RAL-92-024

Science and Engineering Research Council
Rutherford Appleton Laboratory
Chilton DIDCOT Oxon OX11 0QX

RAL-92-024

Nonlinear Response Function Estimation: Moment Formulations and the Effect of Outliers

A D Irving and T Dewson

April 1992

NONLINEAR RESPONSE FUNCTION ESTIMATION: MOMENT FORMULATIONS AND THE EFFECT OF OUTLIERS

A D Irving* and T Dewson+

* Energy Research Unit, Rutherford Appleton Laboratory, Chilton, Didcot, Oxon, OX11 OQX, UK

\+ Department of Mechanical Engineering, Queens Building, The University of Bristol, Bristol BS8 1TR, UK

ABSTRACT

Finite memory nonlinear response functions may be estimated using time series techniques. Absolute, central and product moment values are estimated from the time series data. The nonlinear response values are then deduced using each moment form by a simultaneous algebraic method. The input data $\{x(t)\}$ are drawn from a general stochastic process so that differences between the moment estimators may be highlighted. In the event, no significant differences in the estimated response values and predicted output time series were observed demonstrating that the formalism is insensitive to the moment estimator used. The accuracy of the response estimates is assessed and discussed with reference to the properties of the moment estimators. The effect of additive outliers in the data $\{x(t)\}$ or $\{y(t)\}$ are considered. A simple 'predictor-corrector' methodology for additive outliers is developed and applied to a mixed linear and quadratic nonlinear system, and its ability to reduce the effects of the outliers is demonstrated.

INTRODUCTION

A response function is a characterisation of a relationship between some observed input variable $\{x(t)\}$, such as temperature, and some observed output variable $\{y(t)\}$. In general the form of the relationship is not known and science is concerned with the identification, analysis and interpretation of such a relationship. These relations then form the basis of the physical laws which describe the observed phenomena and our acceptance of such laws relies upon the repeatability of the observations and the predictive power of the laws, given the underlying probabilistic nature of the observations. When estimating such relationships it is important to identify those observables which are essential to the description of the data and which are superfluous. Often experimental situations occur when one does not know which are the important variables and characterisations; consequently some degree of intuition may be necessary. Thus there is a need, not only to characterise the properties of the data, but also determine the minimum amount of information which needs to be extracted from the data in order to adequately describe relevant aspects of that data.

A wide class of physically observable systems are irreversible, time dependent, nonlinear and possess only finite memories. One natural characterisation of the nonlinear input-output problem is the impulse response or Volterra kernel functionals [VO59, VO00]. The Volterra series is known as a functional power series where the functionals are multidimensional convolutions of the impulse response of the system. The Volterra series treats the linear case as a subclass of the nonlinear case so that all the concepts that are used in linear analysis may be carried over into the nonlinear case. The nonlinear response values may be estimated from the input data $\{x(t)\}$ and output data $\{y(t)\}$ sequences.

Wiener [WI42,WI58] was the first to characterise the input-output behaviour of nonlinear systems in terms of Volterra functionals. Wiener expressed the relationship between the input time series sequence $\{x(t)\}$ and output sequence $\{y(t)\}$ as a Volterra series. The Volterra series may be written as

$$y(t) = \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1 \ldots \int d\sigma_n \, h_{yx}n(\sigma_1,\ldots,\sigma_n) \prod_{i=1}^{n} x(t-\sigma_i) \quad (1)$$

where N is the order of the system, where the response functions $h_{yx}n\ (\sigma_1,\ldots,\sigma_n)$ characterise the system, where t denote time and where the σ_i's denotes time delay with respect to the time t.

In the present work the response values are estimated directly from the time series data $\{x(t)\}$ and $\{y(t)\}$. The time delayed moments between $\{x(t)\}$ and $\{y(t)\}$ in terms of the Volterra expansion form a set of simultaneous inhomogeneous equations which may be solved for the unknown response values. Three forms of moments will be considered and the results compared from the analysis of a nonlinear system subjected to the same stochastic input data. In addition the effect of additive outliers, and the use of a predictor-correction method to improve the estimates of the response values in the presence of these outliers, are considered.

GENERAL THEORY

Considering first the second order absolute cross moment between $\{x(t)\}$ and $\{y(t)\}$ of equation (1), we have that [IR92]

$$E[x(t-\tau_1)y(t)] = \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1 \ldots \int d\sigma_n \, h_{yx}n(\sigma_1,\ \ldots,\ \sigma_n) \; E[x(t-\tau_1) \prod_{i=1}^{n} x(t-\sigma_i)] \quad (2)$$

which may be written as

$$A_{xy}(\tau_1) = \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1 \ldots \int d\sigma_n \, h_{yx}n(\sigma_1,\ldots,\sigma_n) A_{xx}n(\tau_1,\sigma_1,\sigma_2,\ldots,\sigma_n) \quad (3)$$

Defining the (n+1)th order absolute auto moment as

$$A_{xx^n}(\tau_1,\sigma_1,\ldots,\sigma_n) = E[x(t-\tau_1)\prod_{i=1}^{n} x(t-\sigma_i)] \tag{4}$$

then by inspection the (r+1)th order cross moment may be written as

$$E[y(t)\prod_{i=1}^{r} x(t-\tau_i)] = \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1\ldots\int d\sigma_n\, h_{yx^n}(\tau_1,\ldots,\tau_r,\sigma_1,\ldots,\sigma_n)$$

$$E[\prod_{i=1}^{r} x(t-\tau_i)\prod_{j=1}^{n} x(t-\sigma_j)] \tag{5}$$

or as

$$A_{x^r y}(\tau_1,\ldots,\tau_r) = \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1\ldots\int d\sigma_n\, h_{yx^n}(\sigma_1,\ldots,\sigma_n)$$

$$A_{x^r x^n}(\tau_1,\tau_2,\ldots,\tau_r,\sigma_1,\ldots,\sigma_n) \tag{6}$$

Moving to central moments now, the average value of the output $\{y(t)\}$ is

$$E[y(t)] = \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1\ldots\int d\sigma_n\, h_{yx^n}(\sigma_1,\ldots,\sigma_n)\; E[\prod_{i=1}^{n} x(t-\sigma_i)] \tag{7}$$

which when subtracted from equation (1) yields

$$\{y(t) - E[y(t)]\} = \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1\ldots\int d\sigma_n\, h_{yx^n}(\sigma_1,\ldots,\sigma_n)$$

$$\{\prod_{i=1}^{n} x(t-\sigma_i) - E[\prod_{i=1}^{n} x(t-\sigma_i)]\}. \tag{8}$$

Multiplying equation (8) by $\{\prod_{j=1}^{r} \{x(t-\tau_j) - E[x(t-\tau_j)]\}\}$ on each side and taking the expectation, by inspection the (r+1)th order cross central moment between the output $\{y(t)\}$ and the input $\{x(t)\}$ is

$$E[\prod_{j=1}^{r} \{x(t-\tau_j) - E[x(t-\tau_j)]\} \{y(t) - E[y(t)]\}]$$

$$= \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1 \ldots \int d\sigma_n \, h_{yx^n}(\sigma_1, \ldots, \sigma_n)$$

$$E[\prod_{j=1}^{r} \{x(t-\tau_j) - E[x(t-\tau_j)]\} \{\prod_{i=1}^{n} x(t-\sigma_i) - E[\prod_{i=1}^{n} x(t-\sigma_i)]\}] \quad (9)$$

Defining the cross central moments as

$$C_{x^r y}(\tau_1, \ldots, \tau_r) = E[\prod_{j=1}^{r} \{x(t-\tau_j) - E[x(t-\tau_j)]\} \{y(t) - E[y(t)]\}] \quad (10)$$

and the auto central moment as

$$C_{x^r x^n}(\tau_1, \tau_2, \ldots, \tau_r, \sigma_1, \ldots, \sigma_n)$$

$$= E[\{\prod_{j=1}^{r} \{x(t-\tau_j) - E[x(t-\tau_j)]\} \{\prod_{i=1}^{n} x(t-\sigma_i) - E[\prod_{i=1}^{n} x(t-\sigma_i)]\}]. \quad (11)$$

then equation (9) may be written as

$$C_{x^r y}(\tau_1, \ldots, \tau_r) = \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1 \ldots \int d\sigma_n \, h_{yx^n}(\sigma_1, \ldots, \sigma_n)$$

$$C_{x^r x^n}(\tau_1, \ldots, \tau_r, \sigma_1, \ldots, \sigma_n). \quad (12)$$

It should be noted that this form of moment is specific to the formalism used here and it is not the same as the usual definition of central moments, for example [RO85]

$$C'_{x^r x^n}(\tau_1,\ldots,\tau_r,\sigma_1,\ldots,\sigma_n)$$

$$= E[\prod_{j=1}^{r} \{x(t-\tau_j)-E[x(t-\tau_j)]\} \prod_{i=1}^{n} \{x(t-\sigma_i)-E[x(t-\sigma_i)]\}]$$

which characterise the properties of the data $\{x(t)\}$.

Operating on the Volterra series with a product moment formulation the (r+1)th term is

$$E[\{\prod_{j=1}^{r} x(t-\tau_j)-E[\prod_{j=1}^{r} x(t-\tau_j)]\}\{y(t)-E[y(t)]\}]$$

$$= \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1 \ldots \int d\sigma_n \, h_{yx^n}(\sigma_1,\ldots,\sigma_n) \, E[\{\prod_{j=1}^{r} x(t-\tau_j)-E[\prod_{j=1}^{r} x(t-\tau_j)]\}$$

$$\{\prod_{i=1}^{n} x(t-\sigma_i)-E[\prod_{i=1}^{n} x(t-\sigma_i)]\}]$$

where now we have operated on equation (8) by

$$E[\{\prod_{j=1}^{r} x(t-\tau j) - E[\prod_{j=1}^{r} x(t-\tau j)]\}.]$$

which may be written as

$$K_{x^r y}(\tau_1,\ldots,\tau_r) = \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1 \ldots \int d\sigma_n \, h_{yx^n}(\sigma_1,\ldots,\sigma_n) \, K_{x^r x^n}(\tau_1,\ldots,\tau_r,\sigma_1,\ldots,\sigma_n) \tag{13}$$

where the cross product moments have been defined as

$$K_{x^r y}(\tau_1,\ldots,\tau_r) = E[\{\prod_{j=1}^{r} x(t-\tau_j) - E[\prod_{j=1}^{r} x(t-\tau_j)]\}\{y(t)-E[y(t)]\}] \tag{14}$$

and the auto product moment as

$$K_{x^r x^n}(\tau_1,\ldots,\tau_r,\sigma_1,\ldots,\sigma_n)$$

$$= E[\{\prod_{j=1}^{r} x(t-\tau_j) - E[\prod_{j=1}^{r} x(t-\tau_j)]\}\{\prod_{i=1}^{n} x(t-\sigma_i) - E[\prod_{i=1}^{n} x(t-\sigma_i)]\}] \tag{15}$$

As can be observed the equation for the (r+1)th order cross moment remains the same for each moment formalism and these can be rewritten by denoting the absolute, central and product moments as $M_{x^r y}(\tau_1,\ldots,\tau_r)$ and $M_{x^r x^n}(\tau_1,\tau_2,\ldots,\tau_r,\sigma_1,\ldots,\sigma_n)$. Note that as the cumulant values are simple rearrangements of the absolute moments that form is not readily amenable for the solution of the response values.

Collecting these equations which are nonlinear and simultaneous in the unknown response function $h_{yx^n}(\tau_1,\ldots,\tau_n)$ then [IR92]

$$M_{xy}(\tau_1) = \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1 \ldots \int d\sigma_n \; h_{yx^n}(\sigma_1,\ldots,\sigma_n) M_{xx^n}(\tau_1,\sigma_1,\sigma_2,\ldots,\sigma_n)$$

$$M_{x^2y}(\tau_1,\tau_2) = \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1 \ldots \int d\sigma_n \; h_{yx^n}(\sigma_1,\ldots,\sigma_n) M_{x^2x^n}(\tau_1,\tau_2,\sigma_1,\ldots,\sigma_n)$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$M_{x^ry}(\tau_1,\ldots,\tau_r) = \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1 \ldots \int d\sigma_n \; h_{yx^r}(\sigma_1,\ldots,\sigma_r)$$

$$\vdots \qquad M_{x^rx^n}(\tau_1,\tau_2,\ldots,\tau_r,\sigma_1,\ldots,\sigma_n)$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$M_{x^Ny}(\tau_1,\ldots,\tau_N) = \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1 \ldots \int d\sigma_n \; h_{yx^n}(\sigma_1,\ldots,\sigma_n)$$

$$M_{x^Nx^n}(\tau_1,\tau_2,\ldots,\tau_N,\sigma_1,\ldots,\sigma_n) \qquad (16)$$

where N is the order of the system.

These form a set of N simultaneous inhomogeneous nonlinear integral equations, from which the values of the N unknown response functions $h_{yx^n}(\sigma_1,\ldots,\sigma_n)$ can be estimated. This may be compactly written as

$$M_{x^Ny} = M_{x^Nx^N} \otimes h_{yx^N} \qquad (17)$$

The matrix equation given by (17), which describes the complete sequence of equations for $\tau_i=0,1,\ldots,\mu$ and where $i=1,\ldots,N$, can be solved algebraically for the unknown response values.

THE EFFECT OF ADDITIVE OUTLIERS

In this section the effect of occasional intermittent faults in the experimental apparatus are considered. These may be considered as additive outliers on the data [BA90]. Such outliers will contaminate the data and may lead to erroneous conclusions about the properties of the process under study.

First consider a single delta functional outlier on the output data $\{y(t)\}$. As before the properties of the process may be characterised using equation (1) with

$$y(t) = \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1 \ldots \int d\sigma_n \, h_{yx^n}(\sigma_1, \ldots, \sigma_n) \prod_{i=1}^{n} x(t-\sigma_i)$$

only now there is an outlier at time t_1 with the new sequence $\{y(t)+\alpha\delta(t=t_1)\}$.

The estimates of the sample moments $M^*_{x^r y}(\tau_1, \ldots, \tau_r)$ will be affected but the sample moments of the input data $M_{x^r x^n}(\tau_1, \ldots, \tau_r, \sigma_1, \ldots, \sigma_n)$ will remain unaffected. Explicitly if an absolute form of moments is used then the cross moments will be

$$M^*_{x^r y}(\tau_1, \ldots, \tau_r) = M_{x^r \alpha\delta(t=t_1)}(\tau_1, \ldots, \tau_r) + M_{x^r y}(\tau_1, \ldots, \tau_r) \qquad (18)$$

which may be called the apparent absolute moment, where $M_{x^r y}(\tau_1, \ldots, \tau_r)$ is the same as for the case when no outliers are present and where $M_{x^r \alpha\delta(t=t_1)}(\tau_1, \ldots, \tau_r)$ is the moment between the input data $\{x(t)\}$ and the outlier. The modified set of equations to be solved is

$$M^*_{xy}(\tau_1) = \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1 \ldots \int d\sigma_n \; h^*_{yx^n}(\sigma_1,\ldots,\sigma_n) M_{xx^n}(\tau_1,\sigma_1,\ldots,\sigma_n)$$

$$\vdots \qquad\qquad \vdots$$

$$M^*_{x^r y}(\tau_1,\ldots,\tau_r) = \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1 \ldots \int d\sigma_n \; h^*_{yx^n}(\sigma_1,\ldots,\sigma_n) \; M_{x^r x^n}(\tau_1,\ldots,\tau_r,\sigma_1,\ldots,\sigma_n)$$

$$\vdots \qquad\qquad \vdots$$

$$M^*_{x^N y}(\tau_1,\ldots,\tau_N) = \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1 \ldots \int d\sigma_n \; h^*_{yx^n}(\sigma_1,\ldots,\sigma_n) \; M_{x^N x^n}(\tau_1,\ldots,\tau_N,\sigma_1,\ldots,\sigma_n) \qquad (19)$$

where the (apparent) kernel values $h^*_{yx^n}(\sigma_1,\ldots,\sigma_n)$ have been modified by the inclusion of the additive outlier at time $t = t_1$.

Given the input data $\{x(t)\}$ and the estimated Volterra kernel values $h^*_{yx^n}(\sigma_1,\ldots,\sigma_n)$ a predicted output sequence $\{y_p(t)\}$ may be generated and compared against the observed data $\{y(t)\}$. The time t_1 at which the additive outlier occurred can be identified and a correction to the observed value by an amount $\{y(k) - y_p(k)\}$ made. So after adjusting the values of $y(t)$, the moment values can be re-estimated using this corrected data, and hence new estimates of the response values are obtained. By adjusting the output data in such a predictor corrector manner the response values may be readjusted until the value of some test statistic between successive predicted time series lie within some predetermined range. So as the magnitude of the corrected outlier tends to zero, ie $\alpha\delta(t=t_1) \rightarrow 0$, so then the corrected values for the estimated moments will converge to their values if no outlier were present, ie $M^*_{x^n x}(\tau_1,\ldots,\tau_n) \rightarrow M_{x^n y}(\tau_1,\ldots,\tau_n)$; consequently the determined Volterra kernel values will converge to their time values, ie $h^*_{yx^n}(\sigma_1,\ldots,\sigma_n) \rightarrow h_{yx^n}(\sigma_1,\ldots,\sigma_n)$.

Next consider the effect of a single delta functional additive outlier $\alpha\delta(k)$ at some time $t=t_1$ that is superimposed onto the values of the input observable $\{x(t) + a\delta(t=t_1)\}$.

All of the estimated sample moments are affected by the outlier and the modified moment equations in this case are

$$M^*_{xy}(\tau_1) = \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1 \ldots \int d\sigma_n\ h^*_{yx}n(\sigma_1,\ldots,\sigma_n) M^*_{xx}n(\tau_1,\sigma_1,\ldots,\sigma_n)$$

$$\vdots \qquad\qquad \vdots$$

$$M^*_{x}r_{y}(\tau_1,\ldots,\tau_r) = \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1 \ldots \int d\sigma_n\ h^*_{yx}n(\sigma_1,\ldots,\sigma_n)\ M^*_{x}r_{x}n(\tau_1,\ldots,\tau_r,\sigma_1,\ldots,\sigma_n)$$

$$\vdots \qquad\qquad \vdots$$

$$M^*_{x}N_{y}(\tau_1,\ldots,\tau_r) = \sum_{n=1}^{N} \frac{1}{n!} \int d\sigma_1 \ldots \int d\sigma_n\ h^*_{yx}n(\sigma_1,\ldots,\sigma_n)\ M^*_{x}N_{x}n(\tau_1,\ldots,\tau_N,\sigma_1,\ldots,\sigma_n) \qquad (20)$$

From the input data $\{x(t) + \alpha\delta(t=t_1)\}$ and the estimated response values $h^*_{yx}n(\sigma_1,\ldots,\sigma_n)$ we may generate a predicted output sequence $\{y_p(t)\}$ and compare this against the observed data $\{y(t)\}$. If we can identify the time t_1 at which the additive outlier occurred we may correct the observed value by a given amount and then re-estimate the response values using the corrected data. In a similar fashion to that applied to the output in the output data case we may thus adjust the input data in a predictor corrector manner until the value of some test statistic between successively predicted time series lies within some predetermined range. So as the magnitude of the corrected outlier tends to zero, ie $\alpha\delta(t=t_1) \to 0$, so then the corrected values for the estimated moments will converge to their values with no outlier were present,

ie

$$M^*_{x}r_{x}n(\tau_1,\ldots,\tau_r,\sigma_1,\ldots,\sigma_n) \to M_{x}r_{x}n(\tau_1,\ldots,\tau_r,\sigma_1,\ldots,\sigma_n) \qquad (21)$$

and

$$M^*_{x}r_{y}(\tau_1,\ldots,\tau_r) \to M_{x}r_{y}(\tau_{1,\ldots,\tau r});$$

and consequently the determined Volterra kernel values will converge to their true values, ie $h^*_{yx}n(\sigma_1,\ldots,\sigma_n) \to h_{yx}n(\sigma_1,\ldots,\sigma_n)$.

Consider a quadratic system where

$$y(t) = \int h_{yx}(\sigma_1)x(t-\sigma_1)d\sigma_1+\int\int h_{yxx}(\sigma_1,\sigma_2)x(t-\sigma_1)x(t-\sigma_2)d\sigma_1 d\sigma_2$$

and introduce an additive outlier $\alpha\delta(t=t_1)$ at $t=t_1$, to the input, then the difference between the observed output $\{y(t)\}$ and $\{y_p(t)\}$ the predicted output when the outlier is present will be given by the characteristic equation

$$\int_0^\mu \{y_o(t_1+\sigma_1) - y(t_1+\sigma_1)\}d\sigma_1$$

$$= \alpha \int_0^\mu g_{yx}(\sigma_1)\, d\sigma_1 + \alpha^2 \int_0^\mu\int_0^\mu g_{yxx}(\sigma_1,\sigma_2)d\sigma_1 d\sigma_2 \qquad (22)$$

where μ is the finite memory of the system and α is the amplitude of the outlier at $t=t_1$. Estimates of $\{y_p(t)\}$ can be obtained by convoluting the input data $\{x(t)\}$ and the estimated response values $h_{yx}(\sigma_1)$ and $h_{yxx}(\sigma_1,\sigma_2)$. The amplitude, α, of the additive outlier at time t_1 may be solved for using the equation

$$\int_0^\mu \{y_p(t_1 + \sigma_1) - y(t_1 + \sigma_1)\}\, d\sigma_1$$

$$= \alpha \int_0^\mu h_{yx}(\sigma_1)d\sigma_1 + \alpha^2 \int_0^\mu\int_0^\mu h_{yx}2(\sigma_1,\sigma_2)d\sigma_1 d\sigma_2 \qquad (23)$$

The value of $x(t_1)$ is now adjusted by an amount α, the moments and response values are recalculated and equation (23) solved for a new value of α until some convergence criteria is satisfied. A similar procedure may be adopted if there are several additive outliers or for higher order nonlinear systems. Demonstration of these algorithms are provided in the examples in the next section.

APPLICATION TO SPECIFIC NUMERICAL EXAMPLES

In the section on moment formulations a set of nonlinear inhomogeneous equations were developed using three separate forms of time series moment estimators. In the present section the formalism is applied to a numerical example where the properties of the system are known in order to assess the ability of and accuracy of the absolute, central and product moment forms to estimate the response values. Each of the time series moment estimators will have a different bias function and similarly each will have different tolerances to any, for example, nonstationary effects or outliers in the data.

In order to make a realistic assessment of the significance of such effects experimental data is used as the input time series sequences $\{x(t)\}$ in the following examples. Time series data were collected at one minute intervals from a solar building for a duration of 30 days in July 1989 at Cranfield. External meteorological measurements were obtained for the dry bulb temperature, wind speed, wind direction, global and diffuse horizontal irradiance and the nett irradiance between the test cell roof and the sky. Internal meteorological measurements were obtained for the air temperature, wall surface temperatures, and the nett irradiance between the north wall and the south facing window. In addition measurements were made of heat flux and temperature time series at a depth of 5mm within the test cell wall's internal and external surfaces.

In the present work the external surface heat flux data are used as the input $\{x(t)\}$. The data $\{x(t)\}$, which is of a general stochastic form, are then convoluted with known response functions to generate the output sequence of data $\{y(t)\}$ for each of the numerical examples. Explicitly using the convolution

$$y(t) = \sum_{\tau_1=0}^{\mu} g_1(\tau_1)\, x(t-\tau_1)$$

$$+ \sum_{\tau_1=0}^{\mu} \sum_{\tau_2=0}^{\mu} g_2(\tau_1,\tau_2)\, x(t-\tau_1)\, x(t-\tau_2) \qquad (24)$$

we obtain the output data values $\{y(t)\}$, where μ is the finite memory of the system and where $g_1(\tau_1)$ and $g_2(\tau_1,\tau_2)$ are the response functions that define the nonlinear system. Given the time series sequences $\{x(t)\}$ and $\{y(t)\}$ we then estimate the absolute, central and product moment values and from these the system response values $h_{yx}(\tau_1)$ and $h_{yx}2(\tau_1,\tau_2)$ are deduced for comparison with the known values.

The accuracy of the response function estimates is determined using the root mean square difference and absolute mean difference with

$$\text{rms} = \left\{ \frac{1}{(\mu+1)} \sum_{\tau_1=0}^{\mu} (h_{yx}(\tau_1) - g_1(\tau_1))^2 \right\}^{½} \tag{25}$$

and

$$\text{abs} = \frac{1}{(\mu+1)} \sum_{\tau_1=0}^{\mu} |h_{yx}(\tau_1) - g_1(\tau_1)| \tag{26}$$

for the first order response and

$$\text{rms} = \left\{ \frac{1}{(\mu+1)^2} \sum_{\tau_1=0}^{\mu} \sum_{\tau_2=0}^{\mu} (h_{yx}2(\tau_1,\tau_2) - g_2(\tau_1,\tau_2)) \right\}^{½} \tag{27}$$

and

$$\text{abs} = \frac{1}{(\mu+1)^2} \sum_{\tau_1=0}^{\mu} \sum_{\tau_2=0}^{\mu} |h_{yx}2(\tau_1,\tau_2) - g_2(\tau_1,\tau_2)| \tag{28}$$

for the second order response values.

When identifying and analysing unknown nonlinear systems the response values are not known a priori. However by generating the predicted output time series values $\{y_p(t)\}$ from the input data $\{x(t)\}$ and the estimated Volterra kernel values $h_{yx}(\tau_1)$ and $h_{yx}2(\tau_1,\tau_2)$, from which we may assess the accuracy of the predicted sequence values using the normalised root mean square difference statistic

$$\text{nrms} = \{\frac{1}{T} \sum_{t=1}^{T} \{\frac{y_p(t) - y(t)}{y(t)}\}^2\}^{½} \tag{29}$$

where we have denoted the sample length as T.

Recently it has been shown that it is possible to isolate individual order response functions and a methodology has been developed which may allow the properties of an unknown nonlinear system to be determined [IR92]. There is no intention to replicate that work here, rather we wish to assess the virtues of utilising the three forms of sample statistics into the formalism.

In this example an input sequence $\{x(t)\}$ is used which consists of 20,000 points of heat flux data measured at the external surface of a wall at intervals of one minute. The sequence of output data $\{y(t)\}$ was generated using $\{x(t)\}$ and the convolution equation (24). The response functions used are

$$g_1(\tau_1) = 10.0 \cos (\pi\tau_1/5)\ e^{-0.25\tau_1} \tag{30}$$

and

$$g_2(\tau_1,\tau_2) = e^{-0.5((\tau_1-6)*(\tau_1-6)/16)}\ e^{-0.5((\tau_2-6)(\tau_2-6)/16)} \tag{31}$$

Figure 1 shows a sample of the input data sequence $\{x(t)\}$ and the corresponding output data sequence $\{y(t)\}$. Estimates of the absolute central and product moment values are obtained and a sample of these are given in Figures 2, 3 and 4 respectively.

Figure 5 shows the estimated linear response $h_{yx}(\tau)$ and the differences $(h_{yx}(\tau)-g_y(\tau))$ for the absolute, central and product moment formalisms. Figure 6 shows the estimated quadratic response values $h_{yxx}(\tau_1,\tau_2)$ and the difference surfaces $(h_{yxx}(\tau_1,\tau_2)-g_2(\tau_1,\tau_2))$ for the absolute, central and product moment cases. In Tables 1, 2 and 3 we present the volume under each response function, the root mean square difference and absolute mean square difference between the estimated and known response values, where the volume is given by

$$\sum_{\tau_1=0}^{\mu} h_{yx}(\tau_1) \text{ for the linear term and}$$

$$\sum_{\tau_1=0}^{\mu} \sum_{\tau_2=0}^{\mu} h_{yx}2(\tau_1,\tau_2) \text{ for the quadratic term.}$$

TABLE 1: ABSOLUTE MOMENT CASE

	Estimated Volume	Theoretical Volume	rms difference	Absolute Mean difference
linear response	10.6747628493	10.6747628496	1.83×10^{-7}	1.44×10^{-7}
quadratic response	90.39153107713	90.39153107751	2.58×10^{-7}	2.06×10^{-7}

The normalised statistic between the known and the predicted output sequence for the absolute moment case is

$$\text{nrms} = 1.74 \times 10^{-7}$$

TABLE 2: CENTRAL MOMENT CASE

	Estimated Volume	Theoretical Volume	rms difference	Absolute Mean difference
linear response	10.6747628457	10.6747628496	7.11×10^{-8}	5.92×10^{-8}
quadratic response	90.3915310774	90.3915310775	1.02×10^{-7}	8.19×10^{-8}

The normalised statistic between the known and the predicted output sequences for the central moment case is

$$\text{nrms} = 1.15\text{x}10^{-7}$$

TABLE 3: PRODUCT MOMENT CASE

	Estimated Volume	Theoretical Volume	rms difference	Absolute Mean difference
linear response	10.6747628484	10.6747628496	$8.83\text{x}10^{-8}$	$6.95\text{x}10^{-8}$
quadratic response	90.30153107750	90.39153107751	$1.91\text{x}10^{-7}$	$1.50\text{x}10^{-7}$

The normalised statistic between the known and the predicted output sequences for the product moment case is

$$\text{nrms} = 1.33 \text{ x } 10^{-7}$$

As can be seen the results demonstrate that the three forms of moments used (absolute, central and product) can all correctly identify the form and order of the system response to a high precision. There are apparently no adverse effects due to the form of moment statistic used, in fact the uncertainties observed seem more related to the conditioning of the matrix than the sample length or bias of each form.

The above results demonstrate the precision of the formalism and that if there are any effects due to, for example, bias, or weak nonstationarity they are not evident even though the input data are drawn from a stochastic distribution and the moment expansions are divergent.

Next, consider the effect of additive outliers, for example due to an instrumentation fault. In the next two examples the same input data $\{x(t)\}$ and output $\{y(t)\}$ as in the previous example are used. However in the first case an additive outlier is included in the input sequence $\{x(t)\}$ and in the second case an additive outlier is included in the output sequence $\{y(t)\}$.

First consider the case of an additive outlier in the sequence $\{x(t)\}$. As before the sequence of output data $\{y(t)\}$ was generated using $\{x(t)\}$ and the convolution equation (24), and the response functions given in equations (30) and (31). Having generated the sequences $\{x(t)\}$ and $\{y(t)\}$ the outlier was then added to $\{x(t)\}$. Figure 7 shows a sample of the input data sequence $\{x(t)\}$ with the outlier marked. From these data estimates are obtained of the absolute moments values. First the time, $t=t_1$, at which the outlier occurred must be identified and then we solve equation (24) for an estimate of the amplitude α of the outlier. The value of $x(t_1)$ is then corrected by an amount α and the moments, responses and output sequence are re-estimated using the corrected value. This 'predictor-corrector' procedure may be repeated until the difference of the predicted time series between successive applications of the procedure converges within some predetermined range. Figure 8 shows the output series $\{y(t)\}$ and the predicted output series $\{y_p(t)\}$. The series $\{y_p(t)\}$ was generated using the input data and the estimated response values $h_{yx}(\tau_1)$ and $h_{yxx}(\tau_1,\tau_2)$ before applying the correction and after two applications. That difference indicates that an outlier may be present at x(1153), this value was then adjusted. Figure 9 shows the estimates of the linear response $h_{yx}(\tau)$ and the differences $(h_{yx}(\tau)-g_{yx}(\tau))$ for the uncorrected case and for the case after the second corrective pass. Figure 10 shows the estimates of the quadratic response $h_{yx}(\tau_1,\tau_2)$ values before application of the method and after two applications of the method.

As can be seen in Figure 8 the predicted output improves as the effect of the outlier is eliminated. An improvement in the estimated response values can be seen in Figures 9 and 10 as the effects of the outlier are eliminated.

The results for the uncorrected case are given in Table 7, Tables 8 and 9 present the results from a single and then two applications of the 'predictor-corrector' method. As can be seen the method works well and yields a good estimate of the true response.

TABLE 7: ABSOLUTE MOMENT CASE FOR OUTLIER IN $\{x(t)\}$: NO CORRECTION

	Estimated Volume	Theoretical Volume	rms difference	Absolute Mean difference
linear response	10.1627	10.6747628	0.919	0.414
quadratic response	90.3874	90.391531	1.004	0.806

The normalised statistic between the known and the predicted output sequence is

nrms = 0.8060

TABLE 8: OUTLIER IN $\{x(t)\}$: RESULTS AFTER THE FIRST PASS CORRECTION

	Estimated Volume	Theoretical Volume	rms difference	Absolute Mean difference
linear response	10.6579	10.67476	3.42×10^{-2}	1.50×10^{-2}
quadratic response	90.39118	90.39153	1.42×10^{-3}	1.42×10^{-3}

The normalised statistic between the known and the predicted output sequences is

nrms = 0.0318

TABLE 9: OUTLIER IN $\{x(t)\}$: RESULTS AFTER THE SECOND PASS CORRECTION

	Estimated Volume	Theoretical Volume	rms difference	Absolute Mean difference
linear response	10.6684	10.67476	1.328×10^{-2}	9.775×10^{-3}
quadratic response	90.39139	90.39153	1.067×10^{-3}	8.362×10^{-4}

The normalised statistic between the known and the predicted output sequences is

$$\text{nrms} = 0.0186$$

In the second example an additive outlier in the output data $\{y(t)\}$ is considered. Again the heat flux data is used as the input data $\{x(t)\}$ and $\{y(t)\}$ is generated using $\{x(t)\}$ convoluted with $g_1(\tau_1)$ and $g_2(\tau_1,\tau_2)$ as defined in equations (30) and (31) with the outlier added at $y(1153)$. The position of the outlier is again identified from the difference set $\{y(t) - y_p(t)\}$, where $\{y_p(t)\}$ is the data predicted using $\{x(t)\}$ and the estimated response values $h_{yx}(\tau_1)$ and $h_{yxx}(\tau_1,\tau_2)$. The difference $(y(1153)-y_p(1153))$ yields an estimate of the amplitude α, and this value is subtracted from $y(1153)$. The moments and response are re-evaluated and a 'predictor-corrector' sequence may be continued until the test statistic nrms falls within a predetermined range.

In Figure 11 are shown sample plots of $y_p(t)$ and $y(t)$ before any correction is applied and after a single application of the predictor-corrector. Figure 12 shows the estimated linear response values $h_{yx}(\tau_1)$ and the differences $(h_{yx}(\tau_1)-g_1(\tau_1))$ and Figure 13 shows the estimated quadratic response surfaces $h_{yxx}(\tau_1,\tau_2)$ before any correction is applied and after a single application of the predictor-corrector. The results for these two cases are presented in Tables 10 and 11 and as can be seen significant improvements in the estimated response values are achieved with only a single application of the method.

TABLE 10: ABSOLUTE MOMENT CASE FOR OUTLIER IN $\{y(t)\}$: NO CORRECTION

	Estimated Volume	Theoretical Volume	rms difference	Absolute Mean difference
linear response	10.723	10.674762	0.408	0.334
quadratic response	90.392	90.391531	8.67×10^{-3}	6.657×10^{-3}

The normalised statistic between the known and the predicted output sequence is

nrms = 0.2015

TABLE 11: OUTLIER IN $\{y(t)\}$: RESULTS AFTER THE FIRST PASS CORRECTION

	Estimated Volume	Theoretical Volume	rms difference	Absolute Mean difference
linear response	10.67478	10.674762	1.977×10^{-4}	1.617×10^{-4}
quadratic response	90.3915315	90.3915311	4.237×10^{-6}	3.245×10^{-4}

The normalised statistic between the known and the predicted output sequences is

nrms = 9.76×10^{-5}

CONCLUSIONS

In this work a formalism [IR92], which can identify the order and form of nonlinear systems, has been used to test if the formulation of the moment estimators is an important factor in the analysis of an unknown process. Three forms of time series moment estimator were used (absolute, central and product) in the formalism and no significant effects due to bias were observed. The accuracy seemed more related to the conditioning of the matrix used in the formalism than to the statistical accuracy of the moments; although when the moments were ill defined the response estimates deteriorated. It is also noted that there is no advantage in using cumulant estimates in the formalism. An absolute moment form was then used to illustrate the use of a simple 'predictor-corrector' method for additive outliers that may be present in the input $\{x(t)\}$ or output $\{y(t)\}$ data sequences. A methodology for the identification and correction of additive outliers was demonstrated and may be used when the set of outliers forms a small fraction of the time series.

ACKNOWLEDGEMENTS

This work was funded by the Building Sub-Committee of the UK Science and Engineering Research Council. The authors are pleased to acknowledge the support of Dr B Day in this project. The authors would like to express their thanks to the Energy Monitoring Company who collected the heat flux data and to the Energy Support Technology Unit of the Department of Energy (UK) for funding the data collection.

RW20

REFERENCES

BA90 Narrett V and Lewis T, Outliers in Statistical Data, John Wiley, Chichester, 1990.

IR92 Irving A D, Dewson T and Hong G, Nonlinear Response Function Estimation: I Inhomogeneous Case, Submitted to J. Phys, A, 1992; see also Rutherford Appleton Laboratory Report RAL-91-067, 1991.

RO85 Rosenblatt M, Stationary Sequences and Random Fields, Birkhauser, Boston, 1985.

VO00 Volterra V, Singli Integrali Lineari dei Moti spontanei a caratheristiche indipendenti, Atti Torino, 35, 1900, p186-192.

VO59 Volterra V, Theory of Functionals and of Integral and Integro-differential Equations, Dover, New York, 1959.

WI42 Wiener N, Response of a Nonlinear Device to Noise, MIT Radiation Laboratory Report No 165, 1942.

WI58 Wiener N, Nonlinear Problems in Random Theory, John Wiley, New York, 1958.

August 1991
updated February 1992
RW20

FIGURE CAPTIONS

1 Sample of a) the heat flux data {x(t)} Wm^{-2} K^{-1} and b) the output data {y(t)} arbitrary units.

2a Third order auto moment of the input data using the absolute moment formalism.

2b Third order cross moment between the input and output data, using the absolute moment formalism.

3a Third order auto moment of the input data using the central moment formalism.

3b Third order cross moment between the input and output data, using the central moment formalism.

4a Third order auto moment of the input data using the product moment formalism.

4b Third order cross moment between the input and output data, using the product moment formalism.

5 a) Estimated first order response using the absolute moment formalism and the differences between the estimated and known first order response using the b) absolute c) central and d) product moment formalisms.

6 a) Estimated second order response using the absolute moment formalism and the difference surfaces between the known and estimated second order response using the b) absolute c) central and d) the product moment formalisms.

7 Sample of a) the heat flux data {x(t)} with an additive outlier and b) the output data {y(t)} arbitrary units.

8 Sample of the output data $\{y(t)\}$ and the output data $\{y_p(t)\}$ predicted using the input data $\{x(t)\}$ and the estimated response values $h_{yx}(\tau_1)$ and $h_{yxx}(\tau_1,\tau_2)$ a) before applying any correction and b) after two applications of the predictor-corrector.

9 The estimated response $h_{yx}(\tau_1)$ and the difference surface $(h_{yx}(\tau_1)-g_1(\tau_1))$ for a) before application of the corrector and b) after the second application of the predictor-corrector.

10 The estimated response surface $h_{yxx}(\tau_1,\tau_2)$ a) before application of the corrector and b) after two applications of the predictor-corrector.

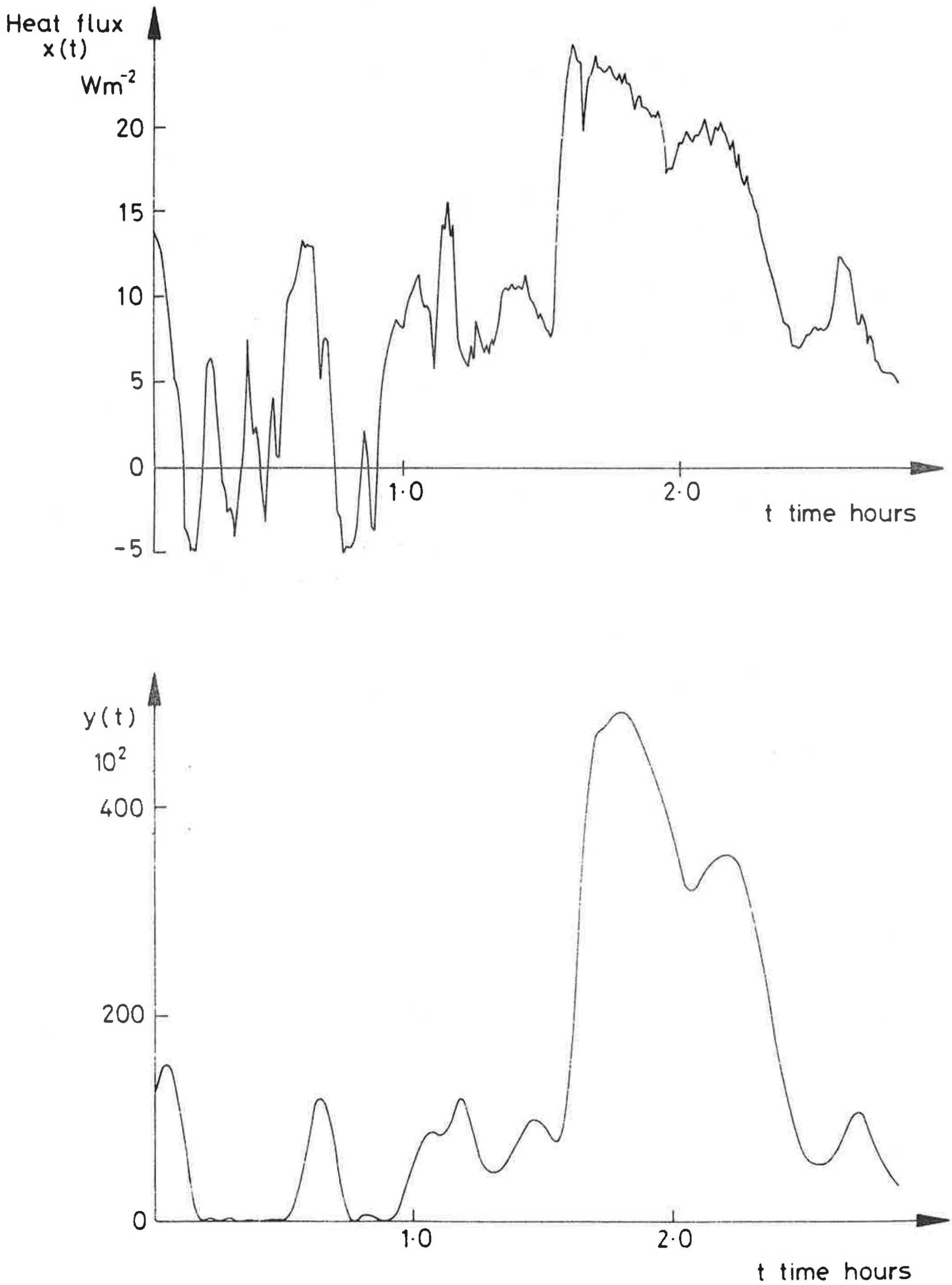
11 Plot of the predicted output $y_p(t)$ and the observed output $y(t)$ in the region of the outlier for a) the uncorrected case and b) after a single application of the predictor-corrector.

12 The estimated response $h_{yx}(\tau_1)$ and the differences $(h_{yxx}(\tau_1)-g_2(\tau_1))$ for a) the uncorrected case and b) after a single application of the predictor-corrector.

13 The estimated response surface $h_{yxx}(\tau_1,\tau_2)$ a) before application of the corrector and b) after a single application of the predictor-corrector.

RW20/ADI2

Figure 1.



Heat flux
x(t)
Wm^-2
20
15
10
5
0
-5
1·0
2·0
t time hours
y(t)
10^2
400
200
0
1·0
2·0
t time hours

Figure 2a.

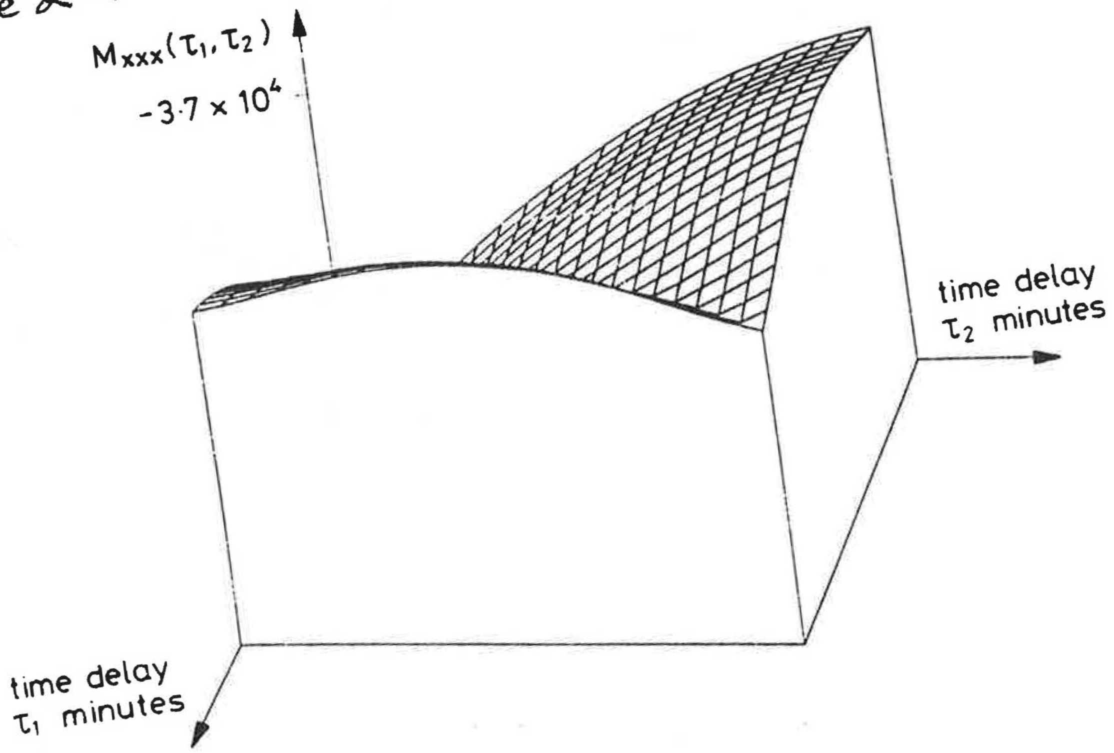


Figure 2b.

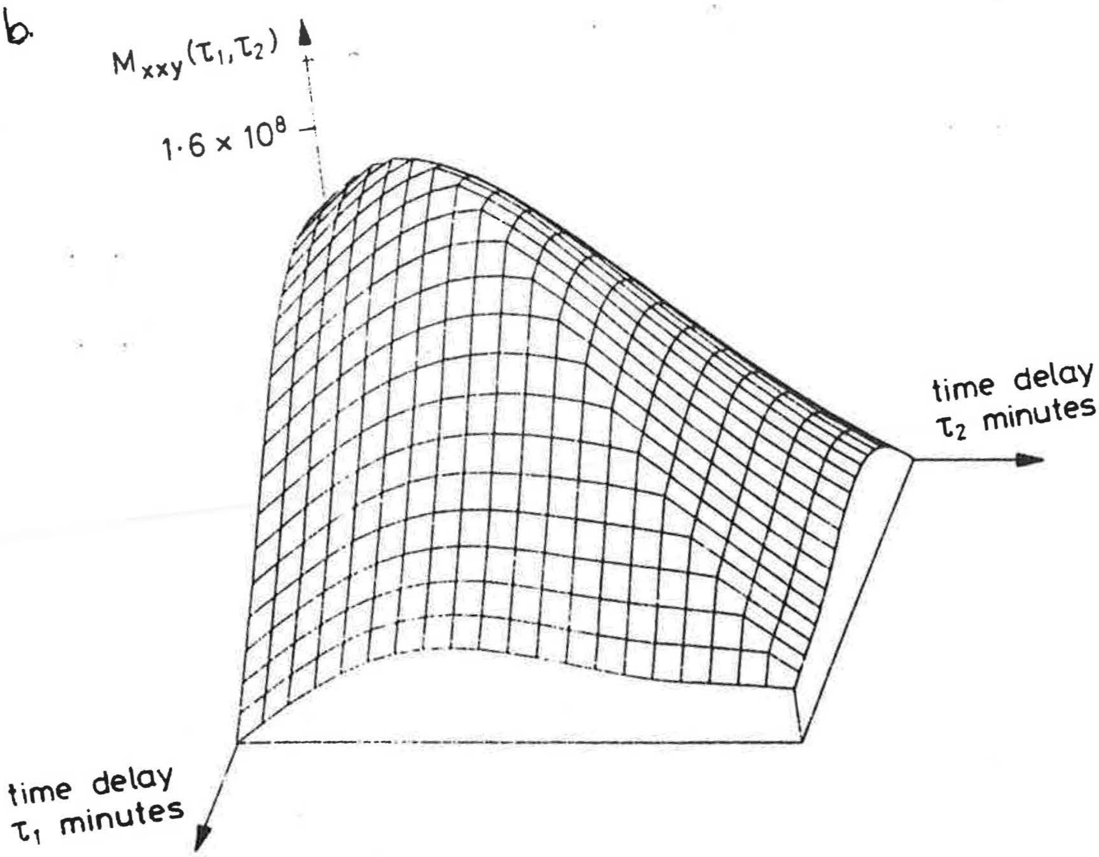


Figure 3a.

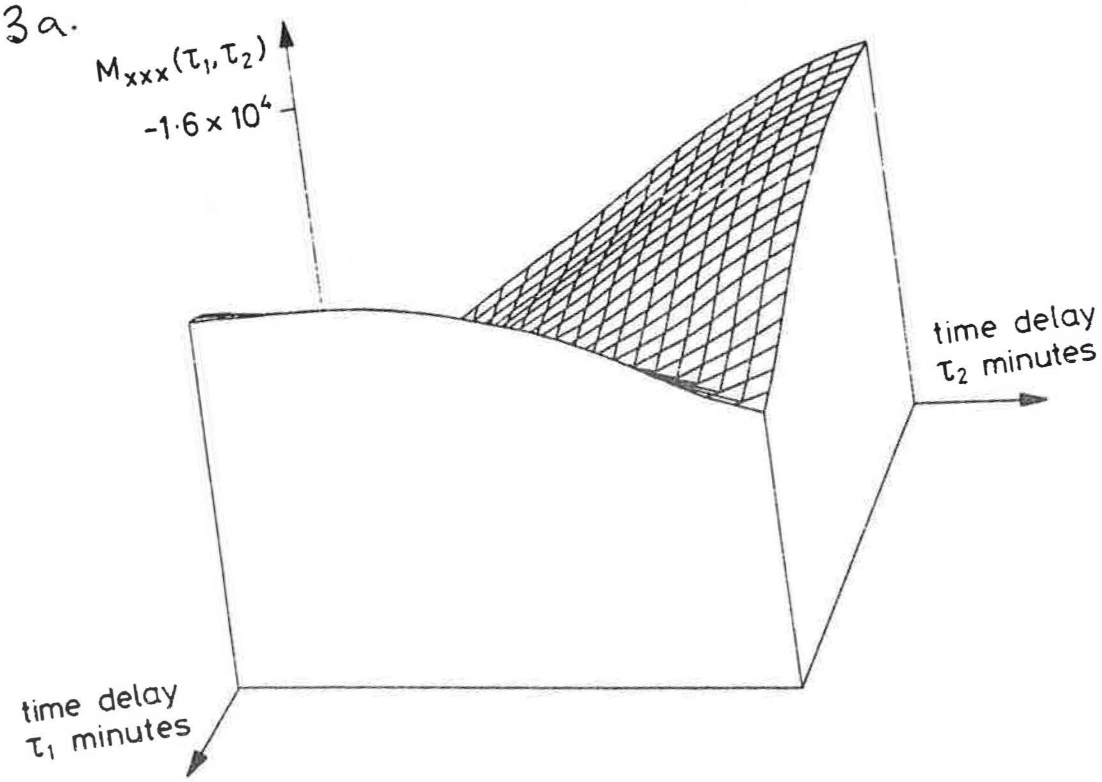


Figure 3b.

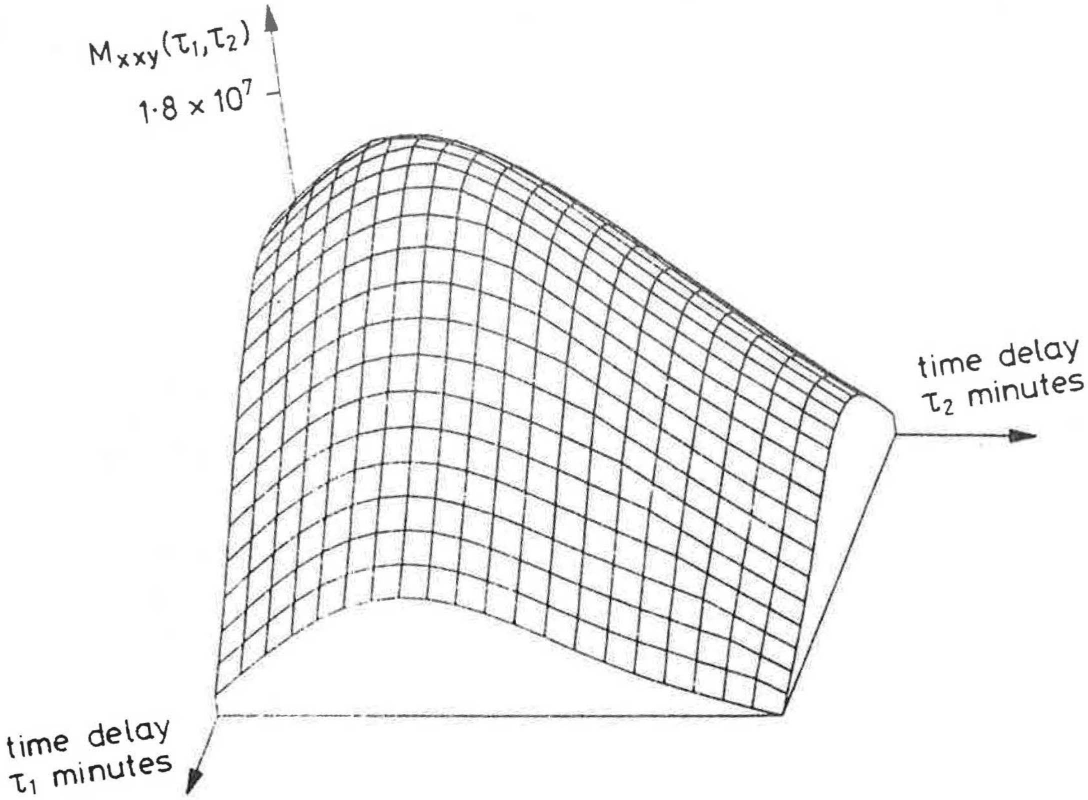


FIGURE 4a

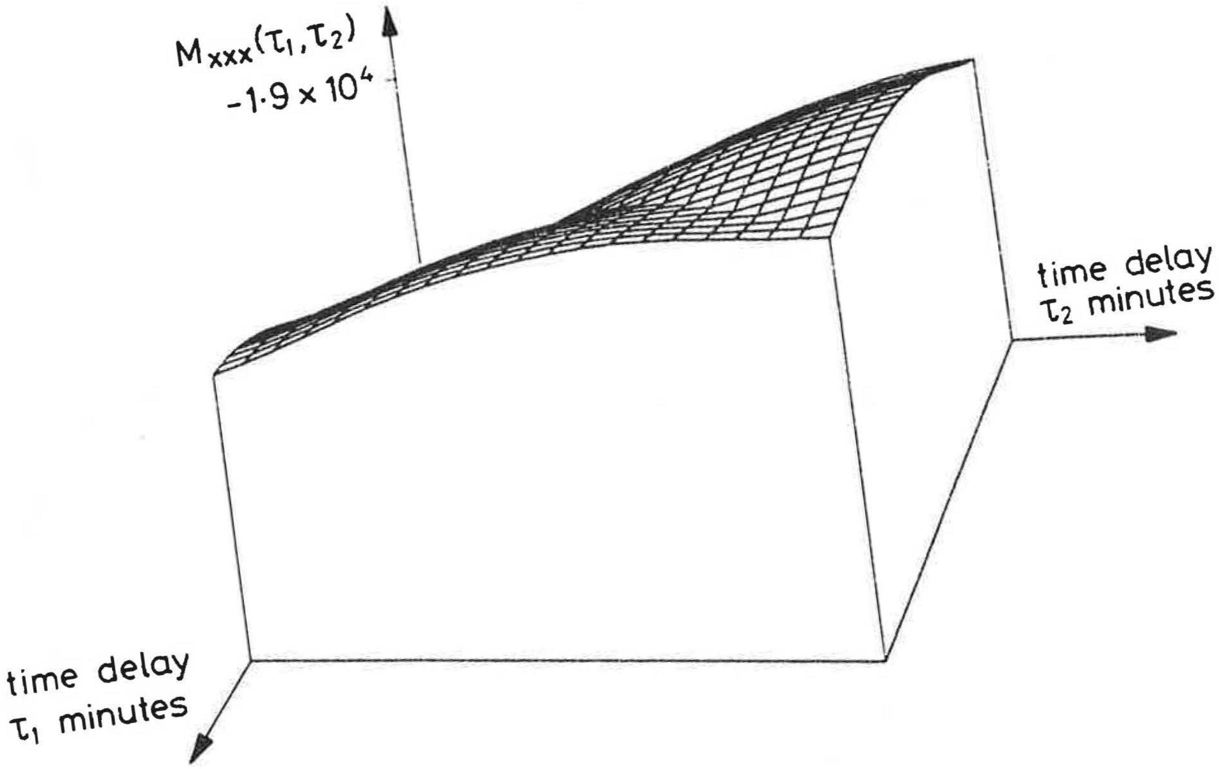


FIGURE 4b

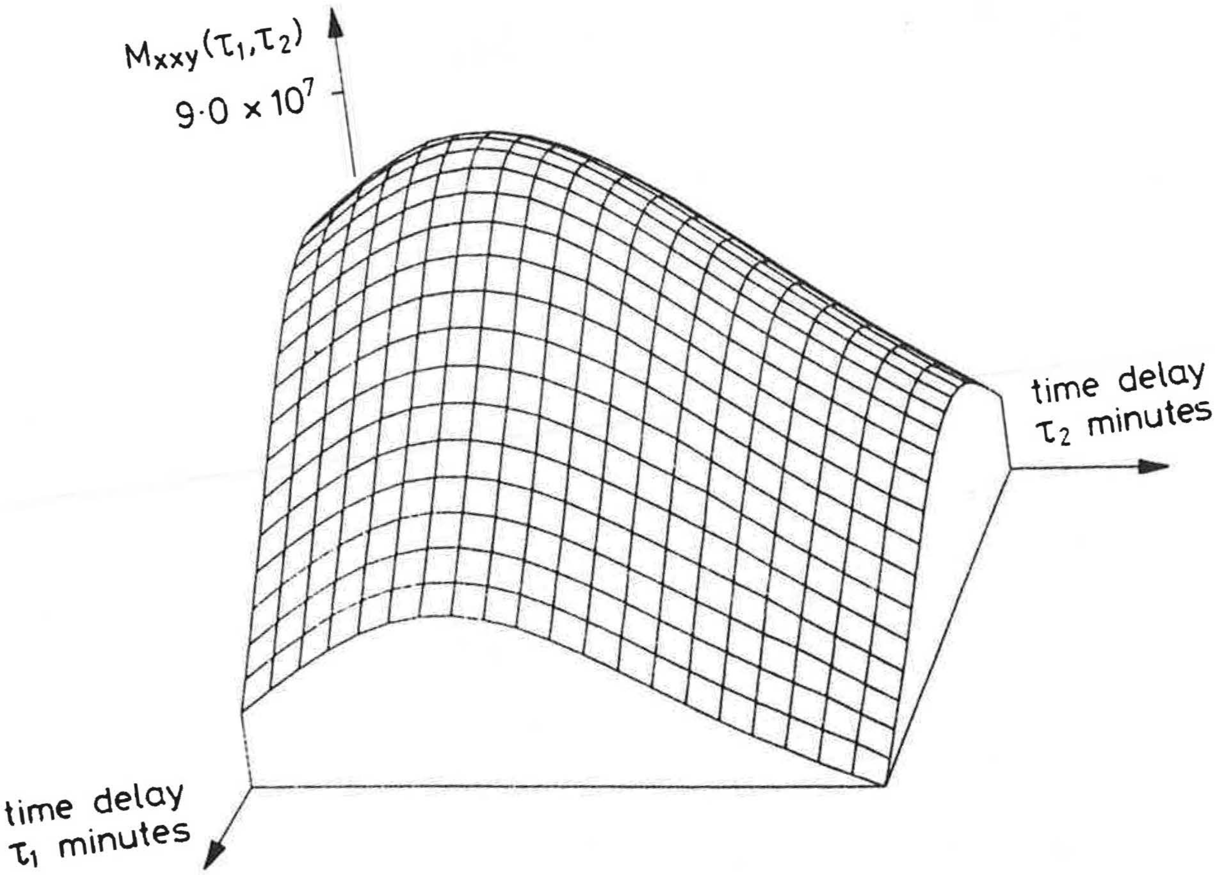


Figure 5a

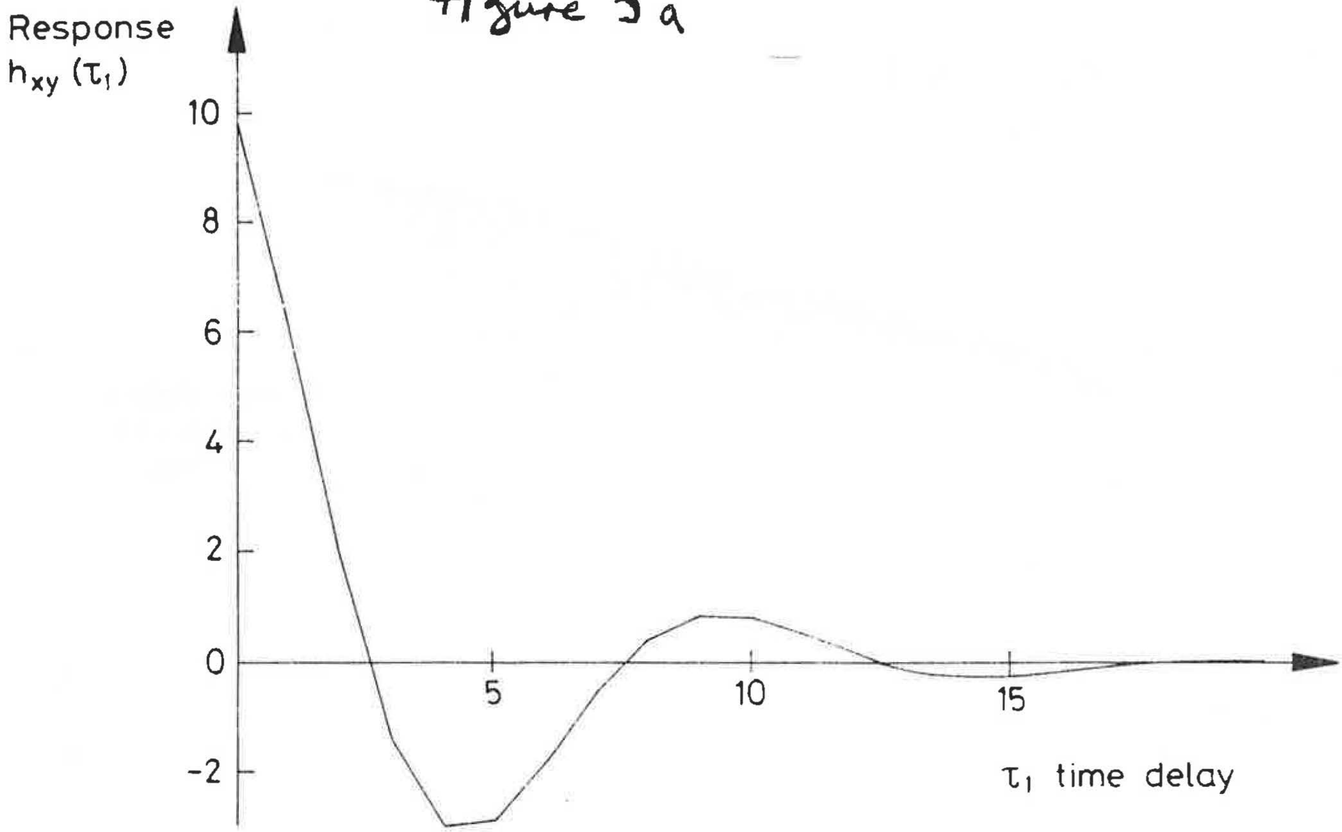


Figure 5b

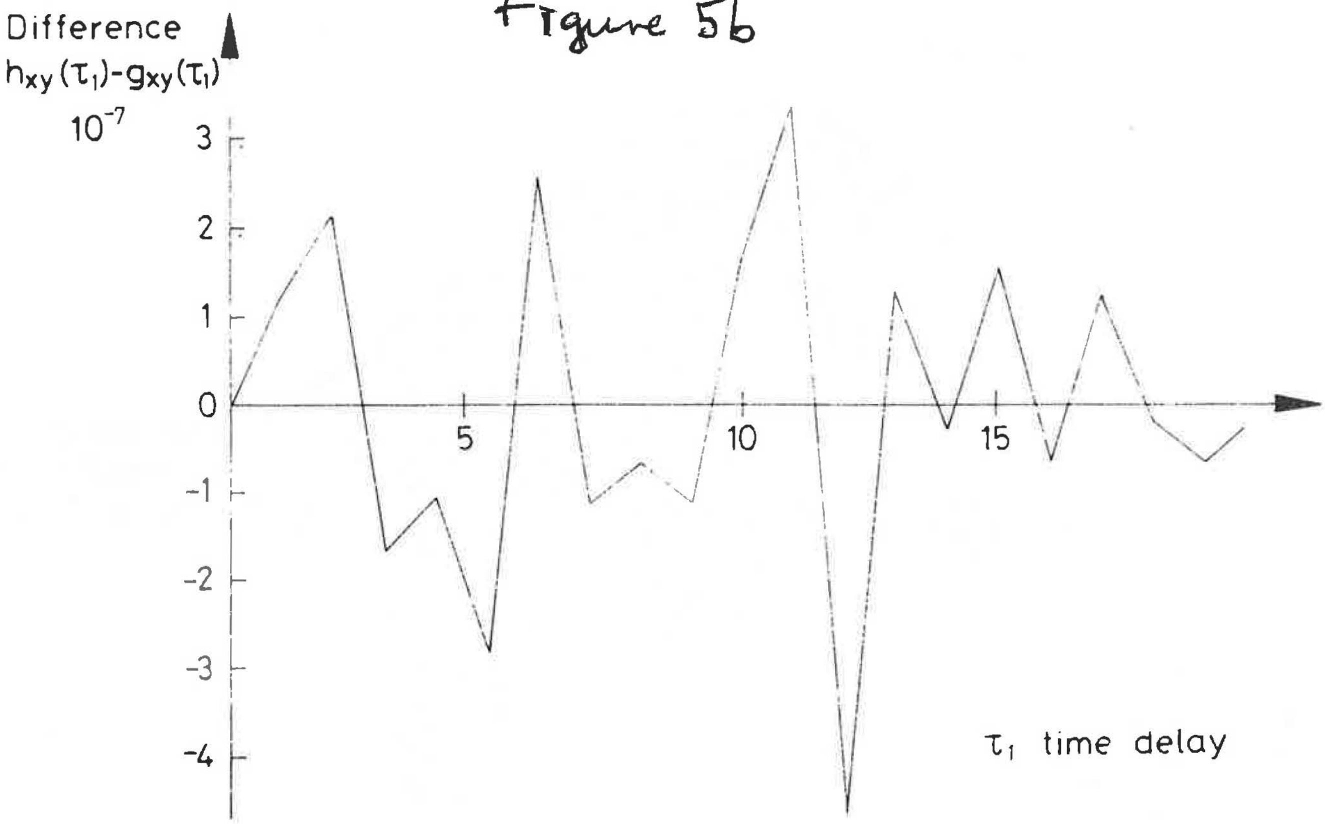


FIGURE 5c

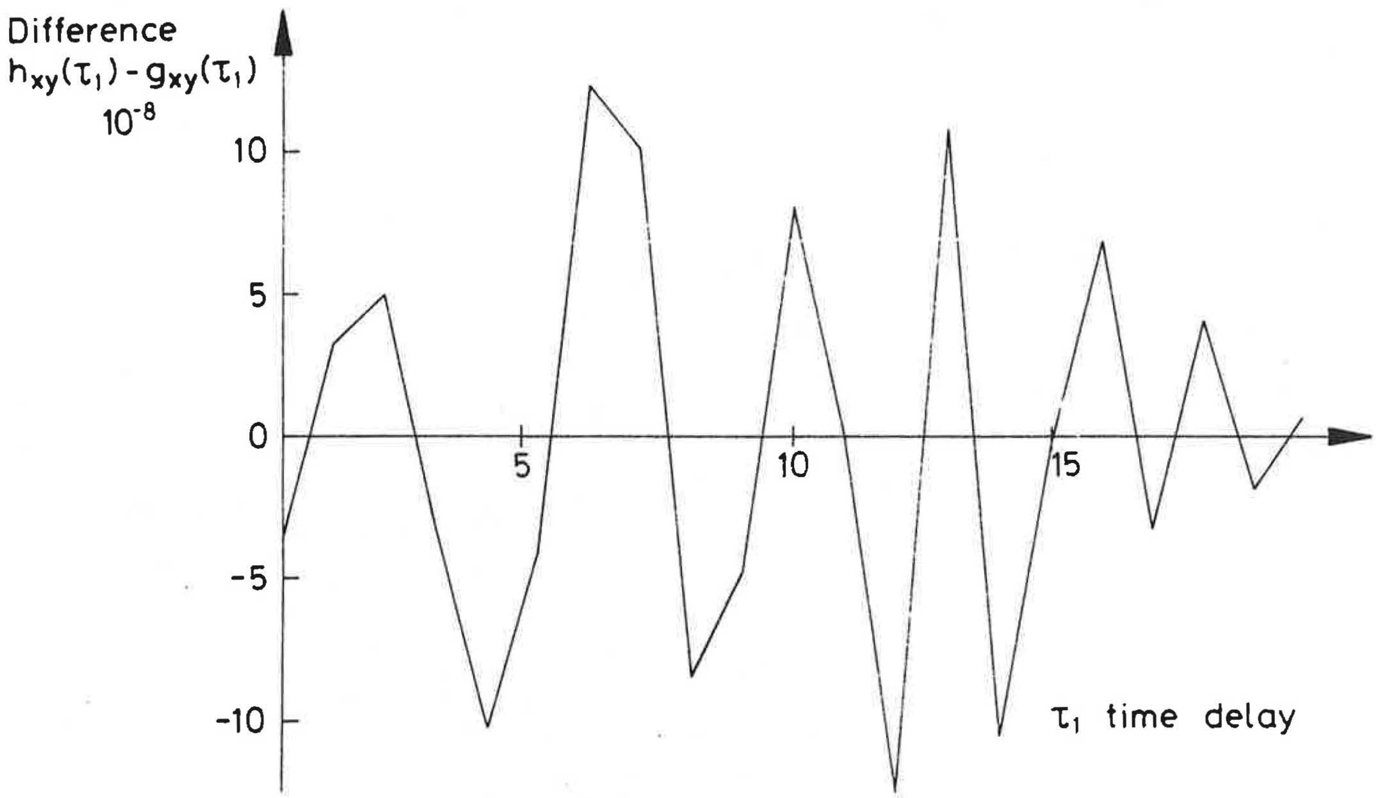


FIGURE 5d

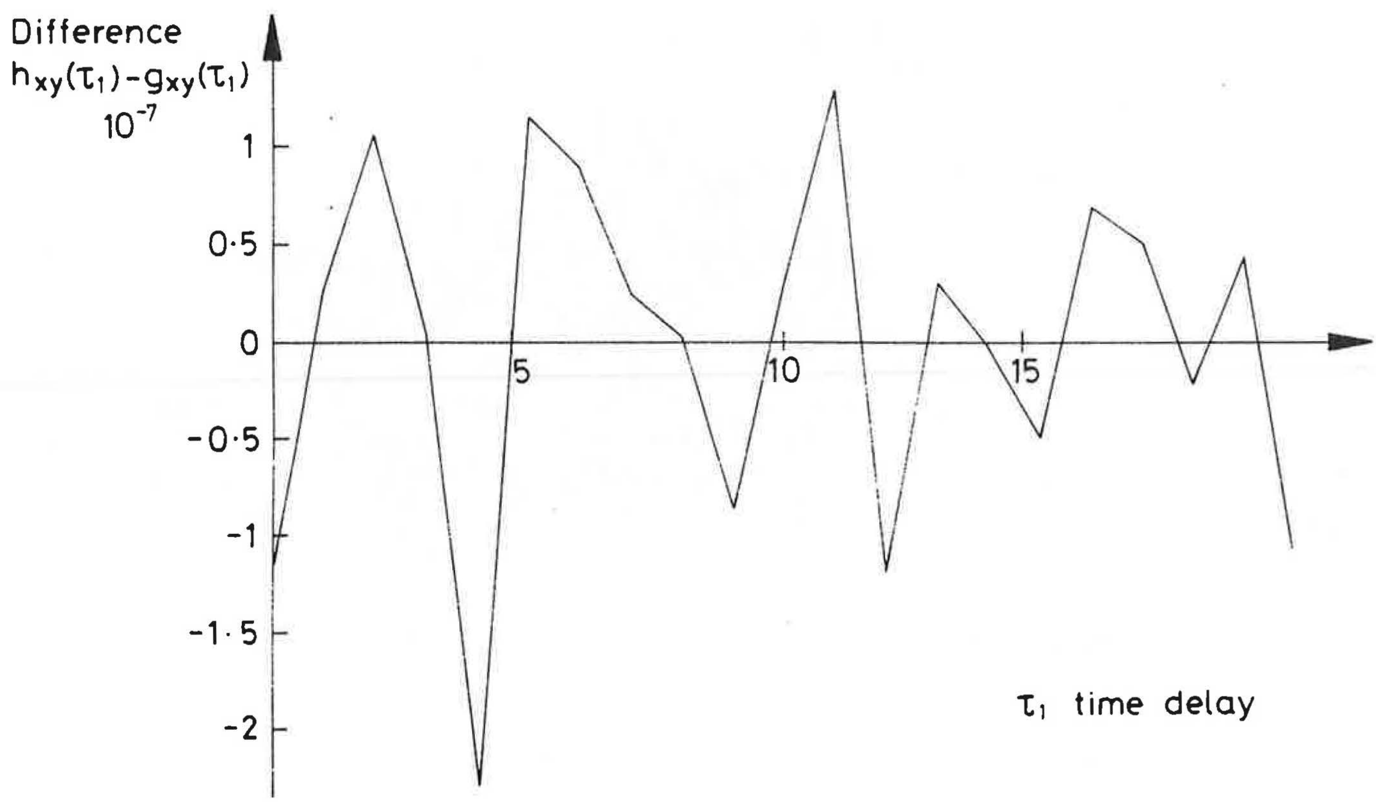


Figure 6a

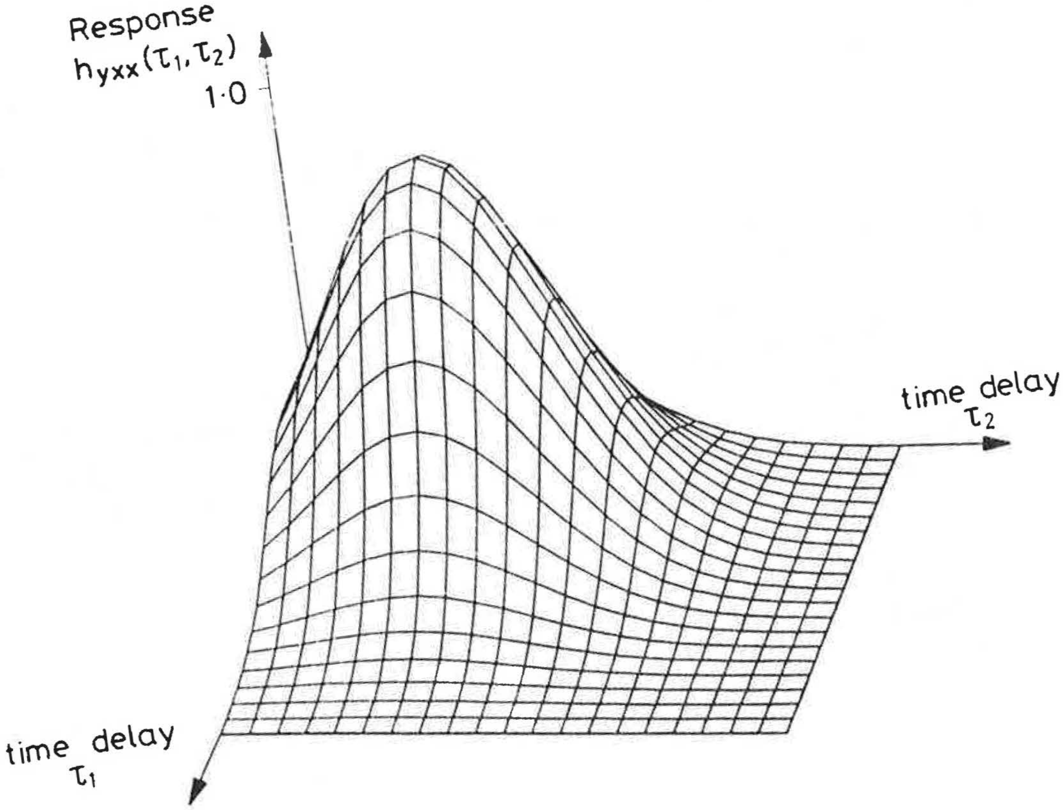
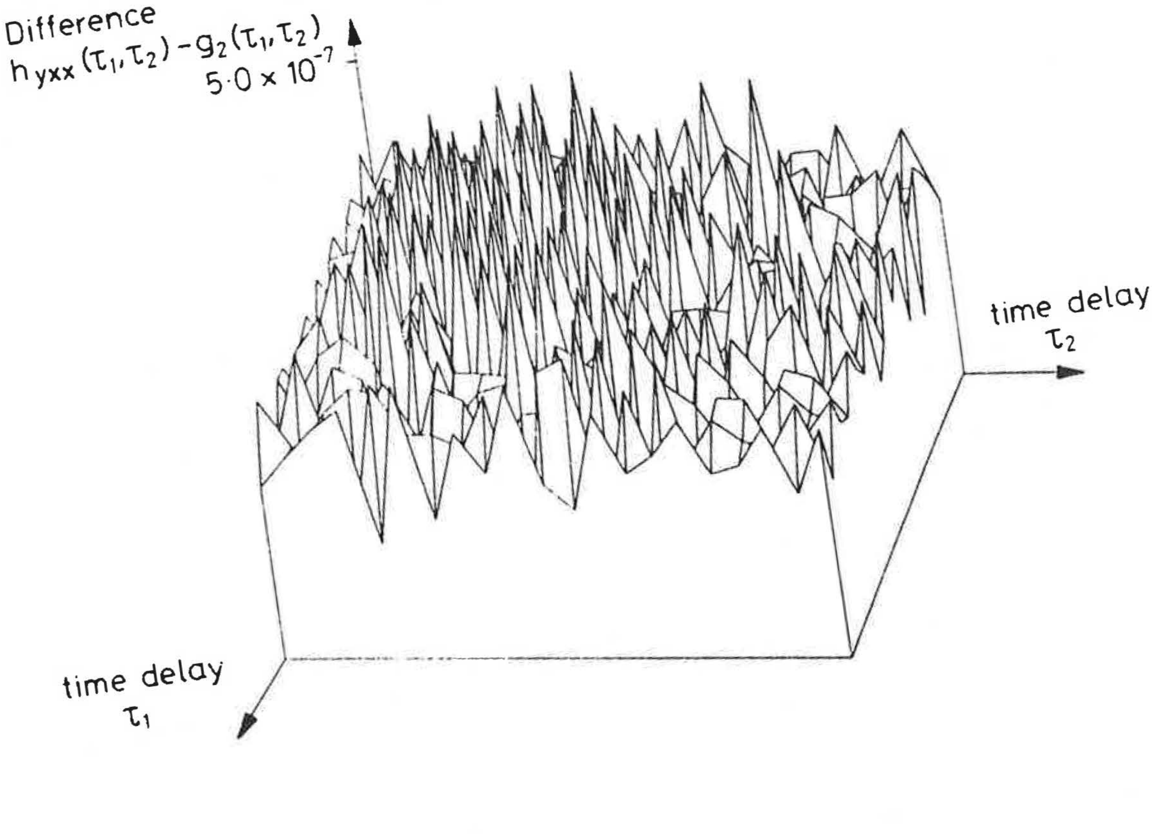


Figure 6b.



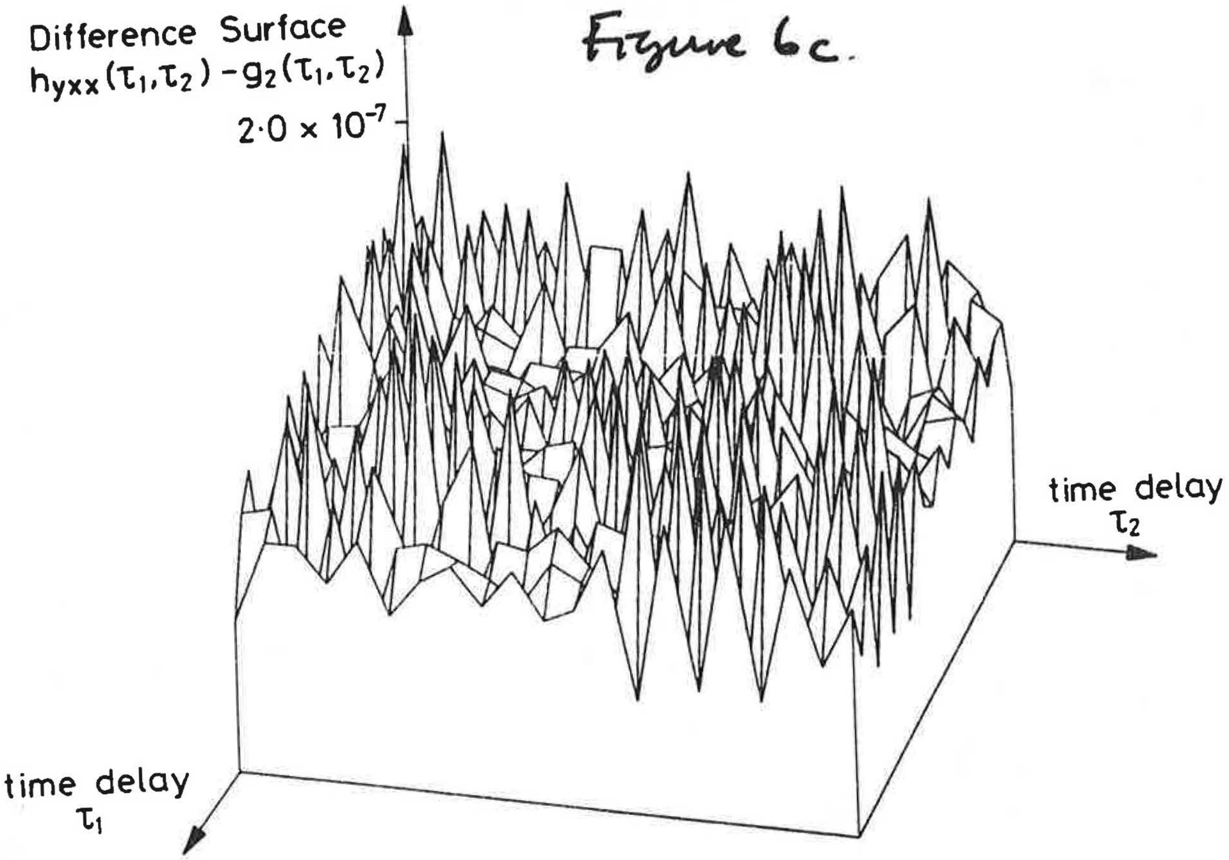


Figure 6c.

Figure 6d.

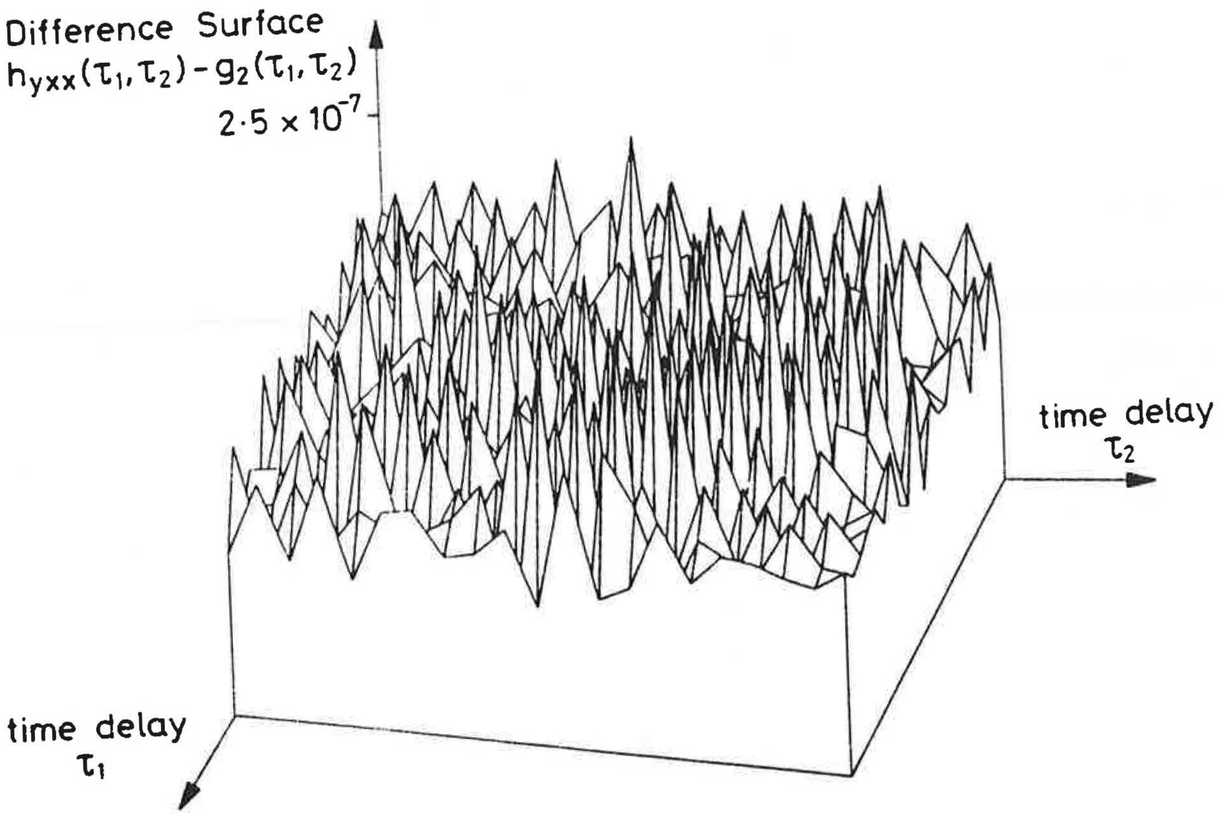


Figure 7

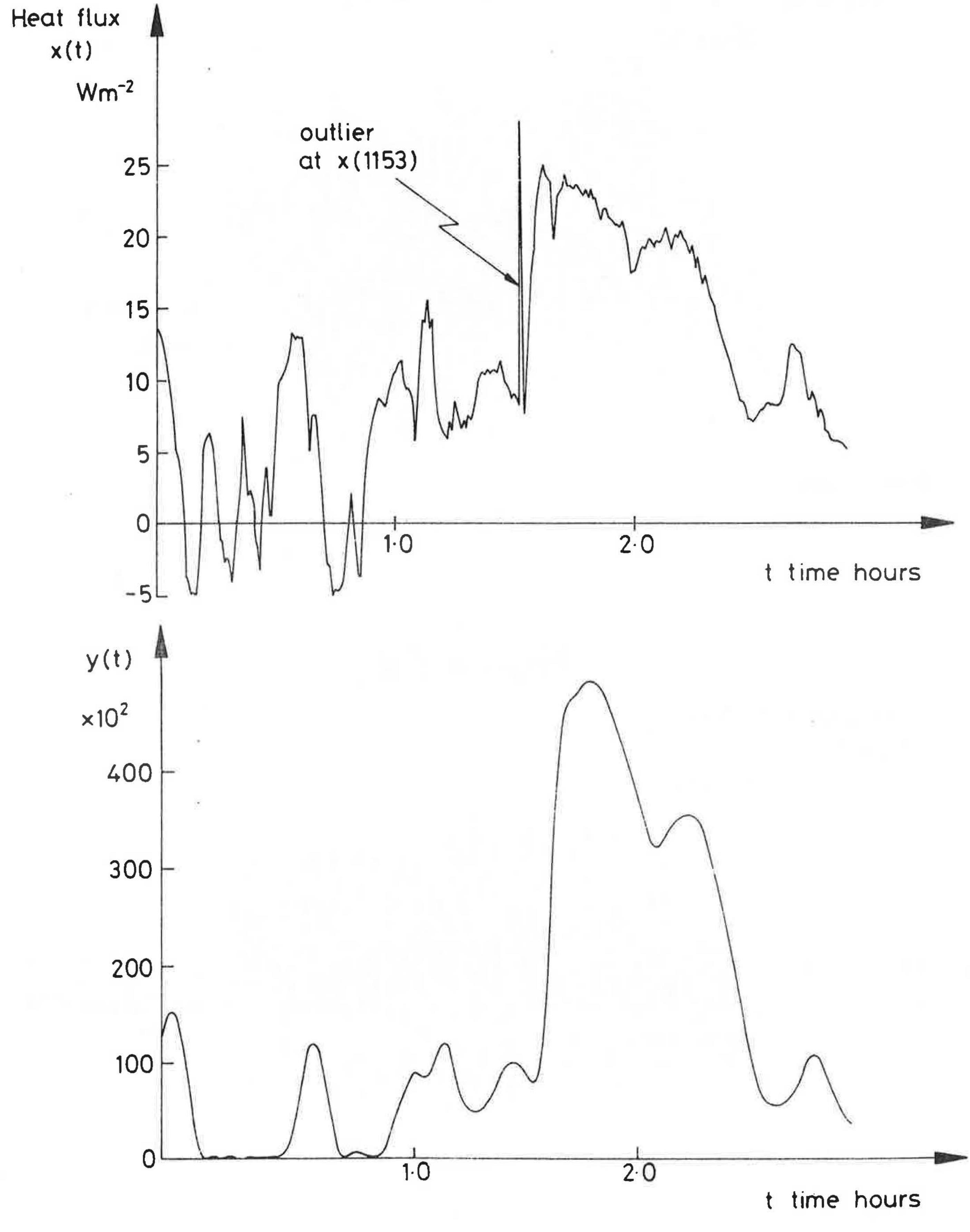

Heat flux
x(t)
Wm-2
outlier
at x(1153)
25
20
15
10
5
0
-5
1·0
2·0
t time hours
y(t)
×10²
400
300
200
100
0
1·0
2·0
t time hours

Figure 8a.

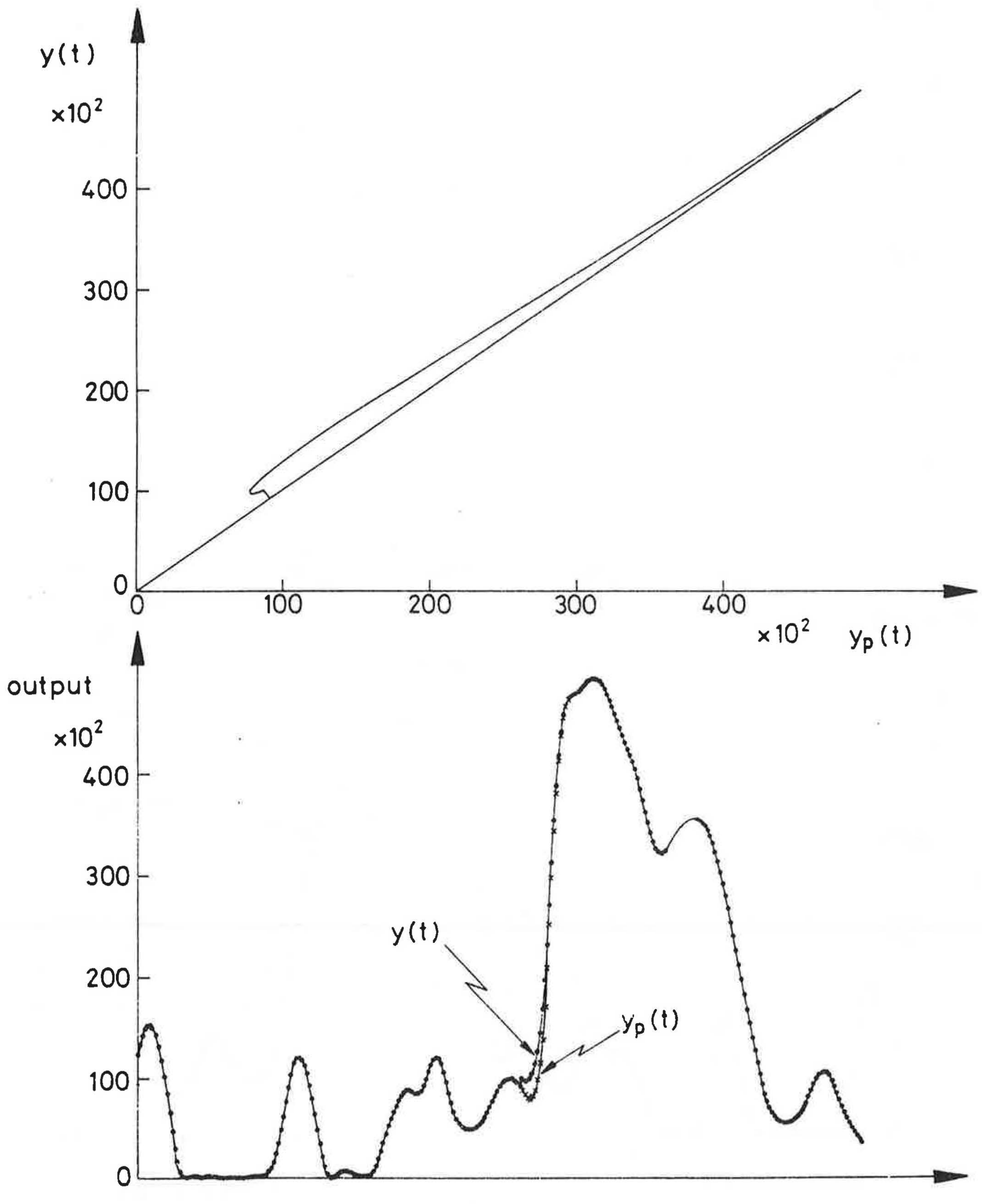

y(t)
×10²
400
300
200
100
0
0
100
200
300
400
×10²
$y_p(t)$
output
×10²
400
300
200
100
0
y(t)
$y_p(t)$

Figure 8b.

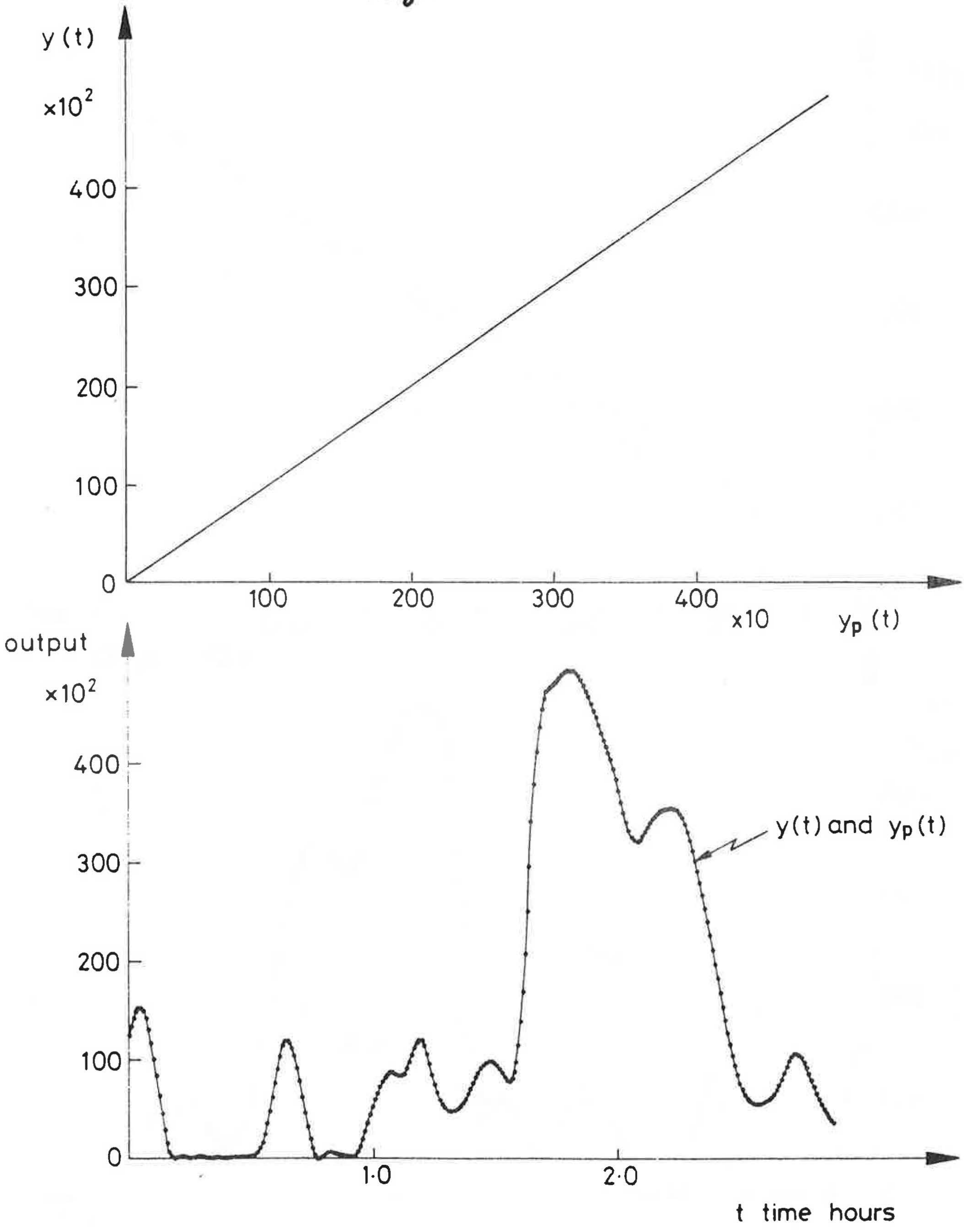
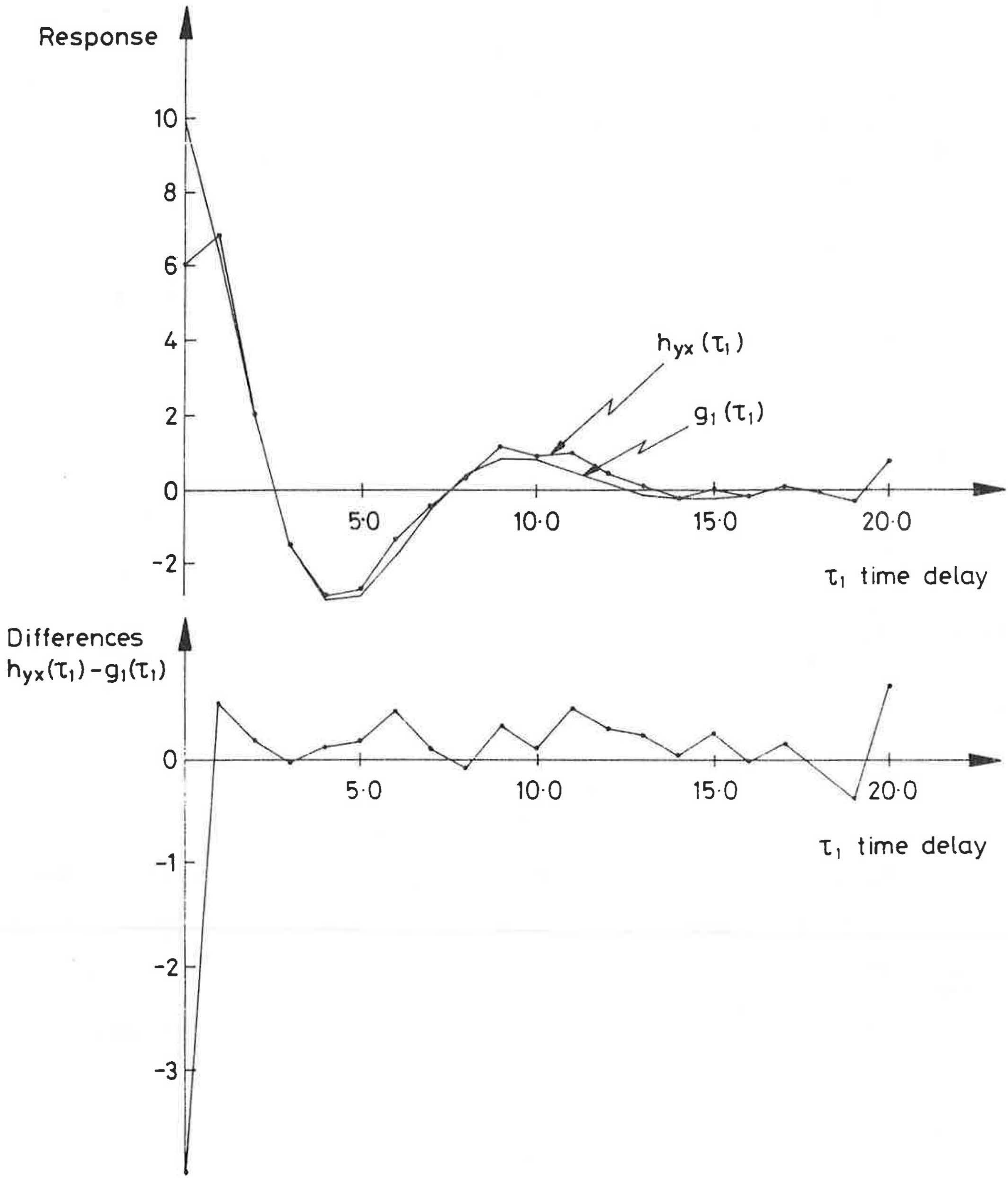
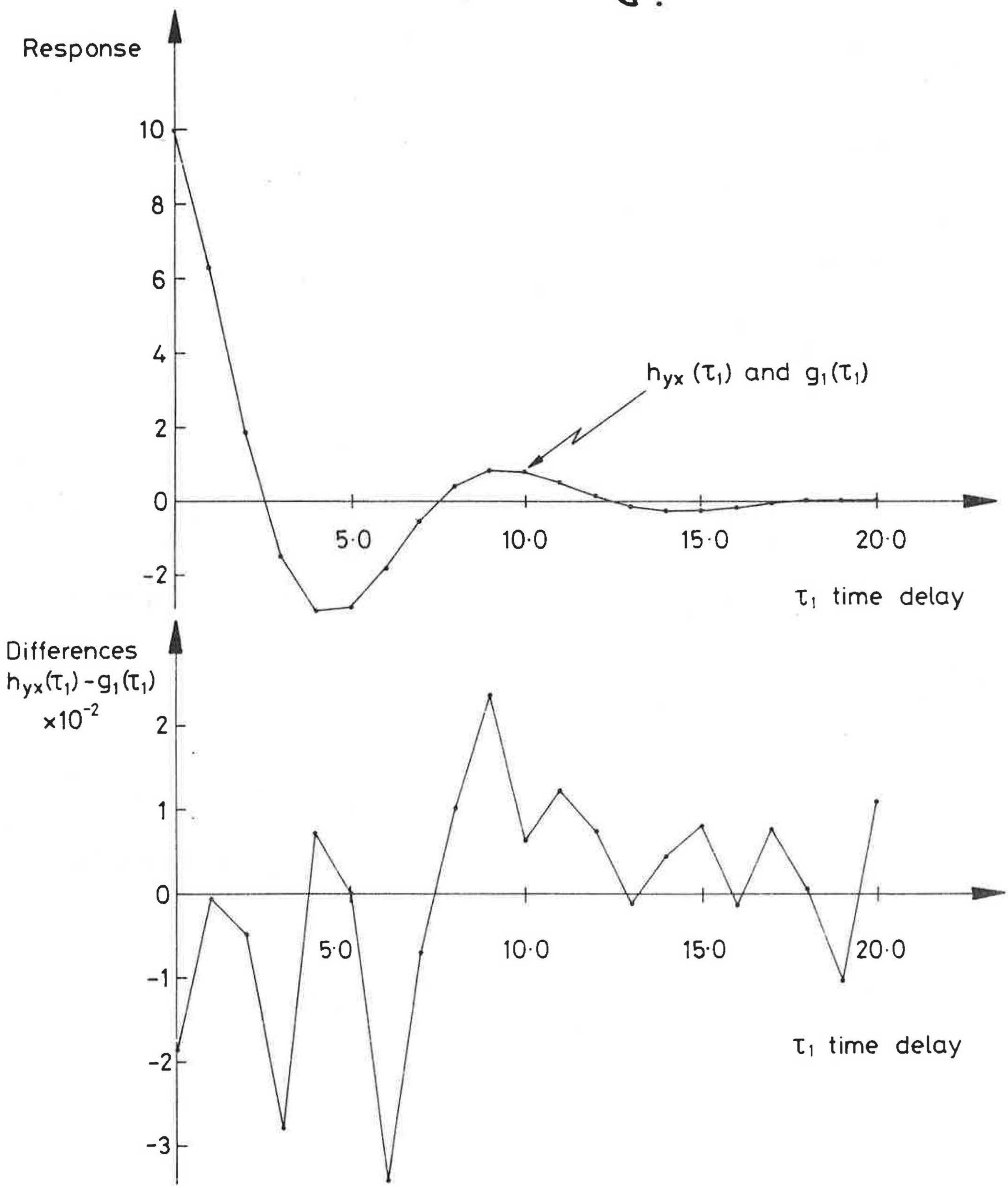

y (t)
x10²
400
300
200
100
0
100
200
300
400
x10
yp (t)
output
x10²
400
300
200
100
0
1·0
2·0
y(t) and yp(t)
t time hours

Figure 9 a.



Response
10
8
6
4
2
0
-2
5·0
10·0
15·0
20·0
$h_{yx}(\tau_1)$
$g_1(\tau_1)$
τ_1 time delay
Differences
$h_{yx}(\tau_1) - g_1(\tau_1)$
0
-1
-2
-3
5·0
10·0
15·0
20·0
τ_1 time delay

Figure 9b.



Response
10
8
6
4
2
0
-2
$h_{yx}(\tau_1)$ and $g_1(\tau_1)$
5·0
10·0
15·0
20·0
τ_1 time delay
Differences
$h_{yx}(\tau_1) - g_1(\tau_1)$
$\times 10^{-2}$
2
1
0
-1
-2
-3
5·0
10·0
15·0
20·0
τ_1 time delay

Figure 10a

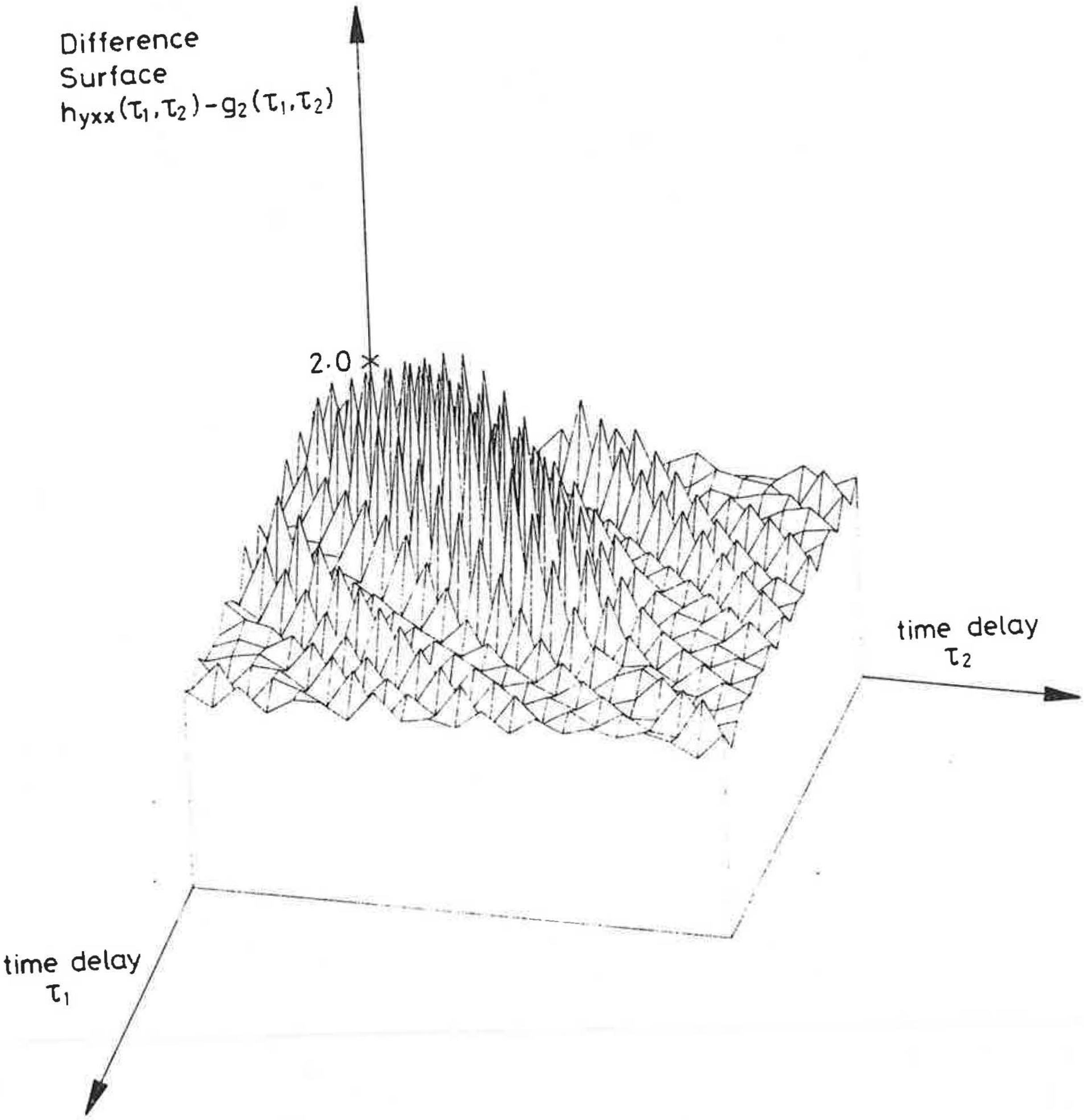

Difference
Surface
$h_{yxx}(\tau_1,\tau_2)-g_2(\tau_1,\tau_2)$
2.0
time delay
τ_2
time delay
τ_1

Figure 10b.

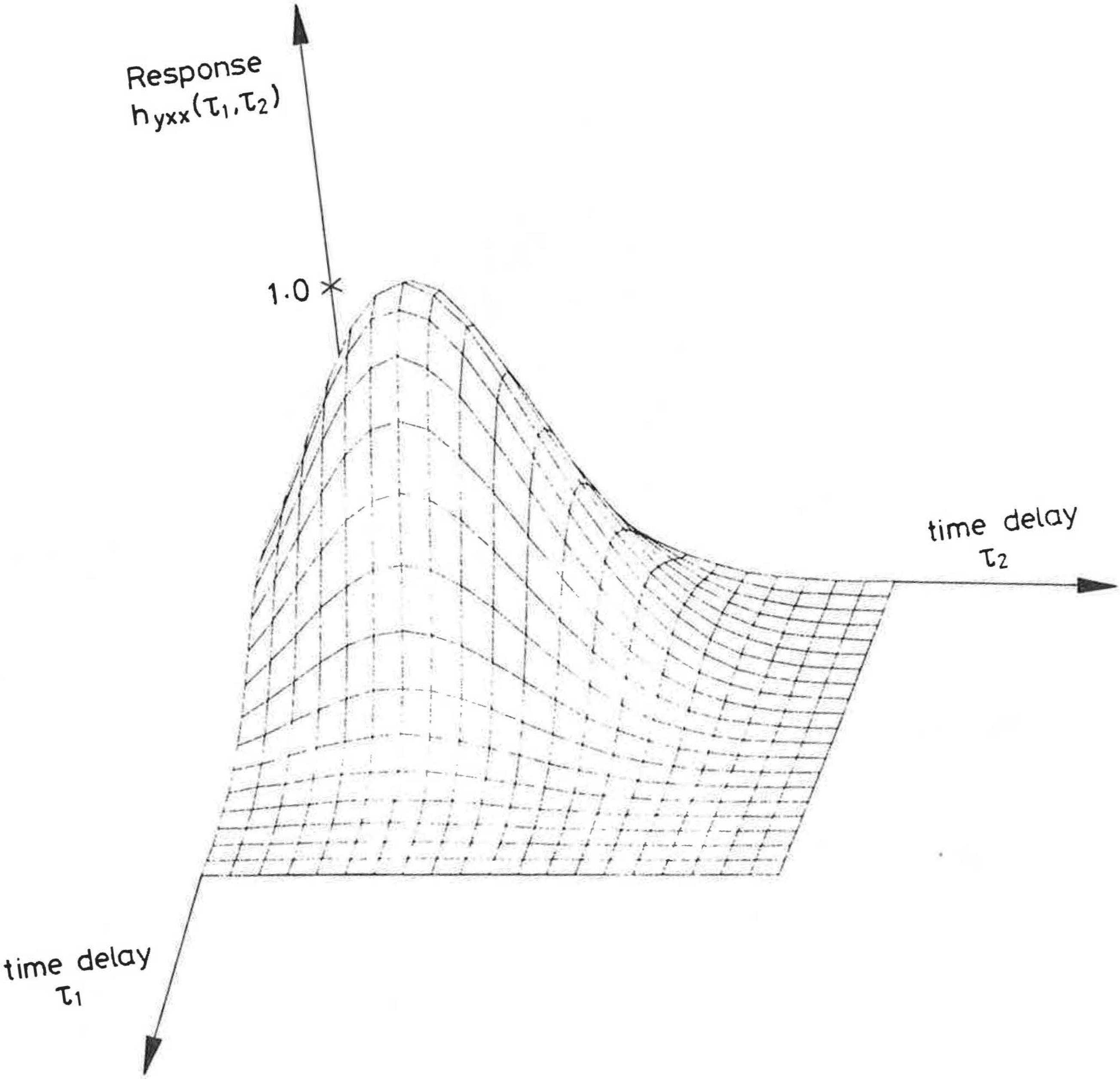
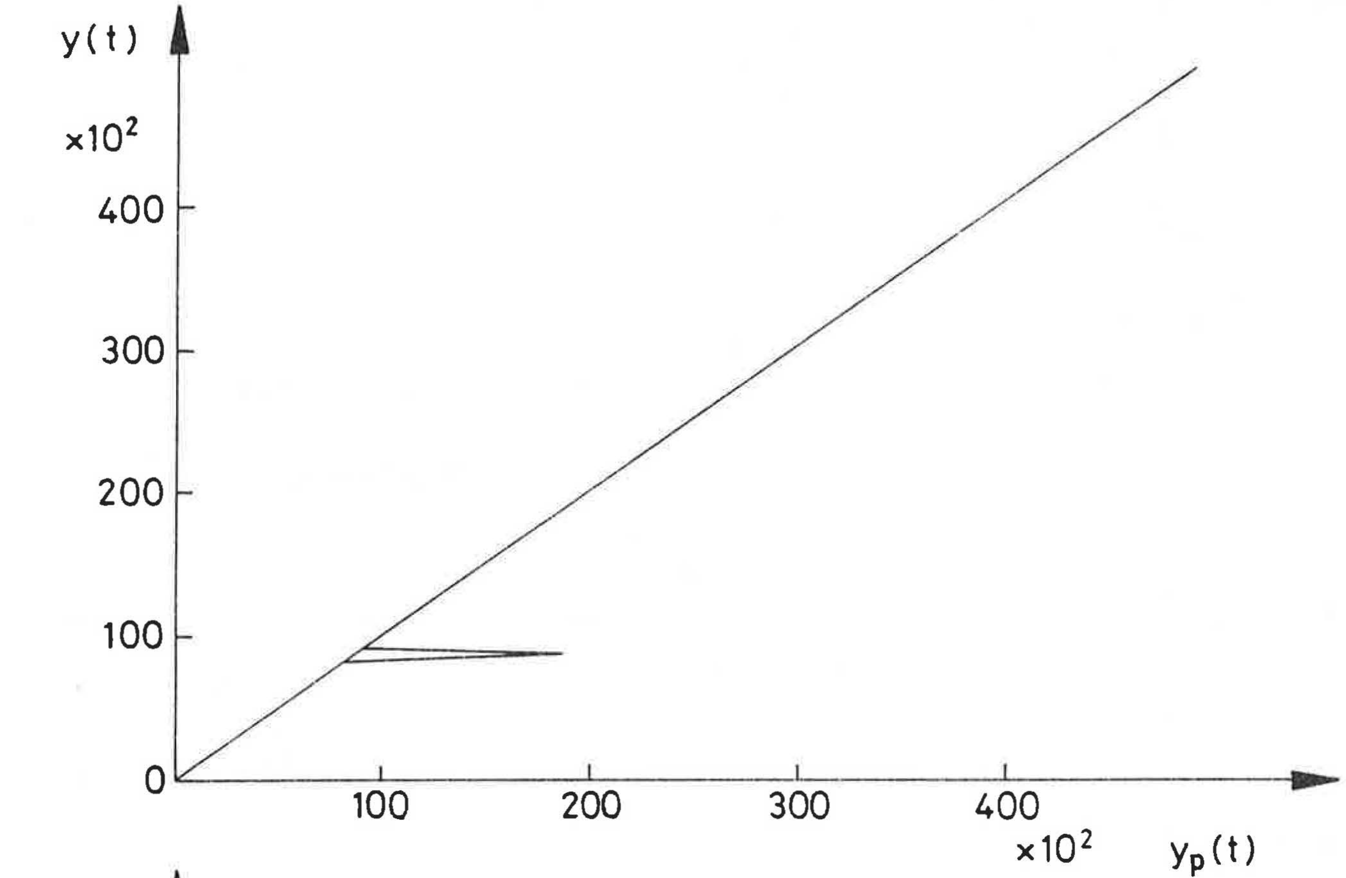


Figure 11.


y(t)
x10²
400
300
200
100
0
100
200
300
400
x10²
yp(t)

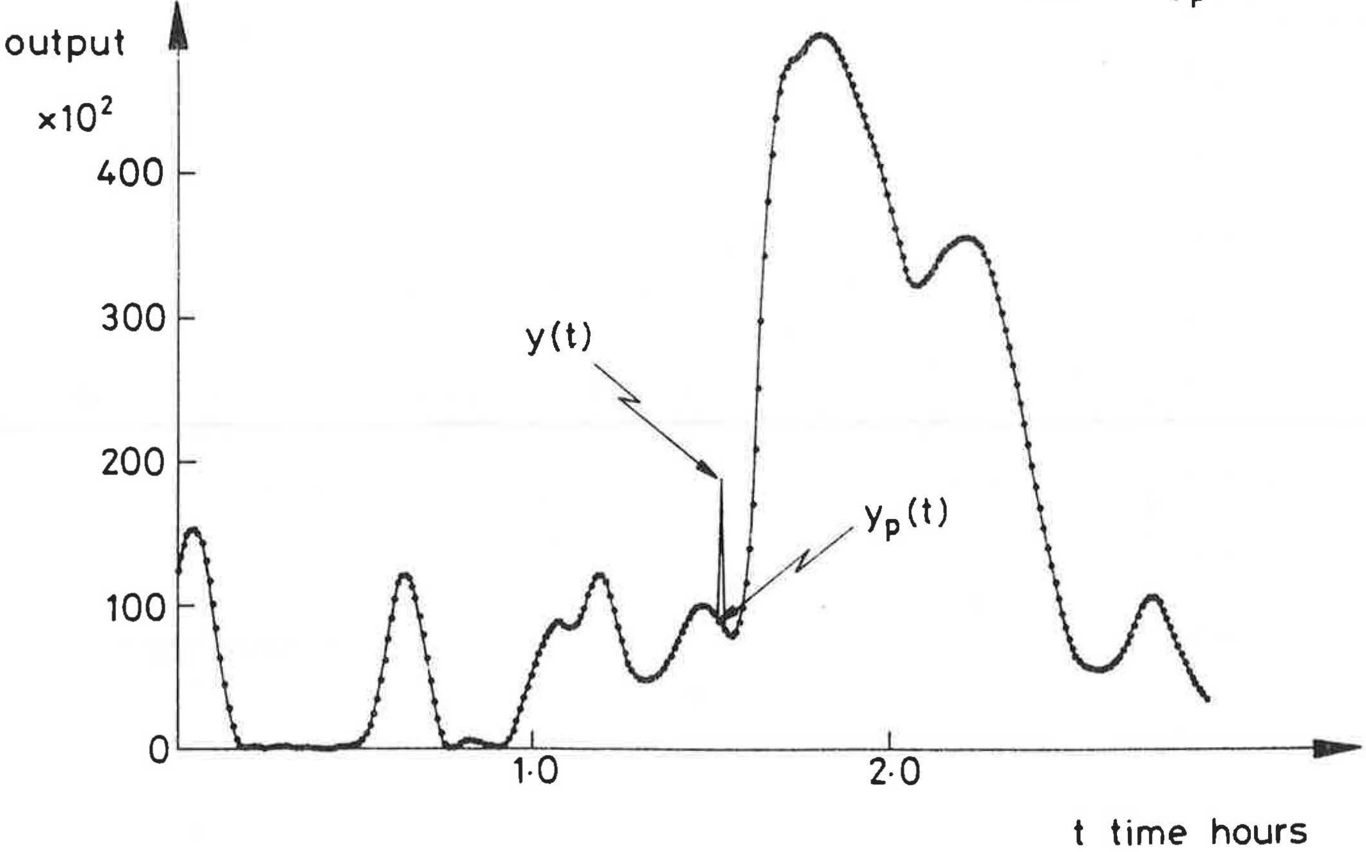
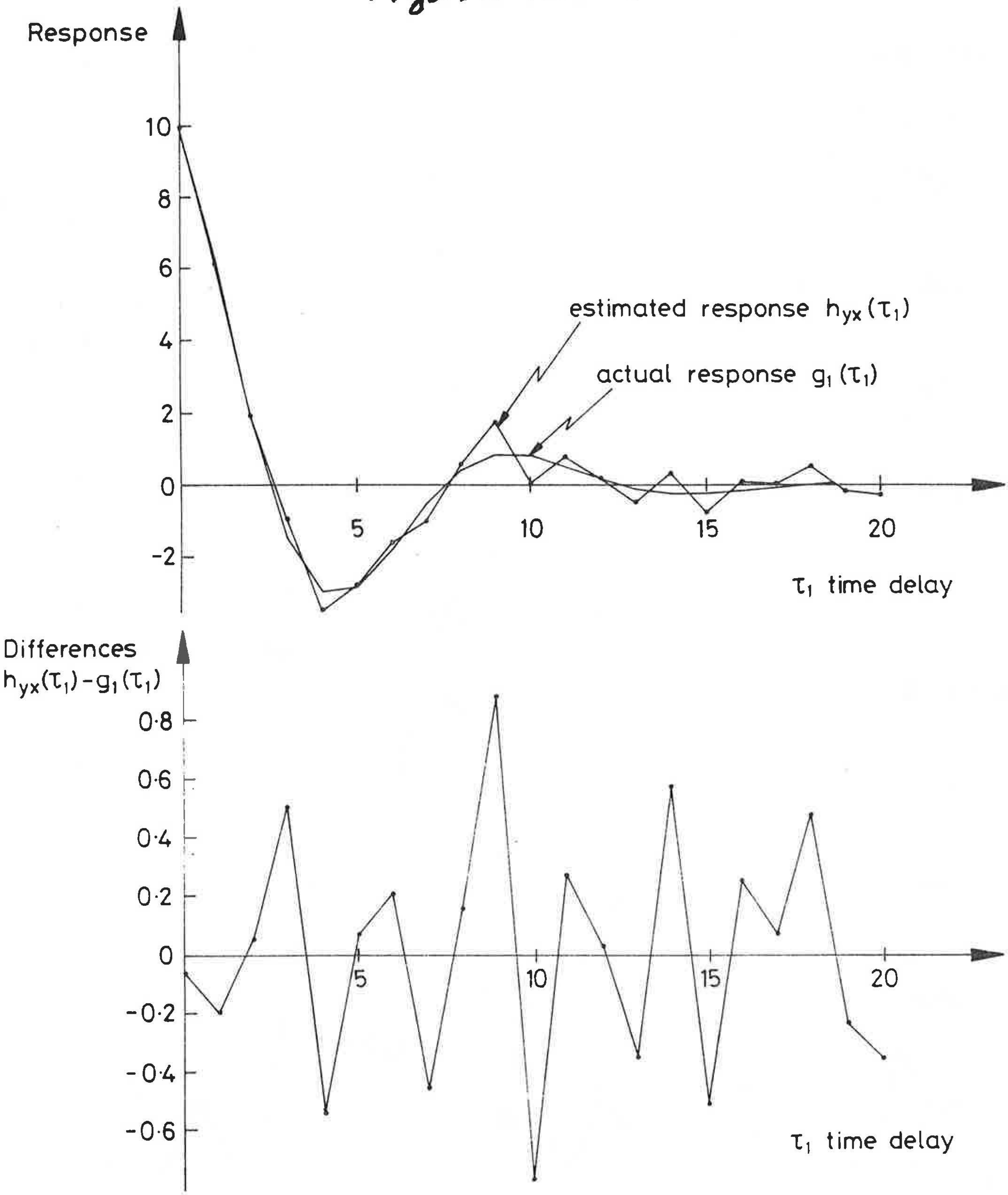

output
x10²
400
300
200
100
0
y(t)
yp(t)
1·0
2·0
t time hours

Figure 12 a.



Response
10
8
6
4
2
0
-2
5
10
15
20
estimated response $h_{yx}(\tau_1)$
actual response $g_1(\tau_1)$
τ_1 time delay
Differences
$h_{yx}(\tau_1) - g_1(\tau_1)$
0·8
0·6
0·4
0·2
0
-0·2
-0·4
-0·6
5
10
15
20
τ_1 time delay

Figure 12b.

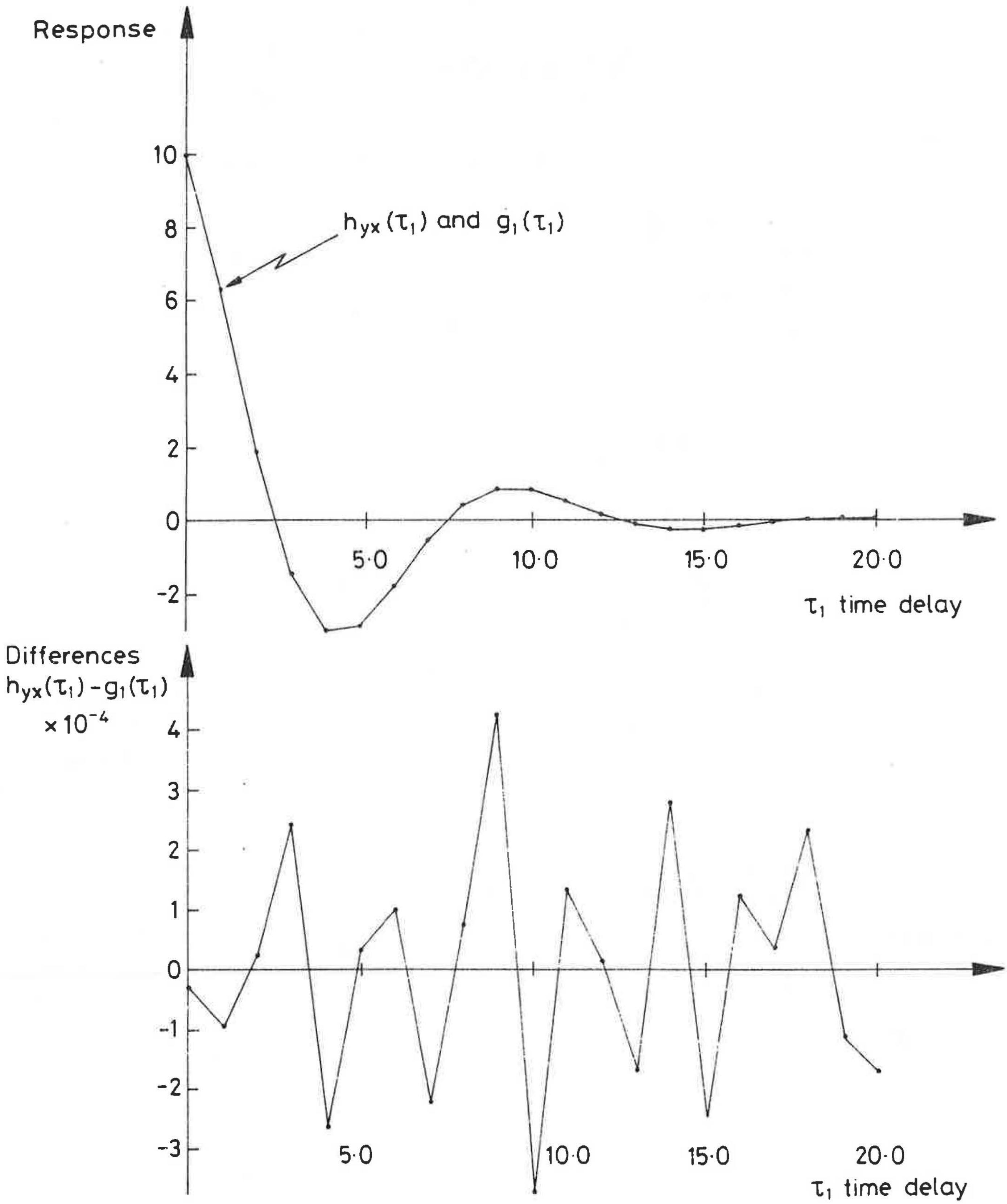

Response
10
8
6
4
2
0
-2
$h_{yx}(\tau_1)$ and $g_1(\tau_1)$
5·0
10·0
15·0
20·0
τ_1 time delay
Differences
$h_{yx}(\tau_1) - g_1(\tau_1)$
$\times 10^{-4}$
4
3
2
1
0
-1
-2
-3
5·0
10·0
15·0
20·0
τ_1 time delay

Figure 13a

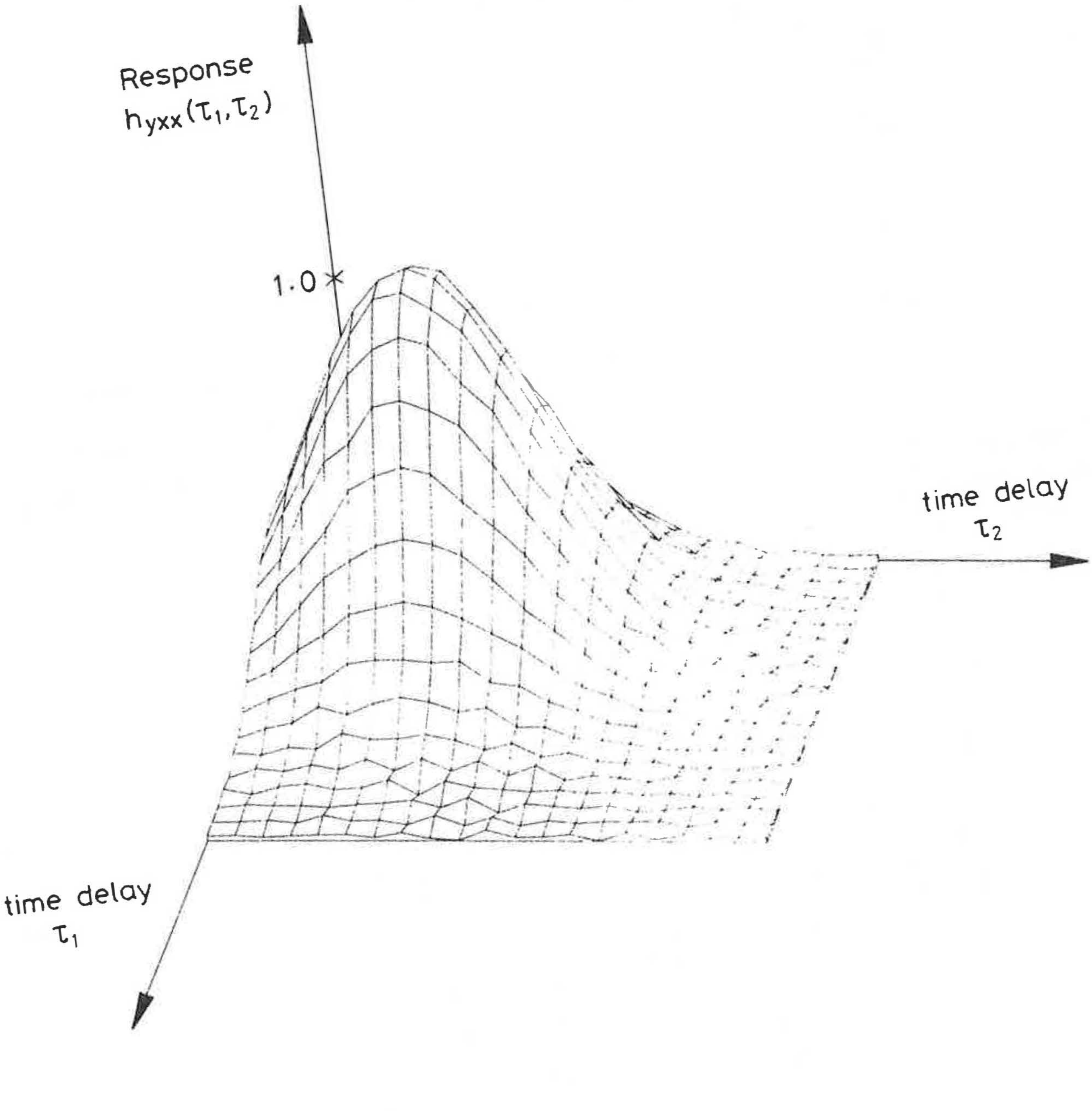


Figure 13 b