+ \dots +\hat{a}_n)\sigma_y^2}{(1+\hat{c}_1+ \dots +\hat{c}_n)^2} \tag{30}$$

$$\frac{\partial}{\partial \hat{c}_i} \longrightarrow - \frac{2(1+\hat{a}_1+ \dots +\hat{a}_n)^2\sigma_y^2}{(1+\hat{c}_1+ \dots +\hat{c}_n)^3} \tag{31}$$

(The derivatives with respect to $\hat{b}_i$ do not contain k.)

In the case $k \to \infty$ the conditions are similar to the previous ones, i.e.

$$1+\hat{a}_1+ \dots + \hat{a}_n = 0 \tag{32}$$

and values of $\hat{c}_i$ are unconcerned.

Writing down the loss function symbolically on the basis of (29)

$$F = \frac{N_1(a_i,b_i,\hat{a}_i,\hat{b}_i,\hat{c}_i)}{D_1(a_i,\hat{c}_i)} + \frac{N_2(a_i,c_i,\hat{a}_i,\hat{c}_i)}{D_2(a_i,\hat{c}_i)} + \frac{(1+\hat{a}_1+\ldots+\hat{a}_n)^2\, k^2 \sigma_y^2}{(1+\hat{c}_1+\ldots+\hat{c}_n)^2} \tag{33}$$

and derivating it:

$$\frac{\partial F}{\partial \hat{a}_i} = \frac{N'_{1\hat{a}_i} D_1 - N_1 D'_{1\hat{a}_i}}{D_1^2} + \frac{N'_{2\hat{a}_i} D_2 - N_2 D'_{2\hat{a}_i}}{D_2^2} + \frac{2(1+\hat{a}_1+\ldots+\hat{a}_n)k^2\sigma_y^2}{(1+\hat{c}_1+\ldots+\hat{c}_n)^2} \tag{34}$$

$$\frac{\partial F}{\partial \hat{b}_i} = \frac{N'_{1\hat{b}_i} D_1 - N_1 D'_{1\hat{b}_i}}{D_1^2} + \frac{N'_{2\hat{b}_i} D_2 - N_2 D_{2\hat{b}_i}}{D_2^2} \tag{35}$$

$$\frac{\partial F}{\partial \hat{c}_i} = \frac{N'_{1\hat{c}_i} D_1 - N_1 D'_{1\hat{c}_i}}{D_1^2} + \frac{N'_{2\hat{c}_i} D_2 - N_2 D'_{2\hat{c}_i}}{D_2^2} - \frac{2(1+\hat{a}_1+\ldots+\hat{a}_n)^2 k^2\sigma_y^2}{(1+\hat{c}_1+\ldots+\hat{c}_n)^3} \tag{36}$$

We can see that the product D_1D_2 must be equal to 0 in order to satisfy the necessary conditions for derivatives in the case of $k \to \infty$. This means for (34) and (36) that either D_1 or D_2 must be 0 but these are also complicated functions of $\hat{c}_i$ and $\hat{a}_i$ so we can not get the exact values of $\hat{c}_i$ at the given values a_i. (Remark: D_1 and D_2 depend only on a_i and $\hat{c}_i$.) But either D_1 or D_2 is 0 then the equation (35)is not fulfilled so the estimates of b_i become unacceptable. The simulation results showed in Table III./d,III./e,III./f and Figs.13,14,15. verify the facts mentioned above.

Summarizing we can say that if an undesirable constant level is in the output values y(t) the $\hat{a}_i$ will be performed on such way that the sum $(1+\hat{a}_1+\ldots+\hat{a}_n)$ tends to 0, the estimates of c_i depend on a_i but their values can not be expressed exactly from the equations. Naturally the smaller k,the less the influence.

The time functions for the case in Table III.,for k=1 are shown on Fig.16. where the form of residuals is significantly different from the usual one.

Para meters	True values	k=1.0		k=5.0	
		LS	ML	LS	ML
a_1	-1.0	-0.878	-0.978	-1.007	-0.927
a_2	0.2	-0.015	0.089	0.022	-0.056
b_1	1.0	1.033	1.041	1.014	1.033
b_2	0.5	0.604	0.534	0.457	0.586
c_1	-1.0	0	-0.251	0	0.091
c_2	0.2	0	0.332	0	0.155
λ	0.4	0.656	0.626	0.963	0.944
$V(\hat{\theta})$	-	107.477	98.234	231.772	223.224
$\bar{y}$	0.63	3.375		14.367	
σ_y	2.75	2.745		2.745	
y_{max}	-	9.801		20.784	
y_{min}	-	-3.141		7.842	

Table III./d. N=500

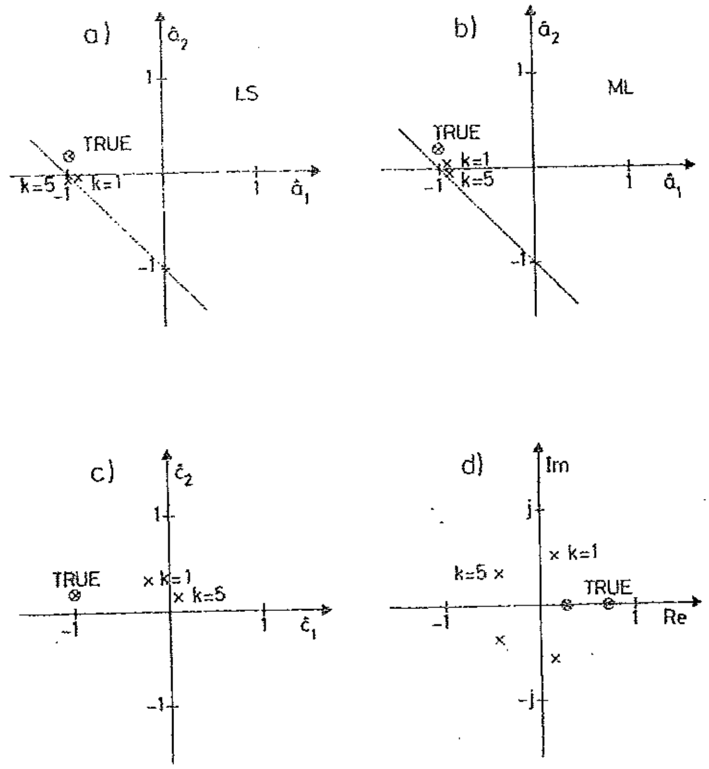

Fig.13. a) Estimates of $\hat{a}_1$ and $\hat{a}_2$ using the LS method and for different k

b) Estimates of $\hat{a}_1$ and $\hat{a}_2$ using the ML method and for different k

c) Estimates of $\hat{c}_1$ and $\hat{c}_2$ for different k

d) Estimates of the roots of $\hat{C}(z^{-1})$ for different k

Para meters	True values	k=1.0		k=5.0	
		LS	ML	LS	ML
a_1	-0.4	-0.254	-0.953	-0.503	-1.864
a_2	-0.4	-0.642	0.022	-0.481	0.864
b_1	1.0	1.032	1.051	1.005	0.996
b_2	0.5	0.634	-0.107	0.372	-0.995
c_1	-1.0	0	-0.906	0	-1.396
c_2	0.2	0	0.583	0	0.491
λ	0.4	0.604	0.551	0.841	0.788
$V(\hat{\Theta})$	-	91.259	75.673	176.744	155.432
$\bar{y}$	0.63	2.65		10.72	
σ_y	2.01	2.02		2.02	
y_{max}	-	7.55		15.62	
y_{min}	-	-2.65		5.42	

Table III./e. N=500

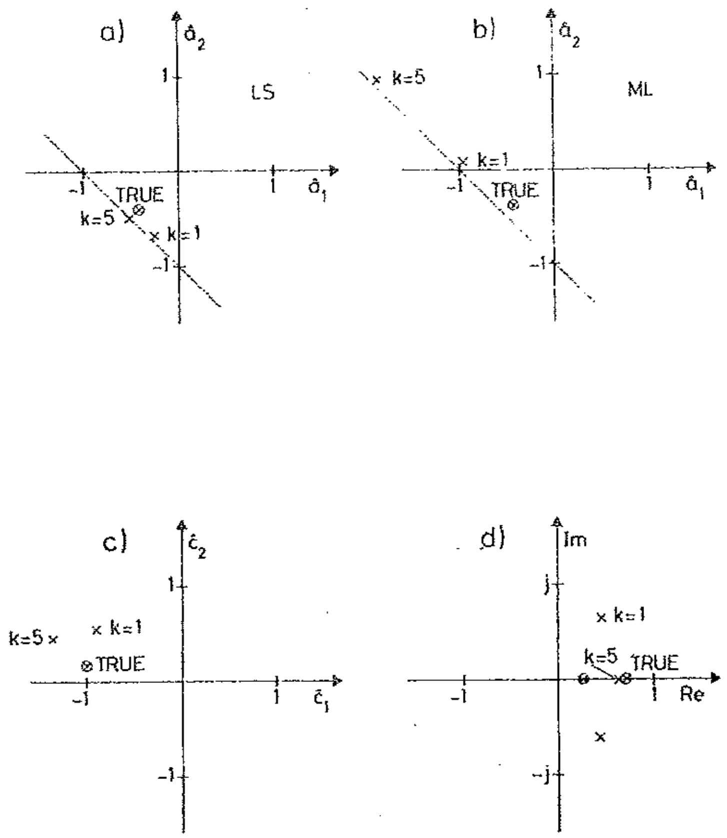

Fig.14. a) Estimates of $\hat{a}_1$ and $\hat{a}_2$ using the LS method and for different k

b) Estimates of $\hat{a}_1$ and $\hat{a}_2$ using the ML method and for different k

c) Estimates of $\hat{c}_1$ and $\hat{c}_2$ for different k

d) Estimates of the roots of $\hat{C}(z^{-1})$ for different k

Para meters	True values	k=1.0		k=5.0	
		LS	ML	LS	ML
a_1	1.5	-0.083	-0.179	-0.161	-0.237
a_2	0.7	-0.786	-0.819	-0.834	-0.763
b_1	1.0	0.923	0.841	0.872	0.794
b_2	0.5	-0.984	-0.896	-1.082	-0.991
c_1	-1.0	0	-1.309	0	-0.604
c_2	0.2	0	0.676	0	0.424
λ	0.4	2.042	0.963	2.240	1.714
$V(\hat{\Theta})$	-	1042.737	232.241	1253.847	736.554
$\bar{y}$	0.04	2.93		14.501	
σ_y	2.89	2.89		2.89	
y_{max}	-	9.98		21.55	
y_{min}	-	-4.49		7.07	

Table III./f.   N=500

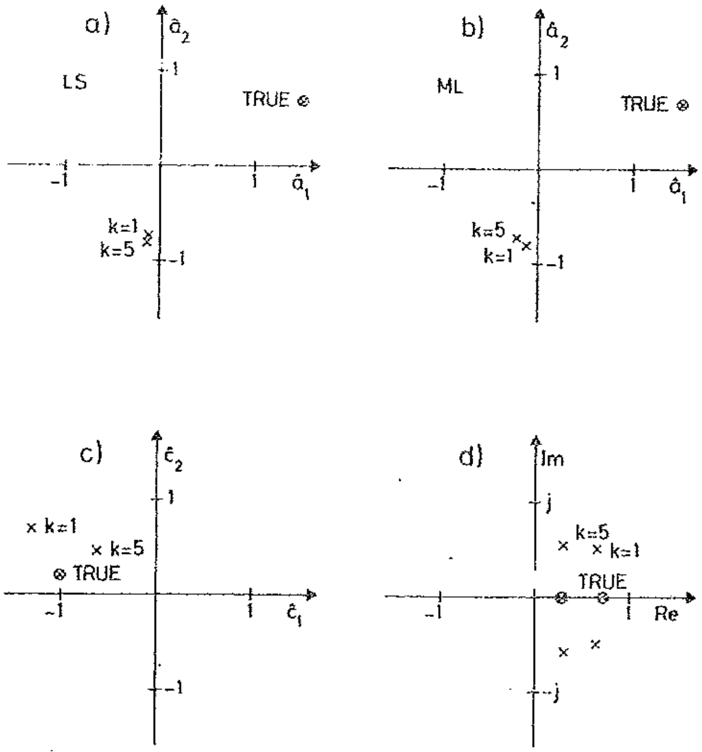

Fig.15. a) Estimates of $\hat{a}_1$ and $\hat{a}_2$ using the LS method and for different k

b) Estimates of $\hat{c}_1$ and $\hat{c}_2$ using the ML method and for different k

c) Estimates of $\hat{c}_1$ and $\hat{c}_2$ for different k

d) Estimates of the roots of $\hat{C}(z^{-1})$ for different k

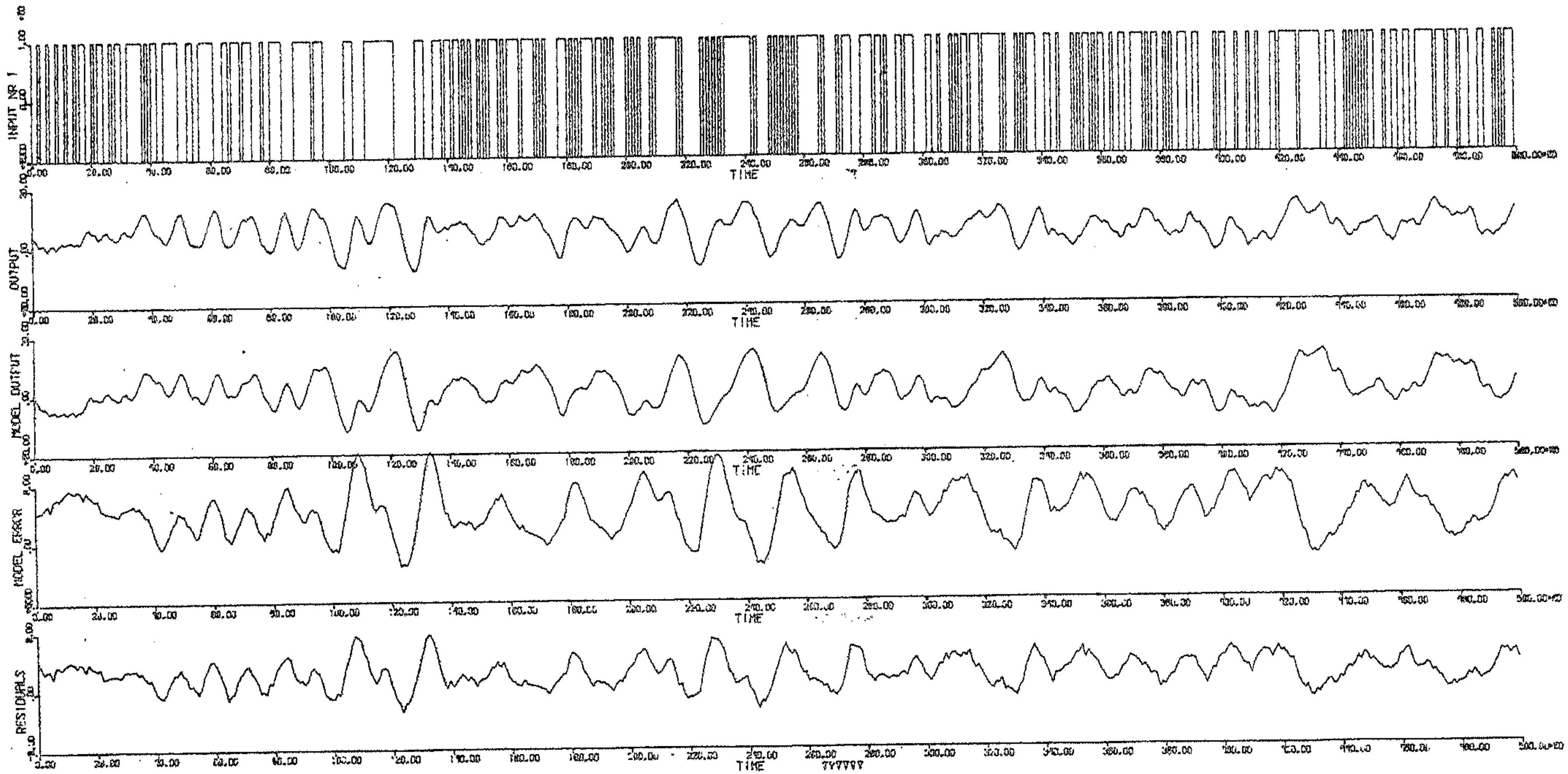

Fig.16. Results of identification for the case k=1 in Table III

4. LINEAR DRIFT

The simulation equation:

$$y_m(t) = y(t) + (t-1)\frac{k}{N-1}\sigma_y \gamma'(t) = y(t) + k\sigma_y \gamma(t) \qquad (37)$$

i.e. the slope of the line was determined as a function of σ_y so after N samples there is a $k\sigma_y$ deformation of the value $y(N)$. See Fig.17.

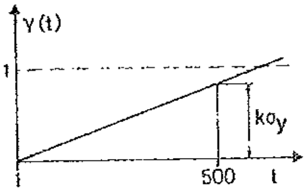

Fig.17. Disturbance γ(t)

The identification results are shown in Table IV. The estimates are getting worse for increasing k. The tendency in the moving of estimates $\hat{c}_i$ for different values of k is similar to the case of constant drift. (See Fig.18.)

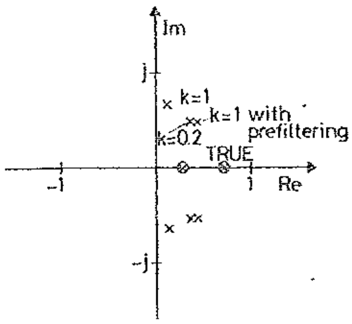

Fig.18. The estimates of the roots of $C(z^{-1})$ with and without prefiltering

A simple strategy can be offered for the improvement of the estimation.

Let us fit a line to the output values and estimate the slope of it and the constant term, i.e.

$$y_e(t) = k_0 + k_1 t \qquad (38)$$

Introducing the following vectors:

$$\underline{f}(t) = [1 \;,\; t] \qquad (39)$$

$$\underline{k} = [k_o \;,\; k_1]$$

$$\underline{y} = \begin{bmatrix} y(1) \\ \vdots \\ y(N) \end{bmatrix} \qquad (40)$$

$$\underline{\underline{F}} = \begin{bmatrix} \underline{f}(1) \\ \vdots \\ \underline{f}(N) \end{bmatrix}$$

the well-known least-squares estimation for $\underline{k}$ is:

$$\hat{\underline{k}} = (\underline{\underline{F}}^T \underline{\underline{F}})^{-1} \underline{\underline{F}} \, \underline{y} = \underline{\underline{G}}^{-1} \underline{w} \qquad (41)$$

where

$$\underline{\underline{G}} = \sum_{t=1}^{N} \underline{f}(t) \, \underline{f}^T(t) \qquad (42)$$

$$\underline{w} = \sum_{t=1}^{N} \underline{f}(t) \, y(t) \qquad (43)$$

Writing it in detail:

$$\underline{\underline{G}} = \begin{bmatrix} \sum_{t=1}^{N} 1 & \sum_{t=1}^{N} t \\ \sum_{t=1}^{N} t & \sum_{t=1}^{N} t^2 \end{bmatrix} = \begin{bmatrix} N & \frac{N(N+1)}{2} \\ \frac{N(N+1)}{2} & \frac{N(N+1)(2N+1)}{6} \end{bmatrix} \qquad (44)$$

$$\underline{\underline{G}}^{-1} = \begin{bmatrix} \frac{4N^2+6N+2}{N(N^2-1)} & -\frac{6(N+1)}{N(N^2-1)} \\ -\frac{6(N+1)}{N(N^2-1)} & \frac{12}{N(N^2-1)} \end{bmatrix} \qquad (45)$$

$$\underline{w} = \begin{bmatrix} \sum_{t=1}^{N} y(t) \\ \sum_{t=1}^{N} t\, y(t) \end{bmatrix} \tag{46}$$

hence the estimates of parameters of the line are the followings:

$$\hat{k}_o = \frac{4N^2+6N+2}{N(N^2-1)} \sum_{t=1}^{N} y(t) - \frac{6N+6}{N(N^2-1)} \sum_{t=1}^{N} t\, y(t) \tag{47}$$

$$\hat{k}_1 = - \frac{6N+6}{N(N^2-1)} \sum_{t=1}^{N} y(t) + \frac{12}{N(N^2-1)} \sum_{t=1}^{N} t\, y(t) \tag{48}$$

The values $\hat{k}_o, \hat{k}_1$ can be estimated easily and their standard deviations are proportional to the diagonal elements of matrix $\underline{\underline{G}}^{-1}$.

After estimating the parameters $\hat{k}_o, \hat{k}_1$ the following filtering strategy can be applied before the identification:

$$y^F(t) = y_m(t) - \hat{k}_o - \hat{k}_1\, t \tag{49}$$

then performing the identification from the values $u(t), y^F(t)$ there are a significant improvement in the estimates, mostly in $\hat{a}_i$ and $\hat{b}_i$, are also shown in Table IV.

The Figs.19.,20. represent the time functions for the case without prefiltering and with prefiltering, respectively.

Parameters	True values	estimations from data without drift		k=0.2		k=1.0		k=1.0 with prefiltering	
		LS	ML	LS	ML	LS	ML	LS	ML
a_1	-1.5	-1.449	-1.502	-1.451	-1.497	-1.484	-1.489	-1.452	-1.501
a_2	0.7	0.653	0.702	0.649	0.694	0.630	0.644	0.656	0.703
b_1	1.0	1.011	1.024	1.018	1.019	1.028	1.019	0.999	0.998
b_2	0.5	0.523	0.473	0.527	0.490	0.501	0.526	0.509	0.469
c_1	-1.0	0	-1.014	0	-0.735	0	-0.273	0	-0.753
c_2	0.2	0	0.268	0	0.337	0	0.460	0	0.365
λ	0.4	0.552	0.391	0.563	0.461	0.707	0.652	0.566	0.464
$V(\hat{\theta})$	-	76.144	38.146	79.200	53.056	125.11	106.35	80.207	53.923
$\bar{y}$	0.63	0.63		1.08		2.89		2.89	
σ_y	4.54	4.54		4.55		4.73		4.73	
y_{max}	10.68	10.68		11.07		13.36		13.36	
y_{min}	-12.58	-12.58		-12.34		-11.41		-11.41	

Table IV. N=500

$\hat{k}_0 = 0.637$; $\hat{k}_1 = 5.153$

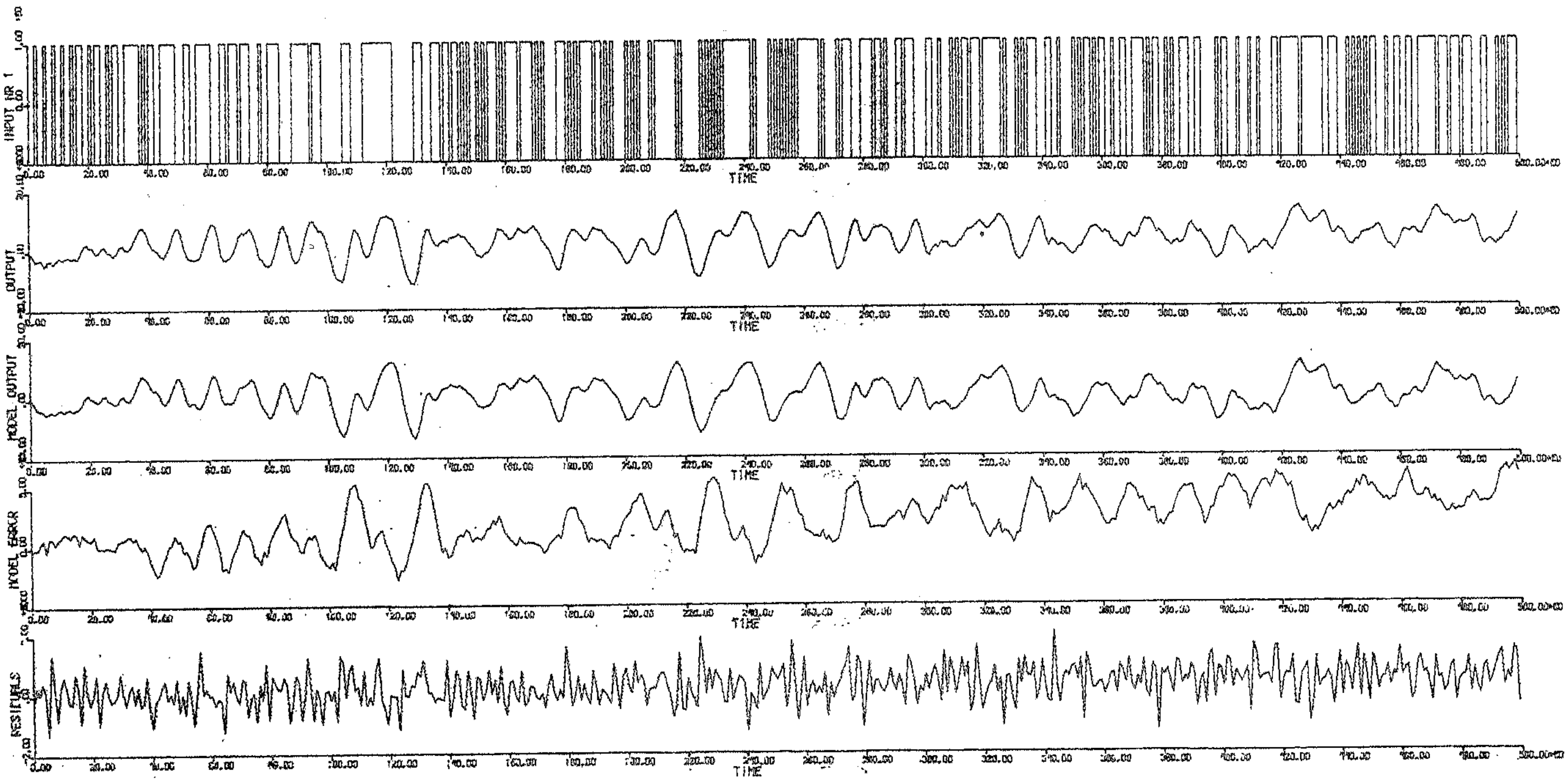

Fig.19. Results of identification for data without prefiltering

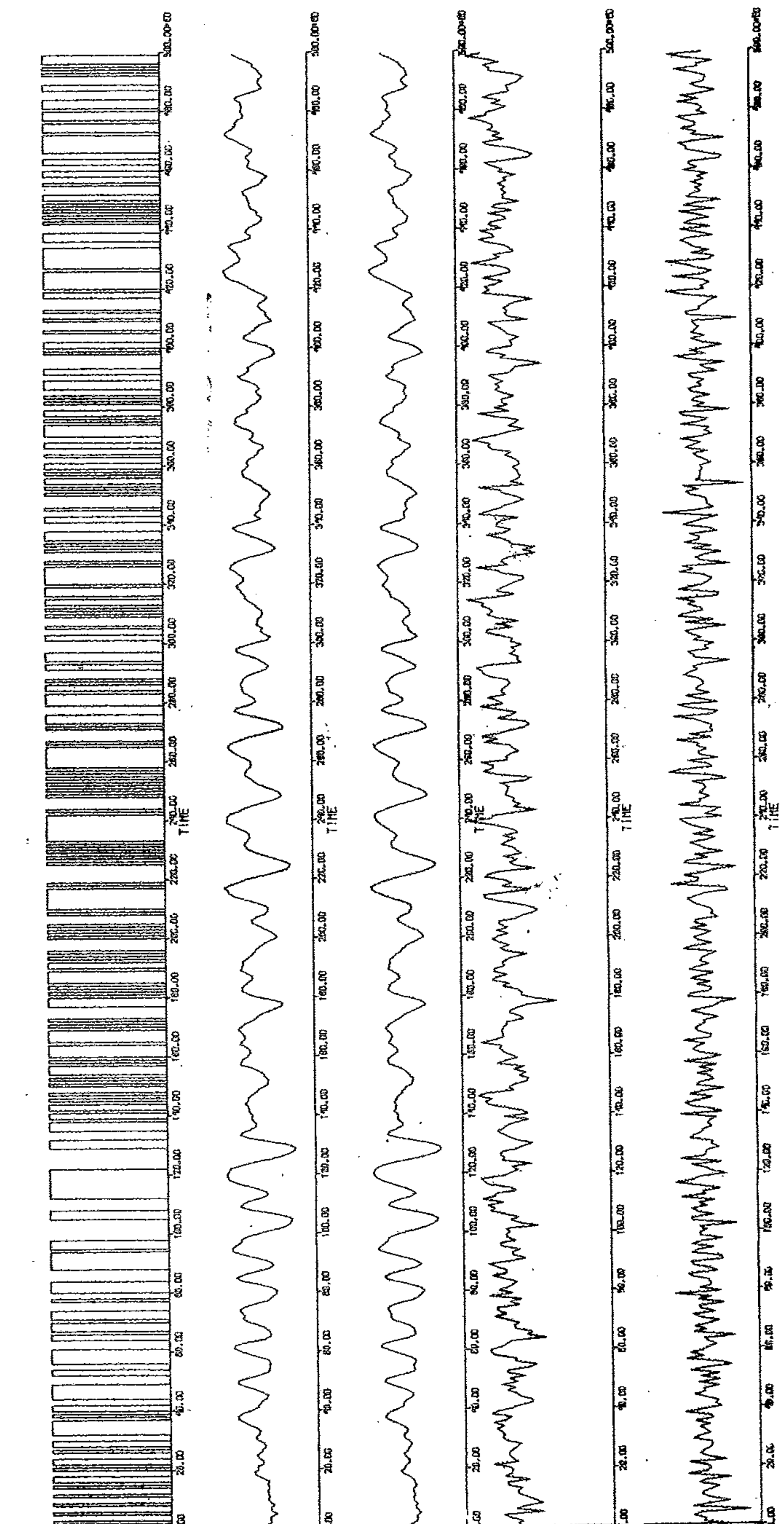

Fig.20. Results of identification of data with prefiltering

5.SINUSOIDAL DRIFT

The simulation was performed according to the equation:

$$y_m(t) = y(t) + k\,\sigma_y \sin \omega t = y(t) + k\,\sigma_y \gamma(t) \qquad (50)$$

where

$$\omega = \frac{2\pi}{T}$$

and T means the number of samplings during a total sinusoidal period.

The identification results are represented in Table V. It can be established from the Table V. that the results are similar to the case 2.,3. for large values of T. The roots of the polynomial $\hat{C}(z^{-1})$ are shown on Figs. 21.,22.

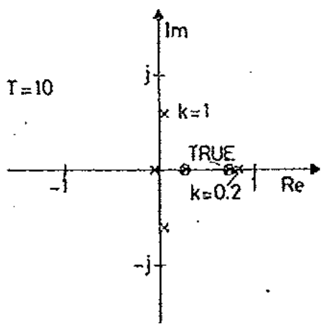

Fig.21. Estimates of the roots of $C(z^{-1})$,T=10

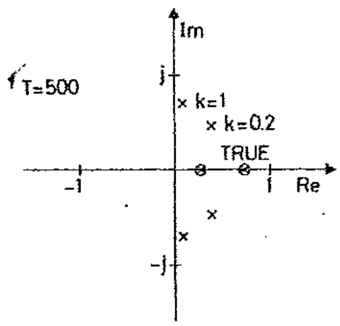

Fig.22. Estimates of the roots of $C(z^{-1})$,T=500

For small values of T the estimates are much worse.

The time functions for k=1, T=10 are shown on Fig.23., for k=1, T=500 on Fig.24.

Para-meters	True values	Estimation from data without drift		T=10				T=500			
				k=0.2		k=1.0		k=0.2		k=1.0	
		LS	ML	LS	ML	LS	ML	LS	ML	LS	ML
a_1	-1.5	-1.449	-1.502	-1.453	-1.499	-1.549	-1.537	-1.451	-1.497	-1.498	-1.485
a_2	0.7	0.653	0.702	0.660	0.701	0.808	0.807	0.649	0.694	0.629	0.628
b_1	1.0	1.011	1.024	1.012	1.031	1.010	1.010	1.015	1.014	1.013	1.010
b_2	0.5	0.523	0.473	0.518	0.483	0.429	0.412	0.523	0.488	0.472	0.531
c_1	-1.0	0.	-1.014	0.	-0.774	0.	-0.064	0.	-0.699	0.	-0.202
c_2	0.2	0.	0.268	0.	-0.030	0.	0.399	0.	0.354	0.	0.476
λ	0.4	0.552	0.391	0.567	0.455	0.799	0.759	0.567	0.473	0.755	0.690
$V(\hat{\theta})$	-	76.144	38.146	80.504	51.783	159.79	144.23	80.410	55.864	142.53	119.01
y	0.63	0.63		0.63		0.63		0.63		0.63	
σ_y	4.54	4.54		4.58		5.52		4.63		5.73	
y_{max}	10.68	10.68		10.51		13.53		11.45		15.08	
y_{min}	-12.58	-12.58		-13.32		-16.82		-11.67		-12.19	

Table V Estimates in the case of sinusoidal drift, N=500.

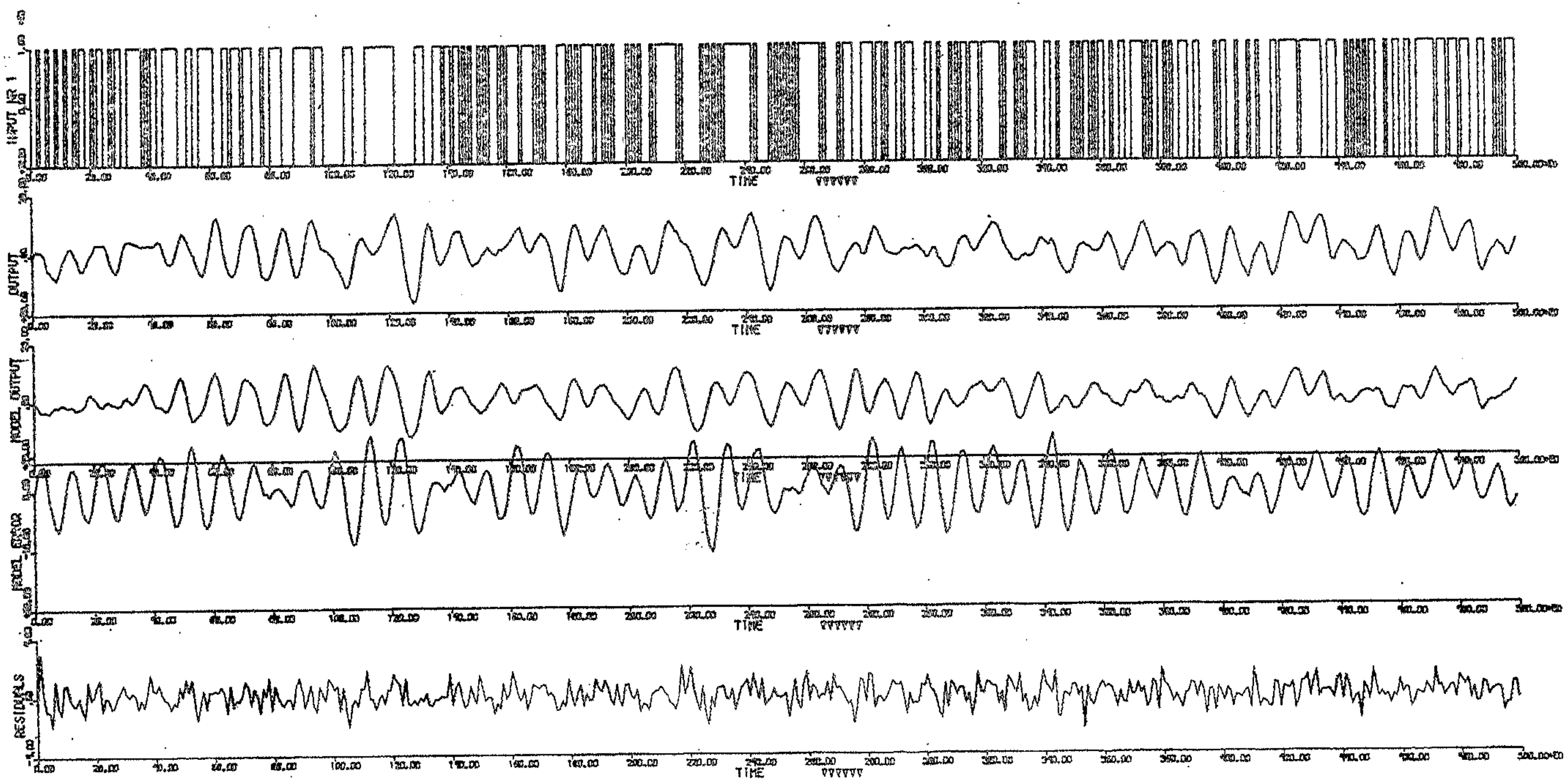

Fig.23 Results of identification, k=1, T=10

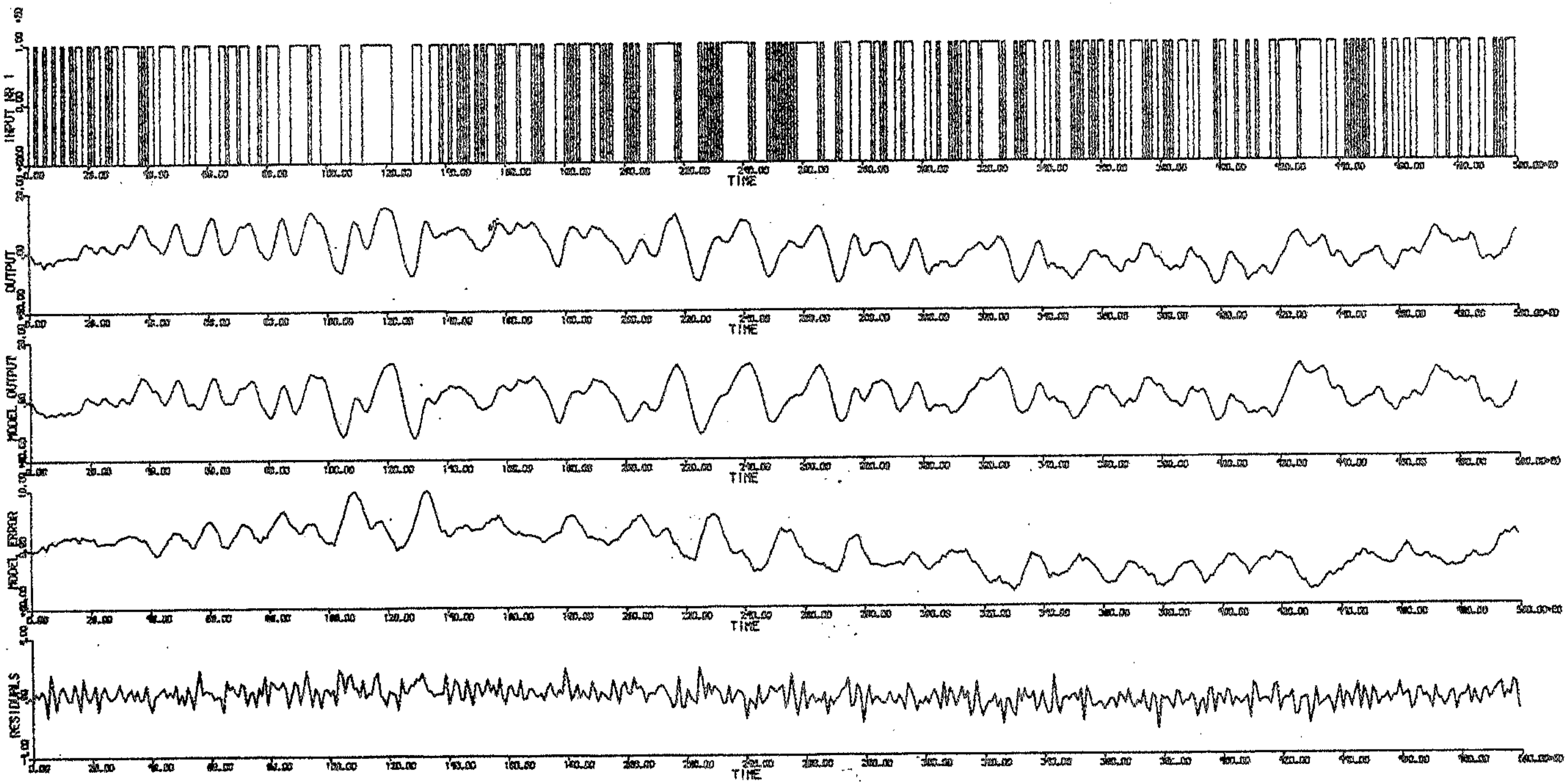

Fig.24. Results of identification, k=1, T=500

6. SUMMARY AND CONCLUSIONS

On the basis of these examinations it can be said that the drift of different character and magnitude can have significant influence on the identification results. Thorough knowledge of the process and technology can help us to form an opinion of existence and character of drift but in spite of this it is worth examining the data.

First of all I suggest to perform a linear regression for the output signal and if it is necessary to filter the data according to the equation (49). After identifying the parameters and plotting the time functions the bad projecting measurements are noticable immediately in the time function of residuals and in this case the estimates $\hat{c}_i$ are unreliable.
The influence of constant drift arises in a complicated way in the estimates. It can be recognized by that the sum $1+\hat{a}_1+\dots+\hat{a}_n$ tends to 0. But in this case the estimates $\hat{c}_i$ and often $\hat{b}_i$, too, are unreliable. Perhaps a prefiltering would be effective in this situation, too.

Referring to the elimination of the influence of drift more exact demands can be created on the measuring conditions or in some cases filtering strategies can be carried out.
The purpose of this paper was not to work out these filtering methods, only to examine the influences of drifts of different types and to interpret the reason of that physically.

7. REFERENCES

[1] Åström K. J. and Bohlin T. :" Numerical Identification of Linear Dynamic System from Normal Operating Records" . Proceedings of the IFAC Conference on Self-Adaptive Control Systems,Teddington (1965).

[2] Gustavsson I.: Parametric Identification of Multiple Input,Single Output Linear Dynamic Systems. Report 6907.July 1969. Lund Institute of Technology,Division of Automatic Control.

[3] Åström K. J. and Eykhoff P.: System Identification - A Survey. IFAC, Prague,1970.

8. ACKNOWLEDGEMENTS

The author wishes to express her gratitude to professor K.J. Åström and I. Gustavsson, for initiating the problem and for valuable suggestions and discussions.